



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ
ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΣΗΜΑΤΩΝ, ΕΛΕΓΧΟΥ ΚΑΙ ΡΟΜΠΟΤΙΚΗΣ
ΕΡΓΑΣΤΗΡΙΟ ΟΡΑΣΗΣ ΥΠΟΛΟΓΙΣΤΩΝ, ΕΠΙΚΟΙΝΩΝΙΑΣ ΛΟΓΟΥ ΚΑΙ
ΕΠΕΞΕΡΓΑΣΙΑΣ ΣΗΜΑΤΩΝ

Θεωρία και Εφαρμογές Μερικών Διαφορικών
Εξισώσεων στην Όραση Υπολογιστών

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

του

Αναστάσιου Ν. Ρούσσου

Επιβλέπων: Πέτρος Α. Μαραγκός
Καθηγητής Ε.Μ.Π.

Αθήνα, Οκτώβριος 2005



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ
ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΣΗΜΑΤΩΝ, ΕΛΕΓΧΟΥ ΚΑΙ ΡΟΜΠΟΤΙΚΗΣ
ΕΡΓΑΣΤΗΡΙΟ ΟΡΑΣΗΣ ΥΠΟΛΟΓΙΣΤΩΝ, ΕΠΙΚΟΙΝΩΝΙΑΣ ΛΟΓΟΥ ΚΑΙ
ΕΠΕΞΕΡΓΑΣΙΑΣ ΣΗΜΑΤΩΝ

Θεωρία και Εφαρμογές Μερικών Διαφορικών
Εξισώσεων στην Όραση Υπολογιστών

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

του

Αναστάσιου Ν. Ρούσου

Επιβλέπων: Πέτρος Α. Μαραγκός
Καθηγητής Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή στις 11 Οκτωβρίου 2005.

.....
Πέτρος Μαραγκός
Καθηγητής Ε.Μ.Π.

.....
Στέφανος Κόλλιας
Καθηγητής Ε.Μ.Π.

.....
Σύρος Τζαφέστας
Καθηγητής Ε.Μ.Π.

Αθήνα, Οκτώβριος 2005.

.....
Αναστάσιος Ν. Ρούσσος

Διπλωματούχος Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών Ε.Μ.Π.

Copyright© Αναστάσιος Ν. Ρούσσος, 2005.

Με επιφύλαξη παντός δικαιώματος.All rights reserved.

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς το συγγραφέα.

Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν το συγγραφέα και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις επίσημες θέσεις του Εθνικού Μετσόβιου Πολυτεχνείου.

Περίληψη

Στην παρούσα διπλωματική εργασία, μελετάται η χρήση των Μερικών Διαφορικών Εξισώσεων (ΜΔΕ) σε τρία σημαντικά προβλήματα της Όρασης Υπολογιστών (ΟΥ) και της Επεξεργασίας Εικόνων: την Απομάκρυνση του Θορύβου, το Inpainting και την Παρεμβολή. Δίνεται ιδιαίτερη έμφαση στην περίπτωση διανυσματικών εικόνων. Στα πρώτα κεφάλαια, παρουσιάζουμε και αναλύουμε θεωρητικά πολλές από τις υπάρχουσες μεθόδους ΜΔΕ διάχυσης για τα παραπάνω προβλήματα. Επίσης, προτείνουμε μία νέα μέθοδο Inpainting διανυσματικών εικόνων (Γενικευμένη CDD μέθοδος) καθώς και μία νέα, μη γραμμική μέθοδο Παρεμβολής διανυσματικών εικόνων, που βασίζεται σε γενικές ΜΔΕ διάχυσης. Στη συνέχεια, διερευνάμε ορισμένα θεωρητικά θέματα, που ανακύπτουν από την ανάγκη αριθμητικής προσέγγισης και υλοποίησης των μεθόδων ΜΔΕ για την ΟΥ. Παραθέτουμε και αποδεικνύουμε κάποια νέα θεωρήματα, που αφορούν την ακρίβεια και την σύγκλιση της προσέγγισης της γενικής Ανισοτροπικής Διάχυσης Κατευθυνόμενων Λαπλασιανών με επαναληπτική τοπική συνέλιξη. Επίσης, διερευνάμε τη σχέση ανάμεσα στην συνεχή και στην εξαρχής διακριτή θεώρηση των φ-συναρτησιακών και αποδεικνύουμε ένα νέο θεώρημα, το οποίο αποσαφηνίζει τη σχέση αυτή. Στη συνέχεια, παρουσιάζουμε διεξοδικά, με έναν αρκετά ενοποιημένο τρόπο, την αριθμητική υλοποίηση πολλών από τις υπάρχουσες μεθόδους ΜΔΕ διάχυσης για την ΟΥ καθώς και των νέων μεθόδων που προτείνουμε. Επιπρόσθετα, προτείνουμε και υλοποιούμε μία νέα μέθοδο αξιολόγησης και σύγκρισης των διαφόρων ΜΔΕ διάχυσης, μέσω της εισαγωγής της έννοιας του Ολικού Φίλτρου της ΜΔΕ. Τέλος, παρουσιάζουμε παραδείγματα εφαρμογής των διαφόρων μεθόδων ΜΔΕ που μελετήσαμε. Συγκρίνουμε μεταξύ τους τις διάφορες μεθόδους, καταλήγοντας σε συμπεράσματα σχετικά με την αποτελεσματικότητά τους, ενώ παράλληλα αιτιολογούμε και θεωρητικά τα πειραματικά αποτελέσματα.

Λέξεις—κλειδιά

όραση υπολογιστών, επεξεργασία εικόνων, μερικές διαφορικές εξισώσεις, απομάκρυνση του θορύβου, inpainting, παρεμβολή, διανυσματικές εικόνες, διάχυ-

ση, συναρτησιακό, καμπύλες ελάχιστης μεταβολής, γενικευμένη καμπυλότητα, συνθήκη αναστρεψιμότητας, τοπική συνέλιξη, υλοποίηση, διακριτοποίηση, ολικό φίλτρο.

Abstract

This diploma thesis deals with the usage of Partial Differential Equations (PDEs) in three critical problems of Computer Vision (CV) and Image Processing: Denoising, Inpainting and Interpolation. Emphasis is put on the case where the input images are vectorial. In the first chapters, we present and theoretically analyze many of the existent diffusion PDE methods for the problems mentioned above. We also propose a new vectorial image Inpainting method (Generalized CDD method) and a new, nonlinear, vectorial image Interpolation method, based on general diffusion PDEs. Afterwards, we examine some theoretical issues that arise from the need of numerical approximation and implementation of PDE methods for CV. We mention and prove some new theorems, concerning the accuracy and convergence of the approximation of the general Oriented Laplacian based Anisotropic Diffusion with iterative local convolutions. Furthermore, we examine the relation between the continuous and totally discrete approach of phi-functionals and we prove a new theorem, which clarifies this relation. Afterwards, we present exhaustively the numerical implementation of many diffusion PDE methods for CV, with a sufficiently unified way. We also present the numerical implementation of the new proposed methods. Moreover, we propose and implement a new method for the evaluation and comparison of the diffusion PDEs. This method is based on the introduction of the concept of PDE Total Filter. Finally, we present application examples of the studied PDE methods. We compare the methods with each other and we conclude about their efficiency. We also give a theoretical justification of the experimental results.

Keywords

computer vision, image processing, partial differential equations, denoising, inpainting, interpolation, vectorial images, diffusion, functional, minimal variation curves, generalized curvature, reversibility condition, local convolution, implementation, discretization, total filter

Πρόλογος

Η Όραση Υπολογιστών (Computer Vision) ασχολείται με την εξαγωγή της συμβολικής πληροφορίας (περιγραφή αντικειμένων) που ενυπάρχει σε μία εικόνα ή σε ένα σύνολο εικόνων (βλ. π.χ. [47]). Απώτερος στόχος της είναι η κατασκευή συστημάτων, τα οποία θα αντιλαμβάνονται τον ορατό κόσμο με ανθρώπινη ευφυΐα. Βέβαια, η σύγχρονη Επιστήμη και Τεχνολογία απέχει ακόμα πολύ από τον στόχο αυτό. Η Όραση Υπολογιστών εμφανίστηκε για πρώτη φορά σαν επιστημονικό πεδίο την δεκαετία του 1960, σαν ένα παρακλάδι της Τεχνητής Νοημοσύνης και έκτοτε έχει γνωρίσει μία ραγδαία ανάπτυξη, συνδυάζοντας πολλά επιστημονικά πεδία.

Η Όραση Υπολογιστών παρουσιάζει μία μεγάλη ποικιλία εφαρμογών, μερικές εκ των οποίων είναι η Επεξεργασία πληροφορίας σε Εικόνες, η Βιοιατρική Έρευνα και Τεχνολογία, το Περιβάλλον και οι Γεωεπιστήμες, η Τηλεπισκόπηση, η Ρομποτική, η Επικοινωνία Ανθρώπου και Υπολογιστή, οι Ψηφιακές Εικαστικές Τέχνες καθώς και τα Ευφυή Συστήματα Ψηφιοποίησης και Ανάλυσης Βιβλίων.

Στην εργασία αυτή, ασχολούμαστε με τις εφαρμογές των Μερικών Διαφορικών Εξισώσεων (ΜΔΕ, Partial Differential Equation - PDEs) στην Όραση Υπολογιστών. Οι ΜΔΕ άρχισαν να χρησιμοποιούνται συστηματικά στην Όραση Υπολογιστών σχετικά πρόσφατα, κατά την δεκαετία του 1980. Από τότε μέχρι σήμερα, έχουν γνωρίσει μία αρκετά ευρεία ανάπτυξη στον τομέα αυτό, με πολλές εφαρμογές και ελπιδοφόρα αποτελέσματα. Οι ΜΔΕ αποτελούν ένα συνεχές εργαλείο για την επεξεργασία εικόνας, σε αντίθεση με τις περισσότερες κλασσικές μεθόδους, που είναι διακριτές.

Αυτή η συνεχής μεθοδολογία για τα προβλήματα της Όρασης Υπολογιστών παρουσιάζει το βασικό χαρακτηριστικό ότι μοντελοποιεί τις εικόνες σαν συνεχή σήματα. Έτσι, ο σχεδιασμός και η θεωρητική αξιολόγηση των διαφόρων μεθόδων επεξεργασίας γίνονται στον συνεχή χώρο. Η διακριτοποίηση επέρχεται μόνο κατά την αριθμητική αναπαράσταση και επίλυση των μοντέλων αυτών στον υπολογιστή, σε αντίθεση με την διακριτή μεθοδολογία, που εξαρχής μοντελοποιεί τις εικόνες σαν διακριτά σήματα.

Ο λόγος που έχει κάνει τις ΜΔΕ δημοφιλείς στην Όραση Υπολογιστών

είναι ότι η συνεχής αυτή μοντελοποίηση παρουσιάζει κάποια σημαντικά πλεονεκτήματα σε σχέση με τη διακριτή. Πρώτα από όλα, η ανάλυση των μοντέλων επεξεργασίας γίνεται πιο απλή και αφαιρετική, διότι αυτά δεν εξαρτώνται από κάποιο διακριτό πλέγμα ορισμού της εικόνας. Επιπρόσθετα, επειδή οι ΜΔΕ εμφανίζονται κατεξοχήν στη μαθηματική μοντελοποίηση φυσικών φαινομένων (π.χ. διάχυση θερμότητας, μεταφορά ύλης, ηλεκτρομαγνητισμός), γίνεται εφικτό να αναπτυχθούν μέθοδοι επεξεργασίας χρησιμοποιώντας τη μίμηση και παραλλαγή φυσικών διαδικασιών. Έτσι, μπορούν να χρησιμοποιηθούν χρήσιμες φυσικές έννοιες (όπως η ενέργεια, το πεδίο ροής, ο συντελεστής και ταυοστής διάχυσης) και να αξιοποιηθούν διάφορα έτοιμα θεωρητικά αποτελέσματα του κλάδου της Μαθηματικής Φυσικής, στον οποίο υπάρχει μία μακρόχρονη ερευνητική δραστηριότητα. Ακόμα, η διακριτοποίηση των συνεχών αυτών μοντέλων μπορεί να γίνει συστηματικά, με θεμελιωμένες μεθόδους της Αριθμητικής Ανάλυσης και έτσι να προκύψουν ακριβείς και ευσταθείς αριθμητικοί αλγόριθμοι επεξεργασίας. Η εξαρχής διακριτή μεθοδολογία θα ήταν δύσκολο να ανακαλύψει τέτοιους αλγορίθμους, αλλά ακόμα και σε αυτή την περίπτωση, η θεωρητική τους θεμελίωση δεν θα ήταν τόσο ισχυρή.

Το περιεχόμενο της παρούσας εργασίας

Το κείμενο της παρούσας διπλωματικής εργασίας είναι οργανωμένο σε 7 κεφάλαια:

Στα **Κεφάλαια 1-3**, παρουσιάζονται αναλυτικά και συγκρίνονται αρκετές από τις υπάρχουσες μεθόδους ΜΔΕ διάχυσης για τρία σημαντικά προβλήματα της Όρασης Υπολογιστών, την Βελτίωση και Απομάκρυνση του Θορύβου (Enhancement and Denoising), το Inpainting και την Παρεμβολή (Interpolation) αντίστοιχα. Δίνεται ιδιαίτερη έμφαση σε μεθόδους ειδικά σχεδιασμένες για διανυσματικές εικόνες (π.χ. έγχρωμες εικόνες). Στα Κεφάλαια 2,3 προτείνουμε επίσης και κάποιες νέες μεθόδους για τα αντίστοιχα προβλήματα.

Πιο συγκεκριμένα, στο **Κεφάλαιο 1**, παρουσιάζονται μέθοδοι ΜΔΕ για το πρόβλημα της Βελτίωσης. Η παρουσίαση αρχίζει από την πρωταρχική ιδέα της Μη Γραμμικής Διάχυσης των Perona-Malik και φτάνει μέχρι κάποιες πιο πρόσφατες μεθόδους, που χρησιμοποιούν, αντί του απλού συντελεστή διάχυσης, ταυοστή διάχυσης. Επίσης, παρουσιάζονται κάποιες μέθοδοι μεταβολών (variational methods) στις οποίες οι ΜΔΕ προκύπτουν από την ελαχιστοποίηση κάποιου κατάλληλα σχεδιασμένου συναρτησιακού, με πιο σημαντική την μέθοδο Ολικής Μεταβολής (Total Variation).

Στο **Κεφάλαιο 2** γίνεται μία εκτενής παρουσίαση των μεθόδων ΜΔΕ που έχουν αναπτυχθεί για το πρόβλημα του Inpainting. Η εφαρμογή των ΜΔΕ για την επίλυση του προβλήματος αυτού είναι μία πολύ πρόσφατη τάση, με ελ-

πιδοφόρα αποτελέσματα. Επιπρόσθετα, προτείνουμε κάποιες νέες ποσότητες (Καμπύλες Ελάχιστης Μεταβολής και Γενικευμένη Καμπυλότητα), οι οποίες είναι ενδεικτικές των γεωμετρικών δομών των διανυσματικών εικόνων και αποτελούν μία αποτελεσματική γενίκευση αντίστοιχων ποσοτήτων που ορίζονται για τις βαθμωτές εικόνες. Οι ποσότητες αυτές μπορούν να χρησιμεύσουν στην γενίκευση διαφόρων αξιόπιστων μεθόδων που έχουν αναπτυχθεί ειδικά για τις βαθμωτές εικόνες. Έτσι, στη συνέχεια χρησιμοποιούμε την Γενικευμένη Καμπυλότητα για την επέκταση στις διανυσματικές εικόνες μίας πρόσφατης μεθόδου Inpainting και προτείνουμε την Γενικευμένη CDD μέθοδο, η οποία καταφέρνει να βελτιώσει την απόδοση άλλων υπάρχουσών μεθόδων (όπως η Vectorial TV).

Στο **Κεφάλαιο 3** ασχολούμαστε με το πρόβλημα της Παρεμβολής, όπου παρουσιάζονται διάφορες υπάρχουσες μη γραμμικές μέθοδοι, που βασίζονται σε ΜΔΕ. Δίνεται ιδιαίτερη έμφαση στην προσέγγιση του προβλήματος με την Συνθήκη Αναστρεψιμότητας. Τέλος, συνδυάζοντας την προσέγγιση αυτή με την χρήση ΜΔΕ διάχυσης στο πρόβλημα αυτό, προτείνουμε μία νέα γενική μέθοδο Παρεμβολής διανυσματικών εικόνων. Κάνοντας κατάλληλες επιλογές, η μέθοδος που προτείνουμε καταφέρνει σε αρκετές περιπτώσεις να καταλήξει σε σημαντικά βελτιωμένα αποτελέσματα, σε σχέση με αρκετές από τις υπάρχουσες γραμμικές αλλά και μη γραμμικές μεθόδους.

Επίσης, στο **Κεφάλαιο 4** διερευνούμε ορισμένα θεωρητικά ζητήματα τα οποία σχετίζονται με την υλοποίηση των ΜΔΕ διάχυσης που εφαρμόζονται στα προβλήματα της Όρασης Υπολογιστών (ΟΥ).

Στην ενότητα 4.1, ασχολούμαστε με την προσέγγιση των ΜΔΕ Γενικής Ανισοτροπικής Διάχυσης Κατευθυνόμενων Λαπλασιανών με επαναληπτική τοπική συνέλιξη. Παραθέτουμε και αποδεικνύουμε κάποια νέα θεωρήματα, που αφορούν την ακρίβεια και την σύγκλιση της προσέγγισης αυτής, παρέχοντας έτσι μία καλύτερη θεωρητική θεμελίωσή της.

Στην ενότητα 4.2, μελετάμε κάποια θεωρητικά ζητήματα, που αφορούν την αριθμητική υλοποίηση των ϕ -συναρτησιακών (ϕ -functionals), μιας γενικής κατηγορίας συναρτησιακών που περιλαμβάνει πολλές από τις μεθόδους μεταβολών (variational methods) της ΟΥ. Αποδεικνύουμε ένα νέο θεώρημα, το οποίο αποσαφηνίζει τη σχέση ανάμεσα στην συνεχή και στην εξαρχής διακριτή θεώρηση των ϕ -συναρτησιακών. Δείχνουμε έτσι ότι οι δύο αυτές θεωρήσεις δεν είναι πάντα ισοδύναμες, αλλά αντίθετα χρειάζεται ειδική μέριμνα στον τρόπο διακριτοποίησης της αντίστοιχης ΜΔΕ για να επιτευχθεί η επιθυμητή ισοδυναμία.

Στο **Κεφάλαιο 5**, κάνουμε μία διεξοδική και ολοκληρωμένη παρουσίαση της αριθμητικής υλοποίησης που αναπτύξαμε για τις περισσότερες μεθόδους ΜΔΕ που παρουσιάσαμε στα κεφάλαια 1-3. Η υλοποίηση που επιλέξαμε είναι σχετικά απλή, αλλά παράλληλα αρκετά αξιόπιστη. Επίσης, προσπαθήσαμε, εντοπίζοντας τις ομοιότητες των μεθόδων, να κάνουμε μία όσο το δυνατόν πιο

ενοποιητική αντιμετώπιση των υλοποιήσεων των διαφόρων μεθόδων. Επιπρόσθετα, παρουσιάζουμε αναλυτικά τους αλγόριθμους των νέων μεθόδων που προτείναμε θεωρητικά στα κεφάλαια 2,3.

Σημειώνουμε τέλος ότι στην ενότητα 5.3, προτείνουμε μία νέα μέθοδο αξιολόγησης και σύγκρισης των διαφόρων ΜΔΕ διάχυσης. Η μέθοδος αυτή βασίζεται σε αυτό που ονομάσαμε Ολικό Φίλτρο της ΜΔΕ, το οποίο φανερώνει την συνολική (για όλες τις επαναλήψεις μαζί) προσαρμοζόμενη εξομάλυνση που επιτελεί μία μέθοδος ΜΔΕ (αφού διακριτοποιηθεί) στην διανυσματική εικόνα εισόδου. Επίσης, προτείνουμε έναν αλγόριθμο, ο οποίος κάνει εφικτό τον δύσκολο υπολογισμό της συνολικής αυτής εξομάλυνσης.

Στο **Κεφάλαιο 6**, παρουσιάζουμε και αναλύουμε παραδείγματα εφαρμογής των διαφόρων μεθόδων ΜΔΕ, τα οποία βασίστηκαν στις αριθμητικές υλοποιήσεις που παρουσιάσαμε στο Κεφάλαιο 5. Σαν εικόνες στα παραδείγματα, χρησιμοποιήσαμε κυρίως έγχρωμες εικόνες του πραγματικού κόσμου. Στα πειράματα αυτά, συγκρίνουμε μεταξύ τους τις διάφορες μεθόδους και εξάγουμε ενδιαφέροντα συμπεράσματα για την αποτελεσματικότητά τους. Παράλληλα όμως, αιτιολογούμε και θεωρητικά τα αποτελέσματα που παρατηρούμε πειραματικά.

Τέλος, στο **Κεφάλαιο 7**, συνοψίζουμε τα σημεία προσφοράς της εργασίας αυτής και παρουσιάζουμε συνοπτικά κάποιες κατευθύνσεις για μελλοντική έρευνα που προκύπτουν από την εργασία αυτή.

Ευχαριστίες

Θέλω να ευχαριστήσω θερμά τον Καθηγητή Πέτρο Μαραγκό, επιβλέποντα της παρούσας διπλωματικής εργασίας, για την σημαντική βοήθειά του και την εμπιστοσύνη που μου έδειξε στη διάρκεια εκπόνησης της εργασίας αυτής. Επίσης, θα ήθελα να τον ευχαριστήσω για την συμβολή του στην επιλογή του θέματος (δίνοντάς μου το περιθώριο να επιλέξω το θέμα εκείνο που μου φαινόταν πιο συναρπαστικό) καθώς και στην διαμόρφωση των ερευνητικών μου ενδιαφερόντων. Θέλω επίσης να ευχαριστήσω όλους τους υποψήφιους διδάκτορες του Εργαστηρίου Ορασης Υπολογιστών, Επικοινωνίας Λόγου και Επεξεργασίας Σημάτων, για την ιδιαίτερα πρόθυμη και άμεση βοήθειά τους σε ό,τι τους ζήτησα.

Ακόμα, θέλω να ευχαριστήσω τα άτομα εκείνα που στάθηκαν δίπλα μου, που μου συμπαραστάθηκαν στις δύσκολες στιγμές με αυταπάρνηση και υπομονή και που ομορφαίνουν τη ζωή μου. Επίσης, θα ήθελα ιδιαίτερα να ευχαριστήσω την Μαριάνθη.

Θα ήθελα να ευχαριστήσω τους γονείς μου, για τα αμέτρητα και ανεκτίμητα που μου έχουν προσφέρει καθώς και για τη βαθειά επιρροή τους στη διαμόρφωση της προσωπικότητάς μου, επιρροή η οποία ήταν καταλυτική για όσα έχω

μέχρι στιγμής πετύχει στη ζωή μου. Θέλω επίσης να τους ευχαριστήσω, εκτός των άλλων, για την στήριξη και ανεκτίμητη βοήθειά τους στην διάρκεια της εκπόνησης της εργασίας αυτής αλλά και γενικότερα των μέχρι τώρα σπουδών μου. Ακόμα, θέλω να ευχαριστήσω τον αδερφό μου, Γιάννη, που, από τότε που ήμασταν μικροί μέχρι σήμερα, με "προστατεύει" και με βοηθάει τόσο πολύ στην πορεία και τις επιλογές μου, κάνοντάς με να τον νιώθω τόσο κοντά μου. Τέλος, θα ήθελα να αφιερώσω την εργασία αυτή στην οικογένειά μου.

Περιεχόμενα

| | | |
|----------|--|-----------|
| 1 | Ανάλυση μεθόδων ΜΔΕ για τη Βελτίωση Εικόνων και την Απομάκρυνση του Θορύβου | 18 |
| 1.1 | Εισαγωγή | 18 |
| 1.1.1 | Χώροι Κλίμακας από τη Γραμμική Διάχυση Θερμότητας | 18 |
| 1.1.2 | Μη Γραμμικές τροποποιήσεις της Διάχυσης Θερμότητας και Μη Γραμμικοί Χώροι Κλίμακας | 20 |
| 1.1.3 | Διανυσματικές Εικόνες και Επεξεργασία τους | 21 |
| 1.1.4 | Η Μήτρα Di Zenzo | 22 |
| 1.2 | Η Μη Γραμμική Διάχυση των Perona-Malik | 24 |
| 1.3 | Ομαλοποιημένη (Regularized) Μη Γραμμική Διάχυση | 25 |
| 1.3.1 | Μία απλή γενίκευση στις διανυσματικές εικόνες | 26 |
| 1.4 | Η Μέθοδος της Ολικής Μεταβολής (Total Variation) | 26 |
| 1.5 | Έγχρωμη Ολική Μεταβολή (Color Total Variation) | 29 |
| 1.6 | Διανυσματικά ϕ -Συναρτησιακά (Vectorial ϕ -Functionals) | 30 |
| 1.7 | Μη Γραμμική Ανισοτροπική Διάχυση | 34 |
| 1.7.1 | Επέκταση στις Διανυσματικές Εικόνες | 36 |
| 1.8 | Η Ροή Beltrami (Beltrami Flow) | 37 |
| 1.9 | Διάχυση Διανυσματικών Εικόνων στην Κατεύθυνση θ_* | 39 |
| 1.10 | Η Ανισοτροπική Εξομάλυνση των Tschumperlé-Deriche | 41 |
| 2 | Ανάλυση μεθόδων ΜΔΕ για το πρόβλημα του Inpainting | 44 |
| 2.1 | Εισαγωγή | 44 |
| 2.2 | Το Μοντέλο BSCB: Μεταφορά Ομαλότητας (Smoothness Transportation) | 45 |
| 2.3 | Total Variation Inpainting | 46 |
| 2.3.1 | Τροποποιήσεις της μεθόδου Total Variation | 48 |
| 2.4 | Διάχυση Οδηγούμενη από την Καμυλότητα (Curvature Driven Diffusion) | 49 |
| 2.5 | Euler's Elastica Inpainting | 50 |
| 2.6 | Από κοινού παρεμβολή διανυσματικών πεδίων και επιπέδων του γκρι | 52 |

| | | |
|----------|--|------------|
| 2.7 | Inpainting με την Ανισοτροπική Εξομάλυνση των Tschumperlé-Deriche | 53 |
| 2.8 | Πρόταση νέων ποσοτήτων για τον χειρισμό των γεωμετρικών δομών των διανυσματικών εικόνων και εφαρμογή τους στο Inpainting | 54 |
| 2.8.1 | Καμπύλες Ελάχιστης Μεταβολής και Γενικευμένη Καμπυλότητα διανυσματικών εικόνων | 55 |
| 2.8.2 | Πρόταση μιας νέας μεθόδου: Γενικευμένη CDD για διανυσματικές εικόνες | 57 |
| 3 | Ανάλυση μεθόδων ΜΔΕ για την Παρεμβολή Εικόνων | 59 |
| 3.1 | Ανασκόπηση των μεθόδων Παρεμβολής | 59 |
| 3.1.1 | Προσέγγιση του προβλήματος με την Συνθήκη Αναστρεψιμότητας (Reversibility Condition) | 61 |
| 3.1.2 | Παρεμβολή Βασισμένη στην Ολική Μεταβολή (Total Variation) | 64 |
| 3.1.3 | Άλλες υπάρχουσες Μέθοδοι Παρεμβολής με ΜΔΕ | 65 |
| 3.2 | Πρόταση μιας νέας γενικής μεθόδου Παρεμβολής διανυσματικών εικόνων με ΜΔΕ | 67 |
| 4 | Θεωρητικά Θέματα Υλοποίησης ΜΔΕ για την ΟΥ | 71 |
| 4.1 | Μελέτη της Ισοδυναμίας της Γενικής Ανισοτροπικής Διάχυσης Κατευθυνόμενων Λαπλασιανών με Επαναληπτική Τοπική Συνέλιξη | 71 |
| 4.1.1 | Εισαγωγή | 71 |
| 4.1.2 | Παρουσίαση και Απόδειξη κάποιων νέων Θεωρημάτων | 74 |
| 4.2 | Αποσαφήνιση της σχέσης ανάμεσα στην συνεχή και στην εξαρχής διακριτή θεώρηση των ϕ -συναρτησιακών | 89 |
| 4.2.1 | Εισαγωγή | 89 |
| 4.2.2 | Ένα νέο θεώρημα για τη σχέση ανάμεσα στην συνεχή και στην εξαρχής διακριτή θεώρηση των ϕ -συναρτησιακών. | 92 |
| 5 | Περιγραφή της Υλοποίησης των μεθόδων ΜΔΕ για την ΟΥ | 108 |
| 5.1 | Υλοποίηση ΜΔΕ για την Βελτίωση διανυσματικών εικόνων | 108 |
| 5.1.1 | Προσέγγιση του όρου $\text{div}(D \nabla u_m)$ | 111 |
| 5.1.2 | Προσέγγιση του όρου $\text{trace}(TH_m)$ | 112 |
| 5.1.3 | Προσέγγιση της νόρμας $\ \nabla \mathbf{u}_\sigma\ $ | 115 |
| 5.1.4 | Υπολογισμός της $\ \nabla \mathbf{u}_\sigma\ $ στα ενδιάμεσα σημεία για πιο ακριβή προσέγγιση του $\text{div}(g \nabla u_m)$ | 116 |
| 5.1.5 | Προσέγγιση του δομικού τανυστή $J_\rho(\nabla \mathbf{u}_\sigma)$ | 117 |
| 5.1.6 | Υλοποίηση των εξομαλύνσεων της εικόνας \mathbf{u} και του τανυστή $J(\nabla \mathbf{u}_\sigma)$ | 118 |

| | | |
|----------|---|------------|
| 5.1.7 | Αλγόριθμοι για την υλοποίηση των ΜΔΕ | 119 |
| 5.1.8 | Επιπλέον Βήματα στην Υλοποίηση κάποιων Μεθόδων | 119 |
| 5.1.9 | Σημείωση σχετικά με την εκτίμηση των τιμών εκτός των ορίων της εικόνας | 123 |
| 5.2 | Υλοποίηση ΜΔΕ για το Inpainting εικόνων | 126 |
| 5.2.1 | Υλοποίηση για Inpainting των ΜΔΕ Βελτίωσης | 128 |
| 5.2.2 | Vectorial DTV Filter | 129 |
| 5.2.3 | Υλοποίηση Διάχυσης Οδηγούμενης από την Καμπυλό- τητα (CDD), για βαθμωτές εικόνες | 134 |
| 5.2.4 | Υλοποίηση της νέας μεθόδου που προτείνουμε: Γενικευ- μένη CDD, για διανυσματικές εικόνες | 135 |
| 5.3 | Πρόταση μιας νέας μεθόδου αξιολόγησης των ΜΔΕ διάχυσης: Υπολογισμός του Ολικού Φίλτρου της ΜΔΕ | 141 |
| 5.3.1 | Εισαγωγή | 141 |
| 5.3.2 | Αλγόριθμος υπολογισμού ενός πυρήνα συσχέτισης του Ολικού Φίλτρου | 144 |
| 5.4 | Υλοποίηση ΜΔΕ για την Παρεμβολή εικόνων | 147 |
| 5.4.1 | Υλοποίηση της μεθόδου Zero-Padding | 149 |
| 5.4.2 | Υπολογισμός της προβολής $P_{u_{0,s}}(\cdot)$ | 152 |
| 6 | Πειράματα και Συγκρίσεις | 156 |
| 6.1 | Πειράματα για το πρόβλημα Βελτίωσης Εικόνων | 156 |
| 6.1.1 | Παράδειγμα 1 | 156 |
| 6.1.2 | Παράδειγμα 2 | 167 |
| 6.1.3 | Παράδειγμα 3 | 170 |
| 6.2 | Πειράματα για το πρόβλημα Inpainting Εικόνων | 172 |
| 6.2.1 | Παράδειγμα 1 | 172 |
| 6.2.2 | Παράδειγμα 2 | 172 |
| 6.2.3 | Παράδειγμα 3 | 174 |
| 6.2.4 | Παράδειγμα 4 | 178 |
| 6.2.5 | Παράδειγμα 5 | 178 |
| 6.3 | Πειράματα για το πρόβλημα Παρεμβολής Εικόνων | 181 |
| 6.3.1 | Παράδειγμα 1 | 181 |
| 6.3.2 | Παράδειγμα 2 | 190 |
| 7 | Συμπεράσματα | 196 |
| 7.1 | Συμβολή της διπλωματικής εργασίας | 196 |
| 7.2 | Κατευθύνσεις για μελλοντική έρευνα | 197 |

Κατάλογος Σχημάτων

| | | |
|------|---|-----|
| 4.1 | Πορεία της αριθμητικής επαναληπτικής μεθόδου, για την μείωση του “μέτρου” $A[k]$, που προτείνουμε. | 87 |
| 4.2 | Ο πυρήνας που προκύπτει, με μειωμένο, σε σχέση με τον Gaussian, “μέτρο” A | 88 |
| 4.3 | Διάγραμμα 1 ^{ης} θεώρησης για την αριθμητική υλοποίηση των μεθόδων ϕ -συναρτησιακού | 90 |
| 4.4 | Διάγραμμα 2 ^{ης} θεώρησης για την αριθμητική υλοποίηση των μεθόδων ϕ -συναρτησιακού | 91 |
| 6.1 | Εικόνα Παράδειγματος 1 | 157 |
| 6.2 | Σύγκριση των μεθόδων βελτίωσης | 159 |
| 6.3 | Εφαρμογή της νέας μεθόδου του Ολικού Φίλτρου | 165 |
| 6.4 | Άλλη μία εφαρμογή της μεθόδου του Ολικού Φίλτρου | 166 |
| 6.5 | Εικόνα Παράδειγματος 2 | 168 |
| 6.6 | Σύγκριση των μεθόδων βελτίωσης | 169 |
| 6.7 | Παράδειγμα Βελτίωσης της Συνοχής, για δημιουργία ειδικών εφέ | 171 |
| 6.8 | Παράδειγμα Inpainting | 173 |
| 6.9 | Παράδειγμα Inpainting για απομάκρυνση κειμένου από έγχρωμη εικόνα | 175 |
| 6.10 | Παράδειγμα για τη σύγκριση της σύγκλισης στο πρόβλημα Inpainting | 176 |
| 6.11 | Σύγκριση της σύγκλισης των VTV και $Vectorial DTVF$ | 177 |
| 6.12 | Παράδειγμα Inpainting για απομάκρυνση κειμένου από έγχρωμη εικόνα | 179 |
| 6.13 | Παράδειγμα Inpainting, με περιοχή inpainting μεγάλου πάχους | 180 |
| 6.14 | Σύγκριση των μεθόδων παρεμβολής | 183 |
| 6.15 | Λεπτομέρειες των αποτελεσμάτων του σχήματος 6.14 | 185 |
| 6.16 | Άλλη μία σύγκριση των μεθόδων παρεμβολής | 192 |
| 6.17 | Λεπτομέρειες των αποτελεσμάτων του σχήματος 6.16 | 194 |

Κατάλογος Πινάκων

| | | |
|------|--|-----|
| 5.1 | Πυρήνας τοπικής συσχέτισης $k_{i,j}(\ell_1, \ell_2)$, για την προσέγγιση του $\text{div}(D \nabla u_m)$ | 113 |
| 5.2 | Πυρήνας τοπικής συσχέτισης $k_{i,j}(\ell_1, \ell_2)$, για την προσέγγιση του $\text{div}(g \nabla u_m)$ | 113 |
| 5.3 | Πυρήνας τοπικής συσχέτισης $k_{i,j}(\ell_1, \ell_2)$, για την προσέγγιση του $\text{trace}(TH_m)$ | 114 |
| 5.4 | Πυρήνας τοπικής συσχέτισης $k_{i,j}(\ell_1, \ell_2)$, για την προσέγγιση του $\text{div}(g \nabla u_m)$, με τιμές στα ενδιάμεσα σημεία | 116 |
| 5.5 | Γενικός αλγόριθμος για ΜΔΕ με βαθμωτό συντελεστή διάχυσης | 120 |
| 5.6 | Περαιτέρω επιλογές για κάθε μέθοδο, στον αλγόριθμο του πίνακα 5.5 | 120 |
| 5.7 | Γενικός αλγόριθμος για ΜΔΕ με τανυστή διάχυσης | 121 |
| 5.8 | Περαιτέρω επιλογές για κάθε μέθοδο, στον αλγόριθμο του πίνακα 5.7 | 122 |
| 5.9 | Πυρήνας τοπικής συσχέτισης $k_{i,j}(\ell_1, \ell_2)$ για την μέθοδο Vectorial DTV Filter | 131 |
| 5.10 | Αλγόριθμος που υλοποιήσαμε για το Vectorial Digital TV Filter | 132 |
| 5.11 | Αλγόριθμος της Γενικευμένης CDD Μεθόδου, που προτείνουμε για το Inpainting διανυσματικών εικόνων | 140 |
| 5.12 | Ο Αλγόριθμος που προτείνουμε, για τον Υπολογισμό του Πυρήνα του Ολικού Φίλτρου της ΜΔΕ | 148 |
| 5.13 | Αλγόριθμος της Γενικής Μεθόδου με ΜΔΕ διάχυσης, που προτείνουμε για την Παρεμβολή διανυσματικών εικόνων | 155 |
| 6.1 | Σηματοθρομβικοί λόγοι (SNR) για τα αποτελέσματα του σχήματος 6.2 | 164 |
| 6.2 | Σηματοθρομβικοί λόγοι (SNR) για τα αποτελέσματα του σχήματος 6.14 | 190 |

Κεφάλαιο 1

Ανάλυση μεθόδων ΜΔΕ για τη Βελτίωση Εικόνων και την Απομάκρυνση του Θορύβου

1.1 Εισαγωγή

Στο κεφάλαιο αυτό, παρουσιάζουμε διάφορες μεθόδους ΜΔΕ διάχυσης για το πρόβλημα της Βελτίωσης Εικόνων και της Απομάκρυνσης του Θορύβου (Image Enhancement and Noise Removal). Πριν από αυτό όμως, προβαίνουμε σε μία απαραίτητη σύντομη εισαγωγή στους χώρους κλίμακας καθώς και στις διανυσματικές εικόνες.

1.1.1 Χώροι Κλίμακας από τη Γραμμική Διάχυση Θερμότητας

Η χρήση ΜΔΕ στην Όραση Υπολογιστών συνδέεται άμεσα με την ανάλυση των εικόνων σε πολλαπλές κλίμακες (*multiscale*), που δημιουργεί τους χώρους κλίμακας (*scale-spaces*). Οι χώροι κλίμακας είναι κατάλληλοι για την επεξεργασία των εικόνων του φυσικού κόσμου, διότι συνήθως αυτές περιέχουν πληροφορίες σε πολλαπλές κλίμακες και τα περισσότερα χαρακτηριστικά τους εντοπίζονται σε μία κλίμακα. Έτσι, η ανάλυση σε πολλαπλές κλίμακες χρησιμεύει σε πολλά προβλήματα της Όρασης Υπολογιστών, όπως η απομάκρυνση του θορύβου, η ανίχνευση χαρακτηριστικών, η κατάτμηση και η εκτίμηση της κίνησης.

Οι χώροι κλίμακας προτάθηκαν και θεμελιώθηκαν θεωρητικά από τον Witkin [66]. Συγκεκριμένα, ο Witkin πρότεινε τη δημιουργία μιας οικογένειας εικόνων $u(x, y, s)$ από την αρχική, προς επεξεργασία, εικόνα $u^0(x, y)$, συνελίσ-

σοντάς την με έναν Gaussian πυρήνα $K_s(x, y)$, διακύμανσης s , δηλαδή:

$$u(x, y, s) = K_s(x, y) * u^0(x, y), \quad s > 0 \quad (1.1)$$

όπου:

$$K_s(x, y) = \frac{1}{2\pi s} \exp\left(-\frac{x^2 + y^2}{2s}\right)$$

Η $u(x, y, s)$ αποτελεί μια σειρά από εικόνες, που αποτελούν θολωμένες εκδοχές της αρχικής $u^0(x, y)$, με το θόλωμα να αυξάνει με την αύξηση της παραμέτρου s . Έτσι, η παράμετρος s δημιουργεί έναν χώρο κλίμακας, με την έννοια ότι για $s = 0$ προκύπτει η αρχική εικόνα και η μεγαλύτερη τιμή της s δημιουργεί εικόνες τραχύτερης κλίμακας.

Ο Koenderink [31] έκανε τη σημαντική παρατήρηση ότι η παραπάνω οικογένεια εικόνων $u(x, y, s)$ μπορεί να μοντελοποιηθεί μέσω της ΜΔΕ της γραμμικής διάχυσης θερμότητας. Η ΜΔΕ αυτή είναι η εξής:

$$\frac{\partial u}{\partial t} = \text{div}(\nabla u) = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \quad (1.2)$$

συμπληρωμένη από την αρχική συνθήκη $u(x, y, 0) = u^0(x, y)$. Όταν οι παράμετροι s και t συνδέονται μέσω της απλής σχέσης $t = s/2$, έχουμε πλήρη ισοδυναμία ανάμεσα στις δύο περιγραφές, δηλαδή, η εξίσωση (1.1) αποτελεί την ολοκληρωτική μορφή της λύσης της ΜΔΕ (1.2). Χάρη στη παρατήρηση αυτή, ο χώρος κλίμακας μπορεί να ειπωθεί σαν μία χρονική εξέλιξη, όπου στην αρχική χρονική στιγμή $t = 0$ έχουμε την εικόνα εισόδου και, καθώς ο χρόνος εξελίσσεται, οδηγούμαστε σε τραχύτερες κλίμακες.

Έχοντας εγκαθιδρύσει την παραπάνω ισοδυναμία, μπορεί να γίνει μία εξίσου σημαντική παρατήρηση, η οποία προκύπτει από τη Φυσική: Όταν η συνάρτηση u εκφράζει θερμοκρασία, η γραμμική διάχυση θερμότητας αποτελεί μία φυσική διαδικασία προς την ελαχιστοποίηση της ακόλουθης ενέργειας:

$$E[u] = \iint_{\Omega} \|\nabla u\|^2 dx \quad (1.3)$$

Το ίδιο βέβαια συνεχίζει να ισχύει και όταν η συνάρτηση u εκφράζει την φωτεινότητα μιας εικόνας, οπότε στο χώρο κλίμακας που προκύπτει, όσο μεγαλύτερη είναι η παράμετρος κλίμακας s , τόσο μικρότερη είναι η τιμή του συναρτησιακού $E[u]$.

Ο παραπάνω χώρος κλίμακας είναι γραμμικός, διότι προκύπτει από γραμμική συνέλιξη και η ΜΔΕ που τον περιγράφει είναι γραμμική. Λόγω της απλότητάς του όμως, παρουσιάζει κάποια σημαντικά μειονεκτήματα. Η συνέλιξη με Gaussian πυρήνες, πέραν των θετικών ότι μειώνει τον θόρυβο και απομακρύνει

χαρακτηριστικά μικρής κλίμακας, θολώνει και εξαπλώνει χωρικά κάποια σημαντικά χαρακτηριστικά της εικόνας (π.χ. ακμές) στην κλίμακα που μας ενδιαφέρει. Αυτό οφείλεται στο ότι η παραπάνω συνέλιξη αντιστοιχεί σε βαθυπερατό φιλτράρισμα και έτσι γρήγορα αποκόπτονται μεγάλες χωρικές συχνότητες της εικόνας, που είναι σημαντικές για τη διατήρηση των δομών (όπως οι ακμές) που παρουσιάζει μία εικόνα του φυσικού κόσμου.

1.1.2 Μη Γραμμικές τροποποιήσεις της Διάχυσης Θερμότητας και Μη Γραμμικοί Χώροι Κλίμακας

Όπως θα δούμε στις επόμενες ενότητες, έχουν προταθεί διάφορες μη γραμμικές παραλλαγές της βασικής ΜΔΕ διάχυσης θερμότητας (1.2), οι οποίες καταφέρνουν και αντιμετωπίζουν τα μειονεκτήματα που αναφέραμε προηγουμένως. Έτσι, δημιουργούνται μη γραμμικοί χώροι κλίμακας, οι οποίοι σέβονται τις δομές που είναι σημαντικές σε μία εικόνα. Για την ανάπτυξη τέτοιων μεθόδων, γίνεται μία απαραίτητη παραδοχή, που σχετίζεται με τις εικόνες που δημιουργούνται από τον φυσικό κόσμο. Σύμφωνα με αυτή, τα πιο σημαντικά χαρακτηριστικά μιας εικόνας είναι τα σχεδόν ομογενή τμήματα που την αποτελούν και τα σύνορά τους. Με τον όρο ομογενή τμήματα, εννοούμε τις χωρικές περιοχές, εντός των οποίων η εικόνα παρουσιάζει σχετικά μικρές μεταβολές (που μπορούν για παράδειγμα να αντιστοιχούν στο εσωτερικό ενός αντικειμένου του φυσικού κόσμου). Τα σύνορα των τμημάτων αυτών, δηλαδή οι ακμές, θεωρούμε ότι είναι αρκετά εντοπισμένα χωρικά και έτσι αντιστοιχούν σε συγκριτικά απότομες μεταβολές της εικόνας.

Με βάση την παραδοχή αυτή λοιπόν, οι μέθοδοι που δημιουργούν μη γραμμικούς χώρους κλίμακας (π.χ. για την απομάκρυνση του θορύβου) σχεδιάζονται με τέτοιο τρόπο, ώστε, καθώς προχωράμε σε μεγαλύτερες κλίμακες s , να επιτελείται μία εξομάλυνση εντός των ομογενών τμημάτων, ενώ οι ακμές να συντηρούνται για αρκετή ώρα αναλλοίωτες.

Έτσι, όσον αφορά την απομάκρυνση του θορύβου, που μας ενδιαφέρει εδώ, τα δύο βασικά κριτήρια για την αξιολόγηση μιας μεθόδου είναι το πόσο καλά καταφέρνει να συντηρήσει τις ακμές αλλά και το πόσο καλά καταφέρνει να εξομαλύνει τις ομογενείς περιοχές, ώστε να απομακρύνει τον θόρυβο.

Τέλος, όπως θα δούμε στη συνέχεια, με τους ίδιους στόχους, έχουν προταθεί διάφορες μέθοδοι που παραλλάσσουν με διαφορετικό τρόπο τη γραμμική διάχυση. Αντί να προτείνουν κατευθείαν μία παραλλαγή της ΜΔΕ (1.2), διατυπώνουν ένα πρόβλημα ελαχιστοποίησης, με παραλλαγή του συναρτησιακού (1.3), ώστε η νέα ελαχιστοποίηση να ταιριάζει περισσότερο στην μετεξέλιξη που επιθυμούμε για μια εικόνα.

1.1.3 Διανυσματικές Εικόνες και Επεξεργασία τους

Στην εργασία αυτή, θα ασχοληθούμε κυρίως με την επεξεργασία διανυσματικών εικόνων (*vectorial ή multivalued images*). Η ανάγκη επεξεργασίας εικόνων αυτής της κατηγορίας, οι οποίες σε κάθε σημείο λαμβάνουν διανυσματικές και όχι βαθμωτές τιμές, εμφανίζεται σε διάφορες εφαρμογές και για πολυάριθμους λόγους, όπως αναφέρουμε στη συνέχεια.

Πρώτα από όλα, για την αναπαράσταση μιας έγχρωμης εικόνας, η απλή αντιστοίχιση μίας τιμής σε κάθε σημείο δεν είναι επαρκής, μια και από αυτή μπορεί να προκύψει μόνο η φωτεινότητα. Αντίθετα, χρειάζονται τουλάχιστον 3 τιμές και έτσι ένας αξιόπιστος τρόπος αναπαράστασης του χρώματος είναι μέσω των συντελεστών του γραμμικού συνδυασμού τριών βασικών χρωμάτων. Έτσι, προκύπτει ένα διάνυσμα διάστασης 3 και η εικόνα μπορεί να ειπωθεί σαν μία απεικόνιση από το επίπεδο, στον 3Δ χώρο ($\mathbb{R}^2 \rightarrow \mathbb{R}^3$). Ένα συνηθισμένο παράδειγμα είναι η επιλογή των χρωμάτων κόκκινο, πράσινο και μπλε σαν βασικά χρώματα, οπότε προκύπτει ο (R, G, B) χώρος χρώματος. Για μία αναλυτική παρουσίαση των χώρων χρώματος παραπέμπουμε στο [47]. Η επεξεργασία έγχρωμων εικόνων είναι φυσικά χρήσιμη όταν αποδέκτης του αποτελέσματος είναι ο άνθρωπος (όπως για παράδειγμα στην βελτίωση μιας έγχρωμης φωτογραφίας), του οποίου το σύστημα όρασης βασίζεται στο χρώμα. Εκτός αυτού όμως, βρίσκει εφαρμογές επίσης στον εντοπισμό ακμών και την αναγνώριση, κι αυτό διότι το χρώμα παρέχει μία χρήσιμη επιπλέον πληροφορία. Για παράδειγμα, η επεξεργασία που βασίζεται στο χρώμα μπορεί να εντοπίσει ακόμα και τις ακμές που σχηματίζονται από αλλαγή του χρώματος αλλά όχι της φωτεινότητας.

Οι δορυφορικές εικόνες, που χρησιμοποιούνται στην τηλεπισκόπηση και τα γεωγραφικά συστήματα πληροφοριών αποτελούν άλλο ένα παράδειγμα διανυσματικών εικόνων. Οι εικόνες αυτές περιέχουν συνήθως μετρήσεις σε 7 φασματικές ζώνες, επομένως είναι διανυσματικές με διάνυσμα διάστασης 7. Επίσης, διανυσματικές εικόνες χρησιμοποιούνται για την αναπαράσταση πεδίων κατεύθυνσης, με διάφορες εφαρμογές, όπως η εκτίμηση της κίνησης.

Όσον αφορά τις μεθόδους επεξεργασίας των διανυσματικών εικόνων, ένας απλός τρόπος είναι η επεξεργασία κάθε καναλιού της εικόνας χωριστά (*channel by channel processing*). Με τον τρόπο αυτό, κάθε κανάλι αντιμετωπίζεται σαν μία βαθμωτή εικόνα, στην οποία εφαρμόζονται κατευθείαν οι μέθοδοι που έχουν αναπτυχθεί για τις βαθμωτές εικόνες. Σε πολλές περιπτώσεις, όμως, η προσέγγιση αυτή δεν είναι ικανοποιητική, αφού τα κανάλια της εικόνας δεν αλληλεπιδρούν μεταξύ τους. Έτσι, έχουν αναπτυχθεί διάφορες μέθοδοι, που επεξεργάζονται την εικόνα σαν διάνυσμα, εκμεταλλευόμενες έτσι με πιο αξιόπιστο τρόπο την υπάρχουσα διανυσματική πληροφορία. Η τελευταία κατηγορία περιλαμβάνει αρκετές ΜΔΕ, ειδικά σχεδιασμένες για διανυσματικές εικόνες, κάποιες από τις οποίες θα παρουσιάσουμε στην εργασία αυτή.

Για να γίνει εφικτή η επεξεργασία της εικόνας σαν διάνυσμα, ιδιαίτερα στην περίπτωση μη γραμμικών μεθόδων όπως οι ΜΔΕ, είναι απαραίτητο να οριστούν ποσότητες ενδεικτικές της μορφής των δομών της εικόνας σε κάθε σημείο. Έτσι, οι ποσότητες αυτές πρέπει να εκφράζουν την τοπική μεταβολή και την τοπική γεωμετρία της εικόνας. Για τον παραπάνω σκοπό, ο Di Zenzo [22] πρότεινε έναν από τους πλέον αξιόπιστους τρόπους περιγραφής της τοπικής μεταβολής μιας διανυσματικής εικόνας, όπως θα παρουσιάσουμε στην ακόλουθη ενότητα. Αναφέρουμε, ότι, όπως θα δούμε πιο κάτω, οι ιδέες αυτές του Di Zenzo χρησιμοποιούνται από τις περισσότερες μεθόδους ΜΔΕ για διανυσματικές εικόνες.

1.1.4 Η Μήτρα Di Zenzo

Όπως αναφέραμε στην προηγούμενη ενότητα, ο Di Zenzo [22], με σκοπό την περιγραφή της τοπικής μεταβολής μιας διανυσματικής εικόνας, πρότεινε την ακόλουθη ανάλυση:

Θεωρούμε το διανυσματικό διαφορικό της εικόνας:

$$d\mathbf{u} = \mathbf{u}_x dx + \mathbf{u}_y dy$$

Επομένως, από το τετράγωνο της ευκλείδειας νόρμας του, προκύπτει η ακόλουθη τετραγωνική μορφή:

$$\begin{aligned} \|d\mathbf{u}\|^2 &= d\mathbf{u}^T d\mathbf{u} \\ &= \|\mathbf{u}_x\|^2 (dx)^2 + 2\mathbf{u}_x^T \mathbf{u}_y dx dy + \|\mathbf{u}_y\|^2 (dy)^2 \\ &= d\mathbf{x}^T G d\mathbf{x} \end{aligned} \quad (1.4)$$

όπου $\mathbf{x} = (x, y)^T$ και η μήτρα G είναι ένας ταυστής που ορίζεται από την σχέση:

$$G = \sum_{i=1}^N \nabla u_i (\nabla u_i)^T \quad (1.5)$$

Από εδώ και στο εξής, για να αναφερθούμε στη μήτρα αυτή, θα χρησιμοποιούμε την ορολογία *μήτρα Di Zenzo*.

Έστω $G = \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix}$, με $g_{21} = g_{12}$, διότι η G είναι συμμετρική μήτρα. Τότε, οι ιδιοτιμές λ_+, λ_- και τα αντίστοιχα μοναδιαία ιδιοδιανύσματα $\boldsymbol{\theta}_+, \boldsymbol{\theta}_-$ της G δίνονται από τις σχέσεις:

$$\lambda_{\pm} = \frac{1}{2} \left(g_{11} + g_{22} \pm \sqrt{\Delta} \right) \quad (1.6)$$

$$\boldsymbol{\theta}_{\pm} \parallel \begin{bmatrix} 2g_{12} \\ g_{22} - g_{11} \pm \sqrt{\Delta} \end{bmatrix}, \quad \text{με: } \|\boldsymbol{\theta}_{\pm}\| = 1 \quad (1.7)$$

όπου: $\Delta = (g_{11} - g_{22})^2 + 4g_{12}^2$ και με “ \parallel ” συμβολίζουμε τη σχέση παραλληλίας δύο διανυσμάτων.

Πρέπει να σημειωθεί ότι, για καθένα από τα $\boldsymbol{\theta}_+$ και $\boldsymbol{\theta}_-$, οι εξισώσεις (1.7) έχουν δύο λύσεις: αν π.χ. το διάνυσμα $\boldsymbol{\theta}_+$ είναι μία λύση, τότε και το αντίθετο διάνυσμα $-\boldsymbol{\theta}_+$ είναι επίσης λύση. Μπορεί όμως να επιλεγεί αυθαίρετα οποιοδήποτε από τα δύο, διότι μας ενδιαφέρει μόνο η διεύθυνση του διανύσματος (δηλ. η ευθεία στην οποία ανήκει) και όχι η φορά του.

Για τις ιδιοτιμές ισχύει: $\lambda_+ \geq \lambda_- \geq 0$, ενώ για τα ιδιοδιανύσματα: $\boldsymbol{\theta}_+^T \boldsymbol{\theta}_- = 0$.

Αν θεωρήσουμε ένα μοναδιαίο διάνυσμα $\hat{\boldsymbol{v}} = (\cos \theta, \sin \theta)^T$, τότε η τετραγωνική μορφή:

$$T(\hat{\boldsymbol{v}}) = \hat{\boldsymbol{v}}^T G \hat{\boldsymbol{v}} \quad (1.8)$$

αποτελεί μέτρο του ρυθμού μεταβολής της διανυσματικής εικόνας \boldsymbol{u} στην κατεύθυνση του $\hat{\boldsymbol{v}}$. Η τετραγωνική αυτή μορφή παρουσιάζει μέγιστο στην κατεύθυνση του ιδιοδιανύσματος $\boldsymbol{\theta}_+$ και ελάχιστο στην κατεύθυνση του ιδιοδιανύσματος $\boldsymbol{\theta}_-$.

Από τα παραπάνω, συμπεραίνουμε ότι το ιδιοδιάνυσμα $\boldsymbol{\theta}_+$ μάς δείχνει την διεύθυνση της μέγιστης μεταβολής του διανύσματος \boldsymbol{u} της εικόνας, και η ιδιοτιμή λ_+ αποτελεί μέτρο της μεταβολής αυτής. Αντίστοιχα, το ιδιοδιάνυσμα $\boldsymbol{\theta}_-$ μάς δείχνει την διεύθυνση της ελάχιστης μεταβολής του \boldsymbol{u} (η οποία, όπως είναι φυσικό, είναι πάντα κάθετη στην διεύθυνση της μέγιστης μεταβολής) και η ιδιοτιμή λ_- αποτελεί μέτρο της ελάχιστης αυτής μεταβολής.

Αξίζει να σημειωθεί ότι, στην απλή περίπτωση όπου $N = 1$ (γκρίζες εικόνες), η μήτρα Di Zenzo γίνεται:

$$G = \nabla u (\nabla u)^T = \|\nabla u\|^2 \hat{\boldsymbol{n}} \hat{\boldsymbol{n}}^T, \quad \text{όπου: } \hat{\boldsymbol{n}} = \frac{\nabla u}{\|\nabla u\|}$$

και οι σχέσεις που δίνουν τις ιδιοτιμές και τα μοναδιαία ιδιοδιανύσματα της G απλοποιούνται στις ακόλουθες:

$$\begin{aligned} \lambda_+ &= \|\nabla u\|^2, & \lambda_- &= 0 \\ \boldsymbol{\theta}_+ &= \hat{\boldsymbol{n}}, & \boldsymbol{\theta}_- &= \hat{\boldsymbol{n}}^\perp \end{aligned}$$

Οι ιδιοτιμές λ_+, λ_- είναι ιδιαίτερα κατάλληλες για τον εντοπισμό των ασυνεχειών της εικόνας. Για τον σκοπό αυτό, μπορούν να οριστούν νόρμες διανυσματικής κλίσης $\mathcal{N} = f(\lambda_+, \lambda_-)$, οι οποίες να γενικεύουν για τις διανυσματικές εικόνες τη νόρμα $\|\nabla u\|$, που υπάρχει για τις γκρίζες εικόνες. Σε αρκετές υπάρχουσες μεθόδους βελτίωσης έγχρωμων εικόνων, χρησιμοποιούνται οι εξής νόρμες \mathcal{N} :

- $\mathcal{N} = \sqrt{\lambda_+}$, η οποία είναι η πιο απλή επιλογή και γενικεύει τη νόρμα $\|\nabla u\|$, υπό την έννοια της τιμής της μέγιστης μεταβολής.
- $\mathcal{N}_- = \sqrt{\lambda_+ - \lambda_-}$, η οποία λαμβάνει υπόψιν της και την ιδιοτιμή λ_- και έτσι μπορεί να εντοπίσει καλύτερα από την προηγούμενη μια διανυσματική ακμή.
- $\mathcal{N}_+ = \sqrt{\lambda_+ + \lambda_-}$, η οποία είναι πιο αξιόπιστη και από τις δύο προηγούμενες, αφού εντοπίζει επιπλέον, εκτός από τις ακμές, τις διανυσματικές γωνίες με ικανοποιητικό τρόπο. Κι αυτό διότι, στις διανυσματικές γωνίες (όπου δηλαδή τέμνονται δύο διανυσματικές ακμές της εικόνας με διαφορετική διεύθυνση), εμφανίζονται πολύ υψηλές μεταβολές σε δύο διαφορετικές διευθύνσεις. Το τελευταίο αντικατοπτρίζεται σε υψηλές μεταβολές ακόμα και στην διεύθυνση ελάχιστης μεταβολής θ_- , δηλαδή $\lambda_+ \approx \lambda_- \gg 0$. Επομένως, στις γωνίες, η νόρμα \mathcal{N}_+ λαμβάνει ακόμα μεγαλύτερη τιμή από τις απλές ακμές (όπου $\lambda_+ \gg \lambda_- \approx 0$), σε αντίθεση με την $\mathcal{N} = \sqrt{\lambda_+}$ (που λαμβάνει ίδιες τιμές με αντίστοιχα σημεία ακμών) και την \mathcal{N}_- (που λαμβάνει πολύ μικρές τιμές, συγχέοντας τις γωνίες με τις ομογενείς περιοχές).

Η νόρμα \mathcal{N}_+ συμβολίζεται επίσης και σαν $\|\nabla \mathbf{u}\|$, διότι, λόγω της (1.6), προκύπτει η εξής απλούστερη σχέση:

$$\mathcal{N}_+ = \sqrt{\lambda_+ + \lambda_-} = \sqrt{g_{11} + g_{22}} = \sqrt{\sum_{i=1}^N \|\nabla u_i\|^2} := \|\nabla \mathbf{u}\| \quad (1.9)$$

1.2 Η Μη Γραμμική Διάχυση των Perona-Malik

Για να ξεπεραστεί το μειονέκτημα του θολώματος των ακμών που συμβαίνει με τη γραμμική ιστροπική διάχυση (βλ. υποενότητα 1.1.1), οι Perona και Malik [40] πρότειναν μία μη γραμμική επέκτασή της:

$$\frac{\partial u}{\partial t} = \operatorname{div}(k \nabla u) \quad (1.10)$$

όπου ο συντελεστής διάχυσης k δεν είναι πλέον σταθερός και ίσος με 1, αλλά μεταβάλλεται (και μπορεί να εξαρτάται από οποιοδήποτε από τα $x, y, t, u, \nabla u$), με σκοπό να ευνοεί το θόλωμα της εικόνας εντός ομογενών περιοχών (intraregion smoothing) και να αποτρέπει την ανάμειξη μεταξύ γειτονικών ετερογενών περιοχών (interregion smoothing). Έτσι, το θόλωμα ελαττώνεται στις έντονες

ακμές, με αποτέλεσμα αυτές να διατηρούνται για αρκετή ώρα κατά την εξέλιξη της διάχυσης.

Τα παραπάνω μπορούν να επιτευχθούν εάν $k = g(\|\nabla u\|)$, οπότε η (1.10) γίνεται:

$$\frac{\partial u}{\partial t} = \operatorname{div}(g(\|\nabla u\|)\nabla u) \quad (1.11)$$

όπου η $g(s)$ είναι μία ομαλή, όχι αύξουσα συνάρτηση με:

$$g(0) = 1 \quad , \quad g(s) \geq 0 \quad \text{και} \quad g(s \rightarrow +\infty) = 0 \quad (1.12)$$

Με την επιλογή αυτή, η διάχυση θα είναι μικρή όταν η κλίση έχει μεγάλο μέτρο (π.χ. στις ακμές) και μεγάλη όταν η κλίση έχει μικρό μέτρο (π.χ. εντός των ομογενών περιοχών).

Στο [40], οι Perona και Malik χρησιμοποίησαν τις εξής δύο επιλογές για την συνάρτηση $g(s)$:

$$g(s) = \exp\left(-\left(\frac{s}{K}\right)^2\right) \quad (1.13)$$

και

$$g(s) = \frac{1}{1 + \left(\frac{s}{K}\right)^2} \quad (1.14)$$

Τα μειονεκτήματα της παραπάνω εξίσωσης διάχυσης είναι ότι σε αρκετές περιπτώσεις, ο υψίσυχνος θόρυβος διατηρείται, διότι δημιουργεί κλίσεις μεγάλου μέτρου $\|\nabla u\|$ στο σήμα, με αποτέλεσμα η διάχυση να επιβραδύνεται σημαντικά στα σημεία με θόρυβο. Επιπλέον, για συγκεκριμένες επιλογές της $g(s)$ (όπως οι (1.13) και (1.14)), η εξίσωση μπορεί να συμπεριφερθεί τοπικά σαν αντίστροφη διάχυση, με αποτέλεσμα το πρόβλημα να είναι γενικά μη καλώς ορισμένο και ευαίσθητο στις αρχικές συνθήκες (παρ' όλα τα αρκετά καλά πειραματικά αποτελέσματα της μεθόδου).

1.3 Ομαλοποιημένη (Regularized) Μη Γραμμική Διάχυση

Για να αντιμετωπίσουν τα παραπάνω προβλήματα της μεθόδου των Perona-Malik, οι Catte, Coll, Lions και Morel [10] πρότειναν ο συντελεστής διάχυσης να μην εξαρτάται από το μέτρο της κλίσης της ίδιας της u , αλλά μιας εξομαλυμένης εκδοχής της:

$$u_\sigma = K_\sigma * u$$

όπου K_σ είναι ένας βαθυπερατός πυρήνας συνέλιξης (συνήθως Gaussian) με τυπική απόκλιση σ .

Έτσι, προκύπτει η ακόλουθη ομαλοποιημένη (regularized) εκδοχή της εξίσωσης (1.11):

$$\frac{\partial u}{\partial t} = \operatorname{div} (g(\|\nabla u_\sigma\|)\nabla u) \quad (1.15)$$

Η παραπάνω εξίσωση πλεονεκτεί έναντι της (1.11) στο ότι το αντίστοιχο πρόβλημα είναι καλώς ορισμένο, για $\sigma > 0$, με αποδεδειγμένη την ύπαρξη και μοναδικότητα της λύσης. Επίσης, από πρακτικής σκοπιάς, η εξίσωση αυτή καταφέρνει να απομακρύνει το θόρυβο με μεγαλύτερη επιτυχία. Κι αυτό διότι η διάχυση στις ομογενείς περιοχές με θόρυβο είναι πλέον ταχύτερη, αφού ο συντελεστής διάχυσης $g(\|\nabla u_\sigma\|)$ είναι αναισθητος στο θόρυβο σε κλίμακες μικρότερες από σ , χάρη στην παραπάνω εξομάλυνση της u . Τέλος, όσον αφορά την αριθμητική προσέγγιση, όταν το σ είναι σχετικά μικρό, με την εξομάλυνση αυτή, επιτυγχάνεται μία πιο αξιόπιστη εκτίμηση της νόρμας $\|\nabla u_\sigma\|$, μειώνοντας σημαντικά την εξάρτηση από τη προσέγγιση που χρησιμοποιείται, ακόμα και στις περιοχές των ισχυρών ακμών.

1.3.1 Μία απλή γενίκευση στις διανυσματικές εικόνες

Η ΜΔΕ (1.15), όπως είδαμε, αναφέρεται σε βαθμωτές εικόνες. Θα ήταν όμως επιθυμητό η μέθοδος αυτή να μπορεί να εφαρμοστεί με αντίστοιχη συμπεριφορά και στις διανυσματικές εικόνες. Ένας απλός τρόπος για να πετύχουμε το παραπάνω είναι να χρησιμοποιήσουμε τη νόρμα $\|\nabla \mathbf{u}_\sigma\|$, που όπως είδαμε (βλ. σχέση (1.9)), αποτελεί μία αξιόπιστη γενίκευση στις διανυσματικές εικόνες της νόρμας $\|\nabla u_\sigma\|$. Όπως θα δούμε παρακάτω, εντελώς παρόμοιες γενικεύσεις έχουν προταθεί στη βιβλιογραφία για διάφορες άλλες μεθόδους ΜΔΕ.

Έτσι, χρησιμοποιώντας την $\|\nabla \mathbf{u}_\sigma\|$, προτείνουμε την εξής άμεση γενίκευση της ΜΔΕ (1.15) στις διανυσματικές εικόνες:

$$\frac{\partial u_i}{\partial t} = \operatorname{div} (g(\|\nabla \mathbf{u}_\sigma\|)\nabla u_i) \quad , \quad i = 1, \dots, N \quad (1.16)$$

όπου $\mathbf{u}_\sigma = K_\sigma * \mathbf{u}$.

1.4 Η Μέθοδος της Ολικής Μεταβολής (Total Variation)

Έστω μια γκριζα εικόνα $u_{cl}(\mathbf{x})$, ορισμένη στο κλειστό χωρίο $\Omega \subset \mathbb{R}^2$, στην οποία έχει προστεθεί θόρυβος $n(\mathbf{x})$, μηδενικής μέσης τιμής και διασποράς σ^2 , οπότε προκύπτει η εικόνα:

$$u^0(\mathbf{x}) = u_{cl}(\mathbf{x}) + n(\mathbf{x}) \quad , \quad \mathbf{x} \in \Omega$$

Για την αποκατάσταση της καθαρής εικόνας, οι Rudin, Osher και Fatemi [41] πρότειναν τη λύση του ακόλουθου προβλήματος ελαχιστοποίησης:

$$\min_u \text{TV}[u] = \iint_{\Omega} \|\nabla u\| \, d\mathbf{x} \quad (1.17)$$

υπό τους περιορισμούς:

$$\iint_{\Omega} (u - u^0) \, d\mathbf{x} = 0 \quad \text{και} \quad \frac{1}{|\Omega|} \iint_{\Omega} (u - u^0)^2 \, d\mathbf{x} = \sigma^2 \quad (1.18)$$

Το συναρτησιακό $\text{TV}[u]$ της εξίσωσης (1.17) ονομάζεται *Ολική Μεταβολή (Total Variation - TV)*. Χρησιμοποιώντας Λογισμό των Μεταβολών (Calculus of Variations), προκύπτουν οι Euler-Lagrange Εξισώσεις για την ελαχιστοποίηση του παραπάνω συναρτησιακού. Εάν δημιουργήσουμε μια χρονική εξέλιξη με τη μέθοδο της ταχύτερης καθόδου (gradient steepest descent flow) για την εύρεση της λύσης των εξισώσεων αυτών, προκύπτει η ακόλουθη ΜΔΕ:

$$\frac{\partial u}{\partial t} = \text{div} \left(\frac{\nabla u}{\|\nabla u\|} \right) - \lambda(t)(u - u^0) \quad (1.19)$$

$$u(x, y, 0) = u^0 + \sigma \frac{u^0 - \bar{u}^0}{\sigma_{u^0}} \quad (1.20)$$

$$\frac{\partial u}{\partial n} = 0, \text{ στο } \partial\Omega \quad (1.21)$$

$$\text{και } \lambda(t) = -\frac{1}{2\sigma} \iint_{\Omega} \left[\|\nabla u\| - \left(\frac{\nabla u \cdot \nabla u^0}{\|\nabla u\|} \right) \right] \, d\mathbf{x} \quad (1.22)$$

όπου σ_{u^0} είναι η τυπική απόκλιση της u^0 . Ο πολλαπλασιαστής Lagrange $\lambda(t)$ βρέθηκε από τις Euler-Lagrange Εξισώσεις (κατάσταση ισορροπίας).

Όπως αποδεικνύεται στο [11], το πρόβλημα ελαχιστοποίησης (1.17,1.18) είναι ισοδύναμο με το ακόλουθο, χωρίς περιορισμούς, πρόβλημα, που επίσης προτάθηκε για πρώτη φορά στο [41] :

$$\min_u \iint_{\Omega} \|\nabla u\| \, d\mathbf{x} + \frac{\lambda}{2} \iint_{\Omega} (u - u^0)^2 \, d\mathbf{x} \quad (1.23)$$

όπου ο πολλαπλασιαστής Lagrange λ θεωρείται σταθερός (για το λ χρησιμοποιείται συχνά στη βιβλιογραφία η ονομασία *όρος πιστότητας - fidelity term*). Για κάθε τιμή της τυπικής απόκλισης σ στον περιορισμό (1.18), υπάρχει κατάλληλη επιλογή της σταθεράς λ στο πρόβλημα (1.23), ώστε αυτό να έχει την ίδια μοναδική λύση με το πρόβλημα (1.17,1.18).

Αντίστοιχα με προηγούμενως, από τις Euler-Lagrange Εξισώσεις και χρησιμοποιώντας τη μέθοδο της ταχύτερης καθόδου (gradient steepest descent flow), προκύπτει η ακόλουθη ΜΔΕ:

$$\frac{\partial u}{\partial t} = \operatorname{div} \left(\frac{\nabla u}{\|\nabla u\|} \right) - \lambda(u - u^0) \quad (1.24)$$

$$u(x, y, 0) = u^0 \quad (1.25)$$

$$\text{και } \frac{\partial u}{\partial n} = 0, \text{ στο } \partial\Omega \quad (1.26)$$

Η διαφορά της προηγούμενης εξίσωσης (1.24) με την (1.19) είναι ότι τώρα ο πολλαπλασιαστής λ είναι δεδομένος και σταθερός καθόλη τη χρονική εξέλιξη. Ο λ πρέπει να επιλέγεται εκ των προτέρων, αντί της τυπικής απόκλισης σ του θορύβου (που ούτως ή άλλως στις περισσότερες περιπτώσεις δεν είναι γνωστή) και εκφράζει την ισορροπία ανάμεσα στο μέγεθος της ομοιότητας της u με την αρχική θορυβώδη u^0 και στην έκταση της διαδικασίας αποθορυβοποίησης της u^0 . Όσο μεγαλύτερο είναι δηλαδή το λ , τόσο λιγότερο μπορεί η u να απομακρυνθεί από την u^0 .

Ιδιαίτερο ενδιαφέρον παρουσιάζει η απλή περίπτωση της (1.23) όπου $\lambda = 0$, οπότε το πρόβλημα ελαχιστοποίησης γίνεται:

$$\min_u \iint_{\Omega} \|\nabla u\| \, d\mathbf{x} \quad (1.27)$$

(χωρίς περιορισμούς) και η αντίστοιχη ΜΔΕ (1.24) γίνεται:

$$\frac{\partial u}{\partial t} = \operatorname{div} \left(\frac{\nabla u}{\|\nabla u\|} \right) \quad (1.28)$$

Κατά τη χρονική εξέλιξη, η Ολική Μεταβολή (TV) της u συνεχώς ελαττώνεται ενώ παράλληλα η απόστασή της από την u^0 συνεχώς αυξάνεται, μέχρι που στην κατάσταση ισορροπίας προκύπτει μία υπερβολικά εξομαλυμένη εικόνα. Εδώ, λοιπόν, δεν μας ενδιαφέρει όπως προηγούμενως η κατάσταση ισορροπίας, αλλά η όλη χρονική ροή. Έτσι, εκτός από την απλότητά της, η περίπτωση αυτή παρουσιάζει το πλεονέκτημα σε σχέση με τα προηγούμενα ότι ο χρόνος t αποτελεί Χώρο Κλίμακας.

Αξίζει να σημειωθεί, τέλος, ότι ο δεύτερος όρος της εξίσωσης (1.28) (TV PDE) αποτελεί την καμπυλότητα κ των ισούψων γραμμών (level lines) της συνάρτησης u . Έτσι, μπορούμε να συσχετίσουμε την TV PDE με την *Κίνηση Μέσης Καμπυλότητας* (Mean Curvature Motion - MCM), που περιγράφεται από την εξίσωση:

$$\text{MCM: } \frac{\partial u}{\partial t} = \kappa \|\nabla u\| = \frac{\partial^2 u}{\partial \xi^2} \quad (1.29)$$

όπου $\xi = \left(\frac{\nabla u}{\|\nabla u\|} \right)^\perp$, δηλαδή το μοναδιαίο διάνυσμα, εφαπτομενικό στην ισοϋψή καμπύλη που διέρχεται από το κάθε σημείο. Για την συσχέτιση με την MCM, γράφουμε την TV PDE (1.28) ως εξής:

$$\text{TV: (1.28)} \Rightarrow \frac{\partial u}{\partial t} = \frac{1}{\|\nabla u\|} \cdot \kappa \|\nabla u\| = \frac{1}{\|\nabla u\|} \cdot \frac{\partial^2 u}{\partial \xi^2} \quad (1.30)$$

Παρατηρούμε λοιπόν ότι η TV PDE μπορεί να ειδωθεί σαν μία σταθμισμένη εκδοχή της MCM. Όπως φαίνεται από τις (1.29,1.30), και στις δύο περιπτώσεις εμφανίζεται η 2^η κατευθυνόμενη παράγωγος ως προς ξ , οπότε η διάχυση γίνεται μόνο κατά την διεύθυνση των ισοϋψών γραμμών. Η διαφορά έγκειται στο ότι, ενώ στην MCM η ποσότητα της κατευθυνόμενης διάχυσης είναι παντού η ίδια, στην TV PDE, η διάχυση, λόγω του όρου $\frac{1}{\|\nabla u\|}$, είναι μεγάλη στις ομογενείς περιοχές και μικρή κοντά στις ακμές.

1.5 Έγχρωμη Ολική Μεταβολή (Color Total Variation)

Η μέθοδος της Ολικής Μεταβολής (Total Variation) της προηγούμενης ενότητας γενικεύθηκε για τις έγχρωμες (και γενικά διανυσματικές) εικόνες από τους Blomgren και Chan [9]. Οι συγγραφείς γενίκευσαν το συναρτησιακό $\text{TV}[u]$ (βλ. εξίσωση (1.17)) στο :

$$\text{CTV}[\mathbf{u}] = \left\{ \sum_{i=1}^N \left(\iint_{\Omega} \|\nabla u_i\| \, d\mathbf{x} \right)^2 \right\}^{\frac{1}{2}} = \left\{ \sum_{i=1}^N (\text{TV}[u_i])^2 \right\}^{\frac{1}{2}} \quad (1.31)$$

όπου $\text{TV}[u_i]$ είναι η (απλή) Ολική Μεταβολή (TV) μόνο του καναλιού u_i και ορίζεται στην εξίσωση (1.17).

Το συναρτησιακό $\text{CTV}[\mathbf{u}]$ ονομάζεται *Έγχρωμη Ολική Μεταβολή (Color Total Variation - CTV)*.

Αν επικεντρώσουμε το ενδιαφέρον μας στην απλή περίπτωση της ελαχιστοποίησης της $\text{CTV}[\mathbf{u}]$ χωρίς περιορισμούς (γενίκευση της (1.27)) και εφαρμόσουμε τη μέθοδο της ταχύτερης καθόδου για την ελαχιστοποίηση, προκύπτει η ακόλουθη ΜΔΕ (γενίκευση της (1.28)):

$$\frac{\partial u_i}{\partial t} = \frac{\text{TV}[u_i]}{\text{CTV}[\mathbf{u}]} \text{div} \left(\frac{\nabla u_i}{\|\nabla u_i\|} \right), \quad i = 1, \dots, N \quad (1.32)$$

Παρατηρούμε ότι, για κάθε κανάλι u_i , η διάχυση είναι παρόμοια με την απλή TV PDE (1.28), ανεξάρτητη δηλαδή από τα άλλα κανάλια, με τη μόνη διαφορά

του πολλαπλασιαστικού όρου:

$$A_i = \frac{\text{TV}[u_i]}{\text{CTV}[\mathbf{u}]} = \left\{ \frac{(\text{TV}[u_i])^2}{\sum_{j=1}^N (\text{TV}[u_j])^2} \right\}^{\frac{1}{2}}, \quad i = 1, \dots, N \quad (1.33)$$

Ο όρος αυτός εκφράζει δηλαδή τη σύζευξη ανάμεσα στα κανάλια της εικόνας και είναι τόσο μεγαλύτερος όσο μεγαλύτερη είναι η συμμετοχή του συγκεκριμένου καναλιού στην CTV. Βλέπουμε επίσης ότι ο όρος αυτός, σε κάθε χρονική στιγμή, είναι χωρικά σταθερός για κάθε κανάλι.

Συμπεραίνουμε λοιπόν ότι η μέθοδος παρουσιάζει το μειονέκτημα ότι τα κανάλια της εικόνας αλληλεπιδρούν μόνο συνολικά σε κάθε χρονική στιγμή και δεν υπάρχει καμμία τοπική αλληλεπίδραση, που να εξαρτάται από τα ιδιαίτερα χαρακτηριστικά της περιοχής του κάθε σημείου της εικόνας (αν βρίσκεται π.χ. σε επίπεδη περιοχή ή κοντά σε ακμή). Αυτό έχει σαν αποτέλεσμα σε πολλές περιπτώσεις, η διάχυση να θολώνει θορυβώδεις διανυσματικές ακμές, που δεν εμφανίζονται καθαρά σε κάθε κανάλι χωριστά.

1.6 Διανυσματικά ϕ -Συναρτησιακά (Vectorial ϕ -Functionals)

Μία άλλη γενίκευση της Ολικής Μεταβολής (1.17) για τις διανυσματικές εικόνες, που ξεπερνάει το μειονέκτημα της Έγχρωμης Ολικής Μεταβολής που είδαμε στην προηγούμενη ενότητα, είναι αυτή που προκύπτει αν στην σχέση (1.17) αντικαταστήσουμε τη νόρμα $\|\nabla u\|$ από την γενίκευσή της $\|\nabla \mathbf{u}\|$ (βλ. ενότητα 1.1.4), οπότε προκύπτει το συναρτησιακό:

$$\text{VTV}[\mathbf{u}] = \iint_{\Omega} \|\nabla \mathbf{u}\| \, d\mathbf{x} = \iint_{\Omega} \sqrt{\sum_{i=1}^N \|\nabla u_i\|^2} \, d\mathbf{x} \quad (1.34)$$

Για το συναρτησιακό αυτό, χρησιμοποιείται συχνά η ορολογία *Διανυσματική Ολική Μεταβολή* (Vectorial Total Variation - VTV). Συγκρίνοντας την VTV με την CTV ((1.31)), βλέπουμε ότι, ενώ στην CTV η άθροιση των όρων που εκφράζουν τη συμμετοχή του κάθε καναλιού γίνεται συνολικά, για όλη την εικόνα, στην VTV η άθροιση των αντίστοιχων όρων (που είναι βέβαια διαφορετικοί) γίνεται σε κάθε σημείο χωριστά. Έτσι, μέσω της νόρμας $\|\nabla \mathbf{u}\|$, υπάρχει τώρα η επιθυμητή τοπική αλληλεπίδραση ανάμεσα στα κανάλια (κάτι που θα φανεί και πιο κάτω, από την προκύπτουσα ΜΔΕ).

Αν αντί της $\|\nabla \mathbf{u}\|$ χρησιμοποιηθεί μία συνάρτηση $\phi(\|\nabla \mathbf{u}\|)$ (βλ. π.χ. [8]),

το συναρτησιακό (1.34) μπορεί να επεκταθεί ακόμα περισσότερο στο:

$$J[\mathbf{u}] = \iint_{\Omega} \phi(\|\nabla \mathbf{u}\|) \, d\mathbf{x} \quad (1.35)$$

Η παραπάνω γενική μορφή συναρτησιακού μελετάται στα [52, 55], ενώ παρόμοιες γενικές μορφές μελετώνται στα [21, 58, 33, 32] (για βαθμωτές εικόνες). Για το γενικό συναρτησιακό (1.35), θα χρησιμοποιήσουμε στην εργασία αυτή, την ορολογία (που χρησιμοποιείται π.χ. στο [52]) *Διανυσματικό ϕ -Συναρτησιακό (Vectorial ϕ -Functional)*.

Η συνάρτηση $\phi(s)$ πρέπει να έχει τέτοια χαρακτηριστικά, ώστε η ελαχιστοποίηση του συναρτησιακού να οδηγεί στην επιθυμητή βελτίωση της εικόνας. Έτσι, πρέπει, όπως και η απλή $\phi(s) = s$ στην Ολική Μεταβολή (TV), να επιφέρει κόστος που να αυξάνεται όσο εντονότερες είναι οι διακυμάνσεις της \mathbf{u} σε κάθε σημείο (άρα να είναι αύξουσα), αλλά παράλληλα να μην εκφράζει προτίμηση ανάμεσα στις θολωμένες και στις απότομες ακμές. Το κίνητρο αυτής της γενίκευσης είναι η χρησιμοποίηση συναρτήσεων $\phi(s)$ πιο κατάλληλων για την βελτίωση εικόνων από την απλή $\phi(s) = s$ της Ολικής Μεταβολής.

Σαν ειδική περίπτωση του (1.35), θα εξετάσουμε εδώ το συναρτησιακό που προκύπτει όταν $N = 1$ (γκρίζες εικόνες), για το οποίο αντίστοιχα, χρησιμοποιείται η ορολογία (απλά) *ϕ -συναρτησιακό (ϕ -functional)* και προφανώς αποτελεί μία πρώτη επέκταση της Ολικής Μεταβολής.

Θεωρώντας την ελαχιστοποίηση του (1.35) χωρίς περιορισμούς, από τις Euler-Lagrange Εξισώσεις και χρησιμοποιώντας ροή ταχύτερης καθόδου, προκύπτει η ΜΔΕ:

$$\frac{\partial u_i}{\partial t} = \operatorname{div} \left(\frac{\phi'(\|\nabla \mathbf{u}\|)}{\|\nabla \mathbf{u}\|} \nabla u_i \right) \quad , \quad i = 1, \dots, N \quad (1.36)$$

Σε κάθε κανάλι, δηλαδή, συμβαίνει μη γραμμική διάχυση με συντελεστή διάχυσης που είναι κοινός για όλα τα κανάλια και σε κάθε σημείο εξαρτάται από την τοπική γεωμετρία της διανυσματικής εικόνας (όπως αυτή εκφράζεται από τη νόρμα $\|\nabla \mathbf{u}\|$). Έτσι, το πλεονέκτημα είναι ότι επιτυγχάνεται με φυσικό τρόπο η τοπική αλληλεπίδραση ανάμεσα στα κανάλια.

Επίσης, αξίζει να σημειωθεί η σύνδεση της ΜΔΕ (1.36) (όταν $N = 1$) με την γενική έκφραση των Perona - Malik για τη ΜΔΕ μη γραμμικής διάχυσης (1.11): Οι δύο αυτές εξισώσεις είναι ισοδύναμες. Η ισοδυναμία αυτή μπορεί να επιτευχθεί, όταν οι συναρτήσεις $g(s)$ και $\phi(s)$ συνδέονται μέσω της σχέσης:

$$g(s) = \frac{\phi'(s)}{s} \quad (1.37)$$

Παρατηρούμε ότι, από τη μία πλευρά, για συγκεκριμένη επιλογή της συνάρτησης $\phi(s)$ στο συναρτησιακό (1.35), προφανώς η προκύπτουσα ΜΔΕ (1.36) είναι

εξίσωση μη γραμμικής διάχυσης. Από την άλλη, παρατηρούμε επίσης ότι, για κάθε επιλογή της συνάρτησης $g(s)$ στη γενική ΜΔΕ μη γραμμικής διάχυσης (1.11), ορίζεται και μία συνάρτηση $\phi(s)$, διότι (αν θεωρήσουμε ότι $\phi(0) = 0$, αφού για την ελαχιστοποίηση δεν παίζουν ρόλο οι προσθετικές σταθερές):

$$(1.37) \Rightarrow \phi(s) = \int_0^s rg(r)dr \quad (1.38)$$

Επομένως, η γενική ΜΔΕ (1.11) αποτελεί ροή προς την ελαχιστοποίηση του συναρτησιακού (1.35) με συνάρτηση $\phi(s)$ που δίνεται από την τελευταία σχέση.

Στη συνέχεια, θα αναφερθούμε σε κάποιες ενδιαφέρουσες ειδικές περιπτώσεις του γενικού διανυσματικού ϕ -συναρτησιακού, που προκύπτουν με συγκεκριμένες επιλογές της συνάρτησης $\phi(s)$:

1. Αν $\phi(s) = s^2$, τότε:

$$\frac{\phi'(s)}{s} = 2$$

και από την ελαχιστοποίηση, όπως είδαμε στην εισαγωγική ενότητα 1.1 (για $N = 1$) προκύπτει η γραμμική ιστροπική διάχυση.

2. Αν $\phi(s) = s$, καταλήγουμε, όπως είδαμε στη διανυσματική Ολική Μεταβολή (1.31) και η προκύπτουσα ΜΔΕ (1.36) γίνεται:

$$\frac{\partial u_i}{\partial t} = \operatorname{div} \left(\frac{1}{\|\nabla \mathbf{u}\|} \nabla u_i \right) \quad , \quad i = 1, \dots, N \quad (1.39)$$

Πρέπει να σημειώσουμε εδώ ότι η εξίσωση αυτή (όπως και οι υπόλοιπες μορφές των ΜΔΕ Ολικής Μεταβολής (TV) που έχουμε παραθέσει πιο πάνω) παρουσιάζει το πρόβλημα ότι, στα σημεία όπου η \mathbf{u} είναι τοπικά σταθερή ή παρουσιάζει τοπικά ακρότατα, ο παρονομαστής $\|\nabla \mathbf{u}\|$ μηδενίζεται, με αποτέλεσμα να έχουμε *ιδιομορφίες* (*singularities*) και η εξίσωση να μην είναι καλώς ορισμένη. Το παραπάνω δημιουργεί πρόβλημα και στην αριθμητική υλοποίηση της ΜΔΕ. Για να αντιμετωπιστούν τα προβλήματα αυτά, γίνεται συχνά (βλ. π.χ. [13]) η εξής εξομάλυνση (regularization) της Ολικής Μεταβολής (TV): στο συναρτησιακό (1.34), αντί της νόρμας $\|\nabla \mathbf{u}\|$, χρησιμοποιείται η τροποποιημένη νόρμα:

$$\|\nabla \mathbf{u}\|_\epsilon := \sqrt{\epsilon^2 + \|\nabla \mathbf{u}\|^2}$$

όπου ϵ είναι συνήθως μία σχετικά μικρή σταθερά. Έτσι, το συναρτησιακό (1.34) μετατρέπεται στο:

$$J[\mathbf{u}] = \iint_{\Omega} \|\nabla \mathbf{u}\|_\epsilon \, d\mathbf{x} = \iint_{\Omega} \sqrt{\epsilon^2 + \|\nabla \mathbf{u}\|^2} \, d\mathbf{x} \quad (1.40)$$

και η αντίστοιχη συνάρτηση $\phi(s)$ γίνεται $\phi(s) = \sqrt{\epsilon^2 + s^2}$, οπότε:

$$\frac{\phi'(s)}{s} = \frac{1}{\sqrt{\epsilon^2 + s^2}}$$

Επομένως, λόγω της (1.36), η ΜΔΕ που προκύπτει είναι η εξής:

$$\frac{\partial u_i}{\partial t} = \operatorname{div} \left(\frac{1}{\|\nabla \mathbf{u}\|_\epsilon} \nabla u_i \right), \quad i = 1, \dots, N \quad (1.41)$$

Για την παραπάνω εξομάλυνση, θα χρησιμοποιήσουμε την ορολογία ϵ -εξομάλυνση (ϵ -regularization). Παρατηρούμε ότι πλέον ο παρονομαστής του συντελεστή διάχυσης δεν μπορεί να μηδενιστεί, μια και η ελάχιστη τιμή που μπορεί να λάβει είναι ίση με ϵ . Με την εξομάλυνση αυτή, λοιπόν, δεν υπάρχουν πλέον ιδιομορφίες στην ΜΔΕ. Παράλληλα, όταν η σταθερά ϵ είναι μικρή, η διαφορά στα αποτελέσματα είναι πρακτικά ασήμαντη.

Αξίζει να σημειωθεί ότι, όταν $N = 1$ (βαθμωτές εικόνες), η παραπάνω εξομαλυσμένη μέθοδος μπορεί να ειδωθεί, από μία άλλη οπτική γωνία, σαν η μέθοδος *Ελαχίστων Επιφανειών (Minimal Surfaces)* (βλ. π.χ. [18]). Πιο συγκεκριμένα, ας θεωρήσουμε την αναπαράσταση της εικόνας $u(x, y)$ σαν γραφική παράσταση, με κλίμακα $1/\epsilon$ στις τιμές, δηλαδή σαν την εξής επιφάνεια 2 Διαστάσεων, ευρισκόμενη στον 3-Διάστατο χώρο:

$$(x, y) \longrightarrow \left(x, y, \frac{1}{\epsilon} u(x, y) \right) \quad (1.42)$$

Το εμβαδό της επιφάνειας αυτής δίνεται από την σχέση:

$$A[\mathbf{u}] = \frac{1}{\epsilon^2} \iint_{\Omega} \sqrt{\epsilon^2 + \|\nabla u\|^2} d\mathbf{x} \quad (1.43)$$

Συγκρίνοντας την (1.43) με την (1.40) (για $N = 1$), διαπιστώνουμε ότι πρόκειται για το ίδιο συναρτησιακό, εκτός από την πολλαπλασιαστική σταθερά $1/\epsilon^2$, που δεν παίζει κανένα ρόλο στην ελαχιστοποίηση. Επομένως, η ΜΔΕ (1.41) για $N = 1$ αποτελεί *χρονική ροή προς την ελαχιστοποίηση του εμβαδού της επιφάνειας* (1.42). Από αυτή την οπτική γωνία, η σταθερά ϵ έχει νόημα ακόμα και όταν δεν είναι μικρή και εκφράζει τη σχέση κλίμακας του άξονα των τιμών $u(x, y)$ με τους άξονες των μεταβλητών x και y , στην γραφική παράσταση της $u(x, y)$.

3. Αν έχουμε τη μη γραμμική διάχυση (1.11) και κάνουμε για την $g(s)$ την επιλογή (1.13) ή (1.14) των Perona και Malik, η ΜΔΕ που προκύπτει

ελαχιστοποιεί το συναρτησιακό (1.36) (για $N = 1$) με την συνάρτηση $\phi(s)$ να δίνεται αντίστοιχα από τις σχέσεις (λόγω της (1.38)):

$$(1.13) \Rightarrow \phi(s) = \int_0^s r \exp\left(-\left(\frac{r}{K}\right)^2\right) dr = \frac{K^2}{2} \left\{1 - \exp\left(-\left(\frac{s}{K}\right)^2\right)\right\} \quad (1.44)$$

$$(1.14) \Rightarrow \phi(s) = \int_0^s r \frac{1}{1 + \left(\frac{r}{K}\right)^2} dr = \frac{K^2}{2} \log\left(1 + \left(\frac{s}{K}\right)^2\right) \quad (1.45)$$

Παρατηρούμε ότι και στις δύο περιπτώσεις, η συνάρτηση $\phi(s)$ είναι αύξουσα, που σημαίνει ότι όσο μεγαλύτερες είναι οι μεταβολές της εικόνας u τόσο τις χρεώνει με μεγαλύτερο κόστος. Έτσι, εξηγείται από την σκοπιά των συναρτησιακών η ικανότητα των μεθόδων των Perona και Malik να βελτιώνουν τις εικόνες με θόρυβο.

1.7 Μη Γραμμική Ανισοτροπική Διάχυση

Για να ξεπεράσει τους περιορισμούς των προηγούμενων μεθόδων μη γραμμικής διάχυσης (βλ. γενική εξίσωση (1.10)), ο Weickert [59, 60, 62] πρότεινε η διάχυση να είναι όχι μόνο μη γραμμική αλλά και ανισοτροπική. Η γενική μορφή της διάχυσης αυτής είναι η ακόλουθη:

$$\frac{\partial u}{\partial t} = \operatorname{div}(D \nabla u) \quad (1.46)$$

$$u(x, y, 0) = u^0 \quad (1.47)$$

$$\langle D \nabla u, n \rangle = 0, \text{ στο } \partial\Omega \quad (1.48)$$

όπου $D = D(x, y, t)$ είναι ένα πεδίο μεταβλητών (χωρικά και χρονικά) τανυστών διάχυσης (δηλ. συμμετρικών και θετικά ημιορισμένων μητρών). Συγκρίνοντας την παραπάνω εξίσωση με την (1.10), παρατηρούμε ότι ο απλός συντελεστής $k(x, y, t)$ έχει αντικατασταθεί από τον τανυστή $D(x, y, t)$. Έτσι, όλες οι περιπτώσεις μη γραμμικής διάχυσης που έχουν αναφερθεί στις προηγούμενες παραγράφους αποτελούν ειδική περίπτωση της (1.46), με ισοτροπικό τανυστή $D(x, y, t) = k(x, y, t) \cdot I$, όπου I ο μοναδιαίος πίνακας.

Για να μπορέσει η διάχυση να επιτύχει την επιθυμητή βελτίωση της εικόνας, οι ιδιοτιμές λ_1, λ_2 και τα ιδιοδιανύσματα $\mathbf{w}_1, \mathbf{w}_2$ του D σχεδιάζονται έτσι, ώστε να προσαρμόζονται στα χαρακτηριστικά της περιοχής του κάθε σημείου. Ο Weickert προτείνει δύο κύριες επιλογές για τον προσδιορισμό του τανυστή D :

1. Με σκοπό την αποκατάσταση μιας θορυβώδους εικόνας, γίνεται η εξής επιλογή:

$$\mathbf{w}_1 \parallel \nabla u_\sigma \quad , \quad \mathbf{w}_2 \perp \nabla u_\sigma \quad (1.49)$$

$$\lambda_1 = g(\|\nabla u_\sigma\|) \quad , \quad \lambda_2 = 1 \quad (1.50)$$

όπου u_σ μία εξομαλυμένη εκδοχή της u (όπως στην ενότητα 1.3) και η $g(s)$ είναι μία συνάρτηση όπως και στην ισοτροπική περίπτωση (βλ. τους περιορισμούς (1.12)). Ένα παράδειγμα τέτοιας συνάρτησης, που παρατίθεται από τον συγγραφέα στο [59], είναι το εξής:

$$g(s) = 1 - \exp\left(-\frac{1}{s^2}\right)$$

Παρατηρούμε ότι στις ομογενείς περιοχές ($\|\nabla u_\sigma\| \simeq 0$), έχουμε $\lambda_1 \simeq \lambda_2 = 1$, οπότε η διάχυση είναι ισοτροπική. Κοντά στις ισχυρές ακμές (μεγάλη $\|\nabla u_\sigma\|$), έχουμε $\lambda_1 \simeq 0$, οπότε η διάχυση είναι, χάρη στον ταυυστή, έντονα ανισοτροπική και με διεύθυνση παράλληλη στις ακμές.

Παρατηρώντας την γενική εξομαλυμένη μέθοδο μη γραμμικής διάχυσης (1.15) της ενότητας 1.3, διαπιστώνουμε ότι ουσιαστικά διαφέρει από τη συγκεκριμένη μέθοδο του Weickert (βλ. (1.49,1.50)) μόνο στην επιλογή της ιδιοτιμής λ_2 : Η ΜΔΕ (1.15) ισοδυναμεί με επιλογή $\lambda_2 = g(\|\nabla u_\sigma\|)$ (οπότε προκύπτει ισοτροπικός ταυυστής), ενώ στη μέθοδο του Weickert επιλέγεται όπως είπαμε $\lambda_2 = 1$.

Με την παρατήρηση αυτή, μπορούμε εύκολα να συγκρίνουμε την μέθοδο της ενότητας 1.3 με την μέθοδο του Weickert και να διαπιστώσουμε το πλεονέκτημα της δεύτερης: Αφού έχουμε υποθέσει ότι η $g(s)$ ικανοποιεί τους περιορισμούς (1.12) ($g(0) = 1, g(s \rightarrow +\infty) = 0$), στις ομογενείς περιοχές (όπου $\|\nabla u_\sigma\| \approx 0$), και στις δύο περιπτώσεις ισχύει $\lambda_1 \approx \lambda_2 \approx 1$, επομένως οι δύο μέθοδοι παρουσιάζουν παρόμοια συμπεριφορά (ισοτροπική διάχυση). Όμως, κοντά στις ακμές (όπου $\|\nabla u_\sigma\| \gg 0$), για τη ΜΔΕ (1.15) ισχύει $\lambda_1 = \lambda_2 \approx 0$, με αποτέλεσμα η διάχυση να επιβραδύνεται σημαντικά, χωρίς να καταφέρει να απομακρύνει τον θόρυβο. Αντίθετα, για τη μέθοδο του Weickert, ισχύει $\lambda_1 \approx 0, \lambda_2 = 1$, με αποτέλεσμα η διάχυση να συνεχίζεται με ανισοτροπικό τρόπο, καταφέροντας να απομακρύνει τον θόρυβο αλλά και να διατηρήσει τις ακμές.

2. Με σκοπό τη βελτίωση των μονοδιάστατων δομών σε εικόνες, όπως αντικείμενα που μοιάζουν με γραμμές, ο Weickert επέλεξε έναν ταυυστή ο οποίος να προσανατολίζεται έντονα προς την κατεύθυνσή τους. Για την

εύρεση του προσανατολισμού των μονοδιάστατων δομών, χρησιμοποιείται ο δομικός τανυστής (*structure tensor*):

$$J_\rho(\nabla u_\sigma) = K_\rho * (\nabla u_\sigma (\nabla u_\sigma)^T) \quad (1.51)$$

Όπως παρατηρούμε, για την εξαγωγή του παραπάνω τανυστή, εφαρμόζονται δύο εξομαλύνσεις, με δύο *Gaussian* πυρήνες K_σ και K_ρ : Η πρώτη εφαρμόζεται στην εικόνα (όπως στη μέθοδο της παρ. 1.3), με σκοπό ο εντοπισμός των ακμών να είναι αναίσθητος στο θόρυβο και σε λεπτομέρειες μικρότερης τάξης μεγέθους από σ . Η δεύτερη εφαρμόζεται στον τανυστή $\nabla u_\sigma (\nabla u_\sigma)^T$, με σκοπό την εύρεση μιας τοπικής μέσης τιμής των κατευθύνσεων που αντιπροσωπεύει ο τανυστής αυτός, ώστε η εκτίμηση της διεύθυνσης των μονοδιάστατων δομών να είναι πιο αξιόπιστη. Έτσι, οι τυπικές αποκλίσεις των πυρήνων σ και ρ αποτελούν παραμέτρους κλίμακας (*local scale* και *integration scale* αντίστοιχα).

Αν θ_+^* , θ_-^* και λ_+^* , λ_-^* είναι οι ιδιοτιμές και τα ιδιοδιανύσματα αντίστοιχα του δομικού τανυστή $J_\rho(\nabla u_\sigma)$, γίνεται η εξής επιλογή για τον τανυστή D :

$$\mathbf{w}_1 = \theta_+^* \quad , \quad \mathbf{w}_2 = \theta_-^* \quad (1.52)$$

$$\lambda_1 = \alpha \quad , \quad \lambda_2 = \alpha + (1 - \alpha) \exp\left(\frac{-1}{(\lambda_+^* - \lambda_-^*)^2}\right) \quad (1.53)$$

όπου α είναι μία μικρή παράμετρος στο διάστημα $(0, 1)$. Ο Weickert ονόμασε την διάχυση αυτή *Διάχυση Βελτίωσης της Συνοχής (Coherence Enhancing Diffusion)*.

Παρατηρούμε ότι στις περιοχές κοντά σε μια έντονη μονοδιάστατη δομή, ο όρος $(\lambda_+^* - \lambda_-^*)^2$ είναι μεγάλος, οπότε $\lambda_2 \simeq 1$ και συντελείται μεγάλη διάχυση κατά την διεύθυνση της μέγιστης συνοχής. Αντίθετα, στις περιοχές όπου δεν υπάρχει μονοδιάστατη δομή, ο όρος $(\lambda_+^* - \lambda_-^*)^2$ είναι μικρός, οπότε $\lambda_2 \simeq \alpha$ και η διάχυση είναι πολύ μικρή.

1.7.1 Επέκταση στις Διανυσματικές Εικόνες

Στο [61] ο Weickert επεκτείνει τη *Διάχυση Βελτίωσης της Συνοχής* στην επεξεργασία διανυσματικών εικόνων: Η γενική μορφή της ανισοτροπικής διά-

χυσης (1.46) επεκτείνεται στο χρώμα ως εξής:

$$\forall i = 1, \dots, N :$$

$$\frac{\partial u_i}{\partial t} = \operatorname{div} (D \nabla u_i) \quad (1.54)$$

$$u_i(x, y, 0) = u_i^0 \quad (1.55)$$

$$\langle D \nabla u_i, n \rangle = 0, \text{ στο } \partial \Omega \quad (1.56)$$

χρησιμοποιείται δηλαδή ένας κοινός τανυστής διάχυσης D για όλα τα κανάλια u_i .

Επίσης, ο δομικός τανυστής (1.51) γενικεύεται ως εξής:

$$J_\rho(\nabla \mathbf{u}_\sigma) = K_\rho * \sum_{i=1}^N \nabla u_{i,\sigma} (\nabla u_{i,\sigma})^T \quad (1.57)$$

Παρατηρούμε ότι η μήτρα:

$$\sum_{i=1}^N \nabla u_{i,\sigma} (\nabla u_{i,\sigma})^T$$

είναι η μήτρα Di Zenzo της εξομαλυμένης εικόνας $\mathbf{u}_\sigma = K_\sigma * \mathbf{u}$. Άρα, ο δομικός τανυστής (1.57) είναι ουσιαστικά μία εξομαλυμένη εκδοχή της μήτρας Di Zenzo της \mathbf{u}_σ .

Με βάση τα παραπάνω, η γενίκευση της μεθόδου είναι άμεση: Ο τανυστής D προκύπτει από τον $J_\rho(\nabla \mathbf{u}_\sigma)$ χρησιμοποιώντας ακριβώς τις ίδιες σχέσεις (1.52,1.53) και η διάχυση περιγράφεται από τις εξισώσεις (1.54-1.56). Η επέκταση της μεθόδου στις διανυσματικές εικόνες παρουσιάζει την ίδια συμπεριφορά με την αντίστοιχη για τις γκριζες εικόνες.

1.8 Η Ποή Beltrami (Beltrami Flow)

Μία άλλη μέθοδος για την επεξεργασία έγχρωμων εικόνων (και γενικότερα πολυδιάστατων διανυσματικών δεδομένων) προτάθηκε από τους Kimmel, Sochen και Malladi [45, 46, 30]. Οι συγγραφείς εμπνεύστηκαν από τη μέθοδο των Ελαχίστων Επιφανειών, που, όπως είδαμε στην ενότητα 1.6, αφορά την απλή περίπτωση βαθμωτών εικόνων και ελαχιστοποιεί το εμβαδόν της γραφικής τους παράστασης. Γενικεύοντας την προσέγγιση αυτή, οι συγγραφείς θεώρησαν μία διανυσματική εικόνα $\mathbf{u} : \Omega \rightarrow \mathbb{R}^n$ σαν μία επιφάνεια 2 Διαστάσεων, ευρισκόμενη σε έναν χώρο $(n+2)$ Διαστάσεων:

$$(x, y) \longrightarrow (x, y, u_1(x, y), u_2(x, y), \dots, u_n(x, y))$$

Οι συγγραφείς πρότειναν την ελαχιστοποίηση της ακόλουθης γενικευμένης δράσης *Polyakov* (*Polyakov action*):

$$\min_{\mathbf{u}: \Omega \rightarrow \mathbb{R}^n} E[\mathbf{u}] = \int_{\Omega} \sigma \sqrt{g} g^{\mu\nu} \partial_{\mu} X^i \partial_{\nu} X^j h_{ij}(X) \, d\Omega \quad (1.58)$$

Το συναρτησιακό (1.58) είναι ένα φυσικό μέτρο του εμβαδού της επιφάνειας και, για την περίπτωση των διανυσματικών εικόνων ορισμένων σε ένα επίπεδο 2-διάστατο χωρίο Ω , μπορεί να γραφεί πιο απλά ως εξής:

$$\min_{\mathbf{u}: \Omega \rightarrow \mathbb{R}^n} E[\mathbf{u}] = \int_{\Omega} \sqrt{\det(I + G)} \, d\Omega \quad (1.59)$$

όπου G είναι η μήτρα Di Zeno (βλ. (1.5)) και I είναι ο μοναδιαίος πίνακας. Παρατηρούμε ότι $\det(I + G) = (1 + \lambda_+)(1 + \lambda_-)$, οπότε το πρόβλημα ελαχιστοποίησης γίνεται τελικά:

$$\min_{\mathbf{u}: \Omega \rightarrow \mathbb{R}^n} E[\mathbf{u}] = \int_{\Omega} \sqrt{(1 + \lambda_+)(1 + \lambda_-)} \, d\Omega \quad (1.60)$$

Η ελαχιστοποίηση του συναρτησιακού και η εφαρμογή καθόδου κλίσης (gradient descent), ως προς τη μετρική της επιφάνειας ($I + G$), καταλήγει στην ακόλουθη ΜΔΕ, που ονομάζεται *Ροή Beltrami* (*Beltrami Flow*):

$$\forall i = 1, \dots, N :$$

$$\frac{\partial u_i}{\partial t} = \frac{1}{\sqrt{\det(I + G)}} \operatorname{div} \left(\sqrt{\det(I + G)} (I + G)^{-1} \nabla u_i \right) \quad (1.61)$$

$$u_i(x, y, 0) = u_i^0 \quad (1.62)$$

Η ροή αυτή ελαχιστοποιεί το συνολικό εμβαδό της επιφάνειας που περιγράφει την εικόνα, ως προς τη μετρική ($I + G$). Όπως αναφέρεται στο [52], η εξίσωση (1.61) μπορεί να ειπωθεί και ως μία ειδική, σταθμισμένη (λόγω του πολλαπλασιαστικού όρου πριν από το div), περίπτωση της ανιστροπικής διάχυσης (1.54), με τανυστή διάχυσης:

$$D = \sqrt{\det(I + G)} (I + G)^{-1}$$

Οι ιδιοτιμές $\lambda_{1,2}$ και τα ιδιοδιανύσματα $\mathbf{w}_{1,2}$ του τανυστή D μπορούν να εκφραστούν με τη βοήθεια των αντίστοιχων στοιχείων της μήτρας Di Zeno:

$$\lambda_1 = \sqrt{\frac{1 + \lambda_-}{1 + \lambda_+}} \quad , \quad \mathbf{w}_1 = \boldsymbol{\theta}_+$$

$$\lambda_2 = \sqrt{\frac{1 + \lambda_+}{1 + \lambda_-}} \quad , \quad \mathbf{w}_2 = \boldsymbol{\theta}_-$$

Έτσι, μπορεί να εξηγηθεί καλύτερα και η συμπεριφορά της μεθόδου αυτής: Ο τανυστής D μεταβάλλεται κατάλληλα στις ομογενείς περιοχές και στις ακμές. Έτσι, διακρίνουμε τις περιπτώσεις:

- Όταν $\lambda_+ \simeq \lambda_-$, το σημείο θεωρείται ότι ανήκει σε ομογενή περιοχή, οπότε $\lambda_1 \simeq \lambda_2 \simeq 1$ με αποτέλεσμα ο τανυστής να γίνεται ισοτροπικός και η διάχυση να είναι μεγάλη.
- Κοντά στις ακμές, όπου $\lambda_+ \gg \lambda_-$, ο τανυστής κατευθύνεται κυρίως από το ιδιοδιάνυσμα θ_- , που είναι παράλληλο προς την διεύθυνση της ακμής.

Όσον αφορά τον πολλαπλασιαστικό όρο στην εξίσωση (1.61), έχουμε:

$$\frac{1}{\sqrt{\det(I + G)}} = \frac{1}{\sqrt{(1 + \lambda_+)(1 + \lambda_-)}}$$

Παρατηρούμε ότι ο όρος αυτός έχει σαν αποτέλεσμα να μειώνει γρήγορα τη διάχυση κοντά στις έντονες μεταβολές του διανύσματος της εικόνας, με αποτέλεσμα οι ακμές να διατηρούνται για πολύ χρόνο στην εξέλιξη της ροής, με το μειονέκτημα όμως ότι διατηρείται επίσης για πολύ χρόνο και ο υψίσυχνος θόρυβος.

1.9 Διάχυση Διανυσματικών Εικόνων στην Κατεύθυνση θ_-

Οι Sapiro και Ringach [42] πρότειναν μία επέκταση της TV PDE (1.30) σε διανυσματικές εικόνες, που βασίζεται στη γραφή της με τη βοήθεια της κατευθυνόμενης παραγώγου. Η ιδέα τους έγκειται στη χρησιμοποίηση του ιδιοδιανύσματος θ_- της μήτρας Di Zenzo (βλ. ενότητα 1.1.4) για να κατευθύνουν τη διάχυση και τις ιδιοτιμές της, λ_+ , λ_- , για τον πολλαπλασιαστικό όρο. Η ΜΔΕ που εισήγαγαν είναι η ακόλουθη:

$$\frac{\partial \mathbf{u}}{\partial t} = g(\lambda_+ - \lambda_-) \frac{\partial^2 \mathbf{u}}{\partial \theta_-^2} \quad (1.63)$$

όπου $g(s)$ είναι μία θετική, όχι αύξουσα συνάρτηση. Για την διάχυση αυτή, θα χρησιμοποιήσουμε την ονομασία *Διάχυση στην Κατεύθυνση θ_-* . Παρατηρούμε ότι:

$$\lambda_+ - \lambda_- = \mathcal{N}_-^2$$

όπου \mathcal{N}_- νόρμα διανυσματικής κλίσης (γενίκευση της $\|\nabla u\|$), βασισμένη στη μήτρα Di Zenzo (βλ. ενότητα 1.1.4).

Μία απλή επιλογή για την $g(\lambda_+ - \lambda_-)$ ¹, αντίστοιχη με την TV PDE, είναι η εξής:

$$g(\lambda_+ - \lambda_-) = \frac{1}{\sqrt{\lambda_+ - \lambda_-}} = \frac{1}{\mathcal{N}_-} \quad (1.64)$$

Συγκεκριμένα, η TV PDE (1.30) θα μπορούσε να ειπωθεί σαν η ειδική περίπτωση της (1.63), όταν η g δίνεται από την (1.64) και $N = 1$ (γκριζες εικόνες), οπότε $\mathcal{N}_- = \|\nabla u\|$ και $\boldsymbol{\theta}_- = \boldsymbol{\xi}$.

Επιπρόσθετα, αν τροποποιήσουμε την παραπάνω επιλογή, χρησιμοποιώντας ϵ -εξομάλυνση, όπως στην TV PDE (βλ. σχέση (1.41)), ώστε να μην μηδενίζεται ο παρονομαστής, καταλήγουμε στην εξής συνάρτηση:

$$g(\lambda_+ - \lambda_-) = \frac{1}{\sqrt{\epsilon^2 + \lambda_+ - \lambda_-}} = \frac{1}{\sqrt{\epsilon^2 + \mathcal{N}_-^2}} \quad (1.65)$$

Η παραπάνω μέθοδος αποτελεί μία από τις πρώτες προσπάθειες δημιουργίας διανυσματικής ΜΔΕ βασισμένης σε κατευθυνόμενες παραγώγους. Η διάχυση συμπεριφέρεται με τον ίδιο τρόπο σε όλα τα κανάλια και έτσι αυτά εξομαλύνονται κατά την ίδια διεύθυνση και με την ίδια ένταση. Η σύζευξη ανάμεσα στα κανάλια επιτυγχάνεται μέσω του διανύσματος $\boldsymbol{\theta}_-$ και της νόρμας \mathcal{N}_- , τα οποία καθορίζουν την διάχυση.

Όπως φαίνεται από την (1.63), η διάχυση γίνεται μόνο κατά την κατεύθυνση της ελάχιστης μεταβολής $\boldsymbol{\theta}_-$ της \mathbf{u} , με αποτέλεσμα οι διανυσματικές ακμές να μην καταστρέφονται. Επίσης, σε ομογενείς περιοχές, η διάχυση είναι μεγάλη, ενώ σε περιοχές μεγάλων διακυμάνσεων του διανύσματος \mathbf{u} (π.χ. κοντά σε ακμές), η διάχυση επιβραδύνεται.

Παράλληλα όμως, η μέθοδος αυτή παρουσιάζει κάποια μειονεκτήματα. Η νόρμα \mathcal{N}_- σε αρκετές περιπτώσεις δεν μπορεί να διακρίνει τις διανυσματικές γωνίες. Πιο κατάλληλη για αυτό θα ήταν η νόρμα $\mathcal{N}_+ = \|\nabla \mathbf{u}\|$ (βλ. ενότητα 1.1.4). Επίσης, το ότι η διάχυση συντελείται μόνο σε μία διεύθυνση δημιουργεί το πρόβλημα ότι σε επίπεδες περιοχές με έντονο θόρυβο, η διάχυση κατευθύνεται κυρίως από το θόρυβο, με αποτέλεσμα η εικόνα να μην καθαρίζει και να δημιουργούνται θορυβώδεις μονοδιάστατες δομές.

Τέλος, αξίζει να σημειωθεί ότι (όπως αναφέρεται στο [52]), επειδή η (1.63) περιέχει κατευθυνόμενη παράγωγο, η διάχυση μπορεί να γραφεί και με τη βοήθεια ενός τανυστή διάχυσης T , χρησιμοποιώντας τον τελεστή ίχνους $\text{trace}(\cdot)$:

$$(1.63) \Leftrightarrow \frac{\partial u_i}{\partial t} = \text{trace}(TH_i) \quad , \quad T = g(\lambda_+ - \lambda_-)\boldsymbol{\theta}_-\boldsymbol{\theta}_-^T \quad (i = 1, \dots, N) \quad (1.66)$$

¹στο [42], οι συγγραφείς δεν αποσαφηνίζουν ποια συγκεκριμένη $g(\lambda_+ - \lambda_-)$ επέλεξαν για τα πειράματα που παραθέτουν, για αυτό στην παρουσίαση αυτή και στα πειράματά μας, κάνουμε μία επιλογή της $g(\lambda_+ - \lambda_-)$, που φαντάζει αρκετά φυσική, με βάση τις υπόλοιπες μεθόδους που έχουμε αναφέρει μέχρι τώρα.

όπου H_i η χωρική *Hessian* Μήτρα της u_i . Όπως φαίνεται από την (1.66), ο τανυστής T έχει ιδιοτιμές $g(\lambda_+ - \lambda_-)$ και 0, και αντίστοιχα ιδιοδιανύσματα $\boldsymbol{\theta}_-$ και $\boldsymbol{\theta}_+$.

1.10 Η Ανισοτροπική Εξομάλυνση των Tschumperlé-Deriche

Οι Tschumperlé και Deriche [52, 55] εισήγαγαν μία άλλη ΜΔΕ που βασίζεται σε κατευθυνόμενες παραγώγους, σε δύο κατευθύνσεις αυτή τη φορά και χρησιμοποιώντας τη νόρμα $\mathcal{N}_+ = \|\nabla \mathbf{u}\|$ (κατάλληλα τροποποιημένη, όπως θα δούμε παρακάτω), με σκοπό τον αποτελεσματικότερο έλεγχο της διαδικασίας της διάχυσης. Η ΜΔΕ αυτή περιγράφεται από την σχέση:

$$\frac{\partial \mathbf{u}}{\partial t} = f_- \left(\sqrt{\lambda_+^* + \lambda_-^*} \right) \frac{\partial^2 \mathbf{u}}{\partial \boldsymbol{\theta}_-^{*2}} + f_+ \left(\sqrt{\lambda_+^* + \lambda_-^*} \right) \frac{\partial^2 \mathbf{u}}{\partial \boldsymbol{\theta}_+^{*2}} \quad (1.67)$$

όπου λ_+^* , λ_-^* και $\boldsymbol{\theta}_+^*$, $\boldsymbol{\theta}_-^*$ οι ιδιοτιμές και τα ιδιοδιανύσματα του δομικού τανυστή (βλ. σχέση (1.57)):

$$J_\rho(\nabla \mathbf{u}) = K_\rho * \sum_{i=1}^N \nabla u_i (\nabla u_i)^T \quad (1.68)$$

Χρησιμοποιείται, δηλαδή, όπως και στη μέθοδο της Ανισοτροπικής Διάχυσης (1.54), ο δομικός τανυστής, με τη διαφορά ότι εδώ γίνεται εξομάλυνση μόνο στη μήτρα Di Zenzo και όχι στην εικόνα \mathbf{u} (έχουμε δηλαδή $\sigma = 0$).

Για τη δεύτερη μερική παράγωγο σε μία κατεύθυνση, όπως οι $\frac{\partial^2 \mathbf{u}}{\partial \boldsymbol{\theta}_-^{*2}}$, $\frac{\partial^2 \mathbf{u}}{\partial \boldsymbol{\theta}_+^{*2}}$ που εμφανίζονται στο δεξιό μέλος της (1.67), χρησιμοποιείται συχνά (βλ. π.χ. [52, 55]) η ορολογία *κατευθυνόμενη Λαπλασιανή (oriented Laplacian)*.

Η εξίσωση (1.67) μπορεί να γραφεί και με τη βοήθεια ενός τανυστή διάχυσης T , χρησιμοποιώντας τον τελεστή ίχνους $\text{trace}(\cdot)$:

$$(1.67) \Leftrightarrow \frac{\partial u_i}{\partial t} = \text{trace}(T H_i) \quad , \quad i = 1, \dots, N \quad (1.69)$$

όπου H_i η χωρική *Hessian* Μήτρα της u_i και:

$$T = f_- \left(\sqrt{\lambda_+^* + \lambda_-^*} \right) \boldsymbol{\theta}_-^* \boldsymbol{\theta}_-^{*T} + f_+ \left(\sqrt{\lambda_+^* + \lambda_-^*} \right) \boldsymbol{\theta}_+^* \boldsymbol{\theta}_+^{*T} \quad (1.70)$$

Η νόρμα $\mathcal{N}_+^* = \sqrt{\lambda_+^* + \lambda_-^*}$ χρησιμοποιείται για τον εντοπισμό των διακυμάνσεων στο διάνυσμα \mathbf{u} της εικόνας. Αποτελεί μία τροποποίηση της νόρμας

$\mathcal{N}_+ = \|\nabla \mathbf{u}\|$ (βλ. ενότητα 1.1.4) και είναι πιο αποτελεσματική, διότι ο τελικός τανυστής λαμβάνει υπόψιν όχι μόνο το σημείο αλλά και την γειτονική του περιοχή και έτσι η εκτίμηση είναι πιο συνεκτική.

Οι συγγραφείς κάνουν τις εξής επιλογές για τις συναρτήσεις f_-, f_+ :

$$f_-(s) = \frac{1}{\sqrt{1+s^2}}$$

$$f_+(s) = \frac{1}{1+s^2}$$

οι οποίες, όπως παρατηρούμε, έχουν τις εξής ιδιότητες:

$$f_-(0) = f_+(0) = 1 \quad \text{και} \quad f_-(s \rightarrow \infty) = f_+(s \rightarrow \infty) = 0$$

$$f_-(s), f_+(s) \text{ φθίνουσες για } s \in [0, +\infty) \quad (1.71)$$

$$\lim_{s \rightarrow \infty} \frac{f_+(s)}{f_-(s)} = 0$$

Οι ιδιότητες αυτές προσδίδουν στη διάχυση συμπεριφορά με επιθυμητά αποτελέσματα για την βελτίωση των εικόνων: Στις σχεδόν ομογενείς περιοχές (μικρή \mathcal{N}_+^*), η διάχυση συντελεί ιστροπική εξομάλυνση, με μεγάλη ένταση, ενώ κοντά στις διανυσματικές ακμές (μεγάλη \mathcal{N}_+^*), η εξομάλυνση γίνεται ανιστροπική, κατά την διεύθυνση των ακμών θ_-^* και με μικρότερη ένταση. Τέλος, κοντά στις διανυσματικές γωνίες (ακόμα μεγαλύτερη \mathcal{N}_+^*), η ένταση της εξομάλυνσης είναι ακόμα μικρότερη. Έτσι, για να έχει η ΜΔΕ τα επιθυμητά αποτελέσματα, πρέπει και κάθε άλλη πιθανή επιλογή των συναρτήσεων $f_-(s), f_+(s)$ να ικανοποιεί τις σχέσεις (1.71).

Όπως παρατηρούμε στις σχέσεις (1.67) και (1.69), χρησιμοποιείται ένας κοινός τανυστής σε όλα τα κανάλια u_i της εικόνας για τον έλεγχο της διάχυσης, δηλαδή η διάχυση γίνεται σε όλα τα κανάλια με τον ίδιο τρόπο (το ίδιο χαρακτηριστικό συναντήσαμε και στην ΜΔΕ (1.63)). Έτσι, αποφεύγεται μη επιθυμητή σύζευξη των καναλιών (η επιθυμητή γίνεται μέσω του υπολογισμού του δομικού τανυστή $J_\rho(\nabla \mathbf{u})$), η οποία θα οδηγούσε σε μεταφορά διάχυσης από το ένα κανάλι στο άλλο (στην περίπτωση των έγχρωμων εικόνων αυτό θα σήμαινε εμφάνιση χρωμάτων άσχετων με την αρχική εικόνα).

Όπως αναφέρεται στα [52, 55], η εξίσωση (1.67) ισοδυναμεί με επαναληπτικές (για κάθε απειροστό χρονικό βήμα) συνελιζεις με προσανατολισμένους Gaussian πυρήνες, οι οποίοι προσανατολίζονται και εκτείνονται σύμφωνα με τον τανυστή T^2 . Πιο συγκεκριμένα, σε κάθε σημείο και κάθε χρονική στιγμή, ο Gaussian πυρήνας συνέλιξης εκτείνεται στην διεύθυνση του ιδιοδιανύσματος θ_-^* (θ_+^*) τόσο περισσότερο, όσο μεγαλύτερη είναι η ιδιοτιμή f_- (f_+). Έτσι,

²στην ενότητα 4.1, θα μελετήσουμε περαιτέρω την ισοδυναμία αυτή.

κάθε χρονική στιγμή, συντελείται ανισοτροπική εξομάλυνση της \mathbf{u} , η οποία καθορίζεται από τον τανυστή T .

Έτσι, από αυτή την οπτική γωνία, η ΜΔΕ (1.67) πλεονεκτεί έναντι της Ανισοτροπικής Διάχυσης (1.54) στο ότι ο τανυστής T μπορεί να ελέγξει ακριβώς την διαδικασία ανισοτροπικής εξομάλυνσης που συντελείται, ενώ ο τανυστής D στην (1.54) δεν μπορεί, διότι αποδεικνύεται (βλ. [52]) ότι στην Ανισοτροπική Διάχυση (1.54), η εξομάλυνση δεν γίνεται τελικά στις κατευθύνσεις που ορίζει ο D .

Επίσης, η μέθοδος αυτή καταφέρνει να ξεπεράσει τα μειονεκτήματα της ΜΔΕ (1.63), που είδαμε στην προηγούμενη ενότητα. Πρώτα από όλα, χρησιμοποιεί τη νόρμα \mathcal{N}_+ , που είναι πιο κατάλληλη από τη νόρμα \mathcal{N}_- , αφού εντοπίζει και τις διανυσματικές γωνίες (βλ. ενότητα 1.1.4). Ακόμα, η διάχυση συντελεί εξομάλυνση σε δύο κάθετες κατευθύνσεις (και όχι όπως πριν μόνο σε μία) και έτσι στις επίπεδες περιοχές μετατρέπεται σε ισοτροπική, με αποτέλεσμα ο θόρυβος να απομακρύνεται πλέον και από τις περιοχές αυτές.

Κεφάλαιο 2

Ανάλυση μεθόδων ΜΔΕ για το πρόβλημα του Inpainting

2.1 Εισαγωγή

Το πρόβλημα του Inpainting συνίσταται στην αποκατάσταση εικόνων, από τις οποίες λείπουν διάφορα τμήματα. Σκοπός του Inpainting είναι να ανακατασκευάσει τα τμήματα αυτά με τέτοιο τρόπο, ώστε να μη γίνεται αντιληπτό από έναν ανθρώπινο παρατηρητή ότι το τελικό αποτέλεσμα έχει προέλθει από επεξεργασία. Η επεξεργασία αυτή μπορεί φυσικά να επιτελεσθεί από τον άνθρωπο, αλλά εδώ θα ασχοληθούμε με το αυτόματο Inpainting, μέσω της ψηφιακής επεξεργασίας με υπολογιστή. Το αυτόματο Inpainting βρίσκει διάφορες εφαρμογές, όπως στην αρχαιολογία (όπου πολύ συχνά σε έργα τέχνης λείπουν τμήματα), στην απομάκρυνση κειμένου από εικόνες, στο ψηφιακό ρετούς φθαρμένων φωτογραφιών και στις τηλεπικοινωνίες (όπου συχνά, στην ασύρματη μετάδοση εικόνας και video χάνονται κάποια μπλοκ).

Τα τελευταία χρόνια έχουν προταθεί διάφορες μέθοδοι, που εφαρμόζουν ΜΔΕ στο πρόβλημα του Inpainting. Η εφαρμογή των ΜΔΕ παρουσιάζει αρκετά ελπιδοφόρα αποτελέσματα και στο πρόβλημα αυτό. Ένα σημαντικό πλεονέκτημα της χρήσης ΜΔΕ για το πρόβλημα αυτό είναι ότι η ανακατασκευή γίνεται αυτόματα, εφαρμόζοντας την ΜΔΕ για την περιοχή που λείπει. Έτσι, δεν υπάρχουν τοπολογικοί περιορισμοί για το σχήμα της περιοχής που λείπει. Στις επόμενες ενότητες, παρουσιάζουμε τις περισσότερες μεθόδους που έχουν αναπτυχθεί για Inpainting με ΜΔΕ, τόσο για βαθμωτές όσο και για έγχρωμες (ή γενικά διανυσματικές) εικόνες. Τέλος, προτείνουμε επίσης μία νέα μέθοδο για Inpainting έγχρωμων εικόνων.

2.2 Το Μοντέλο BSCB: Μεταφορά Ομαλότητας (Smoothness Transportation)

Μία από τις πρώτες προσπάθειες χρήσης ΜΔΕ για το πρόβλημα του Inpainting βαθμωτών εικόνων έγινε από τους Bertalmio, Sapiro, Caselles και Ballester [7] με τη μέθοδο *Μεταφοράς Ομαλότητας (Smoothness Transportation)* κατά μήκος των ισοϋψών γραμμών. Η μέθοδος αυτή ονομάζεται και *Μοντέλο BSCB* (από τα αρχικά των ονομάτων των συγγραφέων). Η ιδέα τους συνίσταται στην διατύπωση μίας ΜΔΕ η οποία να ανακατασκευάζει την εικόνα u^0 στην περιοχή inpainting \mathbb{D} που λείπει, βασισμένη στην μορφή της u^0 γύρω από αυτή και με τέτοιο τρόπο, ώστε οι έντονες ακμές να αποκαθίστανται. Για τον σκοπό αυτό, χρησιμοποίησαν μία συνάρτηση $L(u)$ που μετράει την *ομαλότητα (smoothness)* της εικόνας u . Για παράδειγμα, μία τέτοια συνάρτηση δευτέρου βαθμού μπορεί να γραφεί στην ακόλουθη γενική μορφή:

$$L(u) = f(\nabla u, H)$$

όπου H η χωρική *Hessian* Μήτρα της u . Οι συγγραφείς στο [7] επιλέγουν την Λαπλασιανή για τα πειράματά τους:

$$L(u) = \nabla^2 u = \text{trace}(H)$$

Έτσι, η ΜΔΕ που εισήγαγαν είναι η ακόλουθη:

$$\frac{\partial u}{\partial t} = \|\nabla u\| \boldsymbol{\xi} \cdot \nabla L(u) = \|\nabla u\| \frac{\partial L(u)}{\partial \boldsymbol{\xi}} \quad (2.1)$$

όπου $\boldsymbol{\xi} = \left(\frac{\nabla u}{\|\nabla u\|} \right)^\perp$, δηλαδή το μοναδιαίο διάνυσμα, εφαπτομενικό στην ισοϋψή καμπύλη που διέρχεται από το κάθε σημείο.

Επομένως, η ΜΔΕ (2.1) εξελίσσεται μεταφέροντας την πληροφορία της ομαλότητας (όπως αυτή εκφράζεται από την $L(u)$) από το σύνορο της περιοχής inpainting \mathbb{D} στο εσωτερικό της, κατά μήκος των ανακατασκευασμένων ισοϋψών γραμμών. Όταν η εξέλιξη φτάσει στην κατάσταση ισορροπίας, από την (2.1), έχουμε (εφόσον $\|\nabla u\| \neq 0$):

$$\frac{\partial L(u)}{\partial \boldsymbol{\xi}} = 0 \quad (2.2)$$

που σημαίνει ότι, η εικόνα συμπληρώνεται στο εσωτερικό της \mathbb{D} έτσι ώστε, κατά μήκος μιας ισοϋψούς, η μέτρηση ομαλότητας $L(u)$ να παραμένει σταθερή.

Όμως, λόγω της έλειψης επικοινωνίας μεταξύ των ισοϋψών, η παραπάνω εξέλιξη μεταφοράς μπορεί να καταλήξει σε συγκρούσεις μεταξύ τους και απότομα σπασίματα των καμπυλών. Για τον λόγο αυτό και για να τονιστούν οι

διαφορές μεταξύ των ισοϋψών, στο [7] η εξίσωση (2.1) υλοποιείται με ενδιάμεσες επαναλήψεις μη γραμμικής διάχυσης (βλ. ενότητα 1.2).

Η μέθοδος αυτή, παρά την πρωτοτυπία της και τα σχετικά ικανοποιητικά αποτελέσματά της, παρουσιάζει το μειονέκτημα ότι η ΜΔΕ δεν καταφέρνει από μόνη της να επιτύχει την επιθυμητή συμπεριφορά και για αυτό απαιτείται να παρεμβληθεί από ΜΔΕ μη γραμμικής διάχυσης. Όπως θα δούμε στη συνέχεια, άλλες ΜΔΕ καταφέρνουν με πιο φυσικό τρόπο να εντάξουν τις δύο αυτές διαδικασίες σε μία κοινή ροή. Επίσης, όπως αναφέρεται στο [16], η επιλογή της Λαπλασιανής σαν μέτρο της Ομαλότητας παρουσιάζει κάποια προβλήματα. Πρώτα από όλα, η κατάσταση ισορροπίας δεν είναι μορφολογικά αναλοιώτη. Ακόμα, η απαίτηση σταθερής μέτρησης ομαλότητας $L(u)$ κατά μήκος μιας ισοϋψούς (βλ. σχέση (2.2)) είναι σε πολλές περιπτώσεις υπερβολική, με την έννοια ότι σε φυσικές εικόνες, η Λαπλασιανή συνήθως αλλάζει κατά μήκος μιας ισοϋψούς καμπύλης με σχετικά μεγάλο μήκος.

2.3 Total Variation Inpainting

Οι Chan και Shen [15] πρότειναν την εφαρμογή της ελαχιστοποίησης της Ολικής Μεταβολής (TV), που παρουσιάσαμε στην ενότητα 1.4 για την Βελτίωση εικόνων, και στο πρόβλημα Inpainting βαθμωτών εικόνων. Στη συνέχεια, ο Kang [28] επέκτεινε την ιδέα αυτή στις διανυσματικές εικόνες, χρησιμοποιώντας την Διανυσματική Ολική Μεταβολή (βλ. σχέση (1.34) στην ενότητα 1.6). Εδώ, θα μελετήσουμε κατευθείαν τη μέθοδο για τις διανυσματικές εικόνες, αναφέροντας σαν ειδική περίπτωση τη μέθοδο για τις βαθμωτές εικόνες (όταν $N = 1$).

Η μέθοδος αυτή, λοιπόν, συνίσταται στην ανακατασκευή της εικόνας εντός της περιοχής inpainting, με τέτοιο τρόπο ώστε να ελαχιστοποιείται η Ολική Μεταβολή, κάτι το οποίο δίνει ικανοποιητικά αποτελέσματα για την ανθρώπινη οπτική αντίληψη, όπως και στην περίπτωση της αποθρομβοποίησης (εφόσον βέβαια δεν υπάρχει υφή).

Πιο συγκεκριμένα, θεωρούμε μία εικόνα \mathbf{u}^0 , ορισμένη στο κλειστό χωρίο Ω , που της λείπει ένα σύνολο από κομμάτια \mathbb{D} , αλλά επιπλέον θεωρούμε ότι περιέχει προσθετικό θόρυβο. Με σκοπό το Inpainting της περιοχής \mathbb{D} αλλά και την απομάκρυνση του θορύβου έξω από την \mathbb{D} , η μέθοδος συνίσταται στην (χωρίς περιορισμούς) ελαχιστοποίηση του συναρτησιακού (που συνδέεται άμεσα με το αντίστοιχο (1.23) της ενότητας 1.4):

$$J[\mathbf{u}] = \iint_{\Omega} \|\nabla \mathbf{u}\| \, d\mathbf{x} + \frac{\lambda}{2} \iint_{\Omega \setminus \mathbb{D}} \|\mathbf{u} - \mathbf{u}^0\|^2 \, d\mathbf{x} \quad (2.3)$$

Πρέπει να τονιστεί ότι η δεύτερη ολοκλήρωση γίνεται μόνο στην περιοχή $\Omega \setminus \mathbb{D}$,

όπου υπάρχει αρχική πληροφορία \mathbf{u}^0 και όπως έχουμε δει (βλ. ενότητα 1.4), είναι ένα συναρτησιακό ποινής για την απόσταση από την αρχική εικόνα. Ακόμα, ο όρος πιστότητας λ πρέπει να εξαρτάται από την ισχύ του θορύβου (αντιστρόφως ανάλογα) και εκφράζει την ισορροπία ανάμεσα στους δύο όρους, για την περιοχή (βλ. ενότητα 1.4). Η διαδικασία της αποθορυβοποίησης, παράλληλα με το inpainting, είναι σημαντική ακόμα και όταν η ισχύς του θορύβου είναι μικρή. Κι αυτό διότι αλλιώς (όταν γίνεται μόνο inpainting), όπως έχουμε παρατηρήσει και πειραματικά, ενώ οι γεωμετρικές δομές (περιεχόμενο) μπορεί να συνεχίζονται στην \mathbb{D} με επιτυχία, εμφανίζεται μία απότομη μετάβαση από περιοχές με θόρυβο (εκτός της \mathbb{D}) σε περιοχές χωρίς θόρυβο (εντός της \mathbb{D}). Το γεγονός αυτό κάνει το αποτέλεσμα να είναι λιγότερο ευχάριστο για την ανθρώπινη όραση.

Ελαχιστοποιώντας κατά τα γνωστά την (5.25), από τις Euler-Lagrange Εξισώσεις και χρησιμοποιώντας ροή ταχύτερης καθόδου, προκύπτει η ΜΔΕ:

$$\begin{aligned} \forall i = 1, \dots, N : \\ \begin{cases} \frac{\partial u_i}{\partial t} = \operatorname{div} \left(\frac{1}{\|\nabla \mathbf{u}\|} \nabla u_i \right) + \lambda(u_i - u_i^0) & , \mathbf{x} \in \Omega \setminus \mathbb{D} \\ \frac{\partial u_i}{\partial t} = \operatorname{div} \left(\frac{1}{\|\nabla \mathbf{u}\|} \nabla u_i \right) & , \mathbf{x} \in \mathbb{D} \end{cases} \end{aligned} \quad (2.4)$$

Στην ειδική περίπτωση όπου δεν υπάρχει θόρυβος ($\lambda \rightarrow \infty$ για το συναρτησιακό (5.25)), η εικόνα προφανώς δεν μεταβάλλεται έξω από την \mathbb{D} και έτσι το πρόβλημα ελαχιστοποίησης γίνεται:

$$\min_{\mathbf{u}: \mathbb{D} \rightarrow \mathbb{R}^N} \iint_{\mathbb{D}} \|\nabla \mathbf{u}\| \, d\mathbf{x} \quad (2.5)$$

υπό τον περιορισμό: $\mathbf{u}(\mathbf{x}) = \mathbf{u}^0(\mathbf{x})$, $\mathbf{x} \in \partial\mathbb{D}$. Η ελαχιστοποίηση αυτή καταλήγει στην απλούστερη ΜΔΕ:

$$\frac{\partial u_i}{\partial t} = \operatorname{div} \left(\frac{1}{\|\nabla \mathbf{u}\|} \nabla u_i \right) \quad , \mathbf{x} \in \mathbb{D} \quad (i = 1, \dots, N) \quad (2.6)$$

με συνοριακή συνθήκη:

$$\mathbf{u}(\mathbf{x}) = \mathbf{u}^0(\mathbf{x}) \quad , \mathbf{x} \in \partial\mathbb{D} \quad (2.7)$$

Πρέπει να σημειωθεί ότι στην κατάσταση ισορροπίας της εξίσωσης (2.6), σε αντίθεση με την αντίστοιχη εξίσωση (1.39) που χρησιμοποιείται για απομάκρυνση του θορύβου (βλ. ενότητα 1.6), δεν προκύπτει μία εκφυλισμένη εικόνα (υπερβολικά εξομαλυσμένη), αλλά η ανακατασκευασμένη εικόνα που μας ενδιαφέρει (λόγω της συνοριακής συνθήκης (2.7)).

Αξίζει να σημειωθεί ότι η μέθοδος TV Inpainting έχει αρκετά καλά αποτελέσματα, ιδιαίτερα όταν τα κομμάτια που συνιστούν την περιοχή \mathbb{D} έχουν σχετικά μικρό πάχος. Όμως, παρουσιάζει και κάποια σημαντικά μειονεκτήματα (όπως αναφέρεται και στο [16] για βαθμωτές εικόνες). Πρώτα από όλα, η μέθοδος αυτή συχνά αποτυγχάνει να συνδέσει δύο μέρη ενός αντικειμένου όταν παρεμβάλλεται απόσταση μεγάλου μήκους, λόγω υψηλού κόστους στο συναρτησιακό TV της ένωσης μεγάλης απόστασης. Επίσης, για την περίπτωση inpainting γκριζών εικόνων ($N = 1$), μπορεί να διαπιστωθεί ότι το συναρτησιακό ελαττώνεται όσο μικρότερο είναι το μήκος των ισούψων καμπυλών, με αποτέλεσμα η μέθοδος να επανενώνει ουσιαστικά τις σπασμένες ισούψεις με ευθύγραμμα τμήματα. Για τις διανυσματικές εικόνες ($N \geq 2$), δεν μπορεί να υπάρξει ανάλογη γεωμετρική ερμηνεία, αλλά θα μπορούσαμε να πούμε ότι, όπως διαπιστώσαμε και πειραματικά, η συμπεριφορά της μεθόδου είναι παρόμοια και το μειονέκτημα αυτό επεκτείνεται και στην περίπτωση αυτή.

2.3.1 Τροποποιήσεις της μεθόδου Total Variation

Όπως είδαμε στην ενότητα 1.6 (βλ. σχέση (1.41)), η ΜΔΕ (2.6) μπορεί να τροποποιηθεί, χρησιμοποιώντας ϵ -εξομάλυνση (ϵ -regularization), ώστε στα σημεία όπου $\|\nabla \mathbf{u}\| = 0$ να μην μηδενίζεται ο παρονομαστής. Έτσι, προκύπτει η εξίσωση:

$$\frac{\partial u_i}{\partial t} = \operatorname{div} \left(\frac{1}{\|\nabla \mathbf{u}\|_\epsilon} \nabla u_i \right) \quad , \quad \mathbf{x} \in \mathbb{D} \quad (i = 1, \dots, N), \quad (2.8)$$

όπου: $\|\nabla \mathbf{u}\|_\epsilon := \sqrt{\epsilon^2 + \|\nabla \mathbf{u}\|^2}$.

Μία άλλη τροποποίηση του Total Variation Inpainting, για την ειδική περίπτωση των βαθμωτών εικόνων ($N = 1$) αναφέρεται στο [16]. Συγκεκριμένα, η τροποποίηση αυτή συνίσταται στον πολλαπλασιασμό του δεξιού μέλους της εξίσωσης (2.6) (για $N = 1$) με τη νόρμα $\|\nabla u\|$, οπότε προκύπτει, όπως έχουμε δει, η ΜΔΕ της Κίνησης Μέσης Καμπυλότητας - MCM (βλ. ενότητα 1.4, σχέση (1.29)). Η ΜΔΕ αυτή καταλήγει στην ίδια κατάσταση ισορροπίας με την (2.6) (για $N = 1$), δίνοντας έτσι ίδια αποτελέσματα, ακολουθώντας όμως διαφορετική χρονική ροή. Όπως αναφέρεται στο [16], η ροή αυτή μετά την τροποποίηση παρουσιάζει το πλεονέκτημα ότι είναι μορφολογικά αμετάβλητη (morphologically invariant), δεν εξαρτάται δηλαδή από οποιαδήποτε αλλαγή στην αντίθεση (contrast) της εικόνας. Επίσης, η τροποποίηση αυτή επιταχύνει και την αριθμητική σύγκλιση.¹

¹όπως θα δούμε στην υποενότητα 5.2.2, έχει προταθεί επίσης ο αλγόριθμος του Digital TV

Επίσης, σαν μια άλλη τροποποίηση του Total Variation Inpainting, προτείνουμε εδώ μία άμεση και απλή γενίκευση (μέσω της νόρμας $\|\nabla \mathbf{u}\|$) της Κίνησης Μέσης Καμπυλότητας - MCM στις διανυσματικές εικόνες: πολλαπλασιάζουμε το δεξί μέλος της εξίσωσης (2.6) (για τη γενική περίπτωση $N \geq 1$) με $\|\nabla \mathbf{u}\|$. Έτσι, προτείνουμε την ακόλουθη ΜΔΕ:

$$\frac{\partial u_i}{\partial t} = \|\nabla \mathbf{u}\| \operatorname{div} \left(\frac{1}{\|\nabla \mathbf{u}\|} \nabla u_i \right) , \quad \mathbf{x} \in \mathbb{D} \quad (i = 1, \dots, N), \quad (2.9)$$

για την οποία χρησιμοποιούμε την ονομασία *Σταθμισμένη TV PDE*. Όταν $N = 1$, η ΜΔΕ (2.9) περιγράφει την Κίνηση Μέσης Καμπυλότητας. Προφανώς, όπως και πριν, η ΜΔΕ καταλήγει στην ίδια κατάσταση ισορροπίας με την (2.6) (για $N \geq 1$), ακολουθώντας όμως άλλη χρονική ροή. Για αντίστοιχους λόγους με την απλή περίπτωση, η ροή αυτή συγκλίνει πιο γρήγορα από την (2.6) στην κατάσταση ισορροπίας. Όπως στην περίπτωση της (2.6) και για τους ίδιους λόγους, τροποποιούμε την (2.9) χρησιμοποιώντας ϵ -εξομάλυνση, οπότε προκύπτει η ΜΔΕ:

$$\frac{\partial u_i}{\partial t} = \|\nabla \mathbf{u}\| \operatorname{div} \left(\frac{1}{\|\nabla \mathbf{u}\|_\epsilon} \nabla u_i \right) , \quad \mathbf{x} \in \mathbb{D} \quad (i = 1, \dots, N) \quad (2.10)$$

2.4 Διάχυση Οδηγούμενη από την Καμπυλότητα (Curvature Driven Diffusion)

Για να αντιμετωπιστεί το πρόβλημα της αποτυχίας συνδέσεων μεγάλου μήκους που όπως είδαμε παρουσιάζει το TV Inpainting, οι Chan και Shen [17] τροποποίησαν την εξίσωση της μεθόδου αυτής (για $N = 1$), ώστε ο συντελεστής διάχυσης να αυξάνεται με την αύξηση της καμπυλότητας κ των ισούψων γραμμών. Έτσι, εισήγαγαν την εξής ΜΔΕ για inpainting βαθμωτών εικόνων:

$$\frac{\partial u}{\partial t} = \operatorname{div} \left(\frac{f(|\kappa|)}{\|\nabla u\|} \nabla u \right) , \quad \kappa = \operatorname{div} \left(\frac{\nabla u}{\|\nabla u\|} \right) \quad (2.11)$$

όπου $f(s)$, $s \in [0, +\infty)$ συνεχής, αύξουσα συνάρτηση, με $f(0) = 0$ και $f(s \rightarrow +\infty) = \infty$. Στο [17], οι συγγραφείς κάνουν την εξής επιλογή για την συνάρτηση αυτή: $f(s) = s^p$, $p \geq 1$.

Η διάχυση (2.11) ονομάζεται *Διάχυση Οδηγούμενη από την Καμπυλότητα (Curvature Driven Diffusion - CDD)*.

Filter [13], που συγκλίνει ακόμα γρηγορότερα στην διακριτοποιημένη κατάσταση ισορροπίας της (2.6), αλλά δεν προκύπτει, όπως εδώ, από κάποιο συνεχές μοντέλο.

Η εισαγωγή του όρου $f(|\kappa|)$ βασίζεται στην απλή παρατήρηση ότι, όταν η μέθοδος TV Inpainting αποτυγχάνει να κάνει την επιθυμητή σύνδεση, οι ισοϋψείς των ακμών περιέχουν γωνίες, οι οποίες έχουν πολύ υψηλή καμπυλότητα κ . Έτσι, χάρη σε αυτόν τον όρο, όσο μεγαλύτερη είναι η καμπυλότητα κ , τόσο μεγαλύτερη είναι η ένταση της διάχυσης, με αποτέλεσμα οι ισοϋψείς να μην κόβονται τελικά απότομα και να γίνονται οι επιθυμητές συνδέσεις μεγάλου μήκους.

Όπως και στην περίπτωση της μεθόδου TV Inpainting, μπορούμε να θεωρήσουμε και την περίπτωση που το διαθέσιμο μέρος της εικόνας $\Omega \setminus \mathbb{D}$ περιέχει θόρυβο. Για να απομακρυνθεί ο θόρυβος, παράλληλα με την διάχυση (2.11) στην inpainting περιοχή \mathbb{D} , χρησιμοποιείται η ΜΔΕ της Ολικής Μεταβολής έξω από την \mathbb{D} , οπότε προκύπτει η ΜΔΕ:

$$\begin{cases} \frac{\partial u}{\partial t} = \operatorname{div} \left(\frac{1}{\|\nabla u\|} \nabla u \right) + \lambda(u - u^0) & , \mathbf{x} \in \Omega \setminus \mathbb{D} \\ \frac{\partial u}{\partial t} = \operatorname{div} \left(\frac{f(|\kappa|)}{\|\nabla u\|} \nabla u \right) & , \mathbf{x} \in \mathbb{D} \end{cases} \quad (2.12)$$

όπου λ είναι όπως και προηγουμένως (βλ. ενότητες 1.4 και 2.3) ο όρος πιστότητας, ο οποίος εξαρτάται από την ισχύ του θορύβου (αντιστρόφως ανάλογα) και ελέγχει το μέγεθος απομάκρυνσης της u από την αρχική θορυβώδη u^0 .

Παρόλη τη βελτιωμένη συμπεριφορά της μεθόδου CDD, σε σχέση με τη μέθοδο TV Inpainting, ένα μειονέκτημα του TV Inpainting συνεχίζει να υπάρχει και σε αυτήν: οι σπασμένες ισοϋψείς ενώνονται και εδώ με ευθύγραμμα τμήματα.

2.5 Euler's Elastica Inpainting

Οι *Euler's Elastica* μελετήθηκαν για πρώτη φορά από τον Euler το 1744 για την μοντελοποίηση του σχήματος ισορροπίας μιας λεπτής ράβδου, απουσία στρέψης. Ο Mumford [38] ήταν ο πρώτος που εισήγαγε αυτή την κλάση ομαλών καμπυλών (ή αλλιώς μη γραμμικά *splines*) στην Όραση Υπολογιστών. Αργότερα, οι Masnou και Morel [37] εφάρμοσαν τις καμπύλες Elastica στο πρόβλημα inpainting, για να ανακατασκευάσουν τις σπασμένες ισοϋψείς. Συγκεκριμένα, η μέθοδός τους χρησιμοποιεί έναν αλγόριθμο δυναμικού προγραμματισμού για να προκύψουν ζεύγη από pixels στο σύνορο της περιοχής inpainting και στη συνέχεια ενώνει κάθε ζεύγος με μία καμπύλη Elastica. Τέλος, οι Chan, Kang και Shen [12] μελέτησαν το μοντέλο αυτό για το inpainting από τη σκοπία της ελαχιστοποίησης συναρτησιακού και κατέληξαν σε μία μέθοδο ΜΔΕ, που θα παρουσιάσουμε στην ενότητα αυτή και η οποία κατά μία έννοια, ενοποιεί (για βαθμωτές εικόνες) τις μεθόδους των προηγούμενων παραγράφων.

Οι Chan, Kang και Shen χρησιμοποίησαν την ακόλουθη γενική ενέργεια,

βασισμένη στην καμπυλότητα κ :

$$J[u] = \iint_{\mathbb{D}} \phi(\kappa) \|\nabla u\| \, d\mathbf{x} \quad (2.13)$$

όπου $\phi(s)$ είναι μία ομαλή, όχι αύξουσα συνάρτηση του $|s|$, με $\phi(s \rightarrow \pm\infty) = +\infty$. Η ειδική περίπτωση που χρησιμοποίησαν οι συγγραφείς είναι η ενέργεια της Euler's Elastica, που προκύπτει όταν:

$$\phi(s) = a + bs^2, \quad a, b > 0 \quad (2.14)$$

Από τις Euler-Lagrange Εξισώσεις, η ελαχιστοποίηση της ενέργειας (2.13) οδηγεί στην ΜΔΕ (βλ. [12] για την απόδειξη):

$$\frac{\partial u}{\partial t} = \operatorname{div}(\mathbf{V}), \quad \mathbf{V} = \phi(\kappa)\mathbf{n} - \frac{1}{\|\nabla u\|} \frac{\partial(\phi'(\kappa)\|\nabla u\|)}{\partial \boldsymbol{\xi}} \boldsymbol{\xi} \quad (2.15)$$

όπου $\mathbf{n} = \frac{\nabla u}{\|\nabla u\|}$ και $\boldsymbol{\xi} = \mathbf{n}^\perp$ το κάθετο και εφαπτομενικό, αντίστοιχα, μοναδιαίο διάνυσμα στην ισοϋψή καμπύλη που διέρχεται από το κάθε σημείο. Το διανυσματικό πεδίο \mathbf{V} είναι το πεδίο ροής του μηχανισμού διάχυσης, εάν αυτός ειδωθεί σαν μεταφορά ύλης.

Η ροή \mathbf{V} μπορεί να αναλυθεί σε μία ροή \mathbf{V}_n , κάθετα στις ισοϋψές καμπύλες και μία ροή \mathbf{V}_ξ , παράλληλα σε αυτές, δηλαδή:

$$\mathbf{V} = \mathbf{V}_n + \mathbf{V}_\xi$$

όπου οι δύο αυτές συνιστώσες της \mathbf{V} δίνονται από τις σχέσεις:

$$\begin{aligned} \mathbf{V}_n &= \phi(\kappa)\mathbf{n}, \\ \mathbf{V}_\xi &= -\frac{1}{\|\nabla u\|} \frac{\partial(\phi'(\kappa)\|\nabla u\|)}{\partial \boldsymbol{\xi}} \boldsymbol{\xi} \end{aligned}$$

Παρατηρούμε ότι η ροή \mathbf{V}_n αντιστοιχεί ακριβώς στην μέθοδο CDD (βλ. σχέση (2.11) στην ενότητα 2.4), αν επιλεγεί συνάρτηση $g(\kappa) = \phi(\kappa)$. Κι αυτό διότι η συνεισφορά της ροής \mathbf{V}_n στην ταχύτητα της ΜΔΕ (2.15) γράφεται (λόγω του ότι $\mathbf{n} = \frac{\nabla u}{\|\nabla u\|}$):

$$\operatorname{div}(\mathbf{V}_n) = \operatorname{div}\left(\phi(\kappa) \frac{\nabla u}{\|\nabla u\|}\right)$$

Επίσης, μπορεί εύκολα να διαπιστωθεί ότι η απόκλιση της \mathbf{V}_ξ γράφεται ως εξής:

$$\operatorname{div}(\mathbf{V}_\xi) = \|\nabla u\| \boldsymbol{\xi} \cdot \nabla \left(\frac{-1}{\|\nabla u\|^2} \frac{\partial(\phi'(\kappa)\|\nabla u\|)}{\partial \boldsymbol{\xi}} \right)$$

Συγκρίνοντας την σχέση αυτή με την ΜΔΕ (2.1), παρατηρούμε ότι ο όρος $\operatorname{div}(\mathbf{V}_\xi)$ αντιστοιχεί ακριβώς στην μέθοδο BSCB της ενότητας 2.2, με μέτρηση ομαλότητας:

$$L(u) = \frac{-1}{\|\nabla u\|^2} \frac{\partial(\phi'(\kappa) \|\nabla u\|)}{\partial \xi}$$

Συμπεραίνουμε λοιπόν ότι η μέθοδος αυτή συνδυάζει τον μηχανισμό μεταφοράς ομαλότητας της μεθόδου BSCB και τον μηχανισμό της διάχυσης της μεθόδου CDD.

Στην σχέση (2.14), ο όρος a αντιστοιχεί στο TV Inpainting (το οποίο μπορεί να θεωρηθεί σαν ειδική περίπτωση της μεθόδου αυτής με $b = 0$) ενώ ο όρος b αντιστοιχεί σε ένωση των ισουψών με ομαλές καμπύλες. Έτσι, όταν ο λόγος b/a μεγαλώνει, οι ενώσεις γίνονται ομαλότερες ενώ παράλληλα οι συνδέσεις μεγάλου μήκους πραγματοποιούνται πιο εύκολα (λόγω του ότι το κόστος τους μικραίνει). Βλέπουμε λοιπόν, ότι η μέθοδος αυτή ξεπερνάει τα μειονεκτήματα των προηγούμενων μεθόδων.

2.6 Από κοινού παρεμβολή διανυσματικών πεδίων και επιπέδων του γκρι

Με σκοπό το inpainting βαθμωτών εικόνων, οι Ballester, Bertalmio, Caselles, Sapiro και Verdera [3, 4] πρότειναν ένα άλλο πρόβλημα ελαχιστοποίησης, παρόμοιο με την μέθοδο Euler's Elastica, που και αυτό χρησιμοποιεί την καμπυλότητα κ (μία χαλάρωση της συγκεκριμένα). Αν B μία ζώνη που περιβάλλει περιμετρικά την περιοχή inpainting \mathbb{D} και στην οποία η εικόνα u^0 είναι γνωστή, η μέθοδός τους συνίσταται στην ελαχιστοποίηση του συναρτησιακού:

$$E[\boldsymbol{\theta}, u] = \iint_{\mathbb{D} \cup B} |\operatorname{div}(\boldsymbol{\theta})| (\gamma + \beta \|\nabla K_\sigma * u\|) d\mathbf{x} \quad (2.16)$$

υπό τους περιορισμούς:

$$u = u^0, \boldsymbol{\theta} = \boldsymbol{\theta}^0 \text{ στην } B, \quad \|\nabla u\| - \boldsymbol{\theta} \cdot \nabla u = 0 \text{ στην } B \cup \mathbb{D}, \quad \|\boldsymbol{\theta}\| \leq 1, |u| \leq M \quad (2.17)$$

όπου $p > 1, \gamma > 0, \beta \geq 0, \boldsymbol{\theta}^0$ το διανυσματικό πεδίο των κατευθύνσεων της ∇u^0 και M η μέγιστη απόλυτη τιμή της u^0 στη ζώνη B . Ακόμα, K_σ είναι ένας πυρήνας εξομάλυνσης (συνήθως Gaussian) με μικρή τυπική απόκλιση σ , που χρησιμοποιείται για θεωρητικούς λόγους, ώστε η ύπαρξη λύσης του προβλήματος ελαχιστοποίησης να είναι εξασφαλισμένη. Βλέπουμε ότι, εκτός από την εικόνα, θεωρείται άγνωστο στην περιοχή inpainting και το διανυσματικό πεδίο $\boldsymbol{\theta}$, για αυτό η μέθοδος αυτή ονομάζεται *Από κοινού παρεμβολή διανυσματικών*

πεδίων και επιπέδων του γκρι (*Joint interpolation of vector fields and gray levels*).

Το διανυσματικό πεδίο θ αποτελεί γενίκευση του μοναδιαίου διανύσματος, κάθετου στις ισοϋψείς (χωρίς όμως να είναι παντού ίσο με αυτό) και είναι ενδεικτικό της διεύθυνσης των καμπυλών αυτών. Έτσι, ο όρος $\text{div}(\theta)$ αντιστοιχεί στην καμπυλότητα κ . Η προσέγγιση αυτή για το πεδίο θ είναι ένα από τα σημαντικά σημεία της μεθόδου και παρουσιάζει το πλεονέκτημα ότι το πεδίο αυτό εξελίσσεται με μεγαλύτερη ευελιξία και τα επιπεδοσύνολα μπορούν να συνεχιστούν με ομαλό τρόπο στο εσωτερικό της περιοχής \mathbb{D} .

Χρησιμοποιώντας ροή ταχύτερης καθόδου για την ελαχιστοποίηση του (2.16), προκύπτει το εξής σύστημα δύο συζευγμένων ΜΔΕ (μία για το θ και μία για την u):

$$\begin{cases} \frac{\partial \theta}{\partial t} = -\nabla_{\theta} E[\theta, u] \\ \frac{\partial u}{\partial t} = -\nabla_u E[\theta, u] \end{cases}, \quad \mathbf{x} \in B \cup \mathbb{D}$$

συμπληρωμένο με τα αντίστοιχα συνοριακά δεδομένα και αρχικές συνθήκες. Οι περιορισμοί (2.17) μπορούν να ενσωματωθούν στην διαδικασία είτε προσθέτοντας όρους ποινής στο συναρτησιακό, είτε με προβολή μετά από κάθε χρονικό βήμα.

2.7 Inpainting με την Ανισοτροπική Εξομάλυνση των Tschumperlé-Deriche

Μία άλλη μέθοδος inpainting διανυσματικών εικόνων προτάθηκε από τους Tschumperlé και Deriche [52, 55], οι οποίοι εφάρμοσαν την ανισοτροπική διάχυση που εισήγαγαν (βλ. ενότητα 1.10) και στο πρόβλημα αυτό. Για την εφαρμογή αυτή, επέλεξαν η εξομάλυνση να γίνεται μόνο κατά την διεύθυνση των ακμών θ_-^* (που αντιστοιχεί σε επιλογή $f_+ = 0$), οπότε η ΜΔΕ (1.67) γίνεται:

$$\frac{\partial \mathbf{u}}{\partial t} = f_- \left(\sqrt{\lambda_+^* + \lambda_-^*} \right) \frac{\partial^2 \mathbf{u}}{\partial \theta_-^{*2}}, \quad \mathbf{x} \in \mathbb{D} \quad (2.18)$$

όπου, κατά τα γνωστά, λ_+^*, λ_-^* και θ_+^*, θ_-^* οι ιδιοτιμές και τα ιδιοδιανύσματα του δομικού τανυστή:

$$J_{\rho}(\nabla \mathbf{u}) = K_{\rho} * \sum_{i=1}^N \nabla u_i (\nabla u_i)^T \quad (2.19)$$

Μία κατάλληλη επιλογή για την $f_-(s)$ είναι και πάλι η εξής (βλ. [52]):

$$f_-(s) = \frac{1}{\sqrt{1 + s^2}}$$

Όπως αναφέρεται στο [52], η ΜΔΕ (2.18) εκτελεί εξομάλυνση μόνο κατά την διεύθυνση των ακμών θ_-^* με σκοπό να επανενώνονται οι σπασμένες δομές της διανυσματικής εικόνας. Έτσι, σε αντίθεση με την εφαρμογή της μεθόδου για την απομάκρυνση του θορύβου, εδώ δεν χρησιμοποιείται καθόλου η διεύθυνση θ_+^* , διότι υπάρχει ο κίνδυνος να προκληθεί θόλωμα των δομών, εντός της περιοχής inpainting \mathbb{D} .

Η εξομάλυνση (2.19) στον τανυστή για τον υπολογισμό της νόρμας $\mathcal{N}_+^* = \sqrt{\lambda_+^* + \lambda_-^*}$ έχει το πλεονέκτημα ότι, σε κάθε σημείο λαμβάνεται υπόψιν και η γύρω περιοχή του, με αποτέλεσμα η συμπλήρωση του διανύσματος \mathbf{u} στην \mathbb{D} να είναι πιο συνεκτική. Επίσης, η εξομάλυνση αυτή δίνει έμφαση στις έντονες δομές της εικόνας και βοηθάει την διαδικασία inpainting. Έτσι, η επιλογή της παραμέτρου ρ εξαρτάται από το μέγεθος των κομματιών που λείπουν.

Μειονέκτημα της μεθόδου είναι ότι η διάχυση μόνο κατά την διεύθυνση θ_-^* (παρόλο που είναι πιο αποτελεσματική από την διάχυση και κατά τις δύο κατευθύνσεις) εμφανίζει και αυτή κάποια προβλήματα. Σε κάποιες περιπτώσεις, οι συνδέσεις των δομών της εικόνας δεν γίνονται με ικανοποιητικό τρόπο και δημιουργούνται ψεύτικες ακμές. Ακόμα, στην περίπτωση που λείπουν κομμάτια με μεγάλο πάχος, για την επιθυμητή επανένωση των δομών της εικόνας, απαιτείται μεγάλη τιμή της παραμέτρου σ , κάτι το οποίο όμως οδηγεί σε αποτέλεσμα υπερβολικής συνοχής που φαίνεται έντονα τεχνητό.

2.8 Πρόταση νέων ποσοτήτων για τον χειρισμό των γεωμετρικών δομών των διανυσματικών εικόνων και εφαρμογή τους στο Inpainting

Στις ενότητες που προηγήθηκαν, παρουσιάσαμε αρκετές από τις υπάρχουσες μεθόδους ΜΔΕ διάχυσης για Inpainting βαθμωτών και διανυσματικών εικόνων. Από την παρουσίαση αυτή, γίνεται φανερό ότι το πρόβλημα Inpainting έχει διερευνηθεί σε πολύ μεγαλύτερο βαθμό για τις βαθμωτές εικόνες σε σχέση με τις διανυσματικές εικόνες, κάτι το οποίο έχει οδηγήσει σε περισσότερες και πιο αποτελεσματικές μεθόδους. Για παράδειγμα, οι μέθοδοι *Διάχυση Οδηγούμενη από την Καμπυλότητα (CDD)* και *Euler's Elastica Inpainting*, όπως είδαμε, χρησιμοποιούν για την ανακατασκευή την χρήσιμη μέτρηση της καμπυλότητας των ισοϋψών γραμμών. Με τον τρόπο αυτό, οι ακμές της εικόνας συνεχίζονται στην περιοχή που λείπει με τρόπο ομαλό και σύμφωνο με την οπτική αντίληψη του ανθρώπου. Συμπεραίνουμε λοιπόν ότι θα ήταν επιθυμητό μέθοδοι με αντίστοιχη συμπεριφορά να αναπτυχθούν και για τις διανυσματικές εικόνες.

Με κίνητρο λοιπόν την ανάπτυξη τέτοιων μεθόδων για τις διανυσματικές εικόνες, προτείνουμε στην ενότητα αυτή μια γενίκευση στις διανυσματικές εικόνες της έννοιας της καμπυλότητας των ισοϋψών γραμμών. Στη συνέχεια ασχολούμαστε και με μία εφαρμογή της ιδέας αυτής, αναπτύσσοντας μία μέθοδο που γενικεύει την καμπυλότητα αυτή. Πρέπει να σημειώσουμε ότι η γενίκευση αυτή δεν θα μπορούσε να γίνει άμεσα, μιας και οι διανυσματικές εικόνες ($N \geq 2$) δεν έχουν ισοϋψείς γραμμές. Κι αυτό διότι για τις διανυσματικές εικόνες του παραγματικού κόσμου (π.χ. έγχρωμες φωτογραφίες) το σύνολο των σημείων με σταθερή τιμή δεν είναι πλέον καμπύλη, αλλά μεμονωμένα σημεία. Επιπρόσθετα, αν ακολουθήσουμε μία άμεση και προφανή αντιμετώπιση του προβλήματος Inpainting διανυσματικών εικόνων, εφαρμόζοντας κάποια μέθοδο βασισμένη στην καμπυλότητα σε κάθε κανάλι της χωριστά, τότε τα αποτελέσματα δεν είναι καθόλου ικανοποιητικά. Συγκεκριμένα, υπάρχει ο κίνδυνος σε κάθε κανάλι οι ακμές να ανακατασκευαστούν με διαφορετικό τρόπο (αφού δεν λαμβάνονται υπόψη όλα τα υπόλοιπα κανάλια) με αποτέλεσμα να παρουσιάζεται μία σύγχυση των γεωμετρικών δομών στη τελική διανυσματική εικόνα. Για τους παραπάνω λόγους, χρειάζεται ένας πιο περίπλοκος χειρισμός της μεταβολής και των γεωμετρικών δομών της διανυσματικής εικόνας, για να μπορέσει να γενικευθεί η χρήση της καμπυλότητας, όπως αυτός που προτείνουμε στη συνέχεια.

2.8.1 Καμπύλες Ελάχιστης Μεταβολής και Γενικευμένη Καμπυλότητα διανυσματικών εικόνων

Όπως έχουμε αναφέρει, οι ισοϋψείς γραμμές μιας βαθμωτής εικόνας u έχουν σε κάθε σημείο σαν κάθετο μοναδιαίο διάνυσμα το:

$$\mathbf{n} = \frac{\nabla u}{\|\nabla u\|}$$

Επίσης, το μοναδιαίο διάνυσμα $\boldsymbol{\xi} = \mathbf{n}^\perp$ είναι εφαπτόμενο στις γραμμές αυτές και σε κάθε σημείο ορίζει την διεύθυνση στην οποία η εικόνα παραμένει σταθερή.

Για μία γενίκευση των παραπάνω στην περίπτωση μιας διανυσματικής εικόνας \mathbf{u} , χρησιμοποιούμε τη μήτρα *Di Zenzo* (βλ. ενότητα 1.1.4), η οποία λαμβάνει υπόψη της όλα τα κανάλια της εικόνας και ορίζεται από τη σχέση:

$$G = \sum_{i=1}^N \nabla u_i (\nabla u_i)^T \quad (2.20)$$

Όπως έχουμε σημειώσει (βλ. ενότητα 1.1.4), το μοναδιαίο ιδιοδιάνυσμα $\boldsymbol{\theta}_+$ της μήτρας G ορίζει την διεύθυνση της μέγιστης διανυσματικής μεταβολής της εικόνας, ενώ το μοναδιαίο ιδιοδιάνυσμα $\boldsymbol{\theta}_-$ την διεύθυνση της ελάχιστης μεταβολής.

Αποτελούν συνεπώς γενίκευση στις διανυσματικές εικόνες των διανυσμάτων \mathbf{n} και ξ . Πρέπει να τονιστεί ότι για $N \geq 2$, συνήθως έχουμε μεταβολή ακόμα και στην διεύθυνση θ_- , σε αντίθεση με την περίπτωση $N = 1$ (βαθμωτές εικόνες), όπου η μεταβολή αυτή είναι 0. Αυτό οφείλεται στο ότι εμπλέκονται πολλά κανάλια στην διαμόρφωση του διανύσματος \mathbf{u} , κάθε ένα από τα οποία μπορεί να έχει στο σημείο αυτό τη δική του διεύθυνση μηδενικής μεταβολής. Τα παραπάνω είναι και ο λόγος που, όπως είπαμε, η σταθερή τιμή του διανύσματος \mathbf{u} δεν ορίζει, στις περισσότερες περιπτώσεις, καμπύλη.

Με βάση τα παραπάνω, ορίζουμε καταρχήν μία οικογένεια καμπυλών για μία διανυσματική εικόνα, που αποτελεί γενίκευση των ισοϋψών γραμμών μιας βαθμωτής εικόνας. Η γενίκευση αυτή στηρίζεται στην παρατήρηση ότι, αφού το διάνυσμα ξ είναι σε κάθε σημείο εφαπτόμενο στις ισοϋψείς γραμμές, οι γραμμές αυτές μπορούν να ειδικωθούν σαν οι δυναμικές γραμμές του διανυσματικού πεδίου ξ ². Έτσι, κάνουμε τον εξής ορισμό:

Ορισμός 2.1. *Ορίζουμε σαν Καμπύλες Ελάχιστης Διανυσματικής Μεταβολής μιας διανυσματικής εικόνας \mathbf{u} τις δυναμικές γραμμές του μοναδιαίου ιδιοδιανύσματος θ_- της μήτρας $Di\ Zenzo\ G$, που αντιστοιχεί στη μικρότερη ιδιοτιμή λ_- .*

Οι καμπύλες αυτές είναι χρήσιμες, διότι, όπως οι ισοϋψείς γραμμές των βαθμωτών εικόνων, είναι ενδεικτικές της γεωμετρικής δομής της διανυσματικής εικόνας. Έτσι, στις διανυσματικές ακμές και κοντά σε αυτές, οι καμπύλες αυτές παίρνουν το σχήμα των ακμών και ακολουθούν την μεταβολή τους.

Επομένως, με τη βοήθεια του ορισμού αυτού, μπορούμε τώρα εύκολα να γενικεύσουμε και την καμπυλότητα των ισοϋψών γραμμών, με τέτοιο τρόπο, ώστε η χρησιμότητα της ποσότητας αυτής να μεταφέρεται και στις διανυσματικές εικόνες. Παρατηρούμε όμως πρώτα ότι, όπως ευκολα μπορεί να διαφανεί, το διάνυσμα θ_+ είναι σε κάθε σημείο κάθετο στην Καμπύλη Ελάχιστης Διανυσματικής Μεταβολής που διέρχεται από το σημείο. Έτσι, προχωράμε στον ακόλουθο ορισμό:

Ορισμός 2.2. *Ορίζουμε σαν Γενικευμένη Καμπυλότητα μιας διανυσματικής εικόνας \mathbf{u} την καμπυλότητα των Καμπύλων Ελάχιστης Διανυσματικής Μεταβολής της \mathbf{u} και τη συμβολίζουμε σαν $\hat{\kappa}$. Η $\hat{\kappa}$ δίνεται από τη σχέση:*

$$\hat{\kappa} = \operatorname{div}(\theta_+) \quad (2.21)$$

όπου θ_+ το μοναδιαίο ιδιοδιάνυσμα της μήτρας $Di\ Zenzo\ G$, που αντιστοιχεί στη μεγαλύτερη ιδιοτιμή λ_+ .

²οι δυναμικές γραμμές ενός διανυσματικού πεδίου $\mathbf{F}(\mathbf{x})$ ορίζονται σαν η οικογένεια καμπυλών, της οποίας κάθε καμπύλη $\mathbf{r}(t)$ και σε κάθε σημείο ικανοποιεί την: $\mathbf{r}'(t) \parallel \mathbf{F}(\mathbf{r}(t))$

Σημειώνουμε ότι αν θ είναι η γωνία του μοναδιαίου διανύσματος θ_+ , δηλαδή ισχύει: $\theta_+ = [\cos \theta \quad \sin \theta]^T$, τότε η σχέση (2.21) μπορεί να γραφεί:

$$\hat{k} = -\sin \theta \frac{\partial \theta}{\partial x} + \cos \theta \frac{\partial \theta}{\partial y} \quad (2.22)$$

Όσον αφορά τις εκφράσεις (2.21),(2.22), πρέπει να σημειώσουμε τα εξής: Όπως έχουμε αναφέρει (βλ. ενότητα 1.1.4), το μοναδιαίο διάνυσμα θ_+ εκφράζει την διεύθυνση της μέγιστης διανυσματικής μεταβολής της εικόνας \mathbf{u} . Επομένως, σε κάθε σημείο αντιστοιχούν δύο μοναδιαία διανύσματα, το θ_+ και το $-\theta_+$, αφού αυτά εκφράζουν την ίδια διεύθυνση. Το γεγονός κάνει το διανυσματικό πεδίο θ_+ πλειότιμο, με δύο τιμες σε κάθε σημείο. Επομένως, και η γωνία θ του διανύσματος αυτού, που ούτως ή άλλως είναι πλειότιμη modulo 2π , γίνεται πλειότιμη modulo π .³

Έτσι, θεωρώντας ότι η εικόνα \mathbf{u} δεν παρουσιάζει ασυνέχειες (κάτι που μπορεί εύκολα να επιτευχθεί, ακόμα και στη περίπτωση ασυνεχειών, μέσω της εξομάλυνσης με βαθυπερατό πυρήνα K_σ οσοδήποτε μικρής τυπικής απόκλισης σ), για να ορίζεται η \hat{k} πρέπει στην περιοχή κάθε σημείου να κάνουμε τέτοια επιλογή ανάμεσα στα δύο μοναδιαία διανύσματα, ώστε το διανυσματικό πεδίο θ_+ να είναι συνεχές. Το παραπάνω μπορεί να επιτευχθεί ευκολότερα, μέσω της γωνίας θ και της έκφρασης (2.22), κάνοντας τέτοια επιλογή για τις τιμές της θ στη γειτονιά κάθε σημείου, ώστε η θ να είναι συνεχής στο σημείο αυτό.

Η ποσότητα της Γενικευμένης Καμπυλότητας που εισάγαμε είναι, όπως είπαμε, το ίδιο χρήσιμη με την καμπυλότητα των ισούψων γραμμών. Έτσι, μπορεί να χρησιμεύσει στην αποτελεσματική γενίκευση στις διανυσματικές εικόνες διαφόρων μεθόδων επεξεργασίας βαθμωτών εικόνων. Στην ενότητα που ακολουθεί, θα την χρησιμοποιήσουμε στο πρόβλημα Inpainting και για τη γενίκευση της μεθόδου CDD.

2.8.2 Πρόταση μιας νέας μεθόδου: Γενικευμένη CDD για διανυσματικές εικόνες

Στην ενότητα αυτή γενικεύουμε την μέθοδο *Διάχυσης Οδηγούμενης από την Καμπυλότητα (CDD)* στις διανυσματικές εικόνες. Όπως έχουμε αναφέρει (βλ. ενότητα (2.4)), η CDD τροποποιεί, για την περίπτωση βαθμωτών εικόνων, την ΜΔΕ του TV Inpainting (βλ. ενότητα (2.3)), εισάγοντας έναν όρο $f(|\kappa|)$, ο οποίος εξαρτάται από την καμπυλότητα κ , με σκοπό την αντιμετώπιση της αδυναμίας του TV Inpainting για συνδέσεις μεγάλου μήκους.

³ Διευκρινίζουμε ότι με την έκφραση η συνάρτηση $\theta(\mathbf{x})$ είναι πλειότιμη modulo π , εννοούμε ότι, αν ένα σημείο \mathbf{x} αντιστοιχίζεται στη τιμή θ , τότε μπορεί να αντιστοιχισθεί και στη τιμή $\theta + \nu\pi$, για οποιοδήποτε $\nu \in \mathbb{Z}$.

Η αδυναμία αυτή της μεθόδου Ολικής Μεταβολής (TV) στη βαθμωτή περίπτωση μεταφέρεται και στην διανυσματική περίπτωση (με την Διανυσματική Ολική Μεταβολή - Vectorial TV), που αποτυγχάνει να συνδέσει τις διανυσματικές ακμές, όταν λείπει ένα σχετικά μεγάλο κομμάτι. Αντίστοιχα με την βαθμωτή περίπτωση, μπορούμε να παρατηρήσουμε ότι σε αυτές τις περιπτώσεις αποτυχίας, δημιουργούνται διανυσματικές γωνίες στα σημεία όπου οι ακμές κόβονται, στις οποίες η γενικευμένη καμπυλότητα $\hat{\kappa}$ λαμβάνει πολύ υψηλή τιμή. Για τον σκοπό αυτό, τροποποιούμε την ΜΔΕ της Διανυσματικής Ολικής Μεταβολής - Vectorial TV PDE (βλ. ενότ. 2.3, σχέση (2.6)), εισάγοντας τον όρο $h(|\hat{\kappa}|)$, που αυξάνεται με την αύξηση της γενικευμένης καμπυλότητας $|\hat{\kappa}|$. Έτσι, προτείνουμε την εξής ΜΔΕ:

$$\frac{\partial u_i}{\partial t} = \operatorname{div} \left(\frac{h(|\hat{\kappa}|)}{\|\nabla \mathbf{u}\|} \nabla u_i \right), \quad \mathbf{x} \in \mathbb{D} \quad (i = 1, \dots, N) \quad (2.23)$$

όπου $\hat{\kappa}$ η γενικευμένη καμπυλότητα, που δίνεται από τον ορισμό 2.2 και $h(s)$, $s \in [0, +\infty)$ συνεχής, αύξουσα συνάρτηση, με $h(0) = 0$ και $h(s \rightarrow +\infty) = \infty$. Μία επιλογή που κάνουμε, παρόμοια με την CDD (βλ. ενότ. (2.4)), είναι η εξής:

$$h(s) = as^p \quad (2.24)$$

όπου a, p θετικές σταθερές.

Στη μέθοδο αυτή, ο συντελεστής διάχυσης $g = h(|\hat{\kappa}|)/\|\nabla \mathbf{u}\|$ ελαττώνεται με την αύξηση της νόρμας $\|\nabla \mathbf{u}\|$ (όταν π.χ. πλησιάζουμε στις ανακατασκευασμένες διανυσματικές ακμές), με τον ίδιο τρόπο όπως στην Vectorial TV PDE, αλλά παράλληλα αυξάνεται με την αύξηση της γενικευμένης καμπυλότητας $\hat{\kappa}$. Έτσι, από τη μία όσο πλησιάζουμε στις ανακατασκευασμένες διανυσματικές ακμές, ο συντελεστής διάχυσης μικραίνει, εμποδίζοντας την καταστροφή τους, όπως συμβαίνει και με την μέθοδο TV Inpainting. Από την άλλη, για σημεία με παρόμοια τιμή της νόρμας $\|\nabla \mathbf{u}\|$ (π.χ. κατά μήκος της ίδιας ακμής), η διάχυση είναι πιο μεγάλη στα σημεία με μεγαλύτερη γενικευμένη καμπυλότητα (όπως π.χ. στις ανεπιθύμητες γωνίες), σε αντίθεση με την μέθοδο TV Inpainting, όπου στη περίπτωση αυτή η διάχυση είναι ίδια.

Συμπεραίνουμε λοιπόν ότι η μέθοδος που εισάγαμε παρουσιάζει το πλεονέκτημα ότι βελτιώνει τη συμπεριφορά της μεθόδου TV Inpainting, αντιμετωπίζοντας την αδυναμία της για συνδέσεις μεγάλου μήκους και διατηρώντας παράλληλα τα θετικά της χαρακτηριστικά. Διαπιστώσαμε αυτή τη βελτίωση στα αποτελέσματα και πειραματικά, όπως θα παρουσιάσουμε στην υποενότητα 6.2.

Κεφάλαιο 3

Ανάλυση μεθόδων ΜΔΕ για την Παρεμβολή Εικόνων

3.1 Ανασκόπηση των μεθόδων Παρεμβολής

Το πρόβλημα της *Παρεμβολής Εικόνων (Image Interpolation)* αποτελεί ένα από τα θεμελιώδη προβλήματα της Επεξεργασίας Εικόνων και της Όρασης Υπολογιστών, μιας και ανακύπτει άμεσα από την ψηφιοποίηση των εικόνων, που είναι απαραίτητη για την επεξεργασία τους με υπολογιστές. Η παρεμβολή μιας διακριτής εικόνας συνίσταται στην εύρεση μιας εικόνας υψηλότερης *Χωρικής Ευκρίνειας (Spatial Resolution)*, δηλαδή μιας διακριτής εικόνας που ορίζεται σε πυκνότερο χωρικό πλέγμα. Το πρόβλημα της Παρεμβολής αποτελεί ουσιαστικά την μεταφορά στις δύο διαστάσεις του προβλήματος αύξησης της συχνότητας δειγματοληψίας για τα μονοδιάστατα σήματα (π.χ. σήματα ήχου).

Σημειώνουμε ότι για την Παρεμβολή Εικόνων χρησιμοποιούνται επίσης στην βιβλιογραφία οι ορολογίες *Μεγέθυνση Εικόνων (Image Magnification)* καθώς και *Image Zoom* (ή *Zooming*).

Το πρόβλημα της Παρεμβολής παρουσιάζει ομοιότητες με τα προβλήματα που παρουσιάσαμε στα προηγούμενα κεφάλαια και ιδιαίτερα με το Inpainting. Όπως στο Inpainting προσπαθούμε να ανακατασκευάσουμε την εικόνα στην περιοχή που λείπει, βασιζόμενοι στην πληροφορία της εικόνας γύρω από την περιοχή αυτή, έτσι και στο πρόβλημα αυτό, προσπαθούμε να αυξήσουμε την χωρική ευκρίνεια, βασιζόμενοι στην πληροφορία χαμηλής χωρικής ευκρίνειας. Και στις δύο περιπτώσεις δηλαδή, ζητούμενο είναι, από την υπάρχουσα πληροφορία της αρχικής εικόνας, να καταλήξουμε σε μία εικόνα μεγαλύτερης πληροφορίας.

Για την ανάπτυξη συνεχών μοντέλων αντιμετώπισης της Παρεμβολής, το πρόβλημα μπορεί να θεωρηθεί ότι αποτελείται από δύο υποδιαδικασίες: Στη πρώτη φάση ζητούμενο είναι η εύρεση μιας συνεχούς εικόνας, από την οποία

είναι δυνατό να προήλθε η αρχική διακριτή εικόνα. Πρόκειται δηλαδή για την εύρεση της λύσης ενός αντίστροφου προβλήματος. Στη δεύτερη φάση, εκτελείται δειγματοληψία αυτής της συνεχούς εικόνας στο τελικό πυκνότερο πλέγμα. Βέβαια, πρέπει να σημειωθεί ότι, τελικά στην αριθμητική υλοποίηση, δεν βρίσκεται η συνεχής εικόνα, παρά μόνο μία διακριτή προσέγγισή της, με αποτέλεσμα οι δύο αυτές υποδιαδικασίες να μπορούν να ενσωματωθούν σε μία διαδικασία.

Για την επίλυση του προβλήματος της Παρεμβολής, έχει αναπτυχθεί μία ποικιλία μεθόδων, οι οποίες μπορούν να χωρισθούν σε δύο μεγάλες κατηγορίες: τις γραμμικές και τις μη γραμμικές:

- Οι περισσότερες γραμμικές μέθοδοι θεωρούν σαν ιδανική ανακατασκευή την συνεχή εικόνα, που προκύπτει αν συμπληρώσουμε με 0 (zero-padding) το χαμένο (λόγω δειγματοληψίας) φασματικό περιεχόμενο της αρχικής εικόνας. Έτσι, σύμφωνα με το Θεώρημα της Δειγματοληψίας του Shannon [43], η συνεχής αυτή εικόνα προκύπτει από την διακριτή, μέσω συνέλιξης με πυρήνα την 2D συνάρτηση δειγματοληψίας (2D sinc). Για τη μέθοδο αυτή, χρησιμοποιείται η ονομασία *ιδανική παρεμβολή*, διότι η συνέλιξη με τον πυρήνα αυτό δεν δημιουργεί *aliasing*.

Όμως, η μέθοδος αυτή παρουσιάζει μεγάλη αριθμητική πολυπλοκότητα, διότι η συνάρτηση sinc που χρησιμοποιεί σαν πυρήνα, εκτείνεται υπερβολικά στο χώρο και ταλαντώνεται. Για τον λόγο αυτό, έχουν προταθεί κάποιες μέθοδοι που μειώνουν την πολυπλοκότητα του υπολογισμού της παραπάνω συνέλιξης (βλ. [67]) και αρκετές άλλες που χρησιμοποιούν προσέγγιση της συνάρτησης sinc (βλ. π.χ. [39, 57]). Επίσης, βασιζόμενες σε παρόμοια λογική, υπάρχουν πολλές γραμμικές μέθοδοι που χρησιμοποιούν συνέλιξη της αρχικής εικόνας με άλλους πυρήνες συνέλιξης. Οι πιο διαδεδομένες από αυτές είναι για παράδειγμα οι μέθοδοι μεγέθυνσης *bicubic*, *quadratic* και *b-spline* (βλ. π.χ. [29], [23] και [56, 47] αντίστοιχα).

Όλες οι γραμμικές μέθοδοι χειρίζονται ουσιαστικά την αρχική εικόνα στο πεδίο των συχνοτήτων εκτελώντας βαθυπερατό φιλτράρισμα. Αυτό όμως αποτελεί και ένα σημαντικό μειονέκτημά τους, μιας και καμμία τέτοια μέθοδος δεν μπορεί να αποφύγει συγχρόνως το θόλωμα αλλά και το *aliasing*, κάτι που γίνεται ιδιαίτερα αισθητό στην περίπτωση ισχυρών ακμών. Για παράδειγμα, ακόμη και η ιδανική παρεμβολή (όπως και οι μέθοδοι που την προσεγγίζουν), εμφανίζει σε αρκετές περιπτώσεις θολωμένα αποτελέσματα καθώς και ανεπιθύμητες ταλαντώσεις κοντά στις ισχυρές ακμές (*Gibbs effects*), λόγω του περιορισμένου εύρους και της απότομης αποκοπής στο φασματικό περιεχόμενο της τελικής εικόνας. Για μία αναλυτική μελέτη της συμπεριφοράς και των μειονεκτημάτων των μεθόδων αυτών, παραπέμπουμε στα [35, 36].

- Οι μη γραμμικές μέθοδοι προσαρμόζουν την ανακατασκευή που κάνουν, σύμφωνα με την τοπική συμπεριφορά της αρχικής εικόνας (βλ. π.χ. [1, 2, 50]). Με τον τρόπο αυτό, φέρονται με διαφορετικό τρόπο στις ομογενείς περιοχές και στις περιοχές κοντά στις ακμές. Έτσι, οι μέθοδοι αυτές δίνουν καλά αποτελέσματα και αντιμετωπίζουν τα μειονεκτήματα των γραμμικών μεθόδων. Σε αυτή την κατηγορία ανήκουν και οι μέθοδοι που χρησιμοποιούν ΜΔΕ διάχυσης, με τις οποίες θα ασχοληθούμε στη συνέχεια και στο υπόλοιπο του κεφαλαίου αυτού.

3.1.1 Προσέγγιση του προβλήματος με την Συνθήκη Αναστρεψιμότητας (Reversibility Condition)

Για την ανάπτυξη μιας μεθόδου βασισμένης στην ελαχιστοποίηση συναρτησιακού, οι Malgouyres και Guichard [26, 35, 36] προτείνουν την εξής προσέγγιση του προβλήματος: Η αρχική διακριτή εικόνα χαμηλής χωρικής ευκρίνειας $z(n_1, n_2)$ θεωρείται ότι έχει προέλθει από μία διακριτή εικόνα $u(n_1, n_2)$ μεγαλύτερης χωρικής ευκρίνειας (έστω κατά τον αθέρατο παράγοντα $d \geq 2$), με μία διαδικασία δύο σταδίων: Αρχικά, εφαρμόζεται στην $u(n_1, n_2)$ βαθυπερατό φίλτρο (συνέλιξη με κάποιον πυρήνα s , για την αποφυγή του aliasing) και στη συνέχεια υποδειγματοληψία. Έτσι, η διαδικασία αυτή μπορεί να εκφραστεί ως εξής:

$$z = Q(s * u) \quad (3.1)$$

όπου $Q(\cdot)$ ο τελεστής υποδειγματοληψίας, για την μετάβαση από το διακριτό ορθογωνικό πλέγμα της $u(n_1, n_2)$ στο αραιότερο πλέγμα της $z(n_1, n_2)$. Δηλαδή, αν το πλέγμα αυτό έχει χωρικά βήματα $h_{z,x}$ και $h_{z,y}$, ο τελεστής $Q(\cdot)$ ορίζεται ως εξής: $Q(\tilde{u})(n_1, n_2) = \tilde{u}(n_1 d, n_2 d)$. Επίσης, θεωρούμε ότι ο πυρήνας s έχει αρκετά μικρή χωρική έκταση (τέτοια ώστε το βαθυπερατό φίλτρο να μην έχει υπερβολικά μικρό εύρος ζώνης και έτσι να μην θολώνει υπερβολικά την u).

Έχοντας κάνει την παραδοχή αυτή, η λύση $u(n_1, n_2)$ του προβλήματος της Παρεμβολής αναζητείται μόνο ανάμεσα στις συναρτήσεις που ικανοποιούν τον περιορισμό (3.1), ο οποίος ονομάζεται *Συνθήκη Αναστρεψιμότητας (Reversibility Condition)*. Με άλλα λόγια, στην προσέγγιση αυτή, το πρόβλημα Παρεμβολής συνίσταται στην εύρεση μίας κατάλληλης λύσης της Συνθήκης Αναστρεψιμότητας (3.1).

Πρέπει να σημειωθεί ότι στην παρουσίαση αυτή, ασχολούμαστε μόνο με την περίπτωση όπου οι $u(x, y)$, $z(n_1, n_2)$ ορίζονται εντός ενός ορθογωνικού χωρίου Ω και θεωρούμε ότι επεκτείνονται περιοδικά στο υπόλοιπο επίπεδο (έστω ότι το πλέγμα ορισμού της $z(n_1, n_2)$ μέσα στο Ω είναι διαστάσεων $M_x \times M_y$). Έτσι, μπορούμε ισοδύναμα να θεωρήσουμε ότι η συνέλιξη στην συνθήκη (3.1) είναι *κυκλική συνέλιξη (circular convolution)* (βλ. π.χ. [47] για τον ορισμό της).

Σημειώνουμε επίσης ότι, για την θεωρητική ανάλυση της Συνθήκης Αναστρεψιμότητας και των μεθόδων που βασίζονται σε αυτή, θα μπορούσε να θεωρηθεί (βλ. π.χ. [35, 36]) ότι η άγνωστη, αρχική συνάρτηση u είναι συνεχής και όχι διακριτή. Έτσι, η εύρεση της συνεχούς λύσης $u(x, y)$ μπορεί να θεωρηθεί σαν το πρώτο στάδιο της Παρεμβολής, με δεύτερο στάδιο τη δειγματοληψία της στο πυκνότερο πλέγμα, καταλήγοντας και πάλι σε μία διακριτή εικόνα υψηλότερης ευκρίνειας. Στην παρουσίαση που ακολουθεί, θεωρούμε (όπως στο [26]) ότι η συνάρτηση u είναι διακριτή, αλλά στην περίπτωση όπου η u θεωρείται συνεχής, μπορούν να εξαχθούν αρκετά παρόμοια αποτελέσματα.

Επιλογή του Πυρήνα της Συνθήκης Αναστρεψιμότητας

Πρέπει να σημειωθεί ότι ο πυρήνας s στη συνθήκη αναστρεψιμότητας (3.1) επιλέγεται από την αρχή και επηρεάζει οποιαδήποτε μέθοδο βασίζεται στη συνθήκη αυτή, αφού ουσιαστικά καθορίζει το σύνολο των συναρτήσεων στο οποίο θα αναζητηθεί η λύση.

Όπως αναφέρεται στα [35, 36], δύο απλοί πυρήνες που καταλήγουν σε καλά αποτελέσματα είναι η 2Δ sinc συνάρτηση (που αντιστοιχεί στο Θεώρημα της Δειγματοληψίας του Shannon) και η κανονικοποιημένη κυβική συνάρτηση (2Δ παλμός), που αντιστοιχεί στην εύρεση της μέσης τιμής στην έκταση του τελικού pixel. Βέβαια, η 2Δ sinc συνάρτηση, όπως είπαμε, δεν είναι εύκολο να υλοποιηθεί στη πράξη, οπότε η επιλογή αυτή έχει περισσότερο θεωρητική αξία.

Στην πράξη λοιπόν, επιλέγουμε τον πυρήνα της μέσης τιμής (2Δ παλμός) καθώς και άλλους πυρήνες που περιορίζονται στην έκταση του τελικού pixel. Αυτή η επιλογή απλοποιεί, όπως θα δούμε στην ενότητα 5.4, την αριθμητική υλοποίηση αλλά επιπλέον οδηγεί σε αξιόπιστα αποτελέσματα. Ακόμα, είναι αρκετά αντιπροσωπευτική της πραγματικής διαδικασίας της ψηφιοποίησης, εφόσον αυτή δεν εισάγει κάποιο σημαντικό θόλωμα. Κι αυτό διότι η τιμή σε κάθε pixel δεν επηρεάζεται από το κεντρικό σημείο, αλλά από μία περιοχή γύρω από το σημείο αυτό, μέσω μίας τοπικής “ολοκλήρωσης”. Επιπρόσθετα, εφόσον η ποιότητα της ψηφιοποίησης είναι καλή, οι παραπάνω περιοχές των pixel δεν αλληλοεπικαλύπτονται. Για παράδειγμα, τα παραπάνω ισχύουν όντως στις ψηφιακές φωτογραφικές μηχανές.

Από την άλλη πλευρά όμως, δεν είναι επιθυμητό ούτε ο πυρήνας s να παρουσιάζει πολύ μικρή έκταση, σε σχέση με το τελικό pixel. Κι αυτό διότι, στην περίπτωση αυτή, υπάρχει μία υπερβολικά έντονη διαφορά στον χειρισμό ανάμεσα στην περιοχή γύρω από το κέντρο του κάθε pixel, όπου οι περιορισμοί γίνονται πολύ αυστηροί, και στην υπόλοιπη περιοχή, όπου οι περιορισμοί γίνονται πολύ χαλαροί ή και μηδενικοί. Λόγω της διαφοράς αυτής, υπάρχει ο κίνδυνος το αποτέλεσμα να μην είναι ικανοποιητικό.

Περαιτέρω Ανάλυση της Συνθήκης Αναστρεψιμότητας

Έστω Z και U οι διακριτοί μετασχηματισμοί Fourier (DFT) των z και u αντίστοιχα. Επίσης, όσον αφορά την αντίστοιχη εύρεση του φασματικού περιεχομένου του πυρήνα s , σημειώνουμε τα εξής: Σύμφωνα με όσα έχουμε υποθέσει, οι μη μηδενικές τιμές του πυρήνα s καταλαμβάνουν μία περιοχή πολύ μικρή, σε σχέση με τις διαστάσεις της εικόνας u , με αποτέλεσμα να μπορούμε χωρίς σφάλμα να θεωρήσουμε ότι ο πυρήνας s έχει τις ίδιες διαστάσεις $(dM_x) \times (dM_y)$ με την u . Έστω λοιπόν S ο DFT του s , όταν αυτός θεωρείται σαν διδιάστατο διακριτό σήμα διαστάσεων $(dM_x) \times (dM_y)$.

Με βάση τους παραπάνω συμβολισμούς, η συνθήκη (3.1) μπορεί ισοδύναμα να γραφεί στο πεδίο των συχνοτήτων ως εξής, λόγω της διακριτής μορφής του Θεωρήματος της Δειγματοληψίας (που αφορά δηλαδή υποδειγματοληψία):

$$Z(m_1, m_2) = \frac{1}{d^2} \sum_{\ell_1=0}^{d-1} \sum_{\ell_2=0}^{d-1} S(m_1 + \ell_1 M_x, m_2 + \ell_2 M_y) U(m_1 + \ell_1 M_x, m_2 + \ell_2 M_y) \quad (3.2)$$

όπου m_1, m_2 ακέραιοι.

Αναφερόμενοι στην σχέση (3.2), πρέπει να διευκρινίσουμε ότι οι μετασχηματισμοί Z, U, S έχουν την ίδια έκταση με τα αντίστοιχα σήματα από τα οποία προήλθαν, δηλαδή έκταση $M_x \times M_y$ για τον Z και $(dM_x) \times (dM_y)$ για τους U, S . Επίσης, κατά τα γνωστά της Θεωρίας Σημάτων, μπορούμε (λόγω της χωρικής διακριτοποίησης) να θεωρήσουμε, όπου χρειάζεται, ότι οι DFT αυτοί επεκτείνονται περιοδικά σε όλο το υπόλοιπο διακριτό επίπεδο συχνοτήτων (βέβαια, ο Z παρουσιάζει διαφορετική περιοδικότητα από τους U, S).

Συμβολίζουμε επίσης με $\mathcal{U}_{z,s}$ το σύνολο των συναρτήσεων u , που ικανοποιούν τη συνθήκη αναστρεψιμότητας (3.1). Όπως αναφέρεται στο [26], το $\mathcal{U}_{z,s}$ αποτελεί έναν αφφινικό (affine) υποχώρο, διότι, όπως εύκολα μπορεί να δείχθει:

$$\forall u_1, u_2 \in \mathcal{U}_{z,s}, a \in \mathbb{R}, \text{ ισχύει: } (1-a)u_1 + au_2 \in \mathcal{U}_{z,s}$$

Μέθοδος Zero-Padding για την ικανοποίηση της Συνθήκης Αναστρεψιμότητας

Για την εύρεση μιας εικόνας u που να ικανοποιεί την συνθήκη αναστρεψιμότητας (3.1), δεδομένης της αρχικής εικόνας z και του πυρήνα s , μία απλή λύση είναι η εξής: Αντίστοιχα με την ιδανική παρεμβολή (βλ. εισαγωγή της ενότητας 3.1), κάνουμε την παραδοχή ότι το φασματικό περιεχόμενο $U(m_1, m_2)$ της u περιορίζεται μόνο στη 2Δ ζώνη συχνοτήτων που αντιστοιχεί στην αρχική διακριτή εικόνα $z(n_1, n_2)$, με μηδενικές τις τιμές σε όλες τις υπόλοιπες συχνότητες. Χρησιμοποιούμε δηλαδή και πάλι zero-padding.

Με βάση την παραδοχή αυτή και εφόσον $S(m_1, m_2) \neq 0$ στην παραπάνω ζώνη συχνοτήτων (κάτι το οποίο για τις περισσότερες πρακτικές περιπτώσεις πυρήνων s είναι αληθές), από την σχέση (3.2) μπορεί άμεσα να προκύψει η ζητούμενη εικόνα u (βλ. ενότητα 5.4.1).

Για την μέθοδο αυτή, θα χρησιμοποιήσουμε (όπως στο [26]) την ορολογία *Μέθοδος Zero-Padding*. Πρέπει να τονιστεί ότι το αποτέλεσμα της μεθόδου αυτής εξαρτάται από την επιλογή του πυρήνα s . Η ιδανική παρεμβολή αποτελεί επομένως την ειδική περίπτωση της μεθόδου αυτής, όταν χρησιμοποιείται ο πυρήνας $s = \delta(x, y)$, δηλαδή όταν γίνεται η υπόθεση ότι η $z(n_1, n_2)$ προκύπτει από απλή δειγματοληψία, χωρίς προηγούμενο βαθυπερατό φιλτράρισμα.

Όπως αναφέρεται στα [35, 36], η μέθοδος zero-padding παρουσιάζει και στη γενική περίπτωση (όχι μόνο στην ιδανική παρεμβολή που είδαμε στην προηγούμενη ενότητα) το σημαντικό μειονέκτημα των ανεπιθύμητων ταλαντώσεων κοντά στις ισχυρές ακμές (*Gibbs effects*). Το φαινόμενο αυτό οφείλεται στην απότομη αποκοπή των συχνοτήτων και το διαπιστώσαμε και στα πειράματά μας (βλ. υποενότητα 6.3). Έτσι, η μέθοδος αυτή δεν προσφέρει μία ικανοποιητική λύση, αλλά η λύση αυτή μπορεί να χρησιμοποιηθεί, όπως θα δούμε στη συνέχεια, σαν μία απλή αρχική προσέγγιση, αφού ικανοποιεί τη συνθήκη αναστρεψιμότητας (3.1).

Από τα παραπάνω συμπεραίνουμε ότι πρέπει να ορισθεί κάποιο επιπλέον κριτήριο ή διαδικασία (πιο κατάλληλο από το απλό zero-padding), που να καταλήγει στην τελική επιλογή της συνάρτησης, ανάμεσα σε αυτές που ικανοποιούν τη συνθήκη (3.1). Με το θέμα αυτό θα ασχοληθούμε στα επόμενα, όπου θα παρουσιάσουμε την Μέθοδο Ολικής Μεταβολής (TV) και κάποιες άλλες μεθόδους και στη συνέχεια θα προτείνουμε μία νέα γενική μέθοδο, βασισμένοι στην προσέγγιση αυτή.

3.1.2 Παρεμβολή Βασισμένη στην Ολική Μεταβολή (Total Variation)

Οι Malgouyres και Guichard [26, 35, 36], αφού εισήγαγαν την προσέγγιση μέσω της Συνθήκης Αναστρεψιμότητας (που παρουσιάσαμε προηγουμένως), πρότειναν, σαν ένα κριτήριο για την τελική επιλογή της ζητούμενης εικόνας, την ελαχιστοποίηση της Ολικής Μεταβολής (Total Variation - TV):

$$\text{TV}[u] = \iint_{\Omega} \|\nabla u\| \, d\mathbf{x} \quad (3.3)$$

την οποία έχουμε συναντήσει προηγουμένως τόσο στο πρόβλημα της Βελτίωσης όσο και στο πρόβλημα Inpainting (βλ. ενότητες 1.4 και 2.3). Πρέπει

να σημειωθεί ότι η ελαχιστοποίηση αυτή γίνεται εδώ, όπως είπαμε, υπό τον περιορισμό (3.1). Στα [35, 36] αποδεικνύεται ότι το παραπάνω πρόβλημα ελαχιστοποίησης δεν έχει μοναδική λύση. Παρόλα αυτά όμως, δεν υπάρχουν σημεία τοπικού ελαχίστου, αλλά μόνο πολλαπλά σημεία ολικού ελαχίστου. Επιπλέον, όλες οι λύσεις αυτές βρίσκονται “κοντά” η μία στην άλλη, υπό την έννοια ότι, όπως αποδεικνύεται στο [35], έχουν ακριβώς τις ίδιες ισούψεις γραμμές, επομένως διαφέρουν μόνο κατά μία αλλαγή της αντίθεσης (contrast). Τα παραπάνω καθιστούν οποιαδήποτε λύση ικανοποιητική σαν τελικό αποτέλεσμα.

Στο [26], προτείνεται η ελαχιστοποίηση της Ολικής Μεταβολής (3.3) μέσω ροής ταχύτερης καθόδου με προβολή στους περιορισμούς. Έτσι, από τις Euler-Lagrange εξισώσεις, προκύπτει η ΜΔΕ:

$$\frac{\partial u}{\partial t} = P_{\mathcal{U}_{0,s}} \left(\operatorname{div} \left(\frac{\nabla u}{\|\nabla u\|} \right) \right) \quad (3.4)$$

όπου $P_{\mathcal{U}_{0,s}}(\cdot)$ η προβολή στον χώρο $\mathcal{U}_{0,s}$. Η ΜΔΕ (3.4) συμπληρώνεται με αρχική συνθήκη οποιαδήποτε συνάρτηση $u^0 \in \mathcal{U}_{z,s}$ (που ικανοποιεί δηλαδή τη συνθήκη αναστρεψιμότητας). Οι συγγραφείς προτείνουν σαν μία επιλογή αρχικής συνθήκης το αποτέλεσμα της μεθόδου zero-padding (βλ. πιο πάνω). Πρέπει να σημειώσουμε ότι η προβολή αυτή χρησιμεύει στην παραμονή όλης της χρονικής ροής στον χώρο $\mathcal{U}_{z,s}$, από την στιγμή που η αρχική συνθήκη u^0 βρίσκεται στον χώρο αυτόν. Κι αυτό διότι αν μία συνάρτηση $\delta u \in \mathcal{U}_{0,s}$ προστεθεί σε μία συνάρτηση $u \in \mathcal{U}_{z,s}$ το αποτέλεσμα παραμένει στον $\mathcal{U}_{z,s}$.

Η παραπάνω μέθοδος TV παρουσιάζει το πλεονέκτημα ότι εξαφανίζει τις ανεπιθύμητες ταλαντώσεις κοντά στις ακμές (Gibbs effects) και ανακατασκευάζει τις γραμμικές ακμές με τον επιθυμητό τρόπο. Όμως, παρουσιάζει και στο πρόβλημα αυτό τα μειονεκτήματα που εμφανίζονται κατά την εφαρμογή της στα προβλήματα της Βελτίωσης και του Inpainting. Έτσι, όπως και στο Inpainting, για την ανακατασκευή των ακμών, δείχνει προτίμηση στην ευθύγραμμη, σε σχέση με την καμπυλόγραμμη ανακατασκευή. Τέλος παρουσιάζει το μειονέκτημα ότι εξομαλύνει υπερβολικά τις ομογενείς περιοχές της εικόνας.

3.1.3 Άλλες υπάρχουσες Μέθοδοι Παρεμβολής με ΜΔΕ

Τέλος σημειώνουμε επίσης κάποιες άλλες μεθόδους ΜΔΕ, που προσεγγίζουν με διαφορετικό τρόπο το πρόβλημα της Παρεμβολής:

Πρώτα από όλα, σε κάποιες περιπτώσεις (όπως π.χ. στα [25, 24]), προτείνεται η εύρεση αρχικά μιας εικόνας μεγαλύτερης χωρικής ευκρίνειας μέσω

¹πρέπει να διευκρινίσουμε ότι στη σχέση αυτή (όπως και πιο κάτω), επειδή η εικόνα u είναι όπως υποθέσαμε διακριτή, με τους συμβολισμούς $\operatorname{div}(\cdot)$ και ∇ υπονοούμε μία αξιόπιστη διακριτοποίηση των αντίστοιχων συνεχών τελεστών. Ο λόγος που κρατήσαμε τους συμβολισμούς των συνεχών τελεστών είναι η απλότητα της παρουσίασης.

μίας γραμμικής μεθόδου, όπως η cubic b-spline παρεμβολή, και στη συνέχεια η εφαρμογή στην εικόνα αυτή ΜΔΕ διάχυσης για τη βελτίωσή της. Η παραπάνω διαδικασία γίνεται με στόχο την απομάκρυνση των ανεπιθύμητων χαρακτηριστικών, που εισαγάγει η γραμμική μέθοδος Παρεμβολής (βλ. προηγούμενη υποενότητα 3.1.2). Η προσέγγιση αυτή λοιπόν λύνει το πρόβλημα της Παρεμβολής σαν άμεση εφαρμογή του προβλήματος Βελτίωσης εικόνων. Όμως η προσέγγιση αυτή παρουσιάζει το μειονέκτημα ότι η χρονική εξέλιξη της ΜΔΕ δεν μπορεί να επηρεαστεί άμεσα από την υπάρχουσα πληροφορία των τιμών της αρχικής εικόνας (όπως συμβαίνει στην προσέγγιση της υποενότητας 3.1.1, με τη συνθήκη αναστρεψιμότητας (3.1)). Λόγω αυτού όμως, υπάρχει ο κίνδυνος η τελική εικόνα να είναι υπερβολικά εξομαλυμένη ή να έχει απομακρυνθεί από την αρχική πληροφορία.

Σε κάποιες άλλες περιπτώσεις, όπως στα [14] και [52, 55] (Παρεμβολή με την Ανισοτροπική Εξομάλυνση των Tschumperlé-Deriche), η Παρεμβολή αντιμετωπίζεται από τη σκοπιά του Inpainting. Πιο συγκεκριμένα, η προσέγγιση αυτή θεωρεί ότι η ζητούμενη εικόνα δεν είναι συνεχής, αλλά από την αρχή διακριτή, ορισμένη σε πυκνότερο πλέγμα. Έτσι, στη προσέγγιση αυτή, οι τιμές της αρχικής εικόνας καταλαμβάνουν τις αντίστοιχες θέσεις σε ενδιάμεσα σημεία του πυκνότερου πλέγματος, τα οποία αποτελούν τη γνωστή περιοχή $\Omega \setminus \mathbb{D}$ και τα υπόλοιπα σημεία αποτελούν την περιοχή \mathbb{D} . Σημειώνουμε ότι \mathbb{D} είναι η inpainting περιοχή και Ω το σύνολο ορισμού (εδώ, όλα τα σημεία του πυκνότερου πλέγματος). Προκύπτει δηλαδή ένα πρόβλημα inpainting, το οποίο, όπως έχουμε πει, λύνεται με την χρονική εξέλιξη της ΜΔΕ, μόνο στο σύνολο \mathbb{D} , δηλαδή στα επιπλέον σημεία του νέου πλέγματος. Εδώ λοιπόν διαφαίνεται ακόμα πιο καθαρά η ομοιότητα που αναφέραμε πιο πάνω ανάμεσα στην Παρεμβολή και στο Inpainting, μια και ένας τρόπος λύσης της Παρεμβολής είναι ο μετασχηματισμός της σε πρόβλημα Inpainting. Αξίζει να παρατηρήσουμε ότι η προσέγγιση αυτή μπορεί να ειπωθεί και σαν ειδική περίπτωση της προσέγγισης με την συνθήκη αναστρεψιμότητας (βλ. υποενότητα 3.1.2), όταν στη συνθήκη (3.1) χρησιμοποιηθεί σαν πυρήνας εξομάλυνσης s η συνάρτηση Dirac $\delta(n_1, n_2)$. Το γεγονός αυτό δείχνει όμως ένα σημαντικό μειονέκτημα της μεθόδου αυτής. Ο πυρήνας s παρουσιάζει μηδενική χωρική έκταση, κάτι που, όπως είπαμε (βλ. προηγούμενη υποενότητα 3.1.2), δεν είναι καθόλου επιθυμητό. Κι αυτό διότι δεν υπάρχει η δυνατότητα επεξεργασίας όλων των pixels και παράλληλα, όσον αφορά αυτά που επεξεργάζονται, δεν υπάρχει κανένας περιορισμός στην μεταβολή τους, που να αξιοποιεί τη πληροφορία της αρχικής, χαμηλής ευκρίνειας, εικόνας.

Επίσης, οι Belahmidi και Guichard [6] βασίστηκαν στην προσέγγιση της μεθόδου Ολικής Μεταβολής (TV) και την τροποποίησαν χρησιμοποιώντας, αντί της ΜΔΕ ελαχιστοποίησης της TV (3.4), μία ΜΔΕ διάχυσης με κατευθυνόμενες λαπλασιανές. Η ΜΔΕ αυτή εκτελεί προσαρμοζόμενη ανισοτροπική εξομά-

λυνση, με παρόμοιο σχεδιασμό με μεθόδους που είδαμε στο κεφάλαιο 1, όπως αυτή των Tschumperlé-Deriche (βλ. υποενότητα 1.10). Έκαναν όμως και μια άλλη σημαντική τροποποίηση, αφού δεν απαιτούσαν πλέον η όλη χρονική ροή να ικανοποιεί την συνθήκη αναστρεψιμότητας. Αντί αυτού εισήγαγαν έναν όρο ποινής για την απομάκρυνση από τη συνθήκη αυτή. Η μέθοδος αυτή, λόγω της ΜΔΕ που χρησιμοποιεί, έχει βελτιωμένη συμπεριφορά σε σχέση με την μέθοδο Ολικής Μεταβολής (TV). Όμως και αυτή παρουσιάζει το μειονέκτημα ότι υπάρχει ο κίνδυνος να απομακρυνθεί η λύση από την αρχική πληροφορία, αφού η ικανοποίηση της συνθήκης αναστρεψιμότητας παίζει πλέον δευτερεύοντα ρόλο στην εξέλιξη της ΜΔΕ.

Συμπεραίνουμε λοιπόν ότι οι εναλλακτικές αυτές προσεγγίσεις του προβλήματος με ΜΔΕ, παρά τα σχετικά καλά αποτελέσματά τους και την μειωμένη αλγοριθμική πολυπλοκότητα, υστερούν σε σχέση με τη γενική προσέγγιση της Συνθήκης Αναστρεψιμότητας και τη μέθοδο Ολικής Μεταβολής (TV) που τη χρησιμοποιεί, σύμφωνα με την οπτική γωνία της ανάλυσης που παρουσιάσαμε. Για τον λόγο αυτό, στη συνέχεια, θα βασιστούμε στη προσέγγιση της Συνθήκης Αναστρεψιμότητας και θα προτείνουμε κάποιες γενικεύσεις της μεθόδου Ολικής Μεταβολής τόσο για άλλες ΜΔΕ διάχυσης όσο και για διανυσματικές εικόνες.

3.2 Πρόταση μιας νέας γενικής μεθόδου Παρεμβολής διανυσματικών εικόνων με ΜΔΕ

Στην ενότητα αυτή προτείνουμε μία επέκταση της προσέγγισης της μεθόδου Ολικής Μεταβολής (TV) των Malgouyres και Guichard (βλ. υποενότητες 3.1.1,3.1.2) για τις διανυσματικές εικόνες, χρησιμοποιώντας γενικές ΜΔΕ διάχυσης.

Πρώτα από όλα, λοιπόν, η *Συνθήκη Αναστρεψιμότητας* (3.1) μπορεί άμεσα να γενικευτεί για τις διανυσματικές εικόνες, θεωρώντας ότι ισχύει αυτούσια για κάθε κανάλι. Επομένως, αν z είναι η αρχική διακριτή διανυσματική εικόνα και u η διανυσματική εικόνα υψηλότερης χωρικής ευκρίνειας, από την οποία έχει προέλθει, τότε γενικεύουμε τη συνθήκη αναστρεψιμότητας ως εξής:

$$z_m = Q(s * u_m), \quad m = 1, \dots, N \quad (3.5)$$

Ορίζουμε σαν $\mathbf{U}_{z,s}$ το σύνολο των συναρτήσεων u , που ικανοποιούν τη συνθήκη αναστρεψιμότητας (3.5). Προφανώς, και αυτό το σύνολο αποτελεί έναν αφινικό (affine) υποχώρο, για τους ίδιους λόγους με την βαθμωτή περίπτωση.

Κατά πλήρη αντιστοιχία με την υποενότητα 3.1.1, μία απλή λύση που ικανοποιεί τη συνθήκη αναστρεψιμότητας (3.5) είναι αυτή που προκύπτει από τη

μέθοδο *zero-padding* για κάθε κανάλι χωριστά. Η λύση αυτή όμως, όπως αναφέραμε δεν είναι ικανοποιητική, αλλά μπορούμε και πάλι να τη χρησιμοποιήσουμε σαν μία αρχική τιμή στις μεθόδους που θα προτείνουμε.

Μπορούμε άμεσα να γενικεύσουμε τη μέθοδο της Ολικής Μεταβολής (TV), θέτοντας σαν κριτήριο για την εύρεση της λύσης την Διανυσματική Ολική Μεταβολή (Vectorial TV):

$$\text{VTV}[\mathbf{u}] = \iint_{\Omega} \|\nabla \mathbf{u}\| \, d\mathbf{x} = \iint_{\Omega} \sqrt{\sum_{m=1}^N \|\nabla u_m\|^2} \, d\mathbf{x} \quad (3.6)$$

την οποία συναντήσαμε σαν γενίκευση της Ολικής Μεταβολής για τα προβλήματα της Βελτίωσης αλλά και του Inpainting (βλ. υποενότητες 1.6,2.3)

Αν ελαχιστοποιήσουμε την (3.6) υπό τους περιορισμούς (3.5), χρησιμοποιώντας πάλι ροή ταχύτερης καθόδου με προβολή στους περιορισμούς, καταλήγουμε στη ΜΔΕ:

$$\frac{\partial u_m}{\partial t} = P_{\mathcal{U}_{0,s}} \left(\text{div} \left(\frac{\nabla u_m}{\|\nabla \mathbf{u}\|} \right) \right), \quad m = 1, \dots, N \quad (3.7)$$

όπου $P_{\mathcal{U}_{0,s}}(\cdot)$ η προβολή στον χώρο $\mathcal{U}_{0,s}$ (εφαρμόζεται σε κάθε κανάλι της ταχύτητας της ΜΔΕ). Επιλέγουμε η ΜΔΕ (3.7) να συμπληρωθεί και αυτή με αρχική συνθήκη \mathbf{u}^0 το αποτέλεσμα της *zero-padding* παρεμβολής για κάθε κανάλι.

Η ΜΔΕ (3.7) στην οποία καταλήξαμε, αποτελεί ουσιαστικά την ΜΔΕ Διανυσματικής Ολικής Μεταβολής (Vectorial TV PDE), που εφαρμόζεται στη Βελτίωση και στο Inpainting, τροποποιημένη έτσι, ώστε η ταχύτητά της να προβάλλεται στον χώρο $\mathcal{U}_{0,s}$. Η Διανυσματική Ολική Μεταβολή εφαρμόζεται με αντίστοιχη επιτυχία στα προβλήματα της Βελτίωσης, του Inpainting και της Παρεμβολής, λόγω των κοινών χαρακτηριστικών που παρουσιάζουν τα προβλήματα αυτά.

Βασισμένοι στην παρατήρηση αυτή, προτείνουμε την εφαρμογή και των άλλων μεθόδων ΜΔΕ διάχυσης για Βελτίωση (βλ. ενότητα 1), με τρόπο αντίστοιχο με την TV PDE. Θα μπορούσαν επίσης να εφαρμοστούν και οι ΜΔΕ που έχουν σχεδιαστεί για το Inpainting (βλ. ενότητα 2).

Για το σκοπό αυτό, συμβολίζουμε οποιαδήποτε από αυτές τις ΜΔΕ με τη γενική εξίσωση:

$$\frac{\partial u_m}{\partial t} = v_m(x, y, t), \quad m = 1, \dots, N$$

όπου $v_m(x, y, t)$ το δεύτερο μέλος της ΜΔΕ, που αντιστοιχεί στην ταχύτητα της ΜΔΕ σε κάθε χρονική στιγμή και για κάθε σημείο.

Αν λάβουμε υπόψιν και τη χωρική διακριτοποίηση της ΜΔΕ, η παραπάνω ΜΔΕ

προσεγγίζεται από το εξής ημιδιακριτό σχήμα:

$$\frac{\partial u_m}{\partial t} = v_m(i, j, t), \quad m = 1, \dots, N$$

Η διανυσματική ταχύτητα \mathbf{v} κάθε χρονική στιγμή εξαρτάται από την συνάρτηση της εικόνας \mathbf{u} την ίδια χρονική στιγμή, για αυτό γράφουμε $\mathbf{v} = \mathbf{v}(\mathbf{u})$.

Η συνάρτηση $\mathbf{v}(\cdot)$ περιγράφει ουσιαστικά τη μέθοδο ΜΔΕ που χρησιμοποιείται. Για παράδειγμα, έχουμε:

- $v_m(\mathbf{u}) = \frac{1}{\sqrt{\det(I+G)}} \operatorname{div} \left(\sqrt{\det(I+G)} (I+G)^{-1} \nabla u_m \right)$ (όπου G η μήτρα Di Zenzo της \mathbf{u}), στη μέθοδο *Beltrami Flow* (βλ. ενότητα 1.8).
- $v_m(\mathbf{u}) = \operatorname{trace}(T H_m)$, $T = \frac{1}{1+\lambda_+^*+\lambda_-^*} \boldsymbol{\theta}_+^* \boldsymbol{\theta}_+^{*\top} + \frac{1}{\sqrt{1+\lambda_+^*+\lambda_-^*}} \boldsymbol{\theta}_-^* \boldsymbol{\theta}_-^{*\top}$ (όπου H_m η Hessian Μήτρα της u_m και $\boldsymbol{\theta}_+^*$, $\boldsymbol{\theta}_-^*$, λ_+^* , λ_-^* τα φασματικά στοιχεία του δομικού ταυστή $J_\rho(\nabla \mathbf{u})$), στη μέθοδο *Ανιστροπικής Εξομάλυνσης των Tschumperlé-Deriche* (βλ. ενότητα 1.10).

Έτσι, για την εύρεση της λύσης της Παρεμβολής της \mathbf{z} ανάμεσα στις συναρτήσεις $\mathbf{u} \in \mathbf{U}_{\mathbf{z},s}$, προτείνουμε τη χρησιμοποίηση της εξής γενικής ΜΔΕ:

$$\frac{\partial \mathbf{u}}{\partial t} = P_{\mathbf{u}_{0,s}}(\mathbf{v}(\mathbf{u})) \quad (3.8)$$

Σαν αρχική συνθήκη \mathbf{u}^0 της ΜΔΕ αυτής επιλέγουμε και πάλι την παρεμβολή zero-padding για κάθε κανάλι.

Στη συνέχεια εξηγούμε πιο αναλυτικά τους λόγους εφαρμογής της παραπάνω γενικής μεθόδου ΜΔΕ για την λύση του προβλήματος. Πρώτα από όλα, η αρχική πρόβλεψη $\mathbf{u}^0(x, y)$ που γίνεται μέσω της παρεμβολής zero-padding, όπως έχουμε αναφέρει, περιέχει μία αρκετά ικανοποιητική ανακατασκευή των γεωμετρικών δομών της εικόνας, αλλά συνάμα και κάποια ανεπιθύμητα χαρακτηριστικά (όπως τα Gibbs effects). Επομένως, μπορεί να θεωρηθεί ότι περιέχει τη ζητούμενη λύση, διαταραγμένη με θόρυβο. Έτσι, μία μέθοδος ΜΔΕ για Βελτίωση μπορεί να πετύχει (άλλη λιγότερο και άλλη περισσότερο ικανοποιητικά) την απομάκρυνση των ανεπιθύμητων αυτών χαρακτηριστικών αλλά ταυτόχρονα τη διατήρηση των γεωμετρικών δομών και ακμών. Η ταχύτητα $\mathbf{v}(x, y, t)$ της ΜΔΕ προβάλλεται στον χώρο $\mathbf{U}_{0,s}$, ώστε να συνεχίσουν να ισχύουν οι περιορισμοί (3.5), κάτι το οποίο εξασφαλίζει ακρίβεια στο τελικό αποτέλεσμα. Με άλλα λόγια, με τη μέθοδο αυτή, εκμεταλλευόμαστε σε κάθε χρονικό βήμα την σημαντική και σαφή πληροφορία που διαθέτουμε για τη λύση που ψάχνουμε (μέσω της συνθήκης αναστρεψιμότητας (3.5)). Το πεδίο της ταχύτητας $\mathbf{v}(x, y, t)$ λοιπόν καθορίζει τη μεταβολή που τείνει να προσδώσει στην εικόνα

η διαδικασία της διάχυσης. Η προβολή $P_{\mathbf{u}_{0,s}}(\cdot)$ διατηρεί τελικά μόνο εκείνη τη συνιστώσα της μεταβολής της εικόνας που είναι σύμφωνη με τους περιορισμούς. Από τα παραπάνω συμπεραίνουμε ότι τα χαρακτηριστικά της μεταβολής της εικόνας καθορίζονται από την ταχύτητα \mathbf{v} , με αποτέλεσμα η γενική συμπεριφορά της διάχυσης να μην επηρεάζεται τόσο από την προβολή $P_{\mathbf{u}_{0,s}}(\cdot)$. Αυτός είναι και ο λόγος που η συμπεριφορά της μεθόδου παρουσιάζει ομοιότητα με την συμπεριφορά της αντίστοιχης ΜΔΕ στην Βελτίωση.

Πρέπει να σημειώσουμε επίσης ότι, λόγω της προβολής $P_{\mathbf{u}_{0,s}}(\cdot)$, στη συνθήκη ισορροπίας της ΜΔΕ δεν προκύπτει, όπως στην Βελτίωση, κάποια υπερβολικά εξομαλυμένη εικόνα, αλλά το επιθυμητό αποτέλεσμα (κάτι το οποίο συναντήσαμε και στο Inpainting, εκεί όμως λόγω των συνοριακών συνθηκών). Η συνθήκη ισορροπίας της (3.8) σημαίνει ότι είτε $\mathbf{v} = 0$, οπότε η διάχυση έχει τοπικά ισορροπήσει, είτε η διάχυση τείνει να μεταβάλει την εικόνα με τρόπο που έρχεται σε αντίθεση με τους περιορισμούς. Το τελευταίο σημαίνει ότι ανάμεσα στις γειτονικές συναρτήσεις που ικανοποιούν τη συνθήκη αναστρεψιμότητας (3.5), η συνάρτηση ισορροπίας \mathbf{u} είναι η πιο κατάλληλη λύση, σύμφωνα με τα κριτήρια της συγκεκριμένης μεθόδου διάχυσης.

Λόγω των αντιστοιχιών με το πρόβλημα Βελτίωσης που εξηγήσαμε παραπάνω, συμπεραίνουμε ότι η γενική μέθοδος που εισάγαμε μπορεί να βελτιώσει σημαντικά τη μέθοδο Ολικής Μεταβολής (TV) των Malgouyres και Guichard [26, 35, 36] (βλ. υποενότητα 3.1.2), με τη χρήση κατάλληλης ΜΔΕ. Κι αυτό για τους ίδιους λόγους που διάφορες άλλες μέθοδοι ΜΔΕ παρουσιάζουν βελτιωμένη συμπεριφορά σε σχέση με την μέθοδο Ολικής Μεταβολής στο πρόβλημα Βελτίωσης, όπως είδαμε στο κεφάλαιο 1.

Για παράδειγμα, συγκρίνοντας την ΜΔΕ ελαχιστοποίησης της Διανυσματικής Ολικής Μεταβολής (TV PDE) (που είναι μία άμεση γενίκευση της βαθμωτής περίπτωσης, με αποτελέσματα αντίστοιχης ποιότητας) με την ΜΔΕ Ανισοτροπικής Εξομάλυνσης των Tschumperlé-Deriche (βλ ενότ. 1.10), καταλήγουμε στο ότι η δεύτερη παρουσιάζει σε πολλές περιπτώσεις σημαντικά καλύτερα αποτελέσματα. Κι αυτό διότι, για την προσαρμογή της διάχυσης, ενώ στην πρώτη λαμβάνεται υπόψιν μόνο η τοπική μεταβολή της εικόνας (όπως εκφράζεται από τη νόρμα $\|\nabla \mathbf{u}\|$), στη δεύτερη λαμβάνονται επιπλέον υπόψιν και οι κατευθύνσεις της τοπικής μεταβολής και όλα αυτά με έναν πιο συνεκτικό τρόπο (μέσω του δομικού τανυστή $J_\rho(\nabla \mathbf{u})$). Επίσης, αυτή η προσαρμογή της διάχυσης, ενώ στην πρώτη γίνεται απλά μέσω του βαθμωτού συντελεστή διάχυσης, στην δεύτερη γίνεται μέσω ανισοτροπικού τανυστή, που προσφέρει πιο αποτελεσματικό έλεγχο της διαδικασίας εξομάλυνσης.

Τέλος σημειώνουμε ότι, όπως θα παρουσιάσουμε στην υποενότητα 6.2, διαπιστώσαμε και πειραματικά τα βελτιωμένα αποτελέσματα της εφαρμογής της γενικής μεθόδου που προτείναμε, σε σχέση με την μέθοδο Ολικής Μεταβολής των Malgouyres και Guichard [26, 35, 36], για το πρόβλημα της Παρεμβολής.

Κεφάλαιο 4

Θεωρητικά Θέματα Υλοποίησης ΜΔΕ για την ΟΥ

4.1 Μελέτη της Ισοδυναμίας της Γενικής Ανισοτροπικής Διάχυσης Κατευθυνόμενων Λαπλασιανών με Επαναληπτική Τοπική Συνέλιξη

4.1.1 Εισαγωγή

Στα προηγούμενα κεφάλαια, είδαμε τις ενδιαφέρουσες εφαρμογές των εξισώσεων διάχυσης σε προβλήματα της Όρασης Υπολογιστών. Στις περισσότερες περιπτώσεις, οι ΜΔΕ αυτές μπορούν ισοδύναμα να γραφούν σε μορφή ανισοτροπικής διάχυσης με κατευθυνόμενες λαπλασιανές:

$$\frac{\partial u_i}{\partial t} = c_1 \frac{\partial^2 u_i}{\partial \mathbf{v}_1^2} + c_2 \frac{\partial^2 u_i}{\partial \mathbf{v}_2^2} = \text{trace}(T H_i) \quad , \quad i = 1, \dots, N \quad (4.1)$$

όπου H_i η χωρική Hessian Μήτρα της u_i και $T = c_1 \mathbf{v}_1 \mathbf{v}_1^T + c_2 \mathbf{v}_2 \mathbf{v}_2^T$ ο ανισοτροπικός τανυστής διάχυσης (με ιδιοτιμές c_1, c_2 και ιδιοδιανύσματα $\mathbf{v}_1, \mathbf{v}_2$). Ο τανυστής αυτός ελέγχει την διάχυση και συνήθως εξαρτάται από τις γεωμετρικές ιδιότητες της εικόνας στην περιοχή του κάθε σημείου.

Κάποια παραδείγματα μεθόδων που μπορούν γραφούν στην μορφή (4.1) είναι τα εξής:

- Η γενική μη γραμμική διάχυση (βλ. ενότητα 1.2) μπορεί να γραφεί

(βλ. [19, 34]):

$$\begin{aligned} \frac{\partial u}{\partial t} &= \operatorname{div} (g(\|\nabla u\|)\nabla u) \\ &= g(\|\nabla u\|)\frac{\partial^2 u}{\partial \xi^2} + \left\{ g'(\|\nabla u\|)\|\nabla u\| + g(\|\nabla u\|) \right\} \frac{\partial^2 u}{\partial \mathbf{n}^2} \end{aligned} \quad (4.2)$$

όπου $\mathbf{n} = \frac{\nabla u}{\|\nabla u\|}$ και $\xi = \mathbf{n}^\perp$.

- Αν στην προηγούμενη περίπτωση $g(s) = \phi'(s)/s$, πρόκειται για τη ΜΔΕ ελαχιστοποίησης του γενικού ϕ -συναρτησιακού (ϕ -functional) για βαθμωτές εικόνες (βλ. ενότητα 1.6), η οποία λόγω της (4.2) γράφεται:

$$\frac{\partial u}{\partial t} = \operatorname{div} \left(\frac{\phi'(\|\nabla u\|)}{\|\nabla u\|} \nabla u \right) = \frac{\phi'(\|\nabla u\|)}{\|\nabla u\|} \frac{\partial^2 u}{\partial \xi^2} + \phi''(\|\nabla u\|) \frac{\partial^2 u}{\partial \mathbf{n}^2} \quad (4.3)$$

- Αν $\phi(s) = s$, έχουμε την *TV PDE* για βαθμωτές εικόνες (βλ. ενότητα 1.4), η οποία, όπως ήδη παρουσιάσαμε (βλ. σχέση (1.30)), μπορεί να γραφεί στην ακόλουθη μορφή:

$$\frac{\partial u}{\partial t} = \operatorname{div} \left(\frac{\nabla u}{\|\nabla u\|} \right) = \frac{1}{\|\nabla u\|} \cdot \frac{\partial^2 u}{\partial \xi^2}$$

- Η *Κίνηση Μέσης Καμπυλότητας* (*Mean Curvature Motion - MCM*), όπως είδαμε (βλ. σχέση (1.29)) αντιστοιχεί στην απλή περίπτωση:

$$\frac{\partial u}{\partial t} = \kappa \|\nabla u\| = \frac{\partial^2 u}{\partial \xi^2}$$

- Οι μέθοδοι για διανυσματικές εικόνες της *Διάχυσης στην Κατεύθυνση θ_-* (βλ. ενότητα 1.9) και της *Ανισοτροπικής Εξομάλυνσης των Tschumperlé-Deriche* (βλ. ενότητα 1.10) χρησιμοποιούν ΜΔΕ που είναι εξαρχής σχεδιασμένες στη μορφή κατευθυνόμενων λαπλασιανών (βλ. σχέσεις (1.63) και (1.67) αντίστοιχα).

Όπως θα δούμε στο κεφάλαιο 5, οι ΜΔΕ αυτές μπορούν να υλοποιηθούν με κλασσικά αριθμητικά σχήματα που χρησιμοποιούν κατευθείαν διακριτοποίηση των $2^{\text{ων}}$ χωρικών παραγώγων (της μήτρας Hessian). Όμως, οι Tschumperlé και Deriche [52, 55] πρότειναν μία εναλλακτική υλοποίηση, χρησιμοποιώντας επαναληπτική τοπική συνέλιξη με κατάλληλους κατευθυνόμενους πυρήνες Gauss. Η υλοποίηση αυτή παρουσιάζει το πλεονέκτημα ότι επιτυγχάνει με φυσικό τρόπο και ανεξάρτητα από το χρονικό βήμα την αρχή *μεγίστου-ελαχίστου* (*min-max*

4.1 Μελέτη της Ισοδυναμίας της Γενικής Ανισοτροπικής Διάχυσης Κατευθυνόμενων Λαπλασιανών με Επαναληπτική Τοπική Συνέλιξη ³

principle), αφού ισοδυναμεί με εύρεση τοπικής μέσης τιμής (local averaging). Επίσης, είναι πιο ακριβής, διότι οι υπολογιζόμενοι πυρήνες δεν εξαρτώνται, όπως στα κλασσικά σχήματα, από μία διακριτοποίηση στον οριζόντιο και κάθετο άξονα και έτσι μειώνεται η μη επιθυμητή προτίμηση στις κατευθύνσεις των αξόνων.

Για την αιτιολόγηση της προσέγγισης αυτής, ο Tschumperlé αποδεικνύει (βλ. [52]) ότι, στην ειδική περίπτωση όπου ο τανυστής T είναι χωρικά και χρονικά σταθερός, η λύση της (4.1) εκφράζεται σαν συνέλιξη με πυρήνα *Gauss* που κατευθύνεται σύμφωνα με τον T (το οποίο αποτελεί γενίκευση της ισοδυναμίας της γραμμικής διάχυσης θερμότητας με *Gaussian* συνέλιξη, που πρώτος παρατήρησε ο Koenderink [31], βλ. εισαγωγική ενότητα 1.1). Επιπρόσθετα, στο [52] αναφέρεται, χωρίς απόδειξη ότι, όταν ο τανυστής T είναι μεταβαλλόμενος, η λύση μπορεί να προσεγγιστεί με επαναληπτική τοπική συνέλιξη με κατάλληλους *Gaussian* πυρήνες. Αυτή η έκφραση της λύσης σαν τοπικό φίλτράρισμα εγκαθιδρύει την σύνδεση των ΜΔΕ αυτών με τις τεχνικές του *Bilateral filtering* (βλ. [5, 51]). Μία παρόμοια προσέγγιση τοπικής συνέλιξης με όχι *Gaussian* πυρήνες έχει πρόσφατα προταθεί από τους Sochen κ.ά. [44], για τις εξισώσεις που σχετίζονται με την *Ροή Beltrami* (βλ. ενότητα 1.8), οι οποίες όμως δεν ανήκουν στην κατηγορία των ΜΔΕ που μελετάμε εδώ.

Στην ενότητα αυτή, με κίνητρο την καλύτερη θεωρητική θεμελίωση της προσέγγισης που περιγράψαμε, παρέχουμε κάποια αποτελέσματα που σχετίζονται με την ακρίβεια και τη σύγκλιση της προσέγγισης. Συγκεκριμένα, βρίσκουμε ένα φράγμα του σφάλματος της προσέγγισης της λύσης με επαναληπτική τοπική συνέλιξη, όταν το χρονικό βήμα είναι πεπερασμένο. Από το φράγμα αυτό, εξάγουμε ένα “μέτρο”, χαρακτηριστικό του πυρήνα που χρησιμοποιείται, το οποίο παρέχει μία ένδειξη για την καταλληλότητα της χρήσης του συγκεκριμένου πυρήνα για την προσέγγιση. Επιπρόσθετα, αποδεικνύουμε τη σύγκλιση στη λύση της (4.1), καθώς το χρονικό βήμα γίνεται απειροστό, στην ειδική περίπτωση όπου ο τανυστής T είναι σταθερός, διατηρώντας όμως την γενικότητα ως προς το ότι ο τανυστής T θεωρείται ανισοτροπικός και η συνέλιξη γίνεται με γενικούς κατευθυνόμενους βαθυπερατούς πυρήνες (όχι μόνο *Gaussian*). Εξάγουμε τα αποτελέσματα αυτά για την ανισοτροπική διάχυση για γενικές διανυσματικές συναρτήσεις, που ορίζονται σε έναν πολυδιάστατο χώρο (και όχι μόνο στο επίπεδο, όπως οι εικόνες). Έτσι, τα αποτελέσματα αυτά έχουν αξία όχι μόνο για τις εικόνες, αλλά και για άλλες ενδιαφέρουσες κατηγορίες που εμφανίζονται σε διάφορες εφαρμογές, όπως για παράδειγμα τριδιάστατα βαθμωτά ή διανυσματικά δεδομένα.

Για την ανάλυσή μας, βασιζόμαστε στο [27], όπου αποδεικνύεται η παραπάνω σύγκλιση στην απλή ειδική περίπτωση γραμμικής διάχυσης (που αντιστοιχεί σε σταθερό και ισοτροπικό τανυστή $T = I$).

4.1.2 Παρουσίαση και Απόδειξη κάποιων νέων Θεωρημάτων

Εδώ θα ασχοληθούμε με τη γενική περίπτωση ανισοτροπικής Διάχυσης με Κατευθυνόμενες Λαπλασιανές διανυσματικής συνάρτησης, ορισμένης σε ένα M -διάστατο χώρο. Έστω λοιπόν η συνάρτηση:

$$\mathbf{u} = (u_1, u_2, \dots, u_N) : \mathbb{R}^M \times [0, +\infty) \rightarrow \mathbb{R}^N,$$

η οποία ικανοποιεί την ΜΔΕ:

$$\frac{\partial u_i}{\partial t} = \sum_{j=1}^M c_j \frac{\partial^2 u_i}{\partial \mathbf{v}_j^2} = \text{trace}(TH_i) \quad , i = 1, \dots, N \quad (4.4)$$

$$\mathbf{u}(\mathbf{x}, 0) = \mathbf{u}^0(\mathbf{x})$$

όπου:

- $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_M\}$ ορθομοναδιαία βάση του \mathbb{R}^M και $c_1, c_2, \dots, c_M > 0$. Οι συντελεστές c_j και τα διανύσματα \mathbf{v}_j ($j = 1, \dots, M$) είναι μεταβαλλόμενα και εξαρτώνται κάθε χρονική στιγμή από την συμπεριφορά της ίδιας της συνάρτησης \mathbf{u} στην γειτονιά του κάθε σημείου. Για κάθε σημείο \mathbf{x}_0 και κάθε χρονική στιγμή $t_0 > 0$, είναι, δηλαδή, στη γενική περίπτωση, συναρτησιακά της συνάρτησης $\mathbf{u}(\mathbf{x} - \mathbf{x}_0, t_0)$ (θεωρούμε t_0 σταθερό και μετατοπίζουμε την $\mathbf{u}(\mathbf{x})$ έτσι ώστε η γειτονιά του \mathbf{x}_0 να βρεθεί γύρω από την αρχή των αξόνων), για αυτό γράφουμε:

$$\begin{cases} c_j = c_j[\mathbf{u}(\mathbf{x} - \mathbf{x}_0, t_0)] & = c_j(\mathbf{x}_0, t_0) \\ \mathbf{v}_j = \mathbf{v}_j[\mathbf{u}(\mathbf{x} - \mathbf{x}_0, t_0)] & = \mathbf{v}_j(\mathbf{x}_0, t_0) \end{cases} , j = 1, \dots, M \quad (4.5)$$

Για παράδειγμα, τα c_j, \mathbf{v}_j μπορεί να εξαρτώνται από τις πρώτες μερικές παραγώγους των u_i . Μία τέτοια περίπτωση για $M = 2, N = 1$ είναι π.χ. η ΜΔΕ ελαχιστοποίησης του ϕ -συναρτησιακού (βλ. σχέση (4.3)), στην οποία τα c_j, \mathbf{v}_j προκύπτουν από την κλίση $\nabla u = (u_x, u_y)^T$ ως έξης: $c_1 = \phi'(\|\nabla u\|) / \|\nabla u\|, c_2 = \phi''(\|\nabla u\|), \mathbf{v}_1 = \nabla u / \|\nabla u\|, \mathbf{v}_2 = \mathbf{v}_1^\perp$.

- T ο τανυστής διάχυσης με ιδιοδιανύσματα $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_M$ και ιδιοτιμές c_1, c_2, \dots, c_M :

$$T = V \cdot C \cdot V^{-1} = \sum_{j=1}^M c_j \mathbf{v}_j \mathbf{v}_j^T, \text{ όπου:}$$

$$V = [\mathbf{v}_1 \quad \mathbf{v}_2 \quad \dots \quad \mathbf{v}_M] , C = \begin{bmatrix} c_1 & & \emptyset \\ & \ddots & \\ \emptyset & & c_M \end{bmatrix}$$

με $V^{-1} = V^T$, διότι η $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_M\}$ είναι ορθομοναδιαία βάση του \mathbb{R}^M .

4.1 Μελέτη της Ισοδυναμίας της Γενικής Ανισοτροπικής Διάχυσης Κατευθυνόμενων Λαπλασιανών με Επαναληπτική Τοπική Συνέλιξη 5

Λόγω της (4.5), και για τον ταυοστή T , για κάθε (\mathbf{x}_0, t_0) , ισχύει:

$$T = \tilde{T}[\mathbf{u}(\mathbf{x} - \mathbf{x}_0, t_0)] = T(\mathbf{x}_0, t_0) \quad (4.6)$$

Επομένως, το συναρτησιακό $\tilde{T}[\cdot]$ είναι χαρακτηριστικό της συγκεκριμένης ΜΔΕ και την εκφράζει πλήρως, αφού συμβολίζει την διαδικασία υπολογισμού των συντελεστών και κατευθύνσεων διάχυσης σε κάθε σημείο.

- $H_i = \nabla \otimes \nabla u_i$ η χωρική Hessian Μήτρα της u_i .

Για την προσέγγιση της λύσης της (4.4) με τοπική συνέλιξη, ορίζουμε αρχικά (παρόμοια με το [27]) μία κατηγορία συμμετρικών πυρήνων:

Ορισμός 4.1. Έστω μία ολοκληρώσιμη συνάρτηση $k(\mathbf{x}) : \mathbb{R}^M \rightarrow \mathbb{R}$, την οποία κατανοούμε σαν πυρήνα συνέλιξης. Λέμε ότι η $k(\mathbf{x})$ είναι ακτινική εάν $k(\mathbf{x}) = k(\|\mathbf{x}\|)$. Λέμε ότι η $k(\mathbf{x})$ είναι ψεύδο-ακτινική εάν ικανοποιεί τις ακόλουθες σχέσεις:

$$\int_{\mathbb{R}^M} k(\mathbf{x}) d\mathbf{x} = 1 \quad (4.7)$$

και, για κάθε $i, j = 1, \dots, M$, τέτοια ώστε $i \neq j$:

$$\int_{\mathbb{R}^M} x_i k(\mathbf{x}) d\mathbf{x} = \int_{\mathbb{R}^M} x_i x_j k(\mathbf{x}) d\mathbf{x} = 0 \quad (4.8)$$

$$\int_{\mathbb{R}^M} x_i^2 k(\mathbf{x}) d\mathbf{x} = 2 \quad (4.9)$$

Επίσης, θεωρούμε την οικογένεια πυρήνων $k_h(\mathbf{x})$, που προκύπτει από έναν τέτοιο πυρήνα, με αλλαγές κλίμακας (rescalings) σύμφωνα με τη σταθερά $h > 0$:

$$k_h(\mathbf{x}) = \frac{1}{h^{\frac{M}{2}}} k\left(\frac{\mathbf{x}}{h^{\frac{1}{2}}}\right) \quad (4.10)$$

Για τον πυρήνα $k(\mathbf{x})$, που είναι αντιπροσωπευτικός όλης της οικογένειας αυτής, θα χρησιμοποιούμε την ορολογία *μητρικός πυρήνας*.

Οι $k_h(\mathbf{x})$ ικανοποιούν και αυτές την συνθήκη κανονικοποίησης (4.7) και μηδενικών ροπών (4.8). Ακόμα η σχέση (4.9) που αφορά τη διασπορά του $k(\mathbf{x})$ μετατρέπεται στην:

$$\int_{\mathbb{R}^M} x_i^2 k_h(\mathbf{x}) d\mathbf{x} = 2h$$

Έτσι, καθώς το h μικραίνει, η k_h συγκεντρώνεται όλο και περισσότερο γύρω από την αρχή των αξόνων, ενώ παράλληλα η μεταβολή της γίνεται όλο και πιο απότομη.

Στο [27] παρατίθεται το εξής θεώρημα, που σχετίζεται με την προσέγγιση με συνέλιξη της λύσης της εξίσωσης (4.4), στην ειδική περίπτωση όπου $N = 1$ και $T = I$:

Θεώρημα 4.2. Έστω $k(\mathbf{x})$ ένας πυρήνας, που ικανοποιεί τις συνθήκες του Ορισμού 4.1. Υποθέτουμε επίσης ότι:

$$\int_{\mathbb{R}^M} |k(\mathbf{z})| \|\mathbf{z}\| d\mathbf{z} = A < +\infty \quad (4.11)$$

Τότε, για κάθε φραγμένη και 3 φορές παραγωγίσιμη συνάρτηση $u : \mathbb{R}^M \rightarrow \mathbb{R}$, ισχύει:

$$(k_h * u)(\mathbf{x}) - u(\mathbf{x}) = h \sum_{j=1}^M \frac{\partial^2 u}{\partial x_j^2} + f_i(\mathbf{x}) \quad (4.12)$$

όπου:

$$|f_i(\mathbf{x})| \leq Ah^{\frac{3}{2}} \max_{\mathbf{x}} \|D^3 u(\mathbf{x})\| \quad (4.13)$$

Απόδειξη. Έχουμε:

$$(k_h * u)(\mathbf{x}) - u(\mathbf{x}) = \int_{\mathbb{R}^M} \frac{1}{h^{\frac{M}{2}}} k\left(\frac{\mathbf{y}}{h^{\frac{1}{2}}}\right) (u(\mathbf{x} - \mathbf{y}) - u(\mathbf{x})) d\mathbf{y}$$

Αν κάνουμε την αλλαγή μεταβλητής ολοκλήρωσης $\mathbf{z} = \mathbf{y}/h^{\frac{1}{2}}$, έχουμε:

$$\det\left(\frac{\partial \mathbf{y}}{\partial \mathbf{z}}\right) = h^{\frac{M}{2}}, \text{ οπότε:}$$

$$(k_h * u)(\mathbf{x}) - u(\mathbf{x}) = \int_{\mathbb{R}^M} k(\mathbf{z}) (u(\mathbf{x} - h^{\frac{1}{2}} \mathbf{z}) - u(\mathbf{x})) d\mathbf{z}$$

αναπτύσσοντας κατά Taylor την $u(\mathbf{x} - h^{\frac{1}{2}} \mathbf{z})$, προκύπτει:

$$\begin{aligned} (k_h * u)(\mathbf{x}) - u(\mathbf{x}) &= \int_{\mathbb{R}^M} k(\mathbf{z}) \left(-h^{\frac{1}{2}} \nabla u(\mathbf{x}) \cdot \mathbf{z} + \frac{h}{2} D^2 u(\mathbf{x})(\mathbf{z}, \mathbf{z}) \right) d\mathbf{z} - \\ &\quad - \frac{1}{6} \int_{\mathbb{R}^M} k(\mathbf{z}) D^3 u(\mathbf{x} - h^{\frac{1}{2}} \theta \mathbf{z})(\mathbf{z}, \mathbf{z}, \mathbf{z}) d\mathbf{z} \end{aligned}$$

για κάποιο $\theta = \theta(\mathbf{x}, \mathbf{z}, h) \in [0, 1]$. Χρησιμοποιώντας τις συνθήκες (4.8, 4.9) και το φράγμα (4.11), παίρνουμε τελικά:

$$\left| (k_h * u)(\mathbf{x}) - u(\mathbf{x}) - h \sum_{j=1}^M \frac{\partial^2 u}{\partial x_j^2} \right| \leq Ah^{\frac{3}{2}} \max_{\mathbf{x}} \|D^3 u(\mathbf{x})\|$$

□

4.1 Μελέτη της Ισοδυναμίας της Γενικής Ανισοτροπικής Διάχυσης Κατευθυνόμενων Λαπλασιανών με Επαναληπτική Τοπική Συνέλιξη 7

Πρέπει να σημειώσουμε ότι στο παραπάνω θεώρημα, όπως και πιο κάτω, η θετική σταθερά h αντιστοιχεί στο χρονικό βήμα της προσέγγισης.

Για να επεκτείνουμε το παραπάνω θεώρημα, ώστε να αντιστοιχεί στην γενική περίπτωση της (4.4), ορίζουμε καταρχήν, ένα σύστημα συντεταγμένων $\hat{\mathbf{x}}_T$, που εξαρτάται από έναν ταχυστή $T = V \cdot C \cdot V^{-1}$. Όταν ο ταχυστής αυτός αντιστοιχεί στην (4.4), είναι όπως είπαμε μεταβαλλόμενος, δηλαδή $T = T(\mathbf{x}_0, t_0)$, οπότε το $\hat{\mathbf{x}}_T$ είναι τοπικό σύστημα συντεταγμένων. Ορίζουμε λοιπόν το σύστημα $\hat{\mathbf{x}}_T$ σύμφωνα με τη σχέση:

$$\hat{\mathbf{x}}_T = B^{-1}V^{-1}\mathbf{x} \iff \mathbf{x} = VB\hat{\mathbf{x}}_T \quad (4.14)$$

$$\text{όπου } B = \begin{bmatrix} \sqrt{c_1} & & \emptyset \\ & \ddots & \\ \emptyset & & \sqrt{c_M} \end{bmatrix} = C^{\frac{1}{2}} \quad (4.15)$$

Πρόκειται, δηλαδή, για έναν γραμμικό μετασχηματισμό του \mathbf{x} , που αντιστοιχεί σε περιστροφή των αξόνων, σύμφωνα με τα ιδιοδιανύσματα του T , και σε ανομοιόμορφη αλλαγή κλίμακας των αξόνων, σύμφωνα με τις ιδιοτιμές του T . Όπως θα δούμε στη συνέχεια, με τη βοήθεια αυτού του ανισοτροπικού συστήματος συντεταγμένων, γίνεται δυνατή η απλοποίηση της περιγραφής της (4.4).

Επίσης ορίζουμε, μέσω του $k(\mathbf{x})$, έναν νέο μητρικό πυρήνα:

$$\hat{k}_T(\mathbf{x}) = \frac{1}{\sqrt{\det(T)}} k(B^{-1}V^{-1}\mathbf{x}), \quad T = VB^2V^T \quad (4.16)$$

και, όπως πριν, την αντίστοιχη οικογένεια πυρήνων:

$$\begin{aligned} \hat{k}_{T,h}(\mathbf{x}) &= \frac{1}{h^{\frac{M}{2}}} \hat{k}_T\left(\frac{\mathbf{x}}{h^{\frac{1}{2}}}\right) \\ &= \frac{1}{h^{\frac{M}{2}} \sqrt{\det(T)}} k\left(\frac{1}{h^{\frac{1}{2}}} B^{-1}V^{-1}\mathbf{x}\right), \quad T = VB^2V^T \\ &= \frac{1}{\sqrt{\det(T)}} k_h(B^{-1}V^{-1}\mathbf{x}), \quad \text{λόγω της (4.10)} \end{aligned} \quad (4.17)$$

Βλέπουμε ότι ο νέος μητρικός πυρήνας $\hat{k}_T(\mathbf{x})$ προκύπτει με περιστροφή και ανομοιόμορφη σε κάθε διεύθυνση, μεγέθυνση ή σμίκρυνση του πυρήνα $k(\mathbf{x})$, σύμφωνα με τον ταχυστή T . Όταν $T = T(\mathbf{x}_0, t_0)$, ο πυρήνας είναι χωρικά και χρονικά μεταβαλλόμενος. Συγκεκριμένα, λόγω του ότι ο $k(\mathbf{x})$ εκτείνεται το ίδιο σε όλες τις κατευθύνσεις των αξόνων (βλ. σχέση (4.9)), ο $\hat{k}_T(\mathbf{x})$ εκτείνεται τόσο περισσότερο στην διεύθυνση ενός ιδιοδιανύσματος του T , όσο μεγαλύτερη είναι η αντίστοιχη ιδιοτιμή του T . Για αυτό, χαρακτηρίζουμε τον πυρήνα αυτό σαν *ανισοτροπικό*.

Τέλος, μπορεί εύκολα να αποδειχθεί ότι και ο νέος πυρήνας $\hat{k}_T(\mathbf{x})$ ικανοποιεί και αυτός την συνθήκη κανονικοποίησης (4.7) (λόγω του πολλαπλασιαστικού παράγοντα $1/\sqrt{\det(T)}$ στη σχέση (4.16)) καθώς και την συνθήκη μηδενικών ροπών (4.8). Τα ίδια φυσικά ισχύουν και για τους $\hat{k}_{T,h}(\mathbf{x})$, αφού αυτοί προκύπτουν από τον $\hat{k}_T(\mathbf{x})$, με αλλαγή κλίμακας.

Στη συνέχεια, παραθέτουμε και αποδεικνύουμε την επέκταση του θεωρήματος 4.2 στη γενική περίπτωση, όπου ο ταυστής T είναι ανισοτροπικός:

Θεώρημα 4.3. Έστω $k(\mathbf{x})$ ένας πυρήνας και $w(\mathbf{x})$ μία συνάρτηση που ικανοποιούν τις προϋποθέσεις του θεωρήματος 4.2. Έστω επίσης \mathbf{x}_0 ένα σημείο του \mathbb{R}^M , T ένας ταυστής με ιδιοδιανύσματα $\mathbf{v}_1, \dots, \mathbf{v}_M$ και ιδιοτιμές $c_1, \dots, c_M > 0$ και $\hat{k}_{T,h}(\mathbf{x})$ ο πυρήνας που προκύπτει από αυτόν και τον $k(\mathbf{x})$ μέσω της (4.17). Τότε ισχύει:

$$(\hat{k}_{T,h} * w)(\mathbf{x}_0) - w(\mathbf{x}_0) = h \sum_{j=1}^M c_j \frac{\partial^2 w}{\partial \mathbf{v}_j^2} + f_i(\mathbf{x}_0) \quad (4.18)$$

$$\text{όπου: } |f_i(\mathbf{x}_0)| \leq Ah^{\frac{3}{2}} \max_{\mathbf{x}} \|D^3 w(VB\mathbf{x})\| = O(h^{\frac{3}{2}}) \quad (4.19)$$

όπου η σταθερά A δίνεται από την σχέση:

$$A = \int_{\mathbb{R}^M} |k(\mathbf{z})| \|\mathbf{z}\| d\mathbf{z} \quad (4.20)$$

Απόδειξη. Έστω για το σημείο \mathbf{x}_0 το τοπικό σύστημα συντεταγμένων $\hat{\mathbf{x}}_T$, που ορίζεται από την σχέση (4.14), με $VB^2V^T = T(\mathbf{x}_0)$. Εφαρμόζουμε το θεώρημα 4.2 στο σύστημα συντεταγμένων $\hat{\mathbf{x}}_T$, για τη συνάρτηση $w(VB\hat{\mathbf{x}}_T)$ και γράφουμε τις σχέσεις για το σημείο $\hat{\mathbf{x}}_{T,0} = B^{-1}V^{-1}\mathbf{x}_0$:

$$(k_h(\hat{\mathbf{x}}_T) * w(VB\hat{\mathbf{x}}_T))(\hat{\mathbf{x}}_{T,0}) - w(VB\hat{\mathbf{x}}_{T,0}) = h \sum_{j=1}^M \frac{\partial^2 w}{\partial \hat{x}_j^2} + f(\hat{\mathbf{x}}_{T,0}) \quad (4.21)$$

και

$$|f(\hat{\mathbf{x}}_{T,0})| \leq Ah^{\frac{3}{2}} \max_{\hat{\mathbf{x}}_T} \|D^3 w(VB\hat{\mathbf{x}}_T)\| \quad (4.22)$$

Για τις παραπάνω σχέσεις έχουμε πρώτα από όλα ότι:

$$f(\hat{\mathbf{x}}_{T,0}) = f(\mathbf{x}_0), \quad (4.23)$$

διότι προφανώς το $\hat{\mathbf{x}}_{T,0}$ εξαρτάται από το \mathbf{x}_0 .

Οι υπόλοιποι όροι των παραπάνω εξισώσεων γράφονται ως εξής:

4.1 Μελέτη της Ισοδυναμίας της Γενικής Ανισοτροπικής Διάχυσης Κατευθυνόμενων Λαπλασιανών με Επαναληπτική Τοπική Συνέλιξη 9

$$(1) w(VB\hat{\mathbf{x}}_{T,0}) = w(VBB^{-1}V^{-1}\mathbf{x}_0) = w(\mathbf{x}_0) \quad (4.24)$$

$$(2) \sum_{j=1}^M \frac{\partial^2 w}{\partial \hat{x}_j^2} = \sum_{j=1}^M c_j \frac{\partial^2 w}{\partial \mathbf{v}_j^2} \quad (4.25)$$

διότι για κάθε $j = 1, \dots, M$, έχουμε:

$$\frac{\partial w}{\partial \hat{x}_j} = \sum_{\ell=1}^M \frac{\partial w}{\partial x_\ell} \frac{\partial x_\ell}{\partial \hat{x}_j} = (\nabla w)^T \left(\frac{\partial \mathbf{x}}{\partial \hat{\mathbf{x}}_T} \right)^{(j)} \quad (4.26)$$

όπου $\left(\frac{\partial \mathbf{x}}{\partial \hat{\mathbf{x}}_T} \right)^{(j)}$ η j -οστή στήλη της Ιακωβιανής μήτρας $\partial \mathbf{x} / \partial \hat{\mathbf{x}}_T$. Όμως, λόγω της (4.14), έχουμε:

$$\frac{\partial \mathbf{x}}{\partial \hat{\mathbf{x}}_T} = VB = [\sqrt{c_1} \mathbf{v}_1 \quad \dots \quad \sqrt{c_j} \mathbf{v}_j \quad \dots \quad \sqrt{c_M} \mathbf{v}_M] \Rightarrow \left(\frac{\partial \mathbf{x}}{\partial \hat{\mathbf{x}}_T} \right)^{(j)} = \sqrt{c_j} \mathbf{v}_j \quad (4.27)$$

επομένως:

$$(4.26) \Rightarrow \frac{\partial w}{\partial \hat{x}_j} = \sqrt{c_j} \mathbf{v}_j^T \nabla w \quad (4.28)$$

οπότε, παραγωγίζοντας άλλη μια φορά, παίρνουμε

$$\frac{\partial^2 w}{\partial \hat{x}_j^2} = \sqrt{c_j} \mathbf{v}_j^T \nabla (\sqrt{c_j} \mathbf{v}_j^T \nabla w) = c_j \mathbf{v}_j^T (\nabla \otimes \nabla w) \mathbf{v}_j = c_j \frac{\partial^2 w}{\partial \mathbf{v}_j^2} \quad (4.29)$$

όπου $\nabla \otimes \nabla w = H_i$ η χωρική Hessian Μήτρα της w . Από την προηγούμενη σχέση, προκύπτει άμεσα η (4.25).

$$(3) (k_h(\hat{\mathbf{x}}_T) * w(VB\hat{\mathbf{x}}_T))(\hat{\mathbf{x}}_{T,0}) = \int_{\mathbb{R}^M} k_h(\mathbf{y}) w(VB(\hat{\mathbf{x}}_{T,0} - \mathbf{y})) d\mathbf{y} \quad (4.30)$$

$$= \int_{\mathbb{R}^M} k_h(\mathbf{y}) w(\mathbf{x}_0 - VB\mathbf{y}) d\mathbf{y}$$

Κάνουμε την εξής αλλαγή μεταβλητής ολοκλήρωσης:

$$\hat{\mathbf{y}} = VB\mathbf{y} \Leftrightarrow \mathbf{y} = B^{-1}V^{-1}\hat{\mathbf{y}} \quad (4.31)$$

Η Ιακωβιανή μήτρα του μετασχηματισμού αυτού είναι:

$$\frac{\partial \mathbf{y}}{\partial \hat{\mathbf{y}}} = B^{-1}V^{-1}$$

επομένως:

$$\begin{aligned} \det \left(\frac{\partial \mathbf{y}}{\partial \hat{\mathbf{y}}} \right) &= \frac{1}{\det(B)}, \text{ διότι } V \text{ ορθομοναδιαία, οπότε } \det(V^{-1}) = \det(V^T) = 1 \\ &= \frac{1}{\sqrt{\det(T)}}, \text{ διότι } \det(T) = \det(VB^2V^{-1}) = (\det(B))^2 \end{aligned}$$

Έτσι, χρησιμοποιώντας τη νέα μεταβλητή ολοκλήρωσης $\hat{\mathbf{y}}$, η (4.30) γίνεται:

$$\begin{aligned} (k_h(\hat{\mathbf{x}}_T) * w(VB\hat{\mathbf{x}}_T))(\hat{\mathbf{x}}_{T,0}) &= \int_{\mathbb{R}^M} k_h(B^{-1}V^{-1}\hat{\mathbf{y}})w(\mathbf{x}_0 - \hat{\mathbf{y}}) \frac{1}{\sqrt{\det(T)}} d\hat{\mathbf{y}} \\ &= \left(\frac{1}{\sqrt{\det(T)}} k_h(B^{-1}V^{-1}\mathbf{x}) * w(\mathbf{x}) \right) (\mathbf{x}_0), \end{aligned}$$

από την οποία, λόγω της (4.17), προκύπτει τελικά:

$$(k_h(\hat{\mathbf{x}}_T) * w(VB\hat{\mathbf{x}}_T))(\hat{\mathbf{x}}_{T,0}) = (\hat{k}_{T,h} * w)(\mathbf{x}_0) \quad (4.32)$$

Επομένως, οι σχέσεις (4.21,4.22), μέσω των σχέσεων (4.23,4.24,4.25,4.32), καταλήγουν άμεσα στις αποδεικτές. \square

Από το θεώρημα αυτό, μπορούμε να καταλήξουμε σε ένα φράγμα του σφάλματος της προσέγγισης της (4.4) με τοπική συνέλιξη. Αυτό φαίνεται στο ακόλουθο θεώρημα που αποδεικνύουμε, το οποίο αναφέρεται στην εφαρμογή μίας επανάληψης της τοπικής συνέλιξης:

Θεώρημα 4.4. Έστω $\mathbf{u}(\mathbf{x}, t)$ η λύση της εξίσωσης (4.4), $k(\mathbf{x})$ ένας πυρήνας που ικανοποιεί τις προϋποθέσεις του θεωρήματος 4.2 και $\hat{k}_{T(\mathbf{x},t),h}(\mathbf{x})$ ο πυρήνας που προκύπτει από αυτόν και τον ταυστή διάχυσης $T(\mathbf{x}, t)$ της (4.4), μέσω της (4.17). Τότε, για οποιοδήποτε σημείο \mathbf{x}_0 και χρονική στιγμή $t_0 > 0$ και για κάθε $i = 1, \dots, N$, ισχύει:

$$\left(\hat{k}_{T(\mathbf{x}_0, t_0), h}(\mathbf{x}) * u_i(\mathbf{x}, t_0) \right) (\mathbf{x}_0) - u_i(\mathbf{x}_0, t_0 + h) = e_i(\mathbf{x}_0, t_0) \quad (4.33)$$

όπου:

$$|e_i(\mathbf{x}_0, t_0)| \leq A \max_{\mathbf{x}} \|D^3 u_i(VB\mathbf{x})\| h^{\frac{3}{2}} + \frac{1}{2} \max_{0 \leq \theta \leq 1} |u_{i,tt}(\mathbf{x}_0, t_0 + \theta h)| h^2 = O(h^{\frac{3}{2}}) \quad (4.34)$$

και η σταθερά A δίνεται ξανά από την σχέση:

$$A = \int_{\mathbb{R}^M} |k(\mathbf{z})| \|\mathbf{z}\| d\mathbf{z} \quad (4.35)$$

4.1 Μελέτη της Ισοδυναμίας της Γενικής Ανισοτροπικής Διάχυσης Κατευθυνόμενων Λαπλασιανών με Επαναληπτική Τοπική Συνέλιξη 81

Απόδειξη. Για κάθε $i = 1, \dots, N$, εφαρμόζουμε το θεώρημα 4.3 για τανυστή $T = T(\mathbf{x}_0, t_0)$ αυτόν της εξίσωσης (4.4) και για συνάρτηση $w(\mathbf{x}) = u_i(\mathbf{x}, t_0)$, οπότε προκύπτει:

$$(\hat{k}_{T(\mathbf{x}_0, t_0), h}(\mathbf{x}) * u_i(\mathbf{x}, t_0))(\mathbf{x}_0) - u_i(\mathbf{x}_0, t_0) = h \sum_{j=1}^M c_j \frac{\partial^2 u_i}{\partial \mathbf{v}_j^2} + f_i(\mathbf{x}_0, t_0) \quad (4.36)$$

$$\text{όπου: } |f_i(\mathbf{x}_0)| \leq Ah^{\frac{3}{2}} \max_{\mathbf{x}} \|D^3 u_i(VB\mathbf{x})\| = O(h^{\frac{3}{2}}) \quad (4.37)$$

Από την άλλη, αν αναπτύξουμε κατά Taylor την συνάρτηση $u_i(\mathbf{x}_0, t)$ (θεωρώντας τον t σαν μεταβλητή) στο σημείο t_0 , παίρνουμε:

$$u_i(\mathbf{x}_0, t_0 + h) - u_i(\mathbf{x}_0, t_0) = h \frac{\partial u_i}{\partial t}(\mathbf{x}_0, t_0) + g_i(\mathbf{x}_0, t_0) \quad (4.38)$$

$$\text{όπου: } g_i(\mathbf{x}_0, t_0) = u_{i,tt}(\mathbf{x}_0, t_0 + \theta h) \frac{h^2}{2}, \text{ για κάποιο } \theta \in [0, 1] \quad (4.39)$$

άρα:

$$|g_i(\mathbf{x}_0, t_0)| \leq \frac{1}{2} \max_{0 \leq \theta \leq 1} |u_{i,tt}(\mathbf{x}_0, t_0 + \theta h)| h^2 \quad (4.40)$$

Όμως, επειδή η $\mathbf{u}(\mathbf{x}, t)$ είναι λύση της (4.4), η $\partial u_i / \partial t$ για κάθε $i = 1, \dots, N$ δίνεται από τη σχέση αυτή, οπότε η (4.38) γίνεται:

$$u_i(\mathbf{x}_0, t_0 + h) - u_i(\mathbf{x}_0, t_0) = h \sum_{j=1}^M c_j \frac{\partial^2 u_i}{\partial \mathbf{v}_j^2} + g_i(\mathbf{x}_0, t_0) \quad (4.41)$$

Αν από την (4.36) αφαιρέσουμε κατά μέλη την (4.41), προκύπτει:

$$(\hat{k}_{T(\mathbf{x}_0, t_0), h}(\mathbf{x}) * u_i(\mathbf{x}, t_0))(\mathbf{x}_0) - u_i(\mathbf{x}_0, t_0 + h) = e_i(\mathbf{x}_0, t_0)$$

όπου:

$$e_i(\mathbf{x}_0, t_0) = f_i(\mathbf{x}_0, t_0) - g_i(\mathbf{x}_0, t_0)$$

οπότε:

$$|e_i(\mathbf{x}_0, t_0)| \leq |f_i(\mathbf{x}_0, t_0)| + |g_i(\mathbf{x}_0, t_0)|$$

Άρα, λόγω των (4.37), (4.40), προκύπτει τελικά:

$$|e_i(\mathbf{x}_0, t_0)| \leq A \max_{\mathbf{x}} \|D^3 u_i(VB\mathbf{x})\| h^{\frac{3}{2}} + \frac{1}{2} \max_{0 \leq \theta \leq 1} |u_{i,tt}(\mathbf{x}_0, t_0 + \theta h)| h^2$$

και άρα η απόδειξη ολοκληρώθηκε. □

Για να εκφράσουμε την επαναληπτική διαδικασία της τοπικής συνέλιξης, ορίζουμε έναν μη γραμμικό τελεστή \mathbb{T}_h : Θεωρούμε τη γενική εξίσωση (4.4), το συναρτησιακό του τανυστή $\tilde{T}[\cdot]$ που σχετίζεται με αυτή (βλ. σχέση (4.6)) και μία συνάρτηση $\tilde{\mathbf{u}}$, την οποία κατανοούμε σαν προσέγγιση της λύσης της (4.4) κάποια χρονική στιγμή t . Έτσι, ο τελεστής \mathbb{T}_h^t ορίζεται ως εξής: Για οποιοδήποτε \mathbf{x}_0 , πρέπει να ισχύει:

$$(\mathbb{T}_h \tilde{\mathbf{u}})(\mathbf{x}_0) := \left(\hat{k}_{\tilde{T}[\tilde{\mathbf{u}}(\mathbf{x}-\mathbf{x}_0)],h} * \tilde{\mathbf{u}} \right) (\mathbf{x}_0) \quad (4.42)$$

Πρέπει να τονίσουμε ότι ο τελεστής αυτός προκύπτει εφαρμόζοντας τοπική συνέλιξη με διαφορετικό πυρήνα σε κάθε σημείο (που εξαρτάται από την τιμή του τανυστή $\tilde{T}[\tilde{\mathbf{u}}(\mathbf{x} - \mathbf{x}_0)]$) και παίρνοντας το αποτέλεσμα για όλα τα σημεία. Έτσι, όπως έχουμε πει, η συμπεριφορά της συνάρτησης $\tilde{\mathbf{u}}$ στην γειτονιά του κάθε σημείου \mathbf{x}_0 καθορίζει την τιμή του τανυστή T στο \mathbf{x}_0 και άρα την μορφή της τοπικής συνέλιξης που εφαρμόζει ο τελεστής \mathbb{T}_h .

Συμβολίζουμε με $(\mathbb{T}_h \tilde{\mathbf{u}})_i$ το i -στό κανάλι της εξόδου $(\mathbb{T}_h \tilde{\mathbf{u}})$. Επίσης, συμβολίζουμε με $(\mathbb{T}_h)^n$ την εφαρμογή του τελεστή \mathbb{T}_h σε μία συνάρτηση, n διαδοχικές φορές. Έτσι, η επαναληπτική διαδικασία της τοπικής συνέλιξης μπορεί να εκφραστεί ως εξής:

$((\mathbb{T}_h)^n \mathbf{u}^0)(\mathbf{x})$, για την προσέγγιση της λύσης $\mathbf{u}(\mathbf{x}, nh)$ στο σημείο (\mathbf{x}, nh) .

Βασισμένοι στα παραπάνω, αποδεικνύουμε την σύγκλιση της προσέγγισης αυτής στην λύση της (4.4) (όταν αυτή ορίζεται με σταθερό αλλά ανισοτροπικό τανυστή διάχυσης T), καθώς το χρονικό βήμα h γίνεται απειροστό (το οποίο αποτελεί γενίκευση του αντίστοιχου θεωρήματος του [27]):

Θεώρημα 4.5. Έστω πάλι ότι $\mathbf{u}(\mathbf{x}, t)$ η λύση της εξίσωσης (4.4), στην ειδική περίπτωση όπου ο τανυστής T είναι ανισοτροπικός αλλά σταθερός. Έστω επίσης $k(\mathbf{x})$ ένας γενικός πυρήνας που ικανοποιεί τις προϋποθέσεις του θεωρήματος 4.2 και $\hat{k}_{T,h}(\mathbf{x})$ ο πυρήνας που προκύπτει από αυτόν και τον τανυστή T , μέσω της (4.17). Τότε, ισχύει:

$$((\mathbb{T}_h)^n \mathbf{u}^0)(\mathbf{x}_0) \rightarrow \mathbf{u}(\mathbf{x}_0, t) \text{ καθώς } n \rightarrow \infty, nh \rightarrow t \quad (4.43)$$

Απόδειξη. Εφαρμόζουμε το θεώρημα 4.4 για $T(\mathbf{x}, t) = \text{σταθερό}$, για το σημείο \mathbf{x}_0 και για τη χρονική στιγμή $t_0 > 0$, οπότε, λόγω και του ορισμού (4.42) του τελεστή \mathbb{T}_h , προκύπτει, για κάθε $i = 1, \dots, N$:

$$(\mathbb{T}_h \tilde{\mathbf{u}})_i(\mathbf{x}_0, t_0) - u_i(\mathbf{x}_0, t_0 + h) = O(h^{\frac{3}{2}}) \quad (4.44)$$

Παρατηρούμε ότι εάν $u_i \leq M$, τότε $(\mathbb{T}_h \tilde{\mathbf{u}})_i \leq M$ (αρχή μεγίστου), διότι ο τανυστής \mathbb{T}_h εκφράζει συνέλιξη με τον πυρήνα $\hat{k}_{T,h}(\mathbf{x})$, ο οποίος είναι κανονικοποιημένος. Επομένως, αφού η σχέση (4.44) ισχύει για κάθε \mathbf{x}_0 , η συνάρτηση

4.1 Μελέτη της Ισοδυναμίας της Γενικής Ανισοτροπικής Διάχυσης Κατευθυνόμενων Λαπλασιανών με Επαναληπτική Τοπική Συνέλιξη \mathfrak{B}

$(\mathbb{T}_h(O(h^{\frac{3}{2}})))_i$ είναι επίσης $O(h^{\frac{3}{2}})$. Έτσι, αφού ο τελεστής \mathbb{T}_h είναι γραμμικός (λόγω σταθερού T), αν τον εφαρμόσουμε και στα δύο μέλη της (4.44), προκύπτει:

$$((\mathbb{T}_h)^2 \tilde{\mathbf{u}})_i(\mathbf{x}_0, t_0) - \mathbb{T}_h u_i(\mathbf{x}_0, t_0 + h) = O(h^{\frac{3}{2}}) \quad (4.45)$$

Αν γράψουμε την σχέση (4.44) για την χρονική στιγμή $t_0 + h$, αντί της t_0 , παίρνουμε:

$$(\mathbb{T}_h \tilde{\mathbf{u}})_i(\mathbf{x}_0, t_0 + h) - u_i(\mathbf{x}_0, t_0 + 2h) = O(h^{\frac{3}{2}}) \quad (4.46)$$

Προσθέτοντας κατά μέλη την (4.44) με την (4.45), εξάγεται:

$$((\mathbb{T}_h)^2 \tilde{\mathbf{u}})_i(\mathbf{x}_0, t_0) - u_i(\mathbf{x}_0, t_0 + 2h) = 2O(h^{\frac{3}{2}}) \quad (4.47)$$

Αν επαναλάβουμε n φορές την παραπάνω διαδικασία, καταλήγουμε στη σχέση:

$$((\mathbb{T}_h)^n \tilde{\mathbf{u}})_i(\mathbf{x}_0, t_0) - u_i(\mathbf{x}_0, t_0 + nh) = nO(h^{\frac{3}{2}}) \quad (4.48)$$

Θέτοντας $h = \frac{\tau}{n}$ και θεωρώντας ότι $n \rightarrow \infty$, προκύπτει:

$$((\mathbb{T}_h)^n \tilde{\mathbf{u}})_i(\mathbf{x}_0, t_0) - u_i(\mathbf{x}_0, t_0 + \tau) = O\left(\left(\frac{\tau}{n}\right)^{\frac{1}{2}}\right) \quad (4.49)$$

Επομένως, αν θεωρήσουμε ότι η αρχική συνθήκη της (4.4) είναι τέτοια που η λύση της $\mathbf{u}(\mathbf{x}, t)$ να είναι συνεχής για $t \rightarrow 0$, μπορούμε να γράψουμε την σχέση (4.49) για $t = 0$. Έτσι, εφόσον $n \rightarrow \infty$, έχουμε:

$$((\mathbb{T}_h)^n \tilde{\mathbf{u}})_i(\mathbf{x}_0, t_0) - u_i(\mathbf{x}_0, t_0 + \tau) \rightarrow 0, \text{ για κάθε } i = 1, \dots, N \quad (4.50)$$

από το οποίο προκύπτει άμεσα η αποδεικτέα σχέση. □

Για το παραπάνω θεώρημα πρέπει να παρατηρήσουμε ότι, όταν ο μητρικός πυρήνας είναι Gaussian (βλ. πιο κάτω, σχέση (4.52) για τον ορισμό του), η συνέλιξη με την συνάρτηση \mathbf{u}^0 δίνει ακριβώς τη λύση (μηδενικό σφάλμα), ανεξάρτητα από το χρονικό βήμα h (βλ. [52] για την απόδειξη). Το παραπάνω θεώρημα αποδεικνύει ότι για απειροστό βήμα, καταλήγουμε στη λύση και με οποιοδήποτε άλλο βαθυπερατό πυρήνα, ο οποίος προσεγγίζει τη συμπεριφορά του Gaussian πυρήνα (κάτι το οποίο παρουσιάζει αντιστοιχία με το Κεντρικό Οριακό Θεώρημα).

Συνοψίζοντας, σημειώνουμε ότι τα παραπάνω θεωρήματα προσφέρουν κάποια θεωρητικά αποτελέσματα για την προσέγγιση της λύσης της (4.4) με επαναληπτική συνέλιξη με γενικούς βαθυπερατούς πυρήνες. Εκτός από την απόδειξη της σύγκλισης στη λύση για την ειδική περίπτωση σταθερού ταυστή T , παρέχεται και ένα φράγμα για το σφάλμα της προσέγγισης για την γενική περίπτωση (βλ. θεώρημα (4.4)). Όπως φαίνεται από το θεώρημα αυτό, το σφάλμα

αυτό αναφέρεται σε μόνο μία επανάληψη της τοπικής συνέλιξης και θεωρώντας ότι στη δεδομένη χρονική στιγμή, η διαθέσιμη συνάρτηση ταυτίζεται με τη λύση. Η υπόθεση αυτή δεν είναι βέβαια ρεαλιστική (ιδιαίτερα όταν έχουν προηγηθεί αρκετές επαναλήψεις) αλλά παρ'όλα αυτά, μέσω αυτής, είναι δυνατή η εκτίμηση της συνεισφοράς της κάθε επανάληψης στο ολικό σφάλμα, λόγω της προσθετικής μορφής του. Έτσι, το μερικό αυτό σφάλμα είναι πράγματι αντιπροσωπευτικό της ακρίβειας της προσέγγισης.

Επίσης, από τη σχέση (4.34) παρατηρούμε ότι το φράγμα του σφάλματος περιέχει δύο προσθετικούς όρους. Όμως, όταν το χρονικό βήμα h μικραίνει, το h^2 γίνεται όλο και πιο αμελητέο σε σχέση με το $h^{\frac{3}{2}}$ και έτσι ο πρώτος όρος της (4.34) (που περιέχει το $h^{\frac{3}{2}}$) γίνεται όλο και πιο σημαντικός σε σχέση με τον δεύτερο. Έτσι, ιδιαίτερο ενδιαφέρον παρουσιάζει η πολλαπλασιαστική σταθερά A στον 1ο όρο του φράγματος, που, όπως παρουσιάσαμε, δίνεται από τη σχέση:

$$A = \int_{\mathbb{R}^M} |k(\mathbf{z})| \|\mathbf{z}\| d\mathbf{z} \quad (4.51)$$

Το ενδιαφέρον είναι ότι η A εξαρτάται μόνο από τον μητρικό πυρήνα συνέλιξης $k(\mathbf{x})$ που χρησιμοποιούμε, δηλαδή $A = A[k]$. Αλλά και αντίστροφα, η επιλογή του πυρήνα δεν επηρεάζει στο φράγμα αυτό κανέναν άλλο όρο, παρά μόνο την $A = A[k]$. Θα μπορούσε λοιπόν η $A[k]$ να χρησιμοποιηθεί σαν ένα μέτρο της καταλληλότητας (από την σκοπιά του φράγματος αυτού) του μητρικού πυρήνα συνέλιξης $k(\mathbf{x})$ για την προσέγγιση της λύσης της (4.4). Όσο μικρότερη είναι η $A[k]$ ενός πυρήνα, τόσο περισσότερο εξαναγκάζεται το φράγμα να μικρύνει, ανεξάρτητα από οποιοδήποτε άλλο παράγοντα. Βέβαια, μπορεί να υπάρχουν και άλλοι παράγοντες που να επηρεάζουν το σφάλμα αυτό και έτσι το φράγμα αυτό να μην είναι αντιπροσωπευτικό. Επίσης, από πρακτικής σκοπιάς, πρέπει να ληφθεί υπόψη και η χωρική διακριτοποίηση που είναι απαραίτητη για την αριθμητική υλοποίηση.

Ειδικές περιπτώσεις πυρήνων συνέλιξης

Ιδιαίτερο ενδιαφέρον παρουσιάζει η περίπτωση όπου ο μητρικός πυρήνας $k(\mathbf{x})$ είναι ακτινική συνάρτηση, δηλαδή $k(\mathbf{x}) = k(\|\mathbf{x}\|)$. Στην περίπτωση αυτή, οι ισούψεις του $k(\mathbf{x})$ είναι ομόκεντρες υπερσφαίρες στον \mathbb{R}^M (κύκλοι όταν $M = 2$) με κέντρο την αρχή των αξόνων. Όμως, ο ανισοτροπικός μητρικός πυρήνας $\hat{k}_T(\mathbf{x})$ που προκύπτει από τον $k(\mathbf{x})$ και έναν ανισοτροπικό ταυστή T , μέσω της σχέσης (4.16), γράφεται (λόγω του ότι $\|\mathbf{x}\| = (\mathbf{x}^T \mathbf{x})^{\frac{1}{2}}$):

$$\hat{k}_T(\mathbf{x}) = \frac{1}{\sqrt{\det(T)}} k \left((\mathbf{x}^T V B^{-1} B^{-1} V^{-1} \mathbf{x})^{\frac{1}{2}} \right), \text{ διότι: } V^T = V^{-1}, B^T = B$$

4.1 Μελέτη της Ισοδυναμίας της Γενικής Ανισοτροπικής Διάχυσης Κατευθυνόμενων Λαπλασιανών με Επαναληπτική Τοπική Συνέλιξη 85

$$\Rightarrow \hat{k}_T(\mathbf{x}) = \frac{1}{\sqrt{\det(T)}} k \left((\mathbf{x}^T T^{-1} \mathbf{x})^{\frac{1}{2}} \right), \text{ διότι: } T = VB^2V^{-1}$$

που σημαίνει ότι οι ισοϋψείς του $\hat{k}_T(\mathbf{x})$ είναι υπερελλείψεις (ελλείψεις όταν $M = 2$), αφού η εξίσωση $\mathbf{x}^T T^{-1} \mathbf{x} = c$ παριστάνει υπερέλλειψη. Έτσι, ο προσανατολισμός των υπερελλείψεων και η σχέση μεταξύ των αξόνων τους εξαρτάται από τα ιδιοδιανύσματα και τη σχέση μεταξύ των ιδιοτιμών του ταχυστή T .

Η οικογένεια πυρήνων $k_h(\mathbf{x})$, που προκύπτει από τον μητρικό πυρήνα $\hat{k}_{T,h}(\mathbf{x})$ με αλλαγές κλίμακας (βλ. σχέση (4.17)) γράφεται:

$$\hat{k}_{T,h}(\mathbf{x}) = \frac{1}{h^{\frac{M}{2}} \sqrt{\det(T)}} k \left(\left(\frac{1}{h} \mathbf{x}^T T^{-1} \mathbf{x} \right)^{\frac{1}{2}} \right)$$

Η πιο σημαντική περίπτωση ακτινικού πυρήνα είναι ο Gaussian τύπος. Η M -διάστατη Gaussian συνάρτηση που ικανοποιεί τις προϋποθέσεις του ορισμού 4.1 είναι αυτή με μήτρα συμμεταβλητότητας $2I$. Έτσι, ο Gaussian μητρικός πυρήνας $k(\mathbf{x})$ γράφεται:

$$k(\mathbf{x}) = \frac{1}{(4\pi)^{\frac{M}{2}}} \exp\left(-\frac{1}{4} \|\mathbf{x}\|^2\right) \quad (4.52)$$

Ο αντίστοιχος ανισοτροπικός πυρήνας με ταχυστή T και η προκύπτουσα οικογένεια πυρήνων $k_h(\mathbf{x})$ γράφονται:

$$\hat{k}_T(\mathbf{x}) = \frac{1}{(4\pi)^{\frac{M}{2}} \sqrt{\det(T)}} \exp\left(-\frac{1}{4} \mathbf{x}^T T^{-1} \mathbf{x}\right)$$

$$\hat{k}_{T,h}(\mathbf{x}) = \frac{1}{(4\pi h)^{\frac{M}{2}} \sqrt{\det(T)}} \exp\left(-\frac{1}{4h} \mathbf{x}^T T^{-1} \mathbf{x}\right)$$

Η $\hat{k}_{T,h}$ είναι δηλαδή η M -διάστατη Gaussian συνάρτηση με μήτρα συμμεταβλητότητας $2hT$. Στην ειδική περίπτωση όπου $M = 2$ (όταν η συνάρτηση $\mathbf{u}(\mathbf{x})$ παριστάνει διανυσματική εικόνα), η παραπάνω οικογένεια ανισοτροπικών Gaussian πυρήνων $\hat{k}_{T,h}$ γράφεται:

$$\hat{k}_{T,h}(\mathbf{x}) = \frac{1}{4\pi h \sqrt{\det(T)}} \exp\left(-\frac{1}{4h} \mathbf{x}^T T^{-1} \mathbf{x}\right)^1 \quad (4.53)$$

Από πρακτικής σκοπιάς, η επιλογή ακτινικού μητρικού πυρήνα στην αριθμητική υλοποίηση είναι ιδιαίτερα προτιμητέα, διότι δεν δείχνει εκ των προτέρων

¹όπως έχουμε αναφέρει, η ειδική αυτή περίπτωση Gaussian πυρήνων στο συγκεκριμένο πρόβλημα έχει μελετηθεί από τους Tschumperlé και Deriche [52, 55], οι οποίοι έχουν καταλήξει στην ακόλουθη σχέση για τον πυρήνα $\hat{k}_{T,h}$ (με τον συμβολισμό που χρησιμοποιούμε

προτίμηση σε κάποια κατεύθυνση του χώρου. Επιπρόσθετα, ο Gaussian μητρικός πυρήνας αποτελεί μία αξιόπιστη επιλογή, διότι παρουσιάζει επιθυμητά χαρακτηριστικά με βάση διάφορα κριτήρια. Για την περίπτωση που μελετάται εδώ, το σημαντικότερο πλεονέκτημα του πυρήνα αυτού είναι ότι, όταν ο τανυστής T είναι σταθερός, η χρήση του για την προσέγγιση δίνει, όπως είδαμε, ακριβώς τη λύση (μηδενικό σφάλμα).

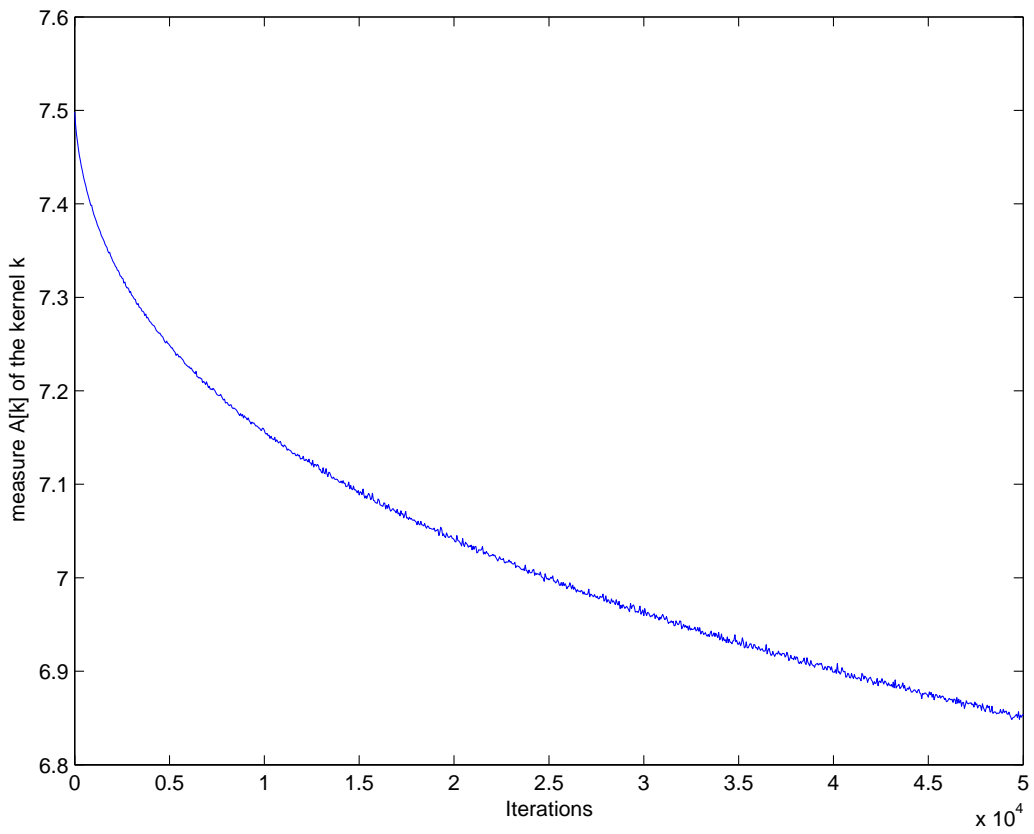
Για την επιλογή του μητρικού πυρήνα που θα χρησιμοποιηθεί, θα μπορούσε να ληφθεί υπόψιν και το “μέτρο” $A[k]$ που εισαγάγαμε πιο πάνω. Διαπιστώνουμε λοιπόν ότι ο Gaussian πυρήνας δεν έχει ελάχιστη τιμή $A[k]$. Συγκεκριμένα, για το διδιάστατο πρόβλημα ($M = 2$) και περιοριζόμενοι σε ακτινικούς μητρικούς πυρήνες, χρησιμοποιήσαμε ροή ταχύτερης καθόδου για την ελαχιστοποίηση του $A[k]$, με προβολή σε κάθε βήμα στους περιορισμούς (4.7-4.9). Σαν αρχική τιμή θεωρήσαμε τον Gaussian πυρήνα και έτσι βρήκαμε τροποποιήσεις του με μικρότερο $A[k]$ (βλ. σχήματα 4.1 και 4.2). Έτσι, μία πρόταση για την προσέγγιση της (4.4) είναι η χρήση τέτοιων πυρήνων με την ελπίδα η προσέγγιση να έχει μεγαλύτερη ακρίβεια. Βέβαια, η καλύτερη προσέγγιση δεν είναι εγγυημένη, διότι υπάρχουν και άλλοι παράγοντες, όπως αυτοί που αναφέραμε πιο πάνω.

εδώ):

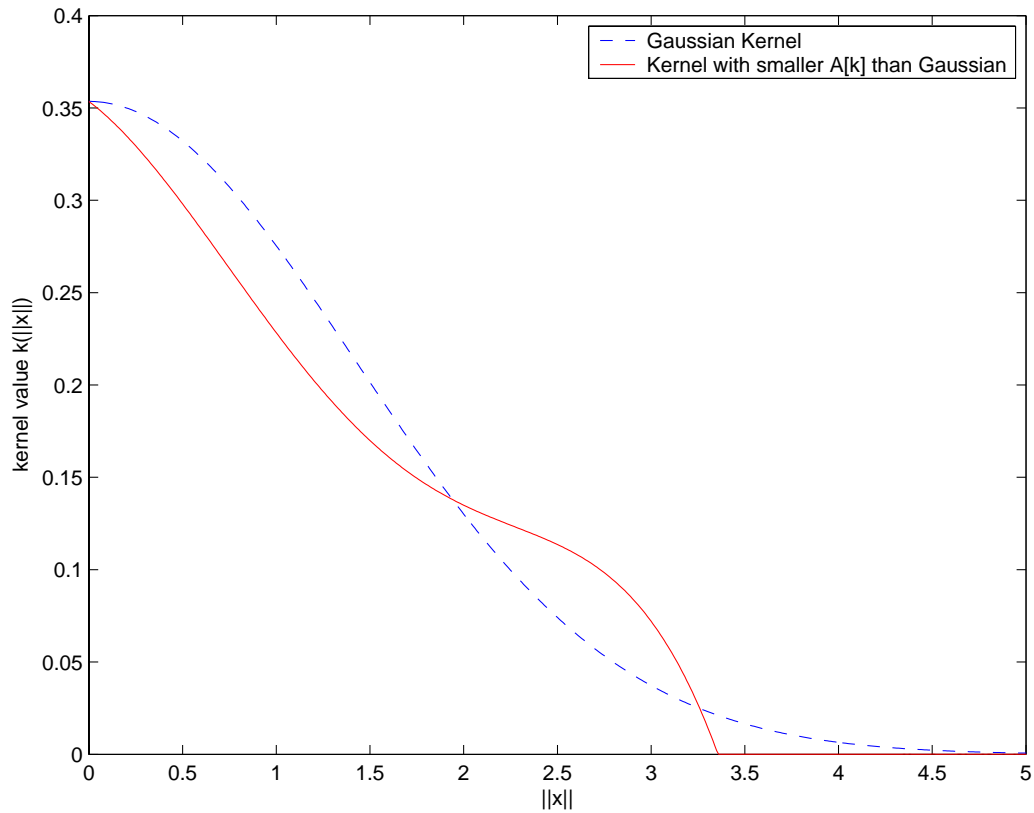
$$\hat{k}_{T,h}(\mathbf{x}) = \frac{1}{4\pi h} \exp\left(-\frac{1}{4h} \mathbf{x}^T T^{-1} \mathbf{x}\right)$$

Παρατηρούμε, λοιπόν μία διαφορά ανάμεσα στην τελευταία έκφραση και στην έκφραση (4.53), που δεν θα έπρεπε να υπάρχει. Πιστεύουμε ότι η έκφραση (4.53) είναι η σωστή, διότι, σε αντίθεση με την τελευταία έκφραση, αντιστοιχεί σε κανονικοποιημένο πυρήνα. Στην σχετική απόδειξη του [52] εντοπίζουμε το πρόβλημα στην φράση ‘...since the Gaussian function is normalized’, η οποία όπως είπαμε δεν ισχύει. Τέλος, σημειώνουμε ότι το πρόβλημα αυτό δεν γίνεται αντιληπτό στην αντίστοιχη αριθμητική υλοποίηση του [52], αφού, μετά τον υπολογισμό των τιμών του πυρήνα στο διακριτό πλέγμα, εφαρμόζεται, για αριθμητικούς λόγους, επιπλέον κανονικοποίηση.

4.1 Μελέτη της Ισοδυναμίας της Γενικής Ανιστροπικής Διάχυσης Κατευθυνόμενων Λαπλασιανών με Επαναληπτική Τοπική Συνέλιξη 87



Σχήμα 4.1: Πορεία της αριθμητικής επαναληπτικής μεθόδου, για την μείωση του “μέτρου” $A[k]$, που προτείνουμε.



Σχήμα 4.2: Τελικό αποτέλεσμα της επαναληπτικής διαδικασίας του σχήμ. 4.1: Ο πυρήνας που προκύπτει μετά από $5 \cdot 10^4$ επαναλήψεις, με μειωμένο, σε σχέση με τον Gaussian, “μέτρο” $A[k]$.

4.2 Αποσαφήνιση της σχέσης ανάμεσα στην συνεχή και στην εξαρχής διακριτή θεώ- ρηση των ϕ -συναρτησιακών

4.2.1 Εισαγωγή

Στην ενότητα αυτή, μελετάμε κάποια θεωρητικά ζητήματα, που αφορούν την αριθμητική υλοποίηση των ϕ -συναρτησιακών (ϕ -functionals), τα οποία παρουσιάσαμε αναλυτικά στην ενότητα 1.6. Για τις βαθμωτές εικόνες, η γενική μορφή ενός ϕ -συναρτησιακού είναι η εξής:

$$J[u] = \iint_{\Omega} \phi(\|\nabla u\|) \, d\mathbf{x} + \frac{\lambda}{2} \iint_{\Omega} (u - u^0)^2 \, d\mathbf{x} \quad (4.54)$$

όπου u^0 είναι η αρχική εικόνα. Η συνάρτηση $\phi(s)$ είναι αύξουσα και με κατάλληλη μορφή, τέτοια ώστε ο πρώτος όρος της (4.54) να επιφέρει μεγάλο κόστος στις ανεπιθύμητες έντονες διακυμάνσεις της εικόνας (που οφείλονται π.χ. στον θόρυβο). Ο δεύτερος όρος της (4.54) επιφέρει κόστος στην απομάκρυνση της εικόνας u^0 από την αρχική και χρησιμεύει στην αποφυγή της υπερβολικής αλλοίωσης της εικόνας. Η σταθερά λ ονομάζεται όρος πιστότητας (fidelity term) και καθορίζει την ισορροπία ανάμεσα σε αυτούς τους δύο όρους του συναρτησιακού $J[u]$ (βλ. και ενότητα 1.4).

Όταν λοιπόν η αρχική εικόνα περιέχει θόρυβο ή κάποια άλλη ανεπιθύμητη διαταραχή, η ελαχιστοποίηση του ϕ -συναρτησιακού $J[u]$ (4.54) δίνει σαν αποτέλεσμα μία βελτιωμένη τελική εικόνα. Με τη γενική μορφή (4.54) του ϕ -συναρτησιακού $J[u]$ μπορεί να περιγραφεί ενοποιητικά ένα αρκετά μεγάλο μέρος μεθόδων μεταβολών (variational methods) που εφαρμόζονται στην Βελτίωση (Enhancement) εικόνων. Η πιο σημαντική ειδική περίπτωση ϕ -συναρτησιακού είναι η μέθοδος της Ολικής Μεταβολής (Total Variation - TV - βλ. 1.4), που αντιστοιχεί στην επιλογή της συνάρτησης $\phi(s) = s$.

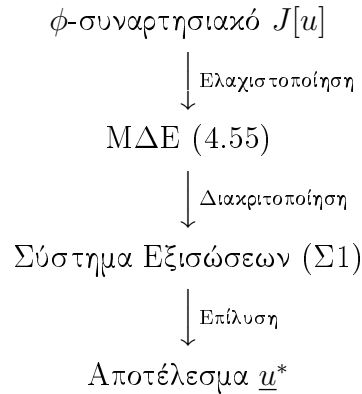
Επομένως, σε μία μέθοδο ϕ -συναρτησιακού, ζητούμενο είναι η εύρεση της εικόνας που ελαχιστοποιεί το συναρτησιακό. Μάλιστα, στην αριθμητική υλοποίηση, τόσο η αρχική όσο και η ζητούμενη εικόνα είναι φυσικά διακριτές. Θα μπορούσαμε λοιπόν να διακρίνουμε δύο κύριες τακτικές (θεωρήσεις), για την αριθμητική λύση του προβλήματος αυτού:

- **Θεώρηση 1^η (Συνεχής): Ελαχιστοποίηση και μετά Διακριτοποίηση.**
Στη θεώρηση αυτή, λύνεται πρώτα από όλα το πρόβλημα της ελαχιστοποίησης του ϕ -συναρτησιακού $J[u]$ (4.54), χρησιμοποιώντας Λογισμό των Μεταβολών (Calculus of Variations). Έτσι, από τις Euler-Lagrange εξισώσεις για την ελαχιστοποίηση του $J[u]$, προκύπτει η εξής ΜΔΕ:

$$\operatorname{div} \left(\frac{\phi'(\|\nabla u\|)}{\|\nabla u\|} \nabla u \right) - \lambda(u - u^0) = 0 \quad (4.55)$$

Το επόμενο βήμα στη θεώρηση αυτή είναι να διακριτοποιήσουμε την ΜΔΕ (4.55), οπότε καταλήγουμε σε ένα σύστημα μη γραμμικών εξισώσεων για τις τιμές της ζητούμενης εικόνας u^* στα pixels.

Τα βήματα αυτά φαίνονται σχηματικά στο διάγραμμα του σχήματος 4.3.



Σχήμα 4.3: Διάγραμμα 1^{ης} θεώρησης για την αριθμητική υλοποίηση των μεθόδων ϕ -συναρτησιακού

Η θεώρηση αυτή παρουσιάζει το πλεονέκτημα ότι εφαρμόζει την διακριτοποίηση μόνο σαν τελικό στάδιο της μεθόδου. Έτσι, έχει όλα τα πλεονεκτήματα των μεθόδων που βασίζονται στις ΜΔΕ, μιας και επιτρέπει την αξιολόγηση της μεθόδου στο συνεχή χώρο, η οποία είναι πιο απλή και αφαιρετική, χωρίς εξάρτηση από κάποιο διακριτό πλέγμα ορισμού της εικόνας. Τέλος, όπως διαφαίνεται και από την παρουσίαση των κεφαλαίων 1, 2 και 3, η θεώρηση αυτή ακολουθείται στις περισσότερες υπάρχουσες μεθόδους μεταβολών (variational methods), που χρησιμοποιούν ΜΔΕ.

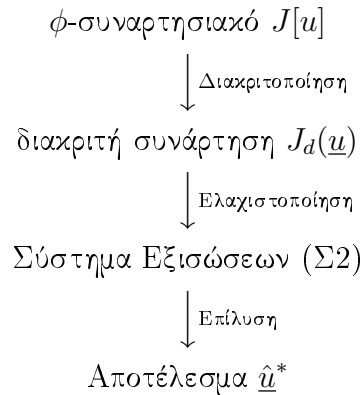
- **Θεώρηση 2^η (Εξαρχής διακριτή): Διακριτοποίηση και μετά Ελαχιστοποίηση.**

Στη θεώρηση αυτή, διακριτοποιείται εξαρχής το ϕ -συναρτησιακό $J[u]$ (4.54). Ένας απλός, αλλά αρκετά αξιόπιστος, τρόπος για την διακριτοποίηση αυτή είναι η προσέγγιση των διπλών ολοκληρωμάτων του (4.54) με άθροισμα Riemann, χρησιμοποιώντας τις τιμές στο πλέγμα ορισμού της διακριτής εικόνας, όποτε προκύπτει μία διακριτή συνάρτηση $J_d(\underline{u})$.

Το επόμενο βήμα στη θεώρηση αυτή είναι η ελαχιστοποίηση της διακριτής συνάρτησης $J_d(\underline{u})$, οπότε από τις αναγκαίες συνθήκες ελαχίστου, καταλή-

γουμε ξανά σε ένα σύστημα μη γραμμικών εξισώσεων για την ζητούμενη εικόνα.

Τα βήματα αυτά φαίνονται σχηματικά στο διάγραμμα του σχήματος 4.4.



Σχήμα 4.4: Διάγραμμα 2^{ης} θεώρησης για την αριθμητική υλοποίηση των μεθόδων ϕ -συναρτησιακού

Η θεώρηση αυτή δεν παρουσιάζει βέβαια τα πλεονεκτήματα της προηγούμενης θεώρησης, αλλά από την άλλη θεμελιώνεται απλά, χρησιμοποιώντας μόνο τα εργαλεία της Αριθμητικής Ανάλυσης για την ελαχιστοποίηση συναρτήσεων πολλών μεταβλητών. Επίσης στην θεώρηση αυτή, σε σύγκριση με την προηγούμενη, είναι πιο εμφανής ο τρόπος που βελτιώνεται η αρχική διακριτή εικόνα, για να προκύψει η τελική. Έτσι, είναι σίγουρο ότι ελαχιστοποιείται μία κατάλληλη διακριτή συνάρτηση, η οποία μεταφέρει με αξιόπιστο τρόπο στον διακριτό χώρο τις ιδιότητες του συνεχούς συναρτησιακού $J[u]$ (4.54). Αυτή η εξαρχής διακριτή θεώρηση χρησιμοποιήθηκε πρόσφατα από τους Chan, Osher και Shen [13], στην ειδική περίπτωση της Ολικής Μεταβολής (TV), για την δημιουργία της μεθόδου του Digital Total Variation Filter (βλ. ενότητα 5.2.2).

Από την παρουσίαση που προηγήθηκε, γεννιέται το ερώτημα αν οι δύο παραπάνω θεωρήσεις είναι ισοδύναμες, αν δηλαδή καταλήγουν στο ίδιο αποτέλεσμα. Από όσα αναφέρονται στο [13], συμπεραίνουμε ότι, στην ειδική περίπτωση που μελετάται στο άρθρο (μέθοδος TV και διακριτοποίηση με συγκεκριμένο τρόπο), η απάντηση είναι καταφατική.

Στην ενότητα αυτή, μελετάμε τη σχέση ανάμεσα στις δύο αυτές θεωρήσεις, στη γενική περίπτωση των ϕ -συναρτησιακών και για διάφορες επιλογές στη διακριτοποίηση (όχι μόνο για τη συγκεκριμένη επιλογή που γίνεται στο [13]). Αποδεικνύουμε ότι, στη γενική περίπτωση οι δύο αυτές θεωρήσεις δεν είναι

πάντα ισοδύναμες, βρίσκοντας τη διακριτή συνάρτηση $\tilde{J}(\underline{u})$, που στην πραγματικότητα ελαχιστοποιεί η λύση \underline{u}^* της 1^{ης} θεώρησης. Η ισοδυναμία δεν ισχύει, διότι η $\tilde{J}(\underline{u})$ δεν είναι πάντα ίδια με την διακριτοποιημένη συνάρτηση $J_d(\underline{u})$ της 2^{ης} θεώρησης.

Η εύρεση της συνάρτησης $\tilde{J}(\underline{u})$ μάς παρέχει επιπρόσθετα, μέσω της σύγκρισής της με την συνάρτηση $J_d(\underline{u})$, καλύτερη εποπτεία της σχέσης ανάμεσα στις δύο προσεγγίσεις. Ακόμα, αποσαφηνίζουμε με τον τρόπο αυτό ότι, για να μπορέσει η 1^η θεώρηση να είναι ισοδύναμη με τη 2^η, πρέπει να δοθεί ειδική μέριμνα στην διακριτοποίηση της ΜΔΕ (4.55). Πιο συγκεκριμένα, δείχνουμε ότι οι επιλογές στην προσέγγιση των χωρικών τελεστών παραγωγίσης $\text{grad}(\cdot)$ και $\text{div}(\cdot)$ επιβάλλουν συγκεκριμένη προσέγγιση της νόρμας $\|\nabla u\|$, εφόσον επιθυμείται η ισοδυναμία αυτή.

Ακόμα, τα βήματα που χρησιμοποιούμε για την απόδειξη των παραπάνω εισάγουν μία μεθοδολογία, η οποία μπορεί να εφαρμοστεί στη διακριτοποίηση οποιασδήποτε ΜΔΕ διάχυσης που εφαρμόζεται στην Όραση Υπολογιστών, με σκοπό να παράσχει μία ένδειξη για την πιθανή σύνδεση της ΜΔΕ αυτής με την ελαχιστοποίηση κάποιου συναρτησιακού. Σημειώνουμε επίσης ότι στην ενότητα αυτή μελετάμε μόνο την περίπτωση βαθμωτών εικόνων, αλλά τα συμπεράσματά μας μπορούν σχετικά εύκολα να γενικευθούν και στην περίπτωση διανυσματικών εικόνων, για τα διανυσματικά ϕ -συναρτησιακά (vectorial ϕ -functionals).

4.2.2 Ένα νέο θεώρημα για τη σχέση ανάμεσα στην συνεχή και στην εξαρχής διακριτή θεώρηση των ϕ -συναρτησιακών.

Προκαταρκτική παρουσίαση

Όπως φαίνεται από την εισαγωγή που προηγήθηκε, βασικό βήμα στην θεώρηση του σχήματος 4.3, για την αριθμητική υλοποίηση του ϕ -συναρτησιακού, είναι η διακριτοποίηση της ΜΔΕ (4.55). Στη συνέχεια, παρουσιάζουμε τα πιο βασικά στοιχεία της διακριτοποίησης της ΜΔΕ (4.55), που μας χρειάζονται στην περαιτέρω ανάλυση. Μια αναλυτική παρουσίαση της διακριτοποίησης ΜΔΕ της κατηγορίας αυτής γίνεται, εκτός των άλλων, στην ενότητα 5.1.

Πρώτα από όλα, πρέπει να σημειωθεί ότι, πριν την διακριτοποίηση της ΜΔΕ (4.55), εφαρμόζεται συνήθως ροή ταχύτερης καθόδου στη ΜΔΕ αυτή. Έτσι, προκύπτει η εξής ΜΔΕ:

$$\frac{\partial u}{\partial t} = \text{div} \left(\frac{\phi'(\|\nabla u\|)}{\|\nabla u\|} \nabla u \right) - \lambda(u - u^0) \quad (4.56)$$

στην οποία, όπως παρατηρούμε, ο συντελεστής διάχυσης g δίνεται από τη σχέση:

$$g(\|\nabla u\|) = \frac{\phi'(\|\nabla u\|)}{\|\nabla u\|} \quad (4.57)$$

Στην ΜΔΕ (4.56), η λύση δεν είναι μία εικόνα, αλλά μία χρονική ροή εικόνων, έχει δηλαδή προστεθεί (σε σχέση με την (4.55)) η εξάρτηση από τον χρόνο. Στην κατάσταση ισορροπίας της ΜΔΕ (4.56) καταλήγουμε στη λύση της ΜΔΕ (4.55). Σημειώνουμε ότι στην ενότητα αυτή, επειδή ενδιαφερόμαστε μόνο για το τελικό αποτέλεσμα, το οποίο προκύπτει από την κατάσταση ισορροπίας, ασχολούμαστε μόνο με την ΜΔΕ (4.55), χωρίς έτσι να έχουμε εξάρτηση από τον χρόνο.

Με σκοπό την διακριτοποίηση της (4.55), λοιπόν, θεωρούμε ένα ομοιόμορφο δισδιάστατο ορθογωνικό πλέγμα, έκτασης $N_x \times N_y$, με μοναδιαία χωρικά βήματα $h_x = h_y = 1$ και τη διακριτή συνάρτηση $u(i, j)$, που προσεγγίζει τις τιμές της $u(x, y)$ στο πλέγμα αυτό. Ακόμα, θεωρούμε την αναπαράσταση της διακριτής εικόνας $u(i, j)$ σαν ένα διάνυσμα \underline{u} , διάστασης $M = N_x N_y$, που προκύπτει από την $u(i, j)$ με κάποια μέθοδο σάρωσης των pixels. Η μέθοδος σάρωσης που χρησιμοποιούμε είναι από πάνω προς τα κάτω και από αριστερά προς τα δεξιά², οπότε $u_{i+jN_x+1} = u(i, j)$, για $0 \leq i \leq N_x - 1, 0 \leq j \leq N_y - 1$ (όπου με $u_\alpha = u_{i+jN_x+1}$ συμβολίζουμε το α -οστό στοιχείο του διανύσματος \underline{u}).

Διευκρινίζουμε ότι, αφού λόγω της παραπάνω σάρωσης αντιστοιχίζεται ένας ακέραιος α σε κάθε pixel, στα επόμενα θα χρησιμοποιούμε την ορολογία “pixel α ” αντί της ορολογίας “pixel (i, j) ”, εφόσον $\alpha = i + jN_x + 1$.

Επίσης, χρησιμοποιούμε τον συμβολισμό $\beta \sim \alpha$ (αντίστοιχα με το [13]), για να εκφράσουμε ότι το pixel β είναι γειτονικό του pixel α , όπου σαν γειτονικά θεωρούμε τα διπλανά σημεία σε κάθε διεύθυνση του πλέγματος (οριζόντια ή κάθετη). Έχουμε δηλαδή:

$$\beta \sim \alpha \Leftrightarrow \beta \in \{\alpha - 1, \alpha + 1, \alpha - N_x, \alpha + N_x\}$$

το οποίο μπορούμε να εκφράσουμε συμβολικά ως εξής:

$$\begin{array}{ccc} & \circ & \\ \circ & \times & \circ \\ & \circ & \end{array}$$

όπου με “ \times ” συμβολίζουμε το εκάστοτε pixel και με “ \circ ” τα γειτονικά του.

²θερούμε ότι η αρχή των αξόνων βρίσκεται πάνω αριστερά στην εικόνα, με τον άξονα των x (ή i στη διακριτή περίπτωση) να είναι κάθετος και τον άξονα των y (ή j) οριζόντιος. Έτσι, η μεταβλητή i αυξάνεται καθώς προχωράμε προς τα κάτω, ενώ η j καθώς προχωράμε προς τα δεξιά.

Στην ΜΔΕ (4.55) πρέπει να διακριτοποιηθούν οι τελεστές $\text{grad}(\cdot)$ και $\text{div}(\cdot)$ καθώς και η νόρμα $\|\nabla u\|$. Για το λόγο αυτό, χρειάζονται προσεγγίσεις στις χωρικές μερικές παραγώγους $\partial/\partial x$ και $\partial/\partial y$. Οι πιο απλές επιλογές για τις προσεγγίσεις αυτές, που θα μελετήσουμε εδώ, είναι οι εξής (παραθέτουμε για την $\partial/\partial x$ αλλά τελείως αντίστοιχα ισχύουν και για την $\partial/\partial y$):

- κεντρική διαφορά βήματος 1 pixel:

$$D_x u(i, j) = u(i + \frac{1}{2}, j) - u(i - \frac{1}{2}, j) \quad (4.58)$$

- κεντρική διαφορά βήματος 2 pixel:

$$D_{2x} u(i, j) = \frac{1}{2} \{u(i + 1, j) - u(i - 1, j)\} \quad (4.59)$$

- εμπρόσθια (forward) διαφορά βήματος 1 pixel:

$$D_{+x} u(i, j) = u(i + 1, j) - u(i, j) \quad (4.60)$$

- οπίσθια (backward) διαφορά βήματος 1 pixel:

$$D_{-x} u(i, j) = u(i, j) - u(i - 1, j) \quad (4.61)$$

Έχοντας ορίσει τους παραπάνω συμβολισμούς, αναφέρουμε τώρα κάποιες απλές επιλογές που μπορούν να γίνουν για τη διακριτοποίηση των τελεστών $\text{grad}(\cdot) = (\frac{\partial}{\partial x}, \frac{\partial}{\partial y}) \cdot$ και $\text{div}(\cdot) = \langle (\frac{\partial}{\partial x}, \frac{\partial}{\partial y}), \cdot \rangle$ στην ΜΔΕ (4.55):

1. Διακριτοποίηση A: Κεντρικές διαφορές βήματος 1 pixel για $\text{grad}(\cdot)$ και $\text{div}(\cdot)$ και γραμμική παρεμβολή για τις τιμές του συντελεστή διάχυσης $g(\|\nabla u\|)$ στα ενδιάμεσα σημεία του πλέγματος:

$$\begin{aligned} \text{grad}(\cdot) &\approx (D_x, D_y) \cdot, \quad \text{div}(\cdot) \approx \langle (D_x, D_y), \cdot \rangle \quad \text{και} \\ g(i \pm \frac{1}{2}, j) &\approx \frac{1}{2} \{g(i, j) + g(i \pm 1, j)\} \\ g(i, j \pm \frac{1}{2}) &\approx \frac{1}{2} \{g(i, j) + g(i, j \pm 1)\} \end{aligned} \quad (4.62)$$

Έτσι, η διακριτοποίηση αυτή της ΜΔΕ (4.55) καταλήγει (ύστερα από κάποιες πράξεις) στο εξής σύστημα μη γραμμικών εξισώσεων:

$$\left\{ \sum_{\beta \sim \alpha} \frac{g_\alpha + g_\beta}{2} (u_\beta - u_\alpha) \right\} - \lambda (u_\alpha - u_\alpha^0) = 0, \quad \text{για κάθε pixel } \alpha \quad (4.63)$$

Η διακριτοποίηση αυτή είναι αρκετά αξιόπιστη και είναι μία από τις επιλογές που κάναμε στην αριθμητική υλοποίηση (βλ. υποενότητα 5.1.1).

2. Διακριτοποίηση Β: Οπίσθιες και εμπρόσθιες διαφορές για $\text{grad}(\cdot)$ και $\text{div}(\cdot)$ αντίστοιχα:

$$\text{grad}(\cdot) \approx (D_{-x}, D_{-y}) \cdot, \text{div}(\cdot) \approx \langle (D_{+x}, D_{+y}), \cdot \rangle \quad (4.64)$$

Η διακριτοποίηση αυτή της ΜΔΕ (4.55) καταλήγει στις εξής εξισώσεις, για κάθε pixel $\alpha = 1, \dots, N_x N_y$:

$$\begin{aligned} &g_{\alpha+1}(u_{\alpha+1} - u_\alpha) + g_\alpha(u_{\alpha-1} - u_\alpha) + \\ &+ g_{\alpha+N_x}(u_{\alpha+N_x} - u_\alpha) + g_\alpha(u_{\alpha-N_x} - u_\alpha) - \lambda(u_\alpha - u_\alpha^0) = 0 \end{aligned} \quad (4.65)$$

3. Διακριτοποίηση Γ: Εμπρόσθιες και οπίσθιες διαφορές για $\text{grad}(\cdot)$ και $\text{div}(\cdot)$ αντίστοιχα:

$$\text{grad}(\cdot) \approx (D_{+x}, D_{+y}) \cdot, \text{div}(\cdot) \approx \langle (D_{-x}, D_{-y}), \cdot \rangle \quad (4.66)$$

Η διακριτοποίηση αυτή της ΜΔΕ (4.55) καταλήγει στις εξής εξισώσεις, για κάθε pixel $\alpha = 1, \dots, N_x N_y$:

$$\begin{aligned} &g_\alpha(u_{\alpha+1} - u_\alpha) + g_{\alpha-1}(u_{\alpha-1} - u_\alpha) + \\ &+ g_\alpha(u_{\alpha+N_x} - u_\alpha) + g_{\alpha-N_x}(u_{\alpha-N_x} - u_\alpha) - \lambda(u_\alpha - u_\alpha^0) = 0 \end{aligned} \quad (4.67)$$

Αναφέρουμε ότι στις παραπάνω επιλογές διακριτοποίησης (4.62,4.64,4.66), έχουμε αφήσει αδιευκρίνιστο ποια προσέγγιση χρησιμοποιούμε για τη νόρμα $\|\nabla u\|$ (που είναι απαραίτητη για τον υπολογισμό του συντελεστή διάχυσης g_α). Κι αυτό διότι, στην ανάλυση που ακολουθεί, θεωρούμε ότι χρησιμοποιούμε μία οποιαδήποτε προσέγγιση της $\|\nabla u\|$ (Du είναι ο γενικός συμβολισμός της προσέγγισης αυτής), με μόνη προϋπόθεση να ικανοποιείται η σχέση:

$$D(h \cdot u)(\alpha) = |h| Du(\alpha), \text{ για κάθε pixel } \alpha, \text{ με } h \in \mathbb{R}^3 \quad (4.68)$$

η οποία είναι εξάλλου μία από τις ιδότητες ορισμού της νόρμας διανύσματος (βλ. π.χ. [48]). Η αντίστοιχη ιδιότητα ισχύει φυσικά για τη συνεχή νόρμα $\|\nabla u\|$ και άρα είναι επιθυμητό να συντηρείται μετά τη διακριτοποίηση.

Κάποιες σχετικά απλές προσεγγίσεις της νόρμας $\|\nabla u\|$, οι οποίες θα μας απασχολήσουν εδώ και ικανοποιούν (όπως μπορεί εύκολα να διαπιστωθεί) την σχέση (4.68), είναι οι ακόλουθες:

³διευκρινίζουμε ότι, με $Du(\alpha)$, συμβολίζουμε την τιμή της προσέγγισης Du στο pixel α (όπου α ο κέραιος που αντιστοιχούμε στο pixel (i, j) , μετά την σάρωση), δηλαδή $Du(\alpha) = Du(i, j)$, όταν $\alpha = i + jN_x + 1$.

- Προσέγγιση με χρήση εμπρόσθιων διαφορών βήματος 1 pixel για τις $\partial u/\partial x, \partial u/\partial y$:

$$\begin{aligned} D^+u(\alpha) &= \{[D_{+x}u(\alpha)]^2 + [D_{+y}u(\alpha)]^2\}^{\frac{1}{2}} \\ &= \{(u_{\alpha+1} - u_\alpha)^2 + (u_{\alpha+N_x} - u_\alpha)^2\}^{\frac{1}{2}} \end{aligned} \quad (4.69)$$

- Προσέγγιση με χρήση οπίσθιων διαφορών βήματος 1 pixel:

$$\begin{aligned} D^-u(\alpha) &= \{[D_{-x}u(\alpha)]^2 + [D_{-y}u(\alpha)]^2\}^{\frac{1}{2}} \\ &= \{(u_\alpha - u_{\alpha-1})^2 + (u_\alpha - u_{\alpha-N_x})^2\}^{\frac{1}{2}} \end{aligned} \quad (4.70)$$

- Προσέγγιση με χρήση κεντρικών διαφορών με βήμα 2 pixels:

$$D^o u(\alpha) = \{[D_{2x}u(\alpha)]^2 + [D_{2y}u(\alpha)]^2\}^{\frac{1}{2}} \quad (4.71)$$

- Προσέγγιση με χρήση εμπρόσθιων και οπίσθιων διαφορών βήματος 1 pixel:

$$\begin{aligned} D^\pm u(\alpha) &= \left\{ \frac{1}{2} \{ [D_{+x}u(\alpha)]^2 + [D_{-x}u(\alpha)]^2 + [D_{+y}u(\alpha)]^2 + [D_{-y}u(\alpha)]^2 \} \right\}^{\frac{1}{2}} \\ &= \left\{ \frac{1}{2} \sum_{\beta \sim \alpha} (u_\beta - u_\alpha)^2 \right\}^{\frac{1}{2}} \end{aligned} \quad (4.72)$$

Το παραπάνω αντιστοιχεί σε προσέγγιση των u_x^2 και u_y^2 με τον μέσο όρο των αντίστοιχων τιμών στα ενδιάμεσα σημεία $(i \pm \frac{1}{2}, j)$ και $(i, j \pm \frac{1}{2})$ αντίστοιχα, στα οποία η προσέγγιση των παραγώγων γίνεται με κεντρικές διαφορές βήματος 1 pixel.

Διατύπωση και απόδειξη του Θεωρήματος

Με βάση την εισαγωγική παρουσίαση που προηγήθηκε, στη συνέχεια παραθέτουμε και αποδεικνύουμε ένα θεώρημα, το οποίο αποσαφηνίζει τη σχέση ανάμεσα στη συνεχή και τη διακριτή θεώρηση των ϕ -συναρτησιακών, θεωρώντας για τις διάφορες επιλογές διακριτοποίησης (4.62,4.64,4.66), μία οποιαδήποτε προσέγγιση της νόρμας $\|\nabla u\|$:

Θεώρημα 4.6. Έστω το συναρτησιακό:

$$J[u] = \iint_{\Omega} \phi(\|\nabla u\|) dx + \frac{\lambda}{2} \iint_{\Omega} (u - u^0)^2 dx \quad (4.73)$$

όπου Ω ένα ορθογωνικό χωρίο ορισμού της $u(x, y)$ (έξω από το οποίο θεωρούμε ότι η $u(x, y)$ επεκτείνεται με ανάκλαση) και $\phi(s) : [0, +\infty) \rightarrow \mathbb{R}$ μία ομαλή συνάρτηση, με:

$$\phi(0) = 0, \quad \phi'(s) > 0, \quad \phi''(s) > 0 \quad (4.74)$$

Έστω επίσης \underline{u} το διάνυσμα (διάστασης $M = N_x N_y$) της διακριτοποίησης της $u(x, y)$ σε ένα ορθογωνικό πλέγμα που καλύπτει το Ω , με έκταση $N_x \times N_y$ και μοναδιαία βήματα $h_x = h_y = 1$ (έξω από το πλέγμα, θεωρούμε ότι η διακριτή συνάρτηση επεκτείνεται με ανάκλαση (*mirroring*)). Τότε, η λύση \underline{u}^* των διακριτοποιημένων συνθηκών ελαχίστου του $J[u]$ (4.73) ελαχιστοποιεί τις εξής διακριτές συναρτήσεις $\tilde{J}(\underline{u})$ (ανάλογα με τις παρακάτω επιλογές διακριτοποίησης των $\text{grad}(\cdot)$ και $\text{div}(\cdot)$):

1. Κεντρικές διαφορές βήματος 1 *pixel* για $\text{grad}(\cdot)$, $\text{div}(\cdot)$ και γραμμική παρεμβολή για τις τιμές του g (βλ. (4.62,4.63)):

$$\tilde{J}(\underline{u}) = \sum_{\alpha=1}^{N_x N_y} \left\{ \phi(Du(\alpha)) \left(\frac{D^\pm u(\alpha)}{Du(\alpha)} \right)^2 \right\} + \frac{\lambda}{2} \sum_{\alpha=1}^{N_x N_y} (u_\alpha - u_\alpha^0)^2 \quad (4.75)$$

2. Οπίσθιες και εμπρόσθιες διαφορές για $\text{grad}(\cdot)$, $\text{div}(\cdot)$ (βλ. (4.64,4.65)):

$$\tilde{J}(\underline{u}) = \sum_{\alpha=1}^{N_x N_y} \left\{ \phi(Du(\alpha)) \left(\frac{D^- u(\alpha)}{Du(\alpha)} \right)^2 \right\} + \frac{\lambda}{2} \sum_{\alpha=1}^{N_x N_y} (u_\alpha - u_\alpha^0)^2 \quad (4.76)$$

3. Εμπρόσθιες και οπίσθιες διαφορές για $\text{grad}(\cdot)$, $\text{div}(\cdot)$ (βλ. (4.66,4.67)):

$$\tilde{J}(\underline{u}) = \sum_{\alpha=1}^{N_x N_y} \left\{ \phi(Du(\alpha)) \left(\frac{D^+ u(\alpha)}{Du(\alpha)} \right)^2 \right\} + \frac{\lambda}{2} \sum_{\alpha=1}^{N_x N_y} (u_\alpha - u_\alpha^0)^2 \quad (4.77)$$

όπου $Du(\alpha)$ είναι μία οποιαδήποτε προσέγγιση της νόρμας $\|\nabla u\|$, που υποθέτουμε ότι έγινε στην διακριτοποίηση των συνθηκών ελαχίστου, τέτοια, ώστε να ικανοποιεί:

$$D(h \cdot u)(\alpha) = |h| Du(\alpha), \quad \text{για κάθε } \textit{pixel} \alpha, \quad \mu \in h \in \mathbb{R} \quad (4.78)$$

Απόδειξη. Αν οι συνθήκες ελαχίστου της $J[u]$ (βλ. ΜΔΕ (4.55)) διακριτοποιηθούν με οποιαδήποτε μέθοδο διακριτοποίησης, προκύπτουν οι εξισώσεις:

$$(\text{div}(g\nabla u))_\alpha - \lambda(u_\alpha - u_\alpha^0) = 0, \quad \text{για κάθε } \alpha = 1, \dots, M \quad (4.79)$$

όπου χρησιμοποιούμε τον γενικό συμβολισμό $(\text{div}(g\nabla u))_\alpha$ για την προσέγγιση του όρου $\text{div}(g\nabla u)$ στο *pixel* α , με μία οποιαδήποτε επιλογή διακριτοποίησης ((4.62),(4.64),(4.66) ή κάποια άλλη).

Η σχέση (4.79) εκφράζει M εξισώσεις, μία για κάθε pixel. Έτσι, πρόκειται για ένα μη γραμμικό σύστημα M εξισώσεων με M αγνώστους (τις τιμές u_α στα pixels). Αν γράψουμε τις εξισώσεις (4.79) σε διανυσματική μορφή (χρησιμοποιώντας διανύσματα διάστασης M), προκύπτει:

$$\underline{G}(\underline{u}) - \lambda(\underline{u} - \underline{u}^0) = 0 \quad (4.80)$$

όπου $\underline{G}(\underline{u})$ το M -διάστατο διάνυσμα, με την α -οστή τιμή του να είναι:

$$G_\alpha(\underline{u}) = (\operatorname{div}(g\nabla u))_\alpha, \alpha = 1, \dots, M \quad (4.81)$$

Έστω επίσης το M -διάστατο διάνυσμα:

$$\underline{F}(\underline{u}) = -\underline{G}(\underline{u}) + \lambda(\underline{u} - \underline{u}^0) \quad (4.82)$$

το οποίο αποτελεί ουσιαστικά την αντίστοιχη διακριτοποίηση της μεταβολικής παραγώγου (variational derivative) $\delta J/\delta u$ του συναρτησιακού $J[u]$. Έτσι, το σύστημα (4.79) γράφεται απλά:

$$-\underline{F}(\underline{u}) = 0 \quad (4.83)$$

Έστω τώρα \underline{u}^* η λύση του παραπάνω συστήματος (βλ. (4.83) ή (4.80) ή (4.79) ισοδύναμα), δηλαδή $-\underline{F}(\underline{u}^*) = 0$. Ψάχνουμε να βρούμε μία διακριτή συνάρτηση $\tilde{J}(\underline{u})$, ορισμένη στον M -διάστατο χώρο \mathbb{R}^M , της οποίας η \underline{u}^* να αποτελεί ελάχιστο. Αναγκαία συνθήκη για το παραπάνω είναι η εξής:

$$-\nabla_M \tilde{J}(\underline{u}^*) = 0 \quad (4.84)$$

όπου ο τελεστής ∇_M δρα σε συναρτήσεις M μεταβλητών, δηλαδή:

$$\nabla_M := \left(\frac{\partial}{\partial u_1}, \frac{\partial}{\partial u_2}, \dots, \frac{\partial}{\partial u_M} \right)$$

με τα u_1, u_2, \dots, u_M να συμβολίζουν τις M μεταβλητές.

Έτσι, η (4.84) αποτελεί άλλο ένα σύστημα M εξισώσεων με M αγνώστους, το οποίο έχει την ίδια λύση με το (4.83). Επιλέγουμε η διακριτή συνάρτηση $\tilde{J}(\underline{u})$ που ψάχνουμε να είναι τέτοια, ώστε τα δύο συστήματα (4.83) και (4.84) να ταυτίζονται (το οποίο αποτελεί μία λογική επιλογή, εφόσον σκοπεύουμε να καταλήξουμε σε μία συνάρτηση $\tilde{J}(\underline{u})$ που να μοιάζει όσο το δυνατόν περισσότερο με το συναρτησιακό $J[u]$). Με την επιλογή αυτή, προκύπτει η σχέση:

$$-\nabla_M \tilde{J}(\underline{u}) = -\underline{F}(\underline{u}) \quad (4.85)$$

από την οποία, όπως είναι γνωστό από την Ανάλυση Συναρτήσεων Πολλών Μεταβλητών (βλ. π.χ. [49]), προκύπτει η εξής σχέση:

$$\tilde{J}(\underline{u}) = \tilde{J}(\underline{0}) + \int_{\underline{0}}^{\underline{u}} \underline{F} \cdot d\underline{r} \quad (4.86)$$

όπου το παραπάνω επικαμπύλιο ολοκλήρωμα είναι ανεξάρτητο του δρόμου και μπορεί να υπολογιστεί επί οποιασδήποτε καμπύλης ενώνει τα σημεία $\underline{0}$ και \underline{u} (στον χώρο \mathbb{R}^M). Επίσης, με $\underline{0}$ συμβολίζουμε το μηδενικό M -διάστατο διάνυσμα (που αντιστοιχεί σε εικόνα με τιμές 0 σε όλα τα pixels).

Αναφέρουμε επίσης ότι η σχέση (4.85) επιτρέπει την αυθαίρετη επιλογή μιας προσθετικής σταθεράς στην $J(\underline{u})$, το οποίο είναι ισοδύναμο με την αυθαίρετη επιλογή της τιμής $\tilde{J}(\underline{0})$ στη σχέση (4.86) (εξάλλου, στην ελαχιστοποίηση, μία προσθετική σταθερά δεν παίζει κανένα ρόλο).

Για τον υπολογισμό του επικαμπύλιου ολοκληρώματος της (4.86), διαλέγουμε λοιπόν την απλή διαδρομή $\underline{r}(q) = q\underline{u}$, $q \in [0, 1]$ (ευθύγραμμο τμήμα με άκρα τα $\underline{0}$ και \underline{u}), για την οποία $d\underline{r}/dq = \underline{u}$, οπότε:

$$(4.86) \Rightarrow \tilde{J}(\underline{u}) = \tilde{J}(\underline{0}) + \int_0^1 \underline{F}(q\underline{u}) \cdot \underline{u} dq = \tilde{J}(\underline{0}) + \int_0^1 \underline{u}^T \underline{F}(q\underline{u}) dq \quad (4.87)$$

επομένως, λόγω της (4.82), έχουμε:

$$\tilde{J}(\underline{u}) = \tilde{J}(\underline{0}) - \int_0^1 \underline{u}^T \underline{G}(q\underline{u}) dq + \int_0^1 \underline{u}^T \lambda (q\underline{u} - \underline{u}^0) dq \quad (4.88)$$

Όμως, το δεύτερο ολοκλήρωμα στην παραπάνω σχέση μπορεί να απλοποιηθεί ως εξής:

$$\begin{aligned} \int_0^1 \underline{u}^T \lambda (q\underline{u} - \underline{u}^0) dq &= \underline{u}^T \lambda \left[\left(\int_0^1 q dq \right) \underline{u} - \left(\int_0^1 dq \right) \underline{u}^0 \right] \\ &= \frac{\lambda}{2} (\underline{u}^T \underline{u} - 2\underline{u}^T \underline{u}^0) \\ &= \frac{\lambda}{2} \left(\|\underline{u}\|^2 - 2\underline{u}^T \underline{u}^0 + \|\underline{u}^0\|^2 - \|\underline{u}^0\|^2 \right) \\ &= -\frac{\lambda}{2} \|\underline{u}^0\|^2 + \frac{\lambda}{2} \|\underline{u} - \underline{u}^0\|^2 \end{aligned}$$

επομένως, λόγω της παραπάνω σχέσης, η (4.88) γίνεται:

$$\tilde{J}(\underline{u}) = \tilde{J}(\underline{0}) - \frac{\lambda}{2} \sum_{\alpha=1}^M (u_\alpha^0)^2 + K(\underline{u}) + \frac{\lambda}{2} \sum_{\alpha=1}^M (u_\alpha - u_\alpha^0)^2 \quad (4.89)$$

όπου η συνάρτηση $K(\underline{u})$ ορίζεται ως εξής:

$$\begin{aligned} K(\underline{u}) &= - \int_0^1 \underline{u}^T \underline{G}(q\underline{u}) \, dq \\ &= - \sum_{\alpha=1}^M \left\{ u_\alpha \int_0^1 \underline{G}_\alpha(q\underline{u}) \, dq \right\} \end{aligned} \quad (4.90)$$

Με σκοπό την απλοποίηση της (4.89), επιλέγουμε σαν αυθαίρετη προσθετική σταθερά $\tilde{J}(\underline{0})$ την εξής:

$$\tilde{J}(\underline{0}) = \frac{\lambda}{2} \sum_{\alpha=1}^M (u_\alpha^0)^2 \quad (4.91)$$

με αποτέλεσμα η (4.89) να γράφεται ως εξής:

$$\tilde{J}(\underline{u}) = K(\underline{u}) + \frac{\lambda}{2} \sum_{\alpha=1}^M (u_\alpha - u_\alpha^0)^2 \quad (4.92)$$

Συνεπώς, από τη σχέση (4.92), βλέπουμε ότι, για τον υπολογισμό της $\tilde{J}(\underline{u})$ αρκεί ο υπολογισμός της $K(\underline{u})$. Όμως, ενώ τα συμπεράσματα στα οποία καταλήξαμε μέχρι τώρα είναι γενικά και δεν εξαρτώνται από την επιλογή διακριτοποίησης, στην συνέχεια πρέπει κάθε περίπτωση διακριτοποίησης να μελετηθεί ξεχωριστά (λόγω διαφορετικής έκφρασης για την $\underline{G}(\underline{u})$ και άρα για την $K(\underline{u})$):

1.Κεντρικές διαφορές βήματος 1 pixel για $\text{grad}(\cdot)$, $\text{div}(\cdot)$ και γραμμική παρεμβολή για τις τιμές του g .

Όπως διαφαίνεται από τις σχέσεις (4.81) και (4.63), στην περίπτωση αυτή, τα στοιχεία του διανύσματος $\underline{G}(\underline{u})$ προκύπτουν από τη σχέση:

$$\underline{G}_\alpha(\underline{u}) = \sum_{\beta \sim \alpha} \frac{g_\alpha + g_\beta}{2} (u_\beta - u_\alpha) \quad (4.93)$$

Επομένως, στην περίπτωση αυτή, η συνάρτηση $K(\underline{u})$ γίνεται:

$$(4.90) \Rightarrow K(\underline{u}) = \sum_{\alpha=1}^M \sum_{\beta \sim \alpha} (u_\alpha - u_\beta) u_\alpha \int_0^1 \frac{g_\alpha(q\underline{u}) + g_\beta(q\underline{u})}{2} q \, dq \quad (4.94)$$

άρα:

$$K(\underline{u}) = \sum_{\alpha=1}^M \sum_{\beta \sim \alpha} (u_\alpha - u_\beta) u_\alpha f_{\alpha\beta}(\underline{u}) \quad (4.95)$$

όπου:

$$f_{ab}(\underline{u}) = \frac{f_\alpha(\underline{u}) + f_\beta(\underline{u})}{2} \quad (4.96)$$

με:

$$f_\alpha(\underline{u}) = \int_0^1 g_\alpha(q\underline{u})q \, dq, \quad f_\beta(\underline{u}) = \int_0^1 g_\beta(q\underline{u})q \, dq \quad (4.97)$$

Στη συνέχεια, βρίσκουμε μία πιο απλή έκφραση για τον όρο $f_\alpha(\underline{u})$. Πρώτα από όλα, λοιπόν, λόγω της σχέσης (4.57) και του ότι χρησιμοποιείται η γενική προσέγγιση $Du(\alpha)$ για τη νόρμα $\|\nabla u\|$, ισχύει, για κάθε pixel α :

$$\begin{aligned} g_\alpha(q\underline{u}) &= \frac{\phi'(D(qu)(\alpha))}{D(qu)(\alpha)} \\ &= \frac{\phi'(qDu(\alpha))}{qDu(\alpha)}, \end{aligned} \quad (4.98)$$

λόγω του περιορισμού (4.78) για την προσέγγιση $Du(\alpha)$. Άρα:

$$\begin{aligned} (4.97) \Rightarrow f_\alpha(\underline{u}) &= \int_0^1 \frac{\phi'(qDu(\alpha))}{Du(\alpha)} \, dq \\ &= \frac{1}{(Du(\alpha))^2} \int_0^{Du(\alpha)} \phi'(qDu(\alpha)) \, d(qDu(\alpha)) \end{aligned} \quad (4.99)$$

Επομένως, ο όρος f_α γράφεται τελικά ως εξής:

$$\begin{aligned} f_\alpha(\underline{u}) &= \frac{1}{(Du(\alpha))^2} [\phi(Du(\alpha)) - \phi(0)] \\ &= \frac{1}{(Du(\alpha))^2} \phi(Du(\alpha)), \end{aligned} \quad (4.100)$$

διότι $\phi(0) = 0$ (βλ. (4.74)).

Επανερχόμαστε τώρα στη σχέση (4.95) και τη γράφουμε ως εξής:

$$K(\underline{u}) = \sum_{\alpha=1}^M \sum_{\beta \sim \alpha} (u_\alpha^2 - u_\alpha u_\beta) f_{\alpha\beta}(\underline{u}) \quad (4.101)$$

Παρατηρώντας τη σχέση (4.101), μπορούμε να διαπιστώσουμε ότι το παραπάνω άθροισμα είναι δυνατό να κατασκευαστεί ως εξής: για κάθε ζεύγος γειτονικών pixels α και β , προσθέτουμε στο άθροισμα τους όρους:

$$(u_\alpha^2 - u_\alpha u_\beta) f_{\alpha\beta}(\underline{u}), \quad (4.102)$$

που συνεισφέρεται από την “πλευρά” του α

$$(u_\beta^2 - u_\beta u_\alpha) f_{\beta\alpha}(\underline{u}), \quad (4.103)$$

που συνεισφέρεται από την “πλευρά” του β .

Από την (4.96), βλέπουμε ότι προφανώς ισχύει $f_{\beta\alpha} = f_{\alpha\beta}$, οπότε, αν προσθέσουμε τους δύο αυτούς όρους, (4.102) και (4.103), προκύπτει:

$$\begin{aligned} (u_\alpha^2 - u_\alpha u_\beta) f_{\alpha\beta}(\underline{u}) + (u_\beta^2 - u_\beta u_\alpha) f_{\beta\alpha}(\underline{u}) &= (u_\alpha - u_\beta)^2 f_{\alpha\beta}(\underline{u}) \\ &= \frac{f_\alpha}{2} (u_\alpha - u_\beta)^2 + \frac{f_\beta}{2} (u_\alpha - u_\beta)^2, \text{ λόγω της (4.96)} \end{aligned} \quad (4.104)$$

Αναφερόμενοι στην τελευταία σχέση, μπορούμε τώρα να αντιστοιχίσουμε τον νέο όρο $\frac{f_\alpha}{2} (u_\alpha - u_\beta)^2$ στο pixel α και τον όρο $\frac{f_\beta}{2} (u_\alpha - u_\beta)^2$ στο pixel β . Έτσι, αφού η παραπάνω διαδικασία γίνεται για κάθε ζεύγος γειτονικών pixels, σε κάθε pixel α αντιστοιχίζεται ένας όρος της μορφής $\frac{f_\alpha}{2} (u_\alpha - u_\beta)^2$ για κάθενα από τα 4 γειτονικά του pixels ⁴. Επομένως, το άθροισμα της σχέσης (4.101) μπορεί να γραφεί και ως εξής:

$$K(\underline{u}) = \sum_{\alpha=1}^M f_\alpha(\underline{u}) \frac{1}{2} \sum_{\beta \sim \alpha} (u_\beta - u_\alpha)^2 \quad (4.105)$$

παρατηρώντας τη σχέση (4.105), βλέπουμε ότι εμφανίστηκε ο όρος $(D^\pm u(\alpha))^2$ (βλ. σχέση (4.72)), οπότε τελικά έχουμε:

$$\begin{aligned} K(\underline{u}) &= \sum_{\alpha=1}^M f_\alpha(\underline{u}) (D^\pm u(\alpha))^2 \\ &= \sum_{\alpha=1}^M \phi(Du(\alpha)) \left(\frac{D^\pm u(\alpha)}{Du(\alpha)} \right)^2, \text{ λόγω της (4.100)} \end{aligned} \quad (4.106)$$

Συνεπώς λοιπόν, λόγω της παραπάνω σχέσης, η συνάρτηση $\tilde{J}(\underline{u})$ γράφεται στην περίπτωση αυτή ως εξής:

$$(4.92) \Rightarrow \tilde{J}(\underline{u}) = \sum_{\alpha=1}^M \phi(Du(\alpha)) \left(\frac{D^\pm u(\alpha)}{Du(\alpha)} \right)^2 + \frac{\lambda}{2} \sum_{\alpha=1}^M (u_\alpha - u_\alpha^0)^2 \quad (4.107)$$

⁴ακόμα και τα pixels στα όρια του πλέγματος, θεωρούμε ότι έχουν 4 γειτονικά pixels, κάτι το οποίο είναι εφικτό λόγω της ανάκλασης στις τιμές της u που θεωρούμε ότι συμβαίνει έξω από το πλέγμα.

η οποία αποτελεί ακριβώς την αποδεικτέα (για την περίπτωση αυτή) σχέση (4.75).

Τέλος, σημειώνουμε ότι λόγω των ιδιοτήτων (4.74) της συνάρτησης $\phi(s)$ (αύξουσα και κυρτή συνάρτηση), μπορεί να αποδειχθεί ότι υπάρχει πράγματι τουλάχιστον μία λύση \underline{u}^* του συστήματος εξισώσεων (4.84) (για την συγκεκριμένη πάντα περίπτωση διακριτοποίησης). Επιπλέον, κάθε τέτοια λύση \underline{u}^* ελαχιστοποιεί πράγματι την συνάρτηση $J(\underline{u})$ (4.107).

2. Οπίσθιες και εμπρόσθιες διαφορές για $\text{grad}(\cdot)$, $\text{div}(\cdot)$.

Από τις σχέσεις (4.81) και (4.65) διαφαίνεται ότι, στην περίπτωση αυτή, τα στοιχεία του $\underline{G}(\underline{u})$ προκύπτουν από τη σχέση:

$$\begin{aligned} \underline{G}_\alpha(\underline{u}) = & g_{\alpha+1}(u_{\alpha+1} - u_\alpha) + g_\alpha(u_{\alpha-1} - u_\alpha) + \\ & + g_{\alpha+N_x}(u_{\alpha+N_x} - u_\alpha) + g_\alpha(u_{\alpha-N_x} - u_\alpha) \end{aligned} \quad (4.108)$$

Επομένως, η $K(\underline{u})$ γίνεται:

$$\begin{aligned} (4.90) \Rightarrow K(\underline{u}) = & \sum_{\alpha=1}^M u_\alpha \left[f_{\alpha+1}(u_\alpha - u_{\alpha+1}) + f_\alpha(u_\alpha - u_{\alpha-1}) + \right. \\ & \left. + f_{\alpha+N_x}(u_\alpha - u_{\alpha+N_x}) + f_\alpha(u_\alpha - u_{\alpha-N_x}) \right] \end{aligned} \quad (4.109)$$

όπου οι συντελεστές f_α ορίζονται ξανά όπως στην προηγούμενη περίπτωση (βλ. σχέση (4.97)) και μπορούν και εδώ να γραφούν στη μορφή που φαίνεται στη σχέση (4.100).

Με ένα σκεπτικό αντίστοιχο με αυτό της προηγούμενη περίπτωσης, μπορούμε να δείξουμε ότι το παραπάνω άθροισμα (4.109) γράφεται ισοδύναμα ως εξής:

$$K(\underline{u}) = \sum_{\alpha=1}^M f_\alpha(\underline{u}) [(u_\alpha - u_{\alpha-1})^2 + (u_\alpha - u_{\alpha-N_x})^2] \quad (4.110)$$

Παρατηρώντας τη σχέση (4.110), βλέπουμε ότι εμφανίστηκε ο όρος $(D^-u(\alpha))^2$ (βλ. σχέση (4.70)), οπότε τελικά έχουμε:

$$\begin{aligned} K(\underline{u}) = & \sum_{\alpha=1}^M f_\alpha(\underline{u})(D^-u(\alpha))^2 \\ = & \sum_{\alpha=1}^M \phi(Du(\alpha)) \left(\frac{D^-u(\alpha)}{Du(\alpha)} \right)^2, \text{ λόγω της (4.100)} \end{aligned} \quad (4.111)$$

Συνεπώς λοιπόν, λόγω της (4.111), η συνάρτηση $\tilde{J}(\underline{u})$ γράφεται στην περίπτωση αυτή ως εξής:

$$(4.92) \Rightarrow \tilde{J}(\underline{u}) = \sum_{\alpha=1}^M \phi(Du(\alpha)) \left(\frac{D^-u(\alpha)}{Du(\alpha)} \right)^2 + \frac{\lambda}{2} \sum_{\alpha=1}^M (u_\alpha - u_\alpha^0)^2 \quad (4.112)$$

η οποία αποτελεί ακριβώς την αποδεικτέα (για την περίπτωση αυτή) σχέση (4.76).

Τέλος, σημειώνουμε ότι, παρόμοια με την προηγούμενη περίπτωση, λόγω των ιδιοτήτων (4.74) της συνάρτησης $\phi(s)$ (αύξουσα και κυρτή), μπορεί να αποδειχθεί η ύπαρξη λύσης \underline{u}^* του συστήματος εξισώσεων (4.84) (για την συγκεκριμένη περίπτωση διακριτοποίησης) και η ελαχιστοποίηση της συνάρτησης $\tilde{J}(\underline{u})$ (4.112) από τη λύση αυτή \underline{u}^* .

3. Εμπρόσθιες και οπίσθιες διαφορές για $\text{grad}(\cdot)$, $\text{div}(\cdot)$.

Η περίπτωση αυτή είναι εντελώς παρόμοια με την προηγούμενη. Έτσι, με τελείως παρόμοια συλλογιστική, μπορούμε να αποδείξουμε τη σχέση (4.77), που αντιστοιχεί στην περίπτωση αυτή. □

Συμπεράσματα

Στη συνέχεια, παρουσιάζουμε κάποια ενδιαφέροντα συμπεράσματα που μπορούν να εξαχθούν από το παραπάνω θεώρημα 4.6.

Για την παρουσίαση των συμπερασμάτων αυτών, πρέπει πρώτα από όλα να παραθέσουμε την συνάρτηση που προκύπτει από την εξαρχής διακριτοποίηση του συνεχούς ϕ -συναρτησιακού $J[u]$ (βλ. σχέση (4.73)): Αν κάνουμε την απλή προσέγγιση των διπλών ολοκληρωμάτων του $J[u]$ με άθροισμα Riemann, χρησιμοποιώντας τις τιμές στο ορθογωνικό πλέγμα σημείων που αναφέρεται στο θεώρημα 4.6, τότε άμεσα καταλήγουμε στην εξής διακριτοποίηση του συναρτησιακού $J[u]$:

$$J_d(\underline{u}) = \sum_{\alpha=1}^{N_x N_y} \phi(Du(\alpha)) + \frac{\lambda}{2} \sum_{\alpha=1}^{N_x N_y} (u_\alpha - u_\alpha^0)^2 \quad (4.113)$$

όπου, όπως στο θεώρημα 4.6, $Du(\alpha)$ είναι μία οποιαδήποτε προσέγγιση της νόρμας $\|\nabla u\|$, που χρειάζεται για την προσέγγιση της ολοκληρωτέας συνάρτησης $\phi(\|\nabla u\|)$ στα σημεία του πλέγματος. Η συνάρτηση $J_d(\underline{u})$ μεταφέρει, σε μεγάλο βαθμό, στον διακριτό χώρο τις ιδιότητες του συνεχούς συναρτησιακού $J[u]$. Κι αυτό διότι η $J_d(\underline{u})$ επιβάλλει κόστος στις υψηλές μεταβολές της εικόνας (καθώς και στην απομάκρυνση από την αρχική εικόνα), με τρόπο εντελώς αντίστοιχο

με το συναρτησιακό $J[u]$. Έτσι, η ελαχιστοποίηση της συνάρτησης $J_d(\underline{u})$ αποτελεί μία ιδιαίτερα κατάλληλη (όσον αφορά τον επιθυμητό τρόπο βελτίωσης της εικόνας) υλοποίηση των μεθόδων που βασίζονται στα ϕ -συναρτησιακά.

Με βάση τα παραπάνω λοιπόν, καταλήγουμε στα εξής συμπεράσματα:

- Συγκρίνοντας την παραπάνω συνάρτηση $J_d(\underline{u})$ με τις συναρτήσεις $\tilde{J}(\underline{u})$ (4.75-4.77) του θεωρήματος 4.6, παρατηρούμε ότι, στην γενική περίπτωση προσέγγισης Du , δεν είναι ίδιες. Αυτό σημαίνει ότι δεν υπάρχει ισοδυναμία ανάμεσα στην ελαχιστοποίηση της συνάρτησης $J_d(\underline{u})$ και στη λύση των διακριτοποιημένων συνθηκών ελαχίστου του $J[u]$, ακόμα και όταν η προσέγγιση της νόρμας $\|\nabla u\|$ γίνει και στις δύο περιπτώσεις με τον ίδιο τρόπο. Με άλλα λόγια, αν πρώτα διακριτοποιήσουμε και μετά ελαχιστοποιήσουμε (βλ. σχήμα 4.4), καταλήγουμε γενικά σε διαφορετικό αποτέλεσμα από όταν πρώτα ελαχιστοποιήσουμε και μετά διακριτοποιήσουμε (βλ. σχήμα 4.3).
- Πιο συγκεκριμένα, παρατηρούμε ότι οι συναρτήσεις $\tilde{J}(\underline{u})$ (4.75), (4.76) και (4.77) διαφέρουν από την $J_d(\underline{u})$ (4.114) στο ότι, κάθε όρος $\phi(Du(\alpha))$ του πρώτου αθροίσματος πολλαπλασιάζεται επιπλέον κατά τον παράγοντα $D^\pm u(\alpha)/Du(\alpha)$, $D^+u(\alpha)/Du(\alpha)$ ή $D^-u(\alpha)/Du(\alpha)$ αντίστοιχα. Όμως, ο πολλαπλασιαστικός αυτός παράγοντας δεν έχει κανένα νόημα και αλλοιώνει το κόστος που είναι επιθυμητό να προσδίδει η συνάρτηση $\tilde{J}(\underline{u})$ στις μεταβολές της διακριτής εικόνας \underline{u} . Το τελευταίο έχει ως συνέπεια το τελικό αποτέλεσμα να μην εκφράζει την πραγματική βελτίωση που μπορεί να συντελέσει το ϕ -συναρτησιακό.

Συμπεραίνουμε λοιπόν ότι θα ήταν ιδιαίτερα επιθυμητό ο πολλαπλασιαστικός αυτός παράγοντας να είναι πάντα ίσος με 1, κάτι το οποίο οδηγεί (όσον αφορά την διακριτοποίηση της ΜΔΕ (4.55)) για κάθε επιλογή προσέγγισης των $\text{grad}(\cdot)$ και $\text{div}(\cdot)$ (βλ. (4.62), (4.64) ή (4.66)), σε αντίστοιχη επιλογή προσέγγισης Du της νόρμας $\|\nabla u\|$ ($D^\pm u$, D^-u ή D^+u αντίστοιχα). Έχοντας επιλέξει, δηλαδή, συγκεκριμένη διακριτοποίηση για τα $\text{grad}(\cdot)$ και $\text{div}(\cdot)$, υποχρεωνόμαστε από το θεώρημα 4.6 να προσεγγίσουμε τη νόρμα $\|\nabla u\|$ με συγκεκριμένο τρόπο, εφόσον επιθυμούμε η λύση \underline{u}^* να ελαχιστοποιεί μία διακριτή συνάρτηση που έχει νόημα (και είναι η διακριτοποίηση του συναρτησιακού $J[u]$).

Στην ειδική περίπτωση λοιπόν της παραπάνω επιλογής για την προσέγγιση Du (δηλ. $Du = D^\pm u, D^-u$ ή D^+u αντίστοιχα), υπάρχει πράγματι ισοδυναμία ανάμεσα στις δύο θεωρήσεις και δεν έχει σημασία για το αποτέλεσμα αν εφαρμόζεται πρώτα η ελαχιστοποίηση ή η διακριτοποίηση.

Πρέπει βέβαια να σημειώσουμε ότι, ακόμα και όταν δεν γίνεται η παραπά-

νω επιλογή που προτείνουμε (δηλ. $Du \neq D^\pm u$, $Du \neq D^-u$ ή $Du \neq D^+u$ αντίστοιχα για τις διακριτοποιήσεις (4.62), (4.64) ή (4.66)), συνήθως στα περισσότερα σημεία ο πολλαπλασιαστικός παράγοντας του $\phi(Du(\alpha))$ στις (4.75-4.77) δεν διαφέρει σημαντικά απο την μονάδα, εφόσον η προσέγγιση Du είναι αρκετά αξιόπιστη. Κι αυτό διότι οι όροι του αριθμητή και του παρονομαστή προσεγγίζουν ουσιαστικά την ίδια ποσότητα, αλλά με διαφορετικό τρόπο. Όμως, στα σημεία όπου η διακριτή εικόνα μεταβάλλεται έντονα (π.χ. στις απότομες ακμές ή σε σημεία διαταραγμένα από υψίσυχο θόρυβο), οι δύο προσεγγίσεις μπορεί να καταλήγουν σε πολύ διαφορετικές τιμές και άρα ο πολλαπλασιαστικός παράγοντας να διαφέρει σημαντικά απο την μονάδα.

- Παρατηρούμε επίσης ότι, για τις επιλογές διακριτοποίησης (4.64) και (4.66), η καλύτερη (σύμφωνα πάντα με τα παραπάνω) δυνατή επιλογή της προσέγγισης Du είναι D^-u και D^+u αντίστοιχα. Ακόμα και τότε, όμως, η συνάρτηση $\tilde{J}(\underline{u}) = J_d(\underline{u})$ που προκύπτει παρουσιάζει το μειονέκτημα ότι υπολογίζει την τοπική μεταβολή της εικόνας (μέσω της Du), με τρόπο, ο οποίος δεν αντιμετωπίζει με συμμετρικό τρόπο τις διευθύνσεις του επιπέδου (βλ. σχέσεις (4.69) και (4.70)). Έτσι, το αποτέλεσμα που προκύπτει από την ελαχιστοποίηση δεν είναι τόσο αξιόπιστο. Το συμπέρασμα αυτό είναι βέβαια κάτι που θα μπορούσε κανείς να περιμένει, παρατηρώντας ήδη τις αντίστοιχες διακριτοποιήσεις των συνθηκών ελαχίστου του $J[u]$ (βλ. σχέσεις (4.65) και (4.67)). Κι αυτό διότι οι διευθύνσεις του επιπέδου δεν αντιμετωπίζονται με συμμετρικό τρόπο, αφού, σε κάθε pixel, λαμβάνονται υπόψιν οι τιμές του συντελεστή g μόνο στα 2 από τα 4 γειτονικά του pixel.
- Όσον αφορά τώρα την επιλογή διακριτοποίησης (4.62) (δηλ. κεντρικές διαφορές βήματος 1 pixel για $\text{grad}(\cdot)$, $\text{div}(\cdot)$ και γραμμική παρεμβολή για το g), συμπεραίνουμε ότι η καλύτερη (σύμφωνα με τα παραπάνω) δυνατή επιλογή της προσέγγισης Du είναι $D^\pm u$, οπότε η συνάρτηση που ελαχιστοποιείται είναι η εξής:

$$\tilde{J}(\underline{u}) = J_d(\underline{u}) = \sum_{\alpha=1}^{N_x N_y} \phi(D^\pm u(\alpha)) + \frac{\lambda}{2} \sum_{\alpha=1}^{N_x N_y} (u_\alpha - u_\alpha^0)^2 \quad (4.114)$$

Η συνάρτηση αυτή είναι μία αξιόπιστη προσέγγιση του συναρτησιακού $J[u]$, μιας και η $D^\pm u$ υπολογίζει με ομοιόμορφο τρόπο την τοπική μεταβολή της διακριτής εικόνας. Έτσι, σε αντίθεση με πριν, το αποτέλεσμα που προκύπτει είναι αξιόπιστο.

Αξίζει τέλος να σημειώσουμε ότι, για αυτήν την επιλογή διακριτοποίησης

(4.62), χωρίς το θεώρημα 4.6 που αποδείξαμε, η επιλογή της προσέγγισης $D^o u$ (βλ. σχέση (4.71)) θα φάνταζε το ίδιο (αν όχι περισσότερο) κατάλληλη με την επιλογή $D^\pm u$. Κι αυτό διότι η $D^o u$ χρησιμοποιεί για τις μερικές παραγώγους προσέγγιση με κεντρικές διαφορές, όπως ακριβώς γίνεται και με τα $\text{grad}(\cdot)$ και $\text{div}(\cdot)$, οπότε θα μπορούσε κάποιος να θεωρήσει το αποτέλεσμα αξιόπιστο, λόγω αυτής της συμμετρίας. Όμως, χάρη στο θεώρημα 4.6, φαίνεται ότι, στην περίπτωση αυτή, η διακριτή συνάρτηση που στην πραγματικότητα ελαχιστοποιείται παρουσιάζει σε κάθε σημείο τον ανεπιθύμητο πολλαπλασιαστικό παράγοντα $D^\pm u(\alpha)/D^o u(\alpha)$, ο οποίος δεν έχει κανένα νόημα. Το γεγονός αυτό κάνει το αποτέλεσμα που προκύπτει με την προσέγγιση $D^o u$ να είναι λιγότερο αξιόπιστο από αυτό που προκύπτει με την προσέγγιση $D^\pm u$, σε αντίθεση με ό,τι θα περίμενε κανείς με μία διασιθητική πρώτη ματιά.

Κεφάλαιο 5

Περιγραφή της Υλοποίησης των μεθόδων ΜΔΕ για την ΟΥ

5.1 Υλοποίηση ΜΔΕ για την Βελτίωση διανυσματικών εικόνων

Στην ενότητα αυτή, θα παρουσιάσουμε την αριθμητική υλοποίηση των ΜΔΕ για την βελτίωση διανυσματικών εικόνων, που παρουσιάσαμε θεωρητικά στην ενότητα 1. Στην συνέχεια, παρουσιάζουμε συνοπτικά τις ΜΔΕ με τις οποίες θα ασχοληθούμε (για αναλυτική παρουσίαση, βλ. ενότητα 1).

Προηγουμένως όμως, αποσαφηνίζουμε τον συμβολισμό που χρησιμοποιούμε. Συμβολίζουμε με: \mathbf{u} την διανυσματική εικόνα και με u_m το m -οστό της κανάλι, K_σ τον Gaussian πυρήνα με τυπική απόκλιση σ , \mathbf{u}_σ την εξομαλυμένη εικόνα $\mathbf{u}_\sigma = K_\sigma * \mathbf{u}$ και H_m την χωρική *Hessian* Μήτρα της u_m . Επίσης, συμβολίζουμε με θ_+, θ_- και λ_+, λ_- τις ιδιοτιμές και τα ιδιοδιανύσματα του δομικού τανυστή:

$$J_\rho(\nabla \mathbf{u}_\sigma) = K_\rho * \sum_{m=1}^N \nabla u_{m,\sigma} (\nabla u_{m,\sigma})^T, \quad (5.1)$$

ενώ με θ_+, θ_- και λ_+, λ_- τις ιδιοτιμές και τα ιδιοδιανύσματα της μήτρας Di Zenzo G (δομικός τανυστής χωρίς εξομαλύνσεις, δηλ. $\sigma, \rho = 0$).

Οι ΜΔΕ που θα ασχοληθούμε στην ενότητα αυτή, είναι λοιπόν οι ακόλουθες:

1. Ομαλοποιημένες μέθοδοι των *Perona-Malik*, γενικευμένες για διανυσμα-

τικές εικόνες (βλ. ενότητες 1.2,1.3):

$$\frac{\partial u_m}{\partial t} = \operatorname{div} (g(\|\nabla \mathbf{u}_\sigma\|) \nabla u_m), \quad g(s) = \exp\left(-\left(\frac{s}{K}\right)^2\right), \quad m = 1, \dots, N \quad (5.2)$$

και

$$\frac{\partial u_m}{\partial t} = \operatorname{div} (g(\|\nabla \mathbf{u}_\sigma\|) \nabla u_m), \quad g(s) = \frac{1}{1 + \left(\frac{s}{K}\right)^2}, \quad m = 1, \dots, N \quad (5.3)$$

2. Μεθόδος ελαχιστοποίησης της Διανυσματικής Ολικής Μεταβολής (*Vectorial Total Variation*) με ϵ -εξομάλυνση (βλ. ενότητα 1.6):

$$\frac{\partial u_m}{\partial t} = \operatorname{div} (g(\|\nabla \mathbf{u}\|) \nabla u_m), \quad g(s) = \frac{1}{\sqrt{\epsilon^2 + s^2}}, \quad m = 1, \dots, N \quad (5.4)$$

3. Διάχυση Βελτίωσης της Συνοχής (*Coherence Enhancing Diffusion*) διανυσματικών εικόνων (βλ. ενότητα 1.7.1):

$$\frac{\partial u_m}{\partial t} = \operatorname{div} (D \nabla u_m), \quad m = 1, \dots, N \quad (5.5)$$

όπου D τανυστής με ιδιοδιανύσματα $\boldsymbol{\theta}_+$, $\boldsymbol{\theta}_-$ και ιδιοτιμές:

$$\lambda_1 = \alpha \quad , \quad \lambda_2 = \alpha + (1 - \alpha) \exp\left(\frac{-1}{(\lambda_+^* - \lambda_-^*)^2}\right)$$

για μία α μικρή θετική σταθερά.

4. Ροή Beltrami (*Beltrami Flow*) (βλ. ενότητα 1.8):

$$\frac{\partial u_m}{\partial t} = \frac{1}{\sqrt{(1 + \lambda_+)(1 + \lambda_-)}} \operatorname{div} (D \nabla u_m), \quad m = 1, \dots, N \quad (5.6)$$

όπου D τανυστής με ιδιοδιανύσματα $\boldsymbol{\theta}_+$, $\boldsymbol{\theta}_-$ και ιδιοτιμές:

$$\lambda_1 = \sqrt{\frac{1 + \lambda_-}{1 + \lambda_+}} \quad , \quad \lambda_2 = \sqrt{\frac{1 + \lambda_+}{1 + \lambda_-}}$$

5. Διάχυση Διανυσματικών Εικόνων στην Κατεύθυνση $\boldsymbol{\theta}_-$ (βλ. ενότητα 1.9):

$$\frac{\partial u_m}{\partial t} = \operatorname{trace}(T H_m), \quad m = 1, \dots, N \quad (5.7)$$

όπου T τανυστής με ιδιοδιανύσματα θ_+, θ_- και ιδιοτιμές:

$$\lambda_1 = 0 \quad , \quad \lambda_2 = g(\lambda_+ - \lambda_-) = \frac{1}{\sqrt{1 + \lambda_+ - \lambda_-}} \quad ^1$$

6. *Ανισοτροπική Εξομάλυνση των Tschumperlé-Deriche* (βλ. ενότητα 1.10):

$$\frac{\partial u_m}{\partial t} = \text{trace}(TH_m), \quad m = 1, \dots, N \quad (5.8)$$

όπου T τανυστής με ιδιοδιανύσματα θ_+^*, θ_-^* και ιδιοτιμές:

$$\lambda_1 = \frac{1}{1 + \lambda_+^* + \lambda_-^*} \quad , \quad \lambda_2 = \frac{1}{\sqrt{1 + \lambda_+^* + \lambda_-^*}}$$

έχοντας επιλέξει $\sigma = 0$ για τον δομικό τανυστή (5.1).

Ένας γενικός συμβολισμός για τις παραπάνω εξισώσεις είναι ο εξής:

$$\frac{\partial u_m}{\partial t} = v_m(x, y, t), \quad m = 1, \dots, N$$

όπου με $v_m(x, y, t)$ συμβολίζουμε το δεύτερο μέλος οποιασδήποτε από τις παραπάνω ΜΔΕ (5.2-5.8), το οποίο αποτελεί ουσιαστικά την ταχύτητα της ΜΔΕ σε κάθε χρονική στιγμή και για κάθε σημείο.

Για την εύρεση μίας διακριτής λύσης που να προσεγγίζει την λύση των παραπάνω ΜΔΕ, θεωρούμε πρώτα από όλα, ένα ομοιόμορφο δισδιάστατο ορθογωνικό πλέγμα, έτσι ώστε κάθε κυψέλη του πλέγματος να ταυτίζεται με ένα pixel της εικόνας. Επίσης, διακριτοποιούμε τον χρόνο με βήμα τ . Χρησιμοποιούμε τον εξής συμβολισμό για την διακριτοποιημένη συνάρτηση u_m : $u_m^n(i, j, n) \approx u_m(ih_x, jh_y, n\tau)$, $1 \leq i \leq N_x, 1 \leq j \leq N_y$, όπου N_x (N_y) ο αριθμός των σημείων του πλέγματος, στην διεύθυνση του x (y). Για τα χωρικά βήματα, θεωρούμε ότι $h_x = h_y = 1$.

Στην προσέγγιση που υλοποιήσαμε, χρησιμοποιούμε πεπερασμένες διαφορές, προσεγγίζοντας την χρονική παράγωγο με εμπρόσθια διαφορά και τις χωρικές παραγώγους με κεντρικές διαφορές. Επίσης, επιλέγουμε ο υπολογισμός της ταχύτητας $v_m(i, j, n) = v_m(\mathbf{u})$ να γίνεται με βάση τις τιμές της εικόνας στην τρέχουσα χρονική στιγμή. Έτσι, προκύπτει το explicit σχήμα:

$$\frac{u_m^{n+1} - u_m^n}{\tau} = v_m(\mathbf{u}^n), \quad m = 1, \dots, N \quad (5.9)$$

¹φυσικά, μπορούν να γίνουν και άλλες επιλογές για την $g(\lambda_+ - \lambda_-)$ στη μέθοδο αυτή, αλλά η παραπάνω ήταν αυτή που χρησιμοποιήσαμε στα πειράματά μας.

5.1 Υλοποίηση ΜΔΕ για την Βελτίωση διανυσματικών εικόνων 111

Επομένως, με βάση την παραπάνω προσέγγιση, το σημαντικό βήμα της υλοποίησης είναι ο προσδιορισμός της ταχύτητας $\mathbf{v}(\mathbf{u}^n)$, διότι εφόσον υπολογιστεί, η τιμή της εικόνας στην επόμενη χρονική στιγμή προκύπτει άμεσα από την τωρινή, μέσω της παραπάνω σχέσης. Με την ανάλυση και περιγραφή του βήματος αυτού θα ασχοληθούμε σε μεγάλο βαθμό στη συνέχεια:

Όπως παρατηρούμε, οι παραπάνω ΜΔΕ (5.2)-(5.8) παρουσιάζουν διάφορα κοινά στοιχεία μεταξύ τους. Η υλοποίησή μας βασίστηκε σε μία όσο το δυνατόν πιο ενοποιητική προσέγγιση, ακολουθώντας τα ίδια βήματα σε όλες τις μεθόδους, όπου αυτό είναι δυνατό. Έτσι, για την ανάλυσή μας κάνουμε μία κατηγοριοποίηση των ΜΔΕ αυτών. Παρατηρώντας τις εξισώσεις αυτές, βλέπουμε ότι μπορούν να χωριστούν στις κατηγορίες αυτών που χρησιμοποιούν τελεστή απόκλισης div (ΜΔΕ (5.2)-(5.6)) και αυτών που χρησιμοποιούν τελεστή ίχνους trace (ΜΔΕ (5.7),(5.8)). Επίσης, ένας άλλος χωρισμός είναι σε αυτές που για την οδήγηση της διάχυσης χρησιμοποιούν βαθμωτό συντελεστή (ΜΔΕ μη γραμμικής ισοτροπικής διάχυσης (5.2)-(5.4)) και σε αυτές που χρησιμοποιούν τανυστή (ΜΔΕ (5.5)-(5.8)). Οι παραπάνω κατηγοριοποιήσεις μπορούν φυσικά να συμπεριλάβουν και άλλες ΜΔΕ που χρησιμοποιούνται στην Όραση Υπολογιστών.

Με βάση τα παραπάνω, παρουσιάζουμε στη συνέχεια τις προσεγγίσεις που κάνουμε για το δεύτερο μέλος των ΜΔΕ, $v_m(x, y, t)$, στις διάφορες περιπτώσεις:

5.1.1 Προσέγγιση του όρου $\text{div}(D \nabla u_m)$

Συμβολίζουμε με w οποιοδήποτε κανάλι της εικόνας \mathbf{u} , δηλαδή:

$$w = u_m, m = 1, \dots, N$$

Επίσης, θεωρούμε ότι τα στοιχεία του τανυστή D δίνονται από την σχέση (λόγω του ότι είναι συμμετρικός):

$$D = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$$

Οι ΜΔΕ με βαθμωτό συντελεστή (ΜΔΕ (5.2)-(5.4)) μπορούν να ειδωθούν σαν ειδική περίπτωση, όπου $D = gI$, οπότε:

$$a = c = g, b = 0$$

Ακολουθώντας παρόμοια διαδικασία με το [63], προσεγγίζουμε τον όρο $\text{div}(D \nabla w)$ στο σημείο (i, j) του πλέγματος, χρησιμοποιώντας κεντρικές δια-

φορές με βήμα 1 ή 2 pixels για τις χωρικές παραγώγους:

$$\begin{aligned} v_m &= \operatorname{div}(D \nabla w) = \operatorname{div} \left(\begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} w_x \\ w_y \end{bmatrix} \right) \\ &= \frac{\partial}{\partial x}(aw_x) + \frac{\partial}{\partial x}(bw_y) + \frac{\partial}{\partial y}(bw_x) + \frac{\partial}{\partial y}(cw_y) \end{aligned}$$

άρα:

$$v_m(i, j) = \left(D_x(aD_x w) + D_{2x}(bD_{2y} w) + D_{2y}(bD_{2x} w) + D_y(cD_y w) \right)(i, j) \quad (5.10)$$

Αναπτύσσουμε τις πιο πάνω διαφορές και, για να συνεχίσουμε, χρησιμοποιούμε απλή γραμμική παρεμβολή για τις τιμές στα ενδιάμεσα σημεία του πλέγματος (παίρνοντας τον μέσο όρο των τιμών στα δύο πιο κοντινά pixels). Αν κάνουμε τις πράξεις, καταλήγουμε στο ότι ο όρος $v_m(i, j)$ προκύπτει σαν γραμμικός συνδυασμός των τιμών της w γύρω από το σημείο (i, j) , σε ένα 3×3 τετραγωνικό πλαίσιο. Οι συντελεστές του γραμμικού αυτού συνδυασμού εξαρτώνται από τις τιμές των στοιχείων a, b, c γύρω από το (i, j) . Ο γραμμικός αυτός συνδυασμός μπορεί να ειπωθεί σαν εφαρμογή διακριτής τοπικής συσχέτισης του καναλιού w της εικόνας με έναν μεταβαλλόμενο 3×3 πυρήνα $k_{i,j}(\ell_1, \ell_2)$, δηλαδή:

$$\begin{aligned} v_m(i, j) &= (k_{i,j} \tilde{*} w)(i, j) \\ &= \sum_{\ell_1=-1}^1 \sum_{\ell_2=-1}^1 k_{i,j}(\ell_1, \ell_2) \cdot w(i + \ell_1, j + \ell_2) \end{aligned} \quad (5.11)$$

όπου με “ $\tilde{*}$ ” συμβολίζουμε την συσχέτιση.

Τα στοιχεία του πυρήνα αυτού, όπως αυτά μπορούν να βρεθούν από τις πράξεις στην (5.10), φαίνονται στον πίνακα 5.1.

Στην απλή υποπερίπτωση με βαθμωτό συντελεστή (ΜΔΕ (5.2)-(5.4)), όπου $a = c = g$ και $b = 0$, ο πυρήνας αυτός απλοποιείται και καταλήγουμε άμεσα στον πυρήνα συσχέτισης του πίνακα 5.2.

5.1.2 Προσέγγιση του όρου $\operatorname{trace}(TH_m)$

Αν για τα στοιχεία του τανυστή T χρησιμοποιήσουμε τον ίδιο συμβολισμό με τον τανυστή D , δηλαδή:

$$T = \begin{bmatrix} a & b \\ b & c \end{bmatrix},$$

5.1 Υλοποίηση ΜΔΕ για την Βελτίωση διανυσματικών εικόνων 113

| | | | | |
|----------|----|--------------------------------------|--|--------------------------------------|
| | | ℓ_2 | | |
| | | -1 | 0 | 1 |
| ℓ_1 | -1 | $\frac{b_{(i-1,j)}+b_{(i,j-1)}}{4}$ | $\frac{a_{(i-1,j)}+a_{(i,j)}}{2}$ | $-\frac{b_{(i-1,j)}+b_{(i,j+1)}}{4}$ |
| | 0 | $\frac{c_{(i,j-1)}+c_{(i,j)}}{2}$ | $-\frac{a_{(i-1,j)}+a_{(i+1,j)}+2a_{(i,j)}+c_{(i,j-1)}+c_{(i,j+1)}+2c_{(i,j)}}{2}$ | $\frac{c_{(i,j+1)}+c_{(i,j)}}{2}$ |
| | 1 | $-\frac{b_{(i+1,j)}+b_{(i,j-1)}}{4}$ | $\frac{a_{(i+1,j)}+a_{(i,j)}}{2}$ | $\frac{b_{(i+1,j)}+b_{(i,j+1)}}{4}$ |

Πίνακας 5.1: Πυρήνας τοπικής συσχέτισης $k_{i,j}(\ell_1, \ell_2)$, για την προσέγγιση του $\text{div}(D \nabla u_m)$

| | | | | |
|----------|----|-----------------------------------|---|-----------------------------------|
| | | ℓ_2 | | |
| | | -1 | 0 | 1 |
| ℓ_1 | -1 | 0 | $\frac{g_{(i-1,j)}+g_{(i,j)}}{2}$ | 0 |
| | 0 | $\frac{g_{(i,j-1)}+g_{(i,j)}}{2}$ | $-\frac{g_{(i-1,j)}+g_{(i+1,j)}+4g_{(i,j)}+g_{(i,j-1)}+g_{(i,j+1)}}{2}$ | $\frac{g_{(i,j+1)}+g_{(i,j)}}{2}$ |
| | 1 | 0 | $\frac{g_{(i+1,j)}+g_{(i,j)}}{2}$ | 0 |

Πίνακας 5.2: Πυρήνας τοπικής συσχέτισης $k_{i,j}(\ell_1, \ell_2)$, για την προσέγγιση του $\text{div}(g \nabla u_m)$

τότε, για οποιοδήποτε κανάλι $w = u_m$ της εικόνας \mathbf{u} , ο όρος $\text{trace}(TH_m)$ γράφεται:

$$v_m = \text{trace}(TH_m) = \text{trace} \left(\begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} w_{xx} & w_{xy} \\ w_{xy} & w_{yy} \end{bmatrix} \right) = aw_{xx} + 2bw_{xy} + cw_{yy} \quad (5.12)$$

Προσεγγίζουμε τις 2^{ος} μερικές παραγώγους χρησιμοποιώντας κεντρικές διαφορές ως εξής:

$$w_{xx}(i, j) = D_x(D_x w(i, j)) = w(i + 1, j) - 2w(i, j) + w(i - 1, j) \quad (5.13)$$

$$w_{yy}(i, j) = D_y(D_y w(i, j)) = w(i, j + 1) - 2w(i, j) + w(i, j - 1) \quad (5.14)$$

$$\begin{aligned} w_{xy}(i, j) &= D_{2y}(D_{2x} w(i, j)) \\ &= \frac{1}{4} \{w(i + 1, j + 1) - w(i - 1, j + 1) - w(i + 1, j - 1) + w(i - 1, j - 1)\} \end{aligned} \quad (5.15)$$

Από τις (5.12-5.15) και κάνοντας τις πράξεις, προκύπτει, αντίστοιχα με την προηγούμενη περίπτωση, ένας μεταβλητός πυρήνας συσχέτισης για την εξαγωγή της προσέγγισης $v_m(i, j)$ του όρου $\text{trace}(TH_m)$. Ο πυρήνας αυτός φαίνεται στον πίνακα 5.3.

| | | | | |
|----------|----|------------------------|-------------------------------|------------------------|
| | | ℓ_2 | | |
| | | -1 | 0 | 1 |
| | -1 | $\frac{b_{(i,j)}}{2}$ | $a_{(i,j)}$ | $-\frac{b_{(i,j)}}{2}$ |
| | 0 | $c_{(i,j)}$ | $-2\{a_{(i,j)} + c_{(i,j)}\}$ | $c_{(i,j)}$ |
| | 1 | $-\frac{b_{(i,j)}}{2}$ | $a_{(i,j)}$ | $\frac{b_{(i,j)}}{2}$ |
| ℓ_1 | | | | |

Πίνακας 5.3: Πυρήνας τοπικής συσχέτισης $k_{i,j}(\ell_1, \ell_2)$, για την προσέγγιση του $\text{trace}(TH_m)$

Από τα παραπάνω διαφαίνεται ότι, για να γίνει εφικτός ο υπολογισμός της ταχύτητας $v_m(i, j)$, απαιτείται για όλα τα σημεία ο υπολογισμός του συντελεστή $g(i, j)$ (στην απλή περίπτωση μη γραμμικής ιστροπικής διάχυσης) ή των τανυστών $D(i, j)$ ή $T(i, j)$ (στις υπόλοιπες περιπτώσεις). Τα στοιχεία όμως αυτά προκύπτουν από τη νόρμα $\|\nabla \mathbf{u}_\sigma\|$ ή τον δομικό τανυστή $J_\rho(\nabla \mathbf{u}_\sigma)$ (βλ. περιγραφή ΜΔΕ ((5.5)-(5.8))) αντίστοιχα, με την προσέγγιση των οποίων θα ασχοληθούμε στη συνέχεια.

5.1.3 Προσέγγιση της νόρμας $\|\nabla u_\sigma\|$

Η νόρμα $\|\nabla u_\sigma\| = \left\{ \sum_{m=1}^N \|\nabla u_{m,\sigma}\|^2 \right\}^{\frac{1}{2}}$ προκύπτει άμεσα από τον υπολογισμό της νόρμας $\|\nabla w\|$ για κάθε εξομαλυμένο κανάλι $w = u_{m,\sigma}$ της εικόνας. Για την προσέγγιση της $\|\nabla w\| = (w_x^2 + w_y^2)^{\frac{1}{2}}$, χρησιμοποιούμε δύο εναλλακτικές επιλογές:

1. Προσέγγιση με χρήση κεντρικών διαφορών με βήμα 2 pixels για τις μερικές παραγώγους w_x, w_y :

$$\|\nabla w\|^2 \approx (D^o w(i, j))^2 = (D_{2x} w(i, j))^2 + (D_{2y} w(i, j))^2 \quad (5.16)$$

2. Προσέγγιση με χρήση εμπρόσθιων και οπίσθιων διαφορών βήματος 1 pixel:

$$\|\nabla w\|^2 \approx (D^\pm w(i, j))^2 = \frac{1}{2} \{ (D_x^+ w)^2 + (D_x^- w)^2 + (D_y^+ w)^2 + (D_y^- w)^2 \} \quad (5.17)$$

Το παραπάνω αντιστοιχεί σε προσέγγιση της w_x^2 με τον μέσο όρο των αντίστοιχων τιμών στα ενδιάμεσα σημεία $(i - \frac{1}{2}, j)$ και $(i + \frac{1}{2}, j)$, δηλαδή:

$$(w_x^2)(i, j) \approx \frac{1}{2} \{ (w_x^2)(i - \frac{1}{2}, j) + (w_x^2)(i + \frac{1}{2}, j) \}$$

και αντίστοιχη προσέγγιση της w_y^2 .

Η προσέγγιση D^\pm πλεονεκτεί έναντι της D^o στο ότι λαμβάνει υπόψη της την τιμή της w και στο σημείο (i, j) και άρα ο υπολογισμός της νόρμας είναι πιο αντιπροσωπευτικός της τοπικής μεταβολής της διακριτής εικόνας. Έτσι, σε αντίθεση με την D^o , μπορούν να εντοπιστούν ακόμα και οι ακμές με το ελάχιστο πάχος του 1 pixel. Επιπρόσθετα, παρουσιάζει το πλεονέκτημα ότι, στις ΜΔΕ ελαχιστοποίησης των ϕ -συναρτησιακών (των οποίων ειδική περίπτωση αποτελούν οι ΜΔΕ (5.2)-(5.4) που υλοποιήσαμε), η διακριτή προσέγγιση της ΜΔΕ που περιγράφουμε εδώ αποτελεί και αυτή ροή προς την ελαχιστοποίηση του αντίστοιχου διακριτού ϕ -συναρτησιακού (όπως αποδείξαμε στην ενότητα ..). Από την άλλη όμως, η υλοποίηση της D^\pm είναι πιο περίπλοκη από την D^o , διότι απαιτεί διπλάσιους υπολογισμούς διαφορών.

Για τους παραπάνω λόγους, στην υλοποίηση που κάναμε χρησιμοποιήσαμε, στις περισσότερες περιπτώσεις, την προσέγγιση D^\pm .

όπου:

$$D_x w(i + \frac{1}{2}, j) = w(i + 1, j) - w(i, j)$$

$$D_{2y} w(i + \frac{1}{2}, j) = \frac{1}{2} \{w(i + \frac{1}{2}, j + 1) - w(i + \frac{1}{2}, j - 1)\}$$

Εδώ, χρησιμοποιούμε γραμμική παρεμβολή μόνο για τις τιμές $w(i + \frac{1}{2}, j + 1)$ και $w(i + \frac{1}{2}, j - 1)$, οπότε τελικά προκύπτει:

$$\|\nabla w\|_{(i+\frac{1}{2},j)} = \left\{ [w(i + 1, j) - w(i, j)]^2 + \frac{1}{4} [w(i, j + 1) + w(i + 1, j + 1) - w(i, j - 1) - w(i + 1, j - 1)]^2 \right\}^{\frac{1}{2}}$$

(5.18)

Για τα σημεία $(i, j + \frac{1}{2})$ καταλήγουμε με παρόμοιο τρόπο στην σχέση:

$$\|\nabla w\|_{(i,j+\frac{1}{2})} = \left\{ [w(i, j + 1) - w(i, j)]^2 + \frac{1}{4} [w(i + 1, j) + w(i + 1, j + 1) - w(i - 1, j) - w(i - 1, j + 1)]^2 \right\}^{\frac{1}{2}}$$

(5.19)

Συμπεραίνουμε λοιπόν ότι στην προσέγγιση αυτή, οι συντελεστές $g(i + \frac{1}{2}, j)$, $g(i, j + \frac{1}{2})$ προσεγγίζονται με μεγαλύτερη ακρίβεια, διότι δεν κάνουμε απευθείας γραμμική παρεμβολή, με αποτέλεσμα οι τιμές της w που λαμβάνονται υπόψιν να είναι περισσότερο συγκεντρωμένες γύρω από το σημείο. Έτσι, έχουμε μία πιο αξιόπιστη προσέγγιση της τοπικής συμπεριφοράς της συνεχούς διάχυσης και η υλοποίηση αυτή παρουσιάζει βελτίωση στα αποτελέσματα, ιδιαίτερα στις ακμές, όπου τις θολώνει λιγότερο. Βέβαια, η προσέγγιση αυτή είναι πιο περίπλοκη, με μεγαλύτερο υπολογιστικό κόστος, διότι απαιτείται ο υπολογισμός του συντελεστή g και της $\|\nabla \mathbf{u}_\sigma\|$ σε διπλάσια σημεία (αφού τα οριζόντια και κάθετα ενδιάμεσα σημεία είναι διπλάσια από τα σημεία του πλέγματος).

5.1.5 Προσέγγιση του δομικού τανυστή $J_\rho(\nabla \mathbf{u}_\sigma)$

Ο δομικός τανυστής $J_\rho(\nabla \mathbf{u}_\sigma)$ υπολογίζεται με εξομάλυνση του τανυστή:

$$J(\nabla \mathbf{u}_\sigma) = \begin{bmatrix} \sum_{m=1}^N (u_{m,\sigma,x}^2) & \sum_{m=1}^N (u_{m,\sigma,x} u_{m,\sigma,y}) \\ \sum_{m=1}^N (u_{m,\sigma,x} u_{m,\sigma,y}) & \sum_{m=1}^N (u_{m,\sigma,y}^2) \end{bmatrix},$$

ο οποίος προκύπτει, σύμφωνα με την παραπάνω σχέση, από τον υπολογισμό των χωρικών παραγώγων w_x, w_y , για κάθε εξομαλυμένο κανάλι $w = u_{m,\sigma}$ της εικόνας. Για την προσέγγιση των w_x, w_y χρησιμοποιούμε κεντρικές διαφορές με βήμα 2 pixels, δηλαδή:

$$\begin{aligned}w_x(i, j) &\approx D_{2x}w(i, j) \\w_y(i, j) &\approx D_{2y}w(i, j)\end{aligned}$$

Στην περίπτωση αυτή, δεν μπορούμε να χρησιμοποιούμε κάτι αντίστοιχο με την πιο αξιόπιστη προσέγγιση D^\pm της νόρμας $\|\nabla w\|$ (βλ. σχέση (5.17)), διότι εδώ χρειαζόμαστε και τις τιμές w_x, w_y (όχι μόνο τις w_x^2, w_y^2 όπως πριν). Όμως, η προσέγγιση D^\pm δεν λαμβάνει υπόψιν της το πρόσημο της μεταβολής, παρά μόνο την έντασή της.

5.1.6 Υλοποίηση των εξομαλύνσεων της εικόνας u και του τανυστή $J(\nabla \mathbf{u}_\sigma)$

Από την μέχρι στιγμής παρουσίαση, γίνεται φανερό ότι, για να υπολογιστούν η νόρμα άλλα και ο δομικός τανυστής $J_\rho(\nabla \mathbf{u}_\sigma)$ απαιτείται, στις περισσότερες μεθόδους, η εξομάλυνση με Gaussian πυρήνες συνέλιξης $K_{\tilde{\sigma}}$ (είτε της εικόνας \mathbf{u} , οπότε $\tilde{\sigma} = \sigma$ είτε του τανυστή $J(\nabla \mathbf{u}_\sigma)$, οπότε $\tilde{\sigma} = \rho$). Υλοποιήσαμε τη συνέλιξη αυτή με κατευθείαν εφαρμογή της στα pixels του πλέγματος. Για την διακριτοποίηση του Gaussian πυρήνα $K_{\tilde{\sigma}}$, χρησιμοποιήσαμε ένα τετραγωνικό πλαίσιο με μεταβλητό όμως μήκος, έτσι, ώστε να εκτείνεται μέχρι που οι τιμές του πυρήνα να γίνονται ασήμαντες (ο λόγος τους προς την μέγιστη τιμή του κέντρου να είναι μικρότερος από μία μικρή θετική σταθερά ϵ_k). Για το σκοπό αυτό, εύκολα μπορούμε να καταλήξουμε (από την έκφραση της 2Δ Gaussian κατανομής) ότι, για τυπική απόκλιση $\tilde{\sigma}$, το τετραγωνικό πλαίσιο πρέπει να έχει πλευρές με μήκος (που φανερώνει την από άκρο σε άκρο έκταση του πυρήνα):

$$N_k = 2\lceil \tilde{\sigma} \sqrt{-2 \ln \epsilon_k} \rceil + 1$$

Αν για παράδειγμα επιλέξουμε $\epsilon_k = 10^{-4}$, όπως στα πειράματά μας, καταλήγουμε:

$$N_k \approx 2\lceil 4.3 \cdot \tilde{\sigma} \rceil + 1 \quad (5.20)$$

Ύστερα από τους υπολογισμούς των τιμών του $K_{\tilde{\sigma}}(i, j)$ στα σημεία του πλαισίου, κανονικοποιούμε τις τιμές, ώστε να έχουν μέση τιμή 1 (λόγω της διακριτοποίησης η αντίστοιχη ιδιότητα του συνεχούς πυρήνα $K_{\tilde{\sigma}}(x, y)$ διαταράσσεται).

Η παραπάνω υλοποίηση που πραγματοποιήσαμε παρουσιάζει το πλεονέκτημα σε σχέση με την επιλογή πλαισίου σταθερού μεγέθους (π.χ. 3×3 ή 5×5 που

χρησιμοποιείται συχνά στην βιβλιογραφία) ότι παρέχει μία πιο αξιόπιστη προσέγγιση του Gaussian βαθυπερατού φιλτραρίσματος, ιδιαίτερα όταν το σ είναι μεγάλο σχετικά με την έκταση του πλαισίου. Πιο συγκεκριμένα, στη χρήση πλαισίου σταθερού μεγέθους, υπάρχει ο κίνδυνος να αποκοπούν και τιμές του πυρήνα που δεν είναι ασήμαντες. Αυτή όμως η απότομη αποκοπή καταλήγει σε επέκταση του φάσματος του φίλτρου, διότι αποτελεί ουσιαστικά πολλαπλασιασμό με κυβικό παλμό, το οποίο ισοδυναμεί με συνέλιξη στο χώρο των συχνοτήτων της Gaussian φασματικής κατανομής με 2Δ συνάρτηση δειγματοληψίας (sinc). Έτσι, το τελικό διακριτό φιλτράρισμα μπορεί να εκτείνεται στις συχνότητες αρκετά περισσότερο από την αντίστοιχη Gaussian φασματική κατανομή, οπότε η συμπεριφορά του φιλτραρίσματος δεν μπορεί να ελεγχθεί με αξιόπιστο τρόπο μέσω της σταθεράς σ .

Από την άλλη βέβαια, η υλοποίηση που πραγματοποιήσαμε υστερεί στην ταχύτητα υπολογισμού, λόγω υψηλού χρονικού κόστους για τον υπολογισμό της συνέλιξης, ιδιαίτερα όταν το πλαίσιο του πυρήνα είναι μεγάλο. Καλύτερη υπολογιστική απόδοση θα μπορούσε όμως να επιτευχθεί με αλγορίθμους επαναληπτικής υλοποίησης της παραπάνω συνέλιξης που έχουν προταθεί στην βιβλιογραφία (βλ. π.χ. [20]).

5.1.7 Αλγόριθμοι για την υλοποίηση των ΜΔΕ

Βασισμένοι στην παραπάνω ανάλυση, παρουσιάζουμε τώρα, όσο το δυνατόν πιο ενοποιημένα, τα βασικά βήματα των αλγορίθμων που υλοποιήσαμε:

1. Ο πρώτος γενικός αλγόριθμος υλοποιεί τις ΜΔΕ με βαθμωτό συντελεστή διάχυσης (ΜΔΕ μη γραμμικής ισοτροπικής διάχυσης (5.2)-(5.4)) και παρουσιάζεται στον πίνακα 5.5. Ο πίνακας 5.6 είναι επεξηγηματικός του πίνακα 5.5 και παρουσιάζει τις περαιτέρω επιλογές που πρέπει να γίνουν για την υλοποίηση των συγκεκριμένων πλέον μεθόδων της κατηγορίας αυτής, με τις οποίες ασχοληθήκαμε πειραματικά.
2. Ο δεύτερος γενικός αλγόριθμος υλοποιεί τις ΜΔΕ με ταυυστή (ΜΔΕ (5.5)-(5.8)) και παρουσιάζεται στον πίνακα 5.7. Αντίστοιχα με πριν, ο πίνακας 5.8 είναι επεξηγηματικός του πίνακα 5.7 για κάθε συγκεκριμένη μέθοδο της κατηγορίας αυτής που υλοποιήσαμε.

5.1.8 Επιπλέον Βήματα στην Υλοποίηση κάποιων Μεθόδων

Όσον αφορά τον αλγόριθμο του πίνακα 5.7, όπως βλέπουμε και από τον πίνακα 5.8, για την υλοποίηση δύο μεθόδων απαιτείται ένα επιπλέον βήμα, το

ΕΙΣΟΔΟΙ: αρχική εικόνα \mathbf{u}^0 (διανυσματική με N κανάλια), αριθμός επαναλήψεων N_{it} , χρονικό βήμα τ , σταθερά εξομάλυνσης σ .

Υπολόγισε τον Gaussian πυρήνα K_σ από το σ , με έκταση σύμφωνα με την (5.20).

Αρχικοποίηση: $\mathbf{u} = \mathbf{u}^0$

For $n = 1, \dots, N_{it}$

Εξομάλυνση της \mathbf{u}

$$\mathbf{u}_\sigma = K_\sigma * \mathbf{u}$$

Υπολόγισε τη νόρμα $\|\nabla \mathbf{u}_\sigma\| = \left\{ \sum_{m=1}^N \|\nabla u_{m,\sigma}\|^2 \right\}^{\frac{1}{2}}$, σύμφωνα με την σχέση

(5.16) ή (5.17).

Υπολογισμός του συντελεστή διάχυσης g

$$g = g(\|\nabla \mathbf{u}_\sigma\|), \text{ ανάλογα με τη μέθοδο (βλ. πίνακ. 5.6)}$$

Υπολόγισε τους τοπικούς πυρήνες συσχέτισης $k_{i,j}$ από τις τιμές του g , σύμφωνα με τον πίνακα 5.2.

Υπολογισμός όλων των καναλιών u_m της ταχύτητας \mathbf{v}

For $m = 1, \dots, N$

Σε κάθε (i, j) , βρες την $u_m(i, j)$, εφαρμόζοντας τοπική συσχέτιση της $k_{i,j}$ με την u_m (βλ. (5.11)).

End for

Υπολογισμός της επόμενης εικόνας

$$\mathbf{u} = \mathbf{u} + \tau \mathbf{v}$$

End for

Πίνακας 5.5: Γενικός αλγόριθμος για ΜΔΕ με βαθμωτό συντελεστή διάχυσης

| Μέθοδος | Εξίσωση | Σταθερά εξομάλυνσης σ | $g(s)$ |
|----------------|---------|------------------------------|-----------------------------|
| Perona-Malik 1 | (5.2) | ≥ 0 | $\exp(-(s/K)^2)$ |
| Perona-Malik 2 | (5.3) | ≥ 0 | $1/\{1 + (s/K)^2\}$ |
| Vectorial TV | (5.4) | 0 | $1/\sqrt{\epsilon^2 + s^2}$ |

Πίνακας 5.6: Περαιτέρω επιλογές για κάθε μέθοδο, στον αλγόριθμο του πίνακα 5.5

5.1 Υλοποίηση ΜΔΕ για την Βελτίωση διανυσματικών εικόνων 121

ΕΙΣΟΔΟΙ: αρχική εικόνα \mathbf{u}^0 (διανυσματική με N κανάλια), αριθμός επαναλήψεων N_{it} , χρονικό βήμα τ (ή παράμετρος β_{\max}), σταθερές εξομαλύνσεων σ, ρ .

Υπολόγισε τους Gaussian πυρήνες K_σ, K_ρ από τα σ, ρ , με έκταση σύμφωνα με την (5.20).

Αρχικοποίηση: $\mathbf{u} = \mathbf{u}^0$

For $n = 1, \dots, N_{it}$

Εξομάλυνση της \mathbf{u}

$$\mathbf{u}_\sigma = K_\sigma * \mathbf{u}$$

$$\text{Υπολογισμός του δομικού τανυστή } J(\nabla \mathbf{u}_\sigma) = \begin{bmatrix} j_1 & j_2 \\ j_2 & j_3 \end{bmatrix}$$

αρχικοποίηση: $j_1 = j_2 = j_3 = 0$

For $m = 1, \dots, N$

$$u_{m,x}(i, j) = \{u_{m,\sigma}(i+1, j) - u_{m,\sigma}(i-1, j)\}/2$$

$$u_{m,y}(i, j) = \{u_{m,\sigma}(i, j+1) - u_{m,\sigma}(i, j-1)\}/2$$

$$j_1 = j_1 + u_{m,x}^2; j_2 = j_2 + u_{m,x}u_{m,y}; j_3 = j_3 + u_{m,y}^2$$

End for

Εξομάλυνση του $J(\nabla \mathbf{u}_\sigma)$ για την εύρεση του τελικού $J_\rho(\nabla \mathbf{u}_\sigma)$

$$j_1 = K_\rho * j_1; j_2 = K_\rho * j_2; j_3 = K_\rho * j_3$$

Υπολογισμός ιδιοτιμών λ_+, λ_- και γωνίας θ του ιδιοδιαν. $\boldsymbol{\theta}_+$ του $J_\rho(\nabla \mathbf{u}_\sigma)$

$$\Delta = (j_1 - j_3)^2 + 4j_2^2$$

$$\lambda_+ = \{(j_1 + j_3) + \sqrt{\Delta}\}/2; \lambda_- = \{(j_1 + j_3) - \sqrt{\Delta}\}/2$$

$$\theta = \text{angle}([2j_2, j_3 - j_1 + \sqrt{\Delta}])$$

Υπολογισμός ιδιοτιμών του τανυστή διάχυσης

$$\lambda_1 = \lambda_1(\lambda_+, \lambda_-); \lambda_2 = \lambda_2(\lambda_+, \lambda_-), \text{ ανάλογα με τη μέθοδο (βλ. πίνακα 5.8)}$$

$$\text{Υπολογισμός στοιχείων του τανυστή διάχυσης } \begin{bmatrix} a & b \\ b & c \end{bmatrix}$$

$$a = \lambda_1 \cos^2 \theta + \lambda_2 \sin^2 \theta; b = (\lambda_1 - \lambda_2) \sin(2\theta)/2; c = \lambda_1 \sin^2 \theta + \lambda_2 \cos^2 \theta$$

Υπολόγισε τους τοπικούς πυρήνες συσχέτισης $k_{i,j}$ από τα a, b, c , ανάλογα με τη μέθοδο (βλ. πίνακα 5.8)

Υπολογισμός όλων των καναλιών v_m της ταχύτητας \mathbf{v}

For $m = 1, \dots, N$

Σε κάθε (i, j) , βρες την $v_m(i, j)$, εφαρμόζοντας τοπική συσχέτιση της $k_{i,j}$ με την u_m (βλ. (5.11)).

End for

Εφάρμοσε κάποιο επιπλέον βήμα ή μην κάνεις τίποτα, ανάλογα με τη μέθοδο (βλ. πίνακα 5.8).

Υπολογισμός της επόμενης εικόνας

$$\mathbf{u} = \mathbf{u} + \tau \mathbf{v}$$

End for

Πίνακας 5.7: Γενικός αλγόριθμος για ΜΔΕ με τανυστή διάχυσης

| Μέθοδος | Εξίσ. | σ | ρ | $\lambda_1(\lambda_+, \lambda_-)$ | $\lambda_2(\lambda_+, \lambda_-)$ | $k_{i,j}$ από: | Επιπλέον βήμα |
|-------------------------|-------|----------|----------|--|---|----------------|---------------|
| Coher.Enhanc. | (5.5) | ≥ 0 | ≥ 0 | α | $\alpha + (1-\alpha)e^{\frac{-1}{(\lambda_+ - \lambda_-)^2}}$ | πίν. 5.1 | όχι |
| Beltrami | (5.6) | 0 | 0 | $\sqrt{\frac{1+\lambda_-}{1+\lambda_+}}$ | $\sqrt{\frac{1+\lambda_+}{1+\lambda_-}}$ | πίν. 5.1 | ναι |
| Διάχυση στην θ_- | (5.7) | 0 | 0 | 0 | $\frac{1}{\sqrt{1+\lambda_+ - \lambda_-}}$ | πίν. 5.3 | όχι |
| Εξομάλ. Tsch.-Der. | (5.8) | 0 | ≥ 0 | $\frac{1}{1+\lambda_+ + \lambda_-}$ | $\frac{1}{\sqrt{1+\lambda_+ + \lambda_-}}$ | πίν. 5.3 | ναι |

Πίνακας 5.8: Περαιτέρω επιλογές για κάθε μέθοδο, στον αλγόριθμο του πίνακα 5.7

οποίο περιγράφουμε στη συνέχεια:

1. Για τη μέθοδο *Beltrami Flow*, όπως φαίνεται από την εξίσωση (5.6) που την περιγράφει, αφού υπολογίσουμε την ταχύτητα v_m σαν τοπική συσχέτιση, για να προκύψει το τελικό αποτέλεσμα, πρέπει να την πολλαπλασιάσουμε με το κατάλληλο βάρος. Το βάρος αυτό εξαρτάται από τα λ_+, λ_- , τα οποία ούτως ή άλλως έχουμε ήδη υπολογίσει. Έτσι, χρειάζεται μόνο να προσθέσουμε το εξής βήμα:

For $m = 1, \dots, N$

$$v_m = \frac{1}{\sqrt{(1+\lambda_+)(1+\lambda_-)}} \cdot v_m$$

 End for

2. Για τη μέθοδο *Ανισοτροπικής Εξομάλυνσης των Tschumperlé-Deriche* χρησιμοποιούμε στην υλοποίηση προσαρμοζόμενο (adaptive) χρονικό βήμα τ , όπως προτείνεται στο [52]. Πιο συγκεκριμένα, ο αλγόριθμος δέχεται σαν είσοδο την παράμετρο β_{\max} (και όχι το τ όπως στις υπόλοιπες μεθόδους), η οποία εκφράζει τη μέγιστη (ανάμεσα στα pixels) διανυσματική μεταβολή $\|\mathbf{u}^{n+1} - \mathbf{u}^n\|$ που επιθυμούμε να συμβαίνει κατά τη διάρκεια 1 επανάληψης. Έτσι, το χρονικό βήμα σε κάθε επανάληψη, προκύπτει από τη σχέση:

$$\tau = \frac{\beta_{\max}}{\max_{(i,j)} \|\mathbf{v}(i, j)\|} \quad (5.21)$$

όπου $\mathbf{v}(i, j)$ η διανυσματική ταχύτητα της διακριτοποιημένης ΜΔΕ, δηλ. $v_m(i, j) = \text{trace}(TH_m)(i, j)$ (διακριτοποίηση του δεύτερου μέλους της ΜΔΕ (5.8)), για κάθε $m = 1, \dots, N$.

Αυτή η επιλογή έχει το πλεονέκτημα ότι, από τη μία περιορίζει τις ασυνεχείς μεταβολές του διανύσματος της εικόνας, κατά τη διάρκεια της χρονικής εξέλιξης και από την άλλη επιταχύνει τη διαδικασία, διότι εξασφαλίζει

ότι σε ένα τουλάχιστον pixel της εικόνας σε κάθε επανάληψη, η διανυσματική μεταβολή έχει μέτρο β_{\max} .

5.1.9 Σημείωση σχετικά με την εκτίμηση των τιμών εκτός των ορίων της εικόνας

Από την παραπάνω περιγραφή των αλγορίθμων, μπορεί να διαφανεί ότι είναι απαραίτητη για τους υπολογισμούς η χρήση των τιμών της εικόνας και διαφόρων μεταβλητών που εμπλέκονται στη διαδικασία σε σημεία εκτός του ορθογωνικού πλαισίου της εικόνας. Πιο συγκεκριμένα, οι τιμές αυτές είναι απαραίτητες για τον υπολογισμό των εξομαλύνσεων σε σημεία κοντά στα όρια της εικόνας καθώς και των μερικών παραγώγων και της τοπικής συσχέτισης στα σημεία του ορίου (ακριανά pixels).

Η επιλογή που κάναμε στην υλοποίηση ήταν *ανάκλαση* για όλες τις παραμέτρους (\mathbf{u} και g ή a, c), εκτός από το στοιχείο b του ταυστή, για το οποίο χρησιμοποιήσαμε *αρνητική ανάκλαση*. Πιο συγκεκριμένα, σαν ενδεικτικό παράδειγμα αναφέρουμε ότι η ανάκλαση για τις τιμές της ίδιας της εικόνας που αφορούν τα pixels που περιβάλλουν το σύνορό της σημαίνει ότι επιλέγουμε:

$$\begin{aligned} \mathbf{u}(0, j) &= \mathbf{u}(1, j) & , & \mathbf{u}(N_x + 1, j) = \mathbf{u}(N_x, j) & , & j = 1, \dots, N_y \\ \mathbf{u}(i, 0) &= \mathbf{u}(i, 1) & , & \mathbf{u}(i, N_y + 1) = \mathbf{u}(i, N_y) & , & i = 1, \dots, N_x \end{aligned}$$

ενώ η αρνητική ανάκλαση του στοιχείου $b(i, j)$ για τα σημεία αυτά σημαίνει ότι:

$$\begin{aligned} b(0, j) &= -b(1, j) & , & b(N_x + 1, j) = -b(N_x, j) & , & j = 1, \dots, N_y \\ b(i, 0) &= -b(i, 1) & , & b(i, N_y + 1) = -b(i, N_y) & , & i = 1, \dots, N_x \end{aligned}$$

Αντίστοιχα προκύπτουν και οι τιμές που χρειαζόμαστε σε σημεία πιο απομακρυσμένα εκτός των ορίων της εικόνας (στα σημεία που βρίσκονται διαγώνια σε σχέση με το πλαίσιο θεωρούμε ότι γίνεται διπλή ανάκλαση).

Στη συνέχεια αιτιολογούμε τους λόγους των παραπάνω επιλογών. Εξετάζουμε το ζήτημα αυτό για τις διάφορες κατηγορίες ΜΔΕ που παρουσιάσαμε:

1. Πρώτα από όλα, στην περίπτωση των ΜΔΕ με τελεστή απόκλισης και βαθμωτό συντελεστή (ΜΔΕ μη γραμμικής ισοτροπικής διάχυσης (5.2)-(5.4)), οι συνεχείς αυτές εξισώσεις συμπληρώνονται από την αδιαβατική συνοριακή συνθήκη:

$$\frac{\partial u_m}{\partial n} = 0, \text{ στο } \partial\Omega, \quad m = 1, \dots, N \quad (5.22)$$

Μετά την διακριτοποίηση των ΜΔΕ, η συνθήκη για ανάκλαση της εικόνας σε κάθε βήμα αποτελεί την διακριτή εκδοχή των συνθηκών αυτών. Έτσι,

όπως η συνθήκη (5.22) συνεπάγεται σταθερή μέση τιμή κάθε καναλιού της συνεχούς εικόνας κατά τη χρονική εξέλιξη, το ίδιο συμβαίνει και με τη συνθήκη για ανάκλαση για την διακριτή εικόνα και στην διακριτή υλοποίηση της ΜΔΕ που περιγράψαμε. Αυτό μπορεί να αποδειχθεί ως εξής: Για οποιοδήποτε κανάλι της εικόνας $w = u_m$, το άθροισμα των τιμών της ταχύτητας $v_m(i, j)$ γράφεται, λόγω της σχέσης (5.10):

$$\begin{aligned} \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} v_m(i, j) &= \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} D_x(gD_x w) + D_y(gD_y w) \\ &= \sum_{j=1}^{N_y} \left\{ \sum_{i=1}^{N_x} D_x(gD_x w) \right\} + \sum_{i=1}^{N_x} \left\{ \sum_{j=1}^{N_y} D_y(gD_y w) \right\} \end{aligned}$$

Παρατηρώντας την παραπάνω σχέση, διαπιστώνουμε ότι τα αθροίσματα εντός των αγκυλών είναι αθροίσματα αλληλοεπικαλυπτόμενων διαφορών, με αποτέλεσμα να παραμένουν τελικά μόνο τα ακραία στοιχεία. Έτσι, η σχέση αυτή συνεχίζεται ως εξής:

$$\begin{aligned} \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} v_m(i, j) &= \sum_{j=1}^{N_y} \left\{ (gD_x w)\left(\frac{1}{2}, j\right) + (gD_x w)\left(N_x + \frac{1}{2}, j\right) \right\} + \\ &\quad + \sum_{i=1}^{N_x} \left\{ (gD_y w)\left(i, \frac{1}{2}\right) + (gD_y w)\left(i, N_y + \frac{1}{2}\right) \right\} \end{aligned} \quad (5.23)$$

Όμως για κάθε $j = 1, \dots, N_y$, ισχύει:

$$D_x w\left(\frac{1}{2}, j\right) = w(1, j) - w(0, j) = 0, \text{ λόγω της ανάκλασης για την } w$$

Αντίστοιχα προκύπτει ότι, λόγω της ανάκλασης της w :

$$\begin{aligned} D_x w\left(N_x + \frac{1}{2}, j\right) &= 0, \quad j = 1, \dots, N_y \\ D_y w\left(i, \frac{1}{2}\right) &= D_y w\left(i, N_y + \frac{1}{2}\right) = 0, \quad i = 1, \dots, N_x \end{aligned}$$

Έτσι, μέσω των παραπάνω σχέσεων, η (5.23) καταλήγει στην:

$$\sum_{i=1}^{N_x} \sum_{j=1}^{N_y} v_m(i, j) = 0$$

Επομένως, λόγω και της (5.9), προκύπτει τελικά:

$$\sum_{i=1}^{N_x} \sum_{j=1}^{N_y} u_m^{n+1}(i, j) = \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} u_m^n(i, j)$$

5.1 Υλοποίηση ΜΔΕ για την Βελτίωση διανυσματικών εικόνων 125

Επίσης, στην περίπτωση αυτή, απαιτείται και ο υπολογισμός του συντελεστή $g = g(\|\nabla \mathbf{u}_\sigma\|)$ εκτός των ορίων (για τη χρήση του στην τοπική συνέλιξη), για τον οποίον επίσης επιλέγουμε ανάκλαση. Κι αυτό διότι η ανάκλαση της εικόνας συνεπάγεται και ανάκλαση της νόρμας $\|\nabla \mathbf{u}_\sigma\|$, διότι το μέτρο των χωρικών παραγώγων δεν επηρεάζεται από την ανάκλαση.

2. Στην πιο περίπλοκη περίπτωση των ΜΔΕ με τελεστή απόκλισης και τανυστή (ΜΔΕ μη γραμμικής ανισοτροπικής διάχυσης (5.5,5.6)), η αντίστοιχη συνεχής συνοριακή συνθήκη είναι η εξής:

$$\langle D\nabla u_m, n \rangle = 0, \text{ στο } \partial\Omega, m = 1, \dots, N, \quad (5.24)$$

μέσω της οποίας εξασφαλίζεται και πάλι σταθερή μέση τιμή κάθε καναλιού. Αντίστοιχα με πριν, μπορεί να αποδειχθεί (μετά από αρκετές πράξεις) ότι στη διακριτή περίπτωση, η συνθήκη για ανάκλαση των τιμών της εικόνας σε συνδυασμό με αρνητική ανάκλαση του στοιχείου $b(i, j)$ του τανυστή D , συνεπάγεται σταθερή μέση τιμή κάθε καναλιού. Επομένως αυτές οι δύο συνθήκες μπορούν να αποτελέσουν την διακριτή εκδοχή της συνθήκης (5.24).

Επίσης, λόγω του τρόπου υπολογισμού του τανυστή D , μπορεί να διαπιστωθεί ότι η ανάκλαση της εικόνας συνεπάγεται ανάκλαση των στοιχείων $a(i, j), c(i, j)$ του τανυστή D (λόγω του ότι η συνέλιξη με συμμετρικούς πυρήνες K_σ και K_ρ διατηρεί την ιδιότητα της ανάκλασης και το μέτρο των χωρικών παραγώγων δεν επηρεάζεται από την ανάκλαση). Από αυτή την οπτική γωνία, μπορούμε να καταλήξουμε σε άλλο ένα επιχείρημα υπέρ της αρνητικής ανάκλασης του στοιχείου $b(i, j)$: Σε κάθε πλευρά, μία από τις δύο μερικές παραγώγους αλλάζει πρόσημο ενώ η άλλη ανακλάται κανονικά, με αποτέλεσμα την αρνητική ανάκλαση όλων των όρων της μορφής $u_{m,\sigma,x}u_{m,\sigma,y}$, το οποίο συνεπάγεται αρνητική ανάκλαση και του b .

3. Τέλος, στην περίπτωση των ΜΔΕ με τελεστή ίχνους (5.7,5.8), τα πράγματα είναι πιο απλά, μια και, όπως παρατηρούμε από τον πίνακα 5.3, ο πυρήνας συσχέτισης απαιτεί για τον υπολογισμό του τις τιμές του τανυστή μόνο στο ίδιο σημείο. Έτσι, οι τιμές που κυρίως απαιτούνται εκτός ορίων είναι της ίδιας της εικόνας. Λόγω της συζήτησης που προηγήθηκε, η επιλογή της ανάκλασης για την εικόνα \mathbf{u} φαντάζει φυσική και σε αυτήν την περίπτωση (βέβαια εδώ δεν μπορεί να αποδειχθεί η ιδιότητα της σταθερή μέση τιμής, παρόλο που πειραματικά διαπιστώσαμε μικρή μόνο διακύμανσή της).

5.2 Υλοποίηση ΜΔΕ για το Inpainting εικόνων

Στην ενότητα αυτή, συνεχίζουμε παρουσιάζοντας την αριθμητική υλοποίηση των ΜΔΕ για το πρόβλημα Inpainting βαθμωτών και κυρίως διανυσματικών εικόνων. Διάφορες μέθοδοι ΜΔΕ για το πρόβλημα αυτό παρουσιάστηκαν θεωρητικά στο κεφάλαιο 2. Στη συνέχεια, αναφέρουμε ξανά με συνοπτικό τρόπο εκείνες τις ΜΔΕ για inpainting που υλοποιήσαμε (χρησιμοποιούμε και εδώ τους συμβολισμούς της ενότητας 5.1 και επίσης συμβολίζουμε με \mathbb{D} την περιοχή inpainting):

1. Διανυσματική *TV PDE* με ϵ -εξομάλυνση (βλ. υποενότητα 2.3.1):

$$\frac{\partial u_m}{\partial t} = \operatorname{div} \left(\frac{1}{\|\nabla \mathbf{u}\|_\epsilon} \nabla u_m \right) , \quad \mathbf{x} \in \mathbb{D} \quad (m = 1, \dots, N), \quad (5.25)$$

$$\text{όπου: } \|\nabla \mathbf{u}\|_\epsilon := \sqrt{\epsilon^2 + \|\nabla \mathbf{u}\|^2}.$$

2. Σταθμισμένη Διανυσματική *TV PDE* με ϵ -εξομάλυνση, που προτείνουμε στην υποενότητα 2.3.1:

$$\frac{\partial u_m}{\partial t} = \|\nabla \mathbf{u}\| \operatorname{div} \left(\frac{1}{\|\nabla \mathbf{u}\|_\epsilon} \nabla u_m \right) , \quad \mathbf{x} \in \mathbb{D} \quad (m = 1, \dots, N) \quad (5.26)$$

3. Ανισοτροπική Εξομάλυνση των *Tschumperlé-Deriche* για *Inpainting* (βλ. ενότητα 2.7):

$$\frac{\partial u_m}{\partial t} = \operatorname{trace}(TH_m) , \quad \mathbf{x} \in \mathbb{D} \quad (m = 1, \dots, N) \quad (5.27)$$

όπου T τανυστής με ιδιοδιανύσματα $\boldsymbol{\theta}_+^*$, $\boldsymbol{\theta}_-^*$ και ιδιοτιμές:

$$\lambda_1 = 0 \quad , \quad \lambda_2 = \frac{1}{\sqrt{1 + \lambda_+^* + \lambda_-^*}}$$

έχοντας επιλέξει $\sigma = 0$ για τον δομικό τανυστή (5.1).

4. Διάχυση Οδηγούμενη από την Καμπυλότητα (*CDD*), για βαθμωτές εικόνες (βλ. ενότητα 2.4):

$$\frac{\partial u}{\partial t} = \operatorname{div} \left(\frac{|\kappa|^p}{\|\nabla u\|} \nabla u \right) , \quad \mathbf{x} \in \mathbb{D} \quad (5.28)$$

όπου:

$$\kappa = \operatorname{div} \left(\frac{\nabla u}{\|\nabla u\|} \right) , \quad p \geq 1$$

5. Διάχυση Οδηγούμενη από την Γενικευμένη Καμπυλότητα (*Gen.CDD*) για διανυσματικές εικόνες, η νέα μέθοδος που προτείναμε στην ενότητα 2.8.2:

$$\frac{\partial u_m}{\partial t} = \operatorname{div} \left(\frac{a |\hat{\kappa}|^p}{\|\nabla \mathbf{u}\|} \nabla u_m \right) \quad , \quad \mathbf{x} \in \mathbb{D} \quad (5.29)$$

όπου: $\hat{\kappa} = \operatorname{div}(\boldsymbol{\theta}_+)$ η γενικευμένη καμπυλότητα (που ορίσαμε στην ενότητα 2.8.1) και a, p θετικές σταθερές.

Όμοια με την προηγούμενη ενότητα, χρησιμοποιούμε τον εξής γενικό συμβολισμό για τις παραπάνω εξισώσεις:

$$\frac{\partial u_m}{\partial t} = v_m(x, y, t) \quad , \quad \mathbf{x} \in \mathbb{D} \quad (m = 1, \dots, N)$$

όπου με $v_m(x, y, t)$ συμβολίζουμε και πάλι το δεύτερο μέλος οποιασδήποτε από τις παραπάνω ΜΔΕ (5.25-5.29), το οποίο αποτελεί ουσιαστικά την ταχύτητα της ΜΔΕ σε κάθε χρονική στιγμή και για κάθε σημείο της περιοχής \mathbb{D} .

Οι παραπάνω ΜΔΕ παρουσιάζουν πολλά κοινά σημεία με τις ΜΔΕ για βελτίωση εικόνων, που είδαμε στην προηγούμενη ενότητα (σχέσεις (5.2-5.8)). Έτσι, τα πιο βασικά βήματα των αριθμητικών προσεγγίσεων που υλοποιήσαμε είναι όμοια με τα αντίστοιχα βήματα που παρουσιάσαμε αναλυτικά στην προηγούμενη ενότητα. Για τον λόγο αυτό, θα είμαστε πιο συνοπτικοί στη συνέχεια, αναφέροντας κυρίως τις διαφοροποιήσεις. Μόνη εξαίρεση αποτελεί η τελευταία μέθοδος *Gen.CDD* (5.29), της οποίας την υλοποίηση θα παρουσιάσουμε πιο αναλυτικά, αφού πρόκειται για καινούργια μέθοδο, που προτείνεται για πρώτη φορά στην εργασία αυτή.

Μόνη βασική διαφορά των παραπάνω ΜΔΕ για inpainting σε σχέση με τις ΜΔΕ για βελτίωση είναι ότι στις ΜΔΕ για inpainting, η εικόνα εξελίσσεται χρονικά μόνο στην περιοχή inpainting \mathbb{D} . Έτσι, στην αριθμητική υλοποίηση προσεγγίζουμε την περιοχή αυτή σαν ένα σύνολο από σημεία (pixels) του ορθογωνικού πλέγματος και μεταβάλλουμε τις τιμές της εικόνας μόνο στα σημεία αυτά. Το παραπάνω σημαίνει ότι, χρειαζόμαστε τον υπολογισμό της ταχύτητας v_m μόνο στο σύνολο \mathbb{D} . Όπως όμως παρουσιάσαμε στην υποενότητα 5.1, για να υπολογιστεί η v_m σε κάποιο σημείο απαιτείται προηγουμένως ο υπολογισμός των τιμών διαφόρων μεταβλητών στο σημείο αυτό, αλλά ενδεχομένως και στα γύρω σημεία. Έτσι, αν λάβουμε υπόψιν μας όλα τα σημεία της \mathbb{D} , προκύπτει, για κάθε μεταβλητή (έστω $\mu(i, j)$) που εμπλέκεται στους υπολογισμούς, το σύνολο των σημείων, στα οποία είναι απαραίτητος ο υπολογισμός της.

Για το σύνολο αυτό, χρησιμοποιούμε την ονομασία μάσκα *inpainting* της μεταβλητής μ και τον συμβολισμό \mathbb{D}_μ .

Στην υλοποίησή μας, υπολογίζουμε κάθε τέτοια μεταβλητή $\mu(i, j)$ μόνο στη μάσκα \mathbb{D}_μ και όχι σε όλη την εικόνα, ώστε να αποφύγουμε τις περιττές πράξεις.

Όσον αφορά τον υπολογισμό της μάσκας \mathbb{D}_μ , μπορεί να προκύψει με διαστολή (*dilation*) της περιοχής inpainting \mathbb{D} με ένα δομικό στοιχείο (*structuring element*), χαρακτηριστικό της μεταβλητής. Αυτό το δομικό στοιχείο φανερώνει το σύνολο των σημείων στα οποία απαιτείται η γνώση της μ για τον υπολογισμό της ταχύτητας $v_m(i, j)$ στο κεντρικό σημείο.

5.2.1 Υλοποίηση για Inpainting των ΜΔΕ Βελτίωσης

Στη συνέχεια, αναφέρουμε συνοπτικά τα βασικά χαρακτηριστικά της υλοποίησης εκείνων των ΜΔΕ που χρησιμοποιούνται τόσο στο πρόβλημα της Βελτίωσης όσο και στο Inpainting:

- Η Διανυσματική *TV PDE* με ϵ -εξομάλυνση για Inpainting (5.25) υλοποιείται όμοια με την αντίστοιχη ΜΔΕ για βελτίωση (βλ. (5.4)) και σύμφωνα με τον αλγόριθμο του πίνακα 5.5 και την επεξήγηση του πίνακα 5.6. Η μόνη τροποποίηση έγκειται στο ότι η ταχύτητα v_m και οι νέες τιμές της εικόνας υπολογίζονται μόνο στην περιοχή inpainting \mathbb{D} . Επίσης, όπως φαίνεται από τον πίνακα 5.2, ο πυρήνας συσχέτισης απαιτεί για τον υπολογισμό του τη γνώση των τιμών της g στα σημεία που φαίνονται από το δομικό στοιχείο:

$$B = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad (5.30)$$

Επομένως, η μάσκα inpainting του g είναι:

$$\mathbb{D}_g = \mathbb{D} \oplus B \quad (5.31)$$

Επίσης, επειδή $g = g(\|\nabla \mathbf{u}\|)$, ισχύει:

$$\mathbb{D}_{\|\nabla \mathbf{u}\|} = \mathbb{D}_g = \mathbb{D} \oplus B \quad (5.32)$$

- Η Σταθμισμένη Διανυσματική *TV PDE* με ϵ -εξομάλυνση (5.26) υλοποιείται όμοια με την Διανυσματική *TV PDE* με ϵ -εξομάλυνση, με την μόνη διαφορά ότι, στον αλγόριθμο μετά τον υπολογισμό της ταχύτητας, προστίθεται το βήμα πολλαπλασιασμού της με την $\|\nabla \mathbf{u}\|$ και ανανέωσης της τιμής της. Οι μάσκες inpainting $\mathbb{D}_g, \mathbb{D}_{\|\nabla \mathbf{u}\|}$ είναι ακριβώς ίδιες (αφού η νόρμα $\|\nabla \mathbf{u}\|$ που χρειάζεται για τον πολλαπλασιασμό που προστέθηκε, υπολογίζεται ούτως ή άλλως, για την εξαγωγή του $g(\|\nabla \mathbf{u}\|)$).

- Η *Ανιστροπική Εξομάλυνση των Tschumperlé-Deriche* για Inpainting υλοποιείται όπως η αντίστοιχη ΜΔΕ για βελτίωση εικόνων, θέτοντας όμως εδώ την ιδιοτιμή $f_+ = 0$. Όπως φαίνεται από τον πίνακα 5.3, ο ταυυστής T χρειάζεται να υπολογιστεί μόνο στην \mathbb{D} , άρα το ίδιο ισχύει και για τον δομικό ταυυστή $J_\rho(\nabla \mathbf{u})$, οπότε:

$$\mathbb{D}_T = \mathbb{D}_{J_\rho(\nabla \mathbf{u})} = \mathbb{D}$$

Αν λάβουμε όμως υπόψιν μας την εξομάλυνση του $J(\nabla \mathbf{u})$ με τον K_ρ που προηγείται, προκύπτει ότι για τον αρχικό ταυυστή $J(\nabla \mathbf{u})$, έχουμε:

$$\mathbb{D}_{J_\rho(\nabla \mathbf{u})} = \mathbb{D} \oplus S_{K_\rho}$$

όπου S_{K_ρ} το σύνολο σημείων στα οποία εκτείνεται η διακριτή προσέγγιση του πυρήνα K_ρ .

5.2.2 Vectorial DTV Filter

Όσον αφορά την Διανυσματική TV PDE (βλ. σχέση (5.25) και υποενότ. 2.3.1), οι Chan, Osher και Shen [13] πρότειναν μία εναλλακτική επαναληπτική διαδικασία η οποία συγκλίνει στην κατάσταση ισορροπίας της ΜΔΕ της Ολικής Μεταβολής (TV PDE), όπως αυτή μετατρέπεται μετά τη διακριτοποίηση που παρουσιάσαμε στην υποενότητα 5.1. Η μέθοδος αυτή αφορά τη γενική περίπτωση της Ολικής Μεταβολής, αλλά εδώ θα ασχοληθούμε ειδικά με την εφαρμογή της στο Inpainting εικόνων χωρίς θόρυβο. Όπως αναφέρεται στο [13], το φίλτρο αυτό θεμελιώνεται και ανεξάρτητα από το αντίστοιχο συνεχές μοντέλο, αφού ελαχιστοποιεί μία διακριτή εκδοχή της Ολικής Μεταβολής. Στην ενότητα 4.2, αποσαφηνίσαμε τη σχέση ανάμεσα στη διακριτοποιημένη συνεχή και στη τελείως διακριτή προσέγγιση, που σε αυτή την ειδική περίπτωση, είναι ισοδύναμες.

Πιο συγκεκριμένα, η διαδικασία που προτάθηκε από τους Chan, Osher και Shen [13], για την ειδική περίπτωση που αντιστοιχεί στην ΜΔΕ (5.25), και με τους συμβολισμούς που χρησιμοποιούμε στην εργασία αυτή, συνίσταται στην επανάληψη του εξής βήματος (μόνο για τα σημεία εντός της περιοχής inpainting \mathbb{D} , αφού τα υπόλοιπα παραμένουν σταθερά):

για $m = 1, \dots, N$ και $(i, j) \in \mathbb{D}$:

$$\begin{aligned}
 u_m^{n+1}(i, j) = \frac{1}{h_{\text{tot}}^n(i, j)} & \left\{ [g^n(i, j) + g^n(i-1, j)] \cdot u_m^n(i-1, j) + \right. \\
 & [g^n(i, j) + g^n(i+1, j)] \cdot u_m^n(i+1, j) + \\
 & \left. + [g^n(i, j) + g^n(i, j-1)] \cdot u_m^n(i, j-1) + [g^n(i, j) + g^n(i, j+1)] \cdot u_m^n(i, j+1) \right\}
 \end{aligned} \tag{5.33}$$

όπου $h_{\text{tot}}^n(i, j)$ το άθροισμα των βαρών, δηλαδή:

$$\begin{aligned}
 h_{\text{tot}}^n(i, j) = [g^n(i, j) + g^n(i-1, j)] & + [g^n(i, j) + g^n(i+1, j)] \\
 + [g^n(i, j) + g^n(i, j-1)] & + [g^n(i, j) + g^n(i, j+1)]
 \end{aligned}$$

επίσης, οι $g^n(i, j)$ είναι ουσιαστικά οι συντελεστές διάχυσης της TV PDE με ϵ -εξομάλυνση (5.25), δηλαδή:

$$g^n(i, j) = \frac{1}{\|\nabla \mathbf{u}^n\|_\epsilon(i, j)} = \frac{1}{\left\{ \epsilon^2 + \sum_{m=1}^N \|\nabla u_m^n\|^2(i, j) \right\}^{1/2}} \tag{5.34}$$

για μία μικρή τιμή της σταθεράς ϵ . Τέλος για τις νόρμες $\|\nabla u_m\|$, χρησιμοποιείται στην (5.34) η προσέγγιση $D^\pm u_m(i, j)$ (βλ. υποενότητα 5.1.3), δηλαδή:

$$\|\nabla u_m\|^2(i, j) = (D^\pm u_m(i, j))^2 = \frac{1}{2} \left\{ (D_x^+ u_m)^2 + (D_x^- u_m)^2 + (D_y^+ u_m)^2 + (D_y^- u_m)^2 \right\} \tag{5.35}$$

Όπως μπορεί εύκολα να διαπιστωθεί, η κάθε επανάληψη της παραπάνω διαδικασίας, που εκφράζεται από τη σχέση (5.33), μπορεί και αυτή να υλοποιηθεί σαν τοπική συσχέτιση με κατάλληλο πυρήνα $k_{i,j}(\ell_1, \ell_2)$, ο οποίος δίνεται από τον πίνακα 5.9, με τους συντελεστές $g(i, j)$ να δίνονται από την (5.34). Βέβαια, στην περίπτωση αυτή, το αποτέλεσμα της συσχέτισης δίνει κατευθείαν την καινούρια εικόνα u_m^{n+1} και όχι την διακριτοποιημένη ταχύτητα. Η επαναληπτική διαδικασία είναι εξαρχής σχεδιασμένη για την διακριτή εικόνα και έτσι δεν υπάρχει, όπως στις υλοποιήσεις που παρουσιάσαμε μέχρι τώρα, χρονικό βήμα ούτε ταχύτητα. Στο πρόβλημα Inpainting που μας ενδιαφέρει εδώ, η παραπάνω εξέλιξη γίνεται μόνο στην περιοχή \mathbb{D} και οι μάσκες inpainting που χρειάζονται είναι ακριβώς ίδιες με την μέθοδο VTV Inpainting (βλ. σχέσεις (5.31), (5.32)). Ο αλγόριθμος που υλοποιήσαμε για το Vectorial DTV Filter φαίνεται στον πίνακα 5.10.

$$\begin{array}{c}
\ell_2 \\
-1 \qquad 0 \qquad 1 \\
\begin{array}{ccc}
-1 & 0 & \frac{g(i-1,j)+g(i,j)}{h_{tot}(i,j)} & 0 \\
\ell_1 \ 0 & \frac{g(i,j-1)+g(i,j)}{h_{tot}(i,j)} & 0 & \frac{g(i,j+1)+g(i,j)}{h_{tot}(i,j)} \\
1 & 0 & \frac{g(i+1,j)+g(i,j)}{h_{tot}(i,j)} & 0
\end{array} \\
\text{όπου } h_{tot}(i,j)=g(i-1,j)+g(i+1,j)+4g(i,j)+g(i,j-1)+g(i,j+1)
\end{array}$$

Πίνακας 5.9: Πυρήνας τοπικής συσχέτισης $k_{i,j}(\ell_1, \ell_2)$ για την μέθοδο Vectorial DTV Filter

Ο αλγόριθμος αυτός παρουσιάζει το πλεονέκτημα σε σχέση με την αριθμητική υλοποίηση της Διανυσματικής TV PDE αλλά και της Σταθμισμένης Διανυσματικής TV PDE (βλ. υποενότητα 5.2.1) ότι συγκλίνει πολύ πιο γρήγορα στην κατάσταση ισορροπίας, η οποία όπως είπαμε είναι ίδια με την κατάσταση ισορροπίας των ΜΔΕ αυτών. Η ακολουθία των εικόνων είναι βέβαια διαφορετική, αλλά όπως έχουμε αναφέρει, στο πρόβλημα Inpainting (σε αντίθεση με το πρόβλημα βελτίωσης), το αποτέλεσμα προκύπτει από την κατάσταση ισορροπίας. Επίσης, η απόδοση της μεθόδου δεν εξαρτάται από κάποιο χρονικό βήμα και έτσι δεν χρειάζεται ειδική μέριμνα για την επιλογή μιας τέτοιας παραμέτρου. Τέλος, ο αλγόριθμος αυτός παρουσιάζει την πολύ σημαντική ιδιότητα ότι είναι εγγυημένη η κατάληξη στην κατάσταση ισορροπίας χωρίς την παραμικρή ταλάντωση (λόγω της αποδεδειγμένης σύγκλισης). Αυτή η ιδιότητα ήταν ο λόγος που χρησιμοποιήσαμε την συνθήκη τερματισμού στον αλγόριθμο του πίνακα 5.10. Είναι σίγουρο, δηλαδή ότι, μετά από κατάλληλο αριθμό επαναλήψεων, η μεταβολή $\mathbf{u}^{new} - \mathbf{u}^{old}$ σε κάθε βήμα θα είναι τόσο μικρή, ώστε η συνθήκη τερματισμού του αλγορίθμου να ικανοποιείται (για οσοδήποτε μικρό κατώφλι ε_{th}). Βέβαια, χρησιμοποιούμε και τον εναλλακτικό τερματισμό, όταν φτάσουμε στον μέγιστο αριθμό επαναλήψεων N_{maxit} , για την περίπτωση όπου απαιτούνται υπερβολικά πολλές επαναλήψεις.

Παρατηρήσεις για την επαναληπτική διαδικασία του Vectorial DTV Filter

Αξίζει να σημειώσουμε μία κάπως διαφορετική συλλογιστική από αυτή του [13], με την οποία μπορούμε και πάλι να καταλήξουμε στην επαναληπτική σχέση (5.33): Αν διακριτοποιήσουμε την κατάσταση ισορροπίας της Διανυσματικής TV PDE (5.26) ή ισοδύναμα αν ελαχιστοποιήσουμε την διακριτή Διανυσματική

ΕΙΣΟΔΟΙ: αρχική εικόνα \mathbf{u}^0 (διανυσματική με N κανάλια), περιοχή inpainting \mathbb{D} , μέγιστος αριθμός επαναλήψεων N_{maxit} , κατώφλι τερματισμού ε_{th} , σταθερά ϵ .

Υπολογισμός μάσκας inpainting

$$\mathbb{D}_g = \mathbb{D} \oplus \begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

Αρχικοποίηση

για $(i, j) \in \Omega \setminus \mathbb{D}$, $\mathbf{u}^{old} = \mathbf{u}^0$

για $(i, j) \in \mathbb{D}$, κάνε μία απλή αρχική πρόβλεψη στην \mathbf{u}^{old} , γεμίζοντας την \mathbb{D} με ομοιόμορφο θόρυβο ή με μία σταθερή τιμή.

For $n = 1, \dots, N_{maxit}$

Υπολόγισε τη νόρμα $\|\nabla \mathbf{u}\| = \left\{ \sum_{m=1}^N \|\nabla u_m\|^2 \right\}^{\frac{1}{2}}$, για $(i, j) \in \mathbb{D}_g$, σύμφωνα

με την σχέση (5.35).

Υπολογισμός του συντελεστή διάχυσης g

$$g = 1/\sqrt{\epsilon^2 + \|\nabla \mathbf{u}\|^2}, \text{ για } (i, j) \in \mathbb{D}_g$$

Υπολόγισε τους τοπικούς πυρήνες συσχέτισης $k_{i,j}$ από τις τιμές του g , σύμφωνα με τον πίνακα 5.9.

Υπολογισμός όλων των καναλιών u_m^{new} της επόμενης εικόνας \mathbf{u}^{new}

For $m = 1, \dots, N$

Σε κάθε $(i, j) \in \mathbb{D}$, βρες την $u_m^{new}(i, j)$, εφαρμόζοντας τοπική συσχέτιση της $k_{i,j}$ με την u_m^{old} .

End for

Συνθήκη τερματισμού

If $\max_{(i,j) \in \mathbb{D}} \|\mathbf{u}^{new}(i, j) - \mathbf{u}^{old}(i, j)\| < \varepsilon_{th}$

$$\mathbf{u}^{out} = \mathbf{u}^{new}$$

Τερμάτισε τον αλγόριθμο

End if

Προετοιμασία επόμενης επανάληψης

$$\mathbf{u}^{old} = \mathbf{u}^{new}$$

End for

$$\mathbf{u}^{out} = \mathbf{u}^{new}$$

ΕΞΟΔΟΣ: ανακατασκευασμένη εικόνα \mathbf{u}^{out}

Πίνακας 5.10: Αλγόριθμος που υλοποιήσαμε για το Vectorial Digital TV Filter

Ολική Μεταβολή², καταλήγουμε στις εξής συνθήκες για την τοπική συσχέτιση:

$$(k_{i,j}(\mathbf{u}) \tilde{*} u_m)(i, j) = 0, \quad m = 1, \dots, N, \quad (5.36)$$

για κάθε σημείο (i, j) του πλέγματος. Στη σχέση αυτή, με “ $\tilde{*}$ ” συμβολίζουμε την συσχέτιση και $k_{i,j}(\mathbf{u})$ είναι ο πυρήνας που χρησιμοποιούμε στην υλοποίηση της VTV PDE. Ο πυρήνας αυτός προκύπτει από τον πίνακα 5.2, με τους συντελεστές διάχυσης $g(i, j)$ να δίνονται από τη σχέση:

$$g(i, j) = \frac{1}{\|\nabla \mathbf{u}\|_\epsilon(i, j)} \quad (5.37)$$

Επίσης, στη διακριτοποίηση αυτή, θεωρούμε ότι, για τον υπολογισμό της $\|\nabla \mathbf{u}\|_\epsilon(i, j)$, επιλέγουμε την προσέγγιση D^\pm (βλ. σχέσεις (5.34),(5.35)).

Επομένως, οι εξισώσεις που περιγράφονται από τη σχέση (5.36), αποτελούν ένα σύστημα μη γραμμικών αλγεβρικών εξισώσεων. Αν επιλύσουμε σε κάθε σημείο (i, j) και κάθε κανάλι m ως προς $u_m(i, j)$, οι συνθήκες (5.36) μετατρέπονται στις:

$$u_m(i, j) = (\hat{k}_{i,j}(\mathbf{u}) \tilde{*} u_m)(i, j), \quad m = 1, \dots, N, \quad (5.38)$$

όπου $\hat{k}_{i,j}(\mathbf{u})$ είναι τώρα ο πυρήνας που προκύπτει από τον πίνακα 5.9, με τους συντελεστές $g(i, j)$ να δίνονται και πάλι από την (5.37).

Επομένως, ο αλγόριθμος του πίνακα 5.10 αποτελεί ουσιαστικά επαναληπτική επίλυση του συστήματος μη γραμμικών εξισώσεων (5.38) (ή των ισοδύναμων εξισώσεων (5.36)), που προκύπτει αν στο αριστερό μέλος θέσουμε $u_m = u_m^{n+1}$ και στο δεξιό $u_m = u_m^n$, $\mathbf{u} = \mathbf{u}^n$ (το οποίο εκφράζει την γενική επαναληπτική μέθοδο της Αριθμητικής Ανάλυσης, βλ. π.χ. [48]), δηλαδή:

$$u_m^{n+1}(i, j) = (\hat{k}_{i,j}(\mathbf{u}^n) \tilde{*} u_m^n)(i, j), \quad m = 1, \dots, N \quad (5.39)$$

Έτσι, εφόσον η παραπάνω ακολουθία \mathbf{u}^n του αλγορίθμου του πίνακα 5.10 συγκλίνει, το όριό της είναι λύση της (5.36), δηλαδή της διακριτοποιημένης συνθήκης ισορροπίας της Διανυσματικής TV PDE.

Από αυτή την οπτική γωνία, η υλοποίηση της Διανυσματικής TV PDE (βλ. υποενότ. 5.2.1) μπορεί να ειπωθεί σαν μία άλλη εναλλακτική αριθμητική επίλυση των μη γραμμικών εξισώσεων (5.36), που προκύπτει ως εξής: Στην (5.36), πολλαπλασιάζουμε με μία θετική σταθερά τ ³ και προσθέτουμε $u_m(i, j)$ και στα δύο μέλη, οπότε προκύπτει το ισοδύναμο σύστημα:

$$u_m(i, j) = u_m(i, j) + \tau \cdot (k_{i,j}(\mathbf{u}) \tilde{*} u_m)(i, j), \quad m = 1, \dots, N, \quad (5.40)$$

²όπως αναφέραμε, για την ισοδυναμία αυτή βλ. ενότητα 4.2

³το τ αντιστοιχεί στο χρονικό βήμα της προσέγγισης της VTV PDE, αλλά, από την οπτική γωνία που αναφέρουμε εδώ, είναι απλά μια σταθερά. Βέβαια, ισχύουν και πάλι αντίστοιχοι περιορισμοί στην τιμή του, που προκύπτουν από τη συνθήκη σύγκλισης της γενικής επαναληπτικής μεθόδου (βλ. π.χ. [48])

Αν ακολουθήσουμε την ίδια με προηγούμενως γενική μέθοδο επίλυσης, θέτωντας $u_m = u_m^{n+1}$ στο αριστερό μέλος και $u_m = u_m^n$, $\mathbf{u} = \mathbf{u}^n$ στο δεξιό μέλος, καταλήγουμε στην επαναληπτική διαδικασία που προσεγγίζει την Διανυσματική TV PDE.

Από όσα παρατηρήσαμε παραπάνω, συμπεραίνουμε ότι μπορούμε να δημιουργήσουμε φίλτρα, αντίστοιχα με το DTV Filter, για την εύρεση της κατάστασης ισορροπίας και άλλων ΜΔΕ οι οποίες υλοποιούνται με τοπική συσχέτιση. Το κίνητρο είναι η ελπίδα αντίστοιχης επιτάχυνσης της σύγκλισης. Όπως θα παρουσιάσουμε στην υποενότητα 5.2.4, υλοποιούμε την μέθοδο *Γενικευμένης CDD διανυσματικών εικόνων* που προτείνουμε με ένα τέτοιο φίλτρο και όπως διαπιστώσαμε πειραματικά, επιτυγχάνεται όντως επιτάχυνση της σύγκλισης.

5.2.3 Υλοποίηση Διάχυσης Οδηγούμενης από την Καμπυλότητα (CDD), για βαθμωτές εικόνες

Η μέθοδος CDD είναι ειδικά σχεδιασμένη για το πρόβλημα του Inpainting (βλ. ενότ. 2.4), για αυτό δεν είχαμε συναντήσει την υλοποίησή της προηγούμενως, στην υποενότητα 5.1. Ανήκει όμως και αυτή στην κατηγορία που είχαμε αναφέρει, των ΜΔΕ με τελεστή απόκλισης div και με βαθμωτό συντελεστή (μη γραμμική ισοτροπική διάχυση), στην ειδική περίπτωση βαθμωτών εικόνων. Ο συντελεστής διάχυσης g δίνεται εδώ από τη σχέση:

$$g = \frac{f(|\kappa|)}{\|\nabla u\|} \quad , \quad \kappa = \text{div} \left(\frac{\nabla u}{\|\nabla u\|} \right) \quad (5.41)$$

όπου κ είναι η καμπυλότητα των ισοϋψών γραμμών της u . Παρατηρούμε ότι ο όρος $\|\nabla u\|$ εμφανίζεται και εδώ στον παρονομαστή, και μάλιστα δύο φορές. Έτσι, στην υλοποίηση, για να μη μηδενίζεται ο παρονομαστής και έχουμε αστάθεια στην αριθμητική επαναληπτική διαδικασία, χρησιμοποιούμε και στην περίπτωση αυτή ϵ -εξομάλυνση στους δύο αυτούς όρους (βλ. π.χ. συζήτηση στην ενότ. 1.6). Τροποποιούμε λοιπόν ελαφρώς τον συντελεστή διάχυσης g ως εξής:

$$g = \frac{f(|\kappa_\epsilon|)}{\|\nabla u\|_\epsilon} \quad , \quad \kappa_\epsilon = \text{div} \left(\frac{\nabla u}{\|\nabla u\|_\epsilon} \right) \quad (5.42)$$

όπου η εξομαλυμένη νόρμα $\|\nabla u\|_\epsilon$, ορίζεται, όπως έχουμε πει, ως εξής:

$$\|\nabla u\|_\epsilon = \sqrt{\epsilon^2 + \|\nabla u\|^2}$$

Επομένως ο αλγόριθμος της υλοποίησης της μεθόδου αυτής είναι παρόμοιος με αυτόν που δίνεται στον πίνακα 5.2, με τους συντελεστές $g(i, j)$ να προκύπτουν από την σχέση (5.42), προσεγγίζοντας την $\|\nabla u\|$ όπως αναφέραμε στην

ενότητα 5.1.3. Η διαφορά έγκειται στο ότι πρέπει να προστεθεί στον αλγόριθμο αυτό ένα βήμα (πριν τον υπολογισμό της τοπικής συσχέτισης για την εξαγωγή της ταχύτητας), για τον υπολογισμό της καμπυλότητας κ και της συνάρτησης $f(|\kappa|)$. Όπως έχουμε αναφέρει (βλ. ενότ. 1.4), η καμπυλότητα αυτή είναι ακριβώς η ταχύτητα της ΜΔΕ της Ολικής Μεταβολής (TV PDE), επομένως μπορεί να βρεθεί με την εφαρμογή άλλης μίας τοπικής συσχέτισης με πυρήνα όπως του πίνακα 5.2, θέτοντας όμως $g = 1/\|\nabla u\|$ αυτή τη φορά. Έτσι, σε κάθε επανάληψη, εφαρμόζονται δύο τοπικές συνελιξίες.

Με βάση όσα είπαμε στην υποενότητα 5.2 και ακολουθώντας ανάλυση αντίστοιχη με της υποενότητας 5.2.1, καταλήγουμε ότι στον παραπάνω αλγόριθμο, οι μεταβλητές υπολογίζονται μόνο στα σημεία των αντίστοιχων μασκών inpainting, όπως αυτές δίνονται από τις σχέσεις:

$$\begin{aligned}\mathbb{D}_\kappa &= \mathbb{D}_g = \mathbb{D} \oplus B \\ \mathbb{D}_{\|\nabla u\|} &= \mathbb{D}_g \cup \mathbb{D}_\kappa \oplus B = \mathbb{D}_\kappa \oplus B\end{aligned}$$

όπου B είναι το δομικό στοιχείο:

$$B = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

5.2.4 Υλοποίηση της νέας μεθόδου που προτείνουμε: Γενικευμένη CDD, για διανυσματικές εικόνες

Στην υποενότητα 2.8.2, προτείνουμε μία νέα μέθοδο Inpainting για διανυσματικές εικόνες, την Γενικευμένη CDD, η οποία περιγράφεται από την ακόλουθη ΜΔΕ:

$$\frac{\partial u_i}{\partial t} = \operatorname{div} \left(\frac{h(|\hat{\kappa}|)}{\|\nabla \mathbf{u}\|} \nabla u_i \right), \quad \mathbf{x} \in \mathbb{D} \quad (i = 1, \dots, N) \quad (5.43)$$

όπου $\hat{\kappa}$ είναι η Γενικευμένη Καμπυλότητα, που ορίσαμε στην υποενότητα 2.8.1 σαν την καμπυλότητα των Καμπυλών Ελάχιστης Διανυσματικής Μεταβολής (δηλ. των δυναμικών γραμμών του μοναδιαίου διανύσματος $\boldsymbol{\theta}_-$ που εκφράζει την διεύθυνση της ελάχιστης μεταβολής της εικόνας \mathbf{u}). Όπως αναφέραμε στην υποενότητα αυτή, η $\hat{\kappa}$ δίνεται από τις εξής ισοδύναμες σχέσεις:

$$\hat{\kappa} = \operatorname{div}(\boldsymbol{\theta}_+) \quad (5.44)$$

και

$$\hat{\kappa} = -\sin \theta \frac{\partial \theta}{\partial x} + \cos \theta \frac{\partial \theta}{\partial y} \quad (5.45)$$

όπου θ_+ το μοναδιαίο ιδιοδιάνυσμα της μήτρας Di Zenzo G της \mathbf{u} , που αντιστοιχεί στη μεγαλύτερη ιδιοτιμή λ_+ και θ η γωνία του διανύσματος αυτού.

Η ΜΔΕ (5.43) ανήκει και αυτή στην κατηγορία των ΜΔΕ με τελεστή απόκλισης div και με βαθμωτό συντελεστή (μη γραμμική ιστροπική διάχυση) (βλ. εισαγωγή της υποενότ. 5.1). Στη συγκεκριμένη περίπτωση, ο συντελεστής διάχυσης g δίνεται από την σχέση:

$$g = \frac{h(|\hat{\kappa}|)}{\|\nabla \mathbf{u}\|} \quad (5.46)$$

Έτσι, βασικό σημείο της υλοποίησης, που την διαφοροποιεί από τις προηγούμενες μεθόδους, είναι ο υπολογισμός της γενικευμένης καμπυλότητας $\hat{\kappa}$. Στη συνέχεια λοιπόν, θα παρουσιάσουμε καταρχήν τον υπολογισμό της $\hat{\kappa}$:

Για τον υπολογισμό της $\hat{\kappa}$, χρησιμοποιούμε την έκφραση (5.45), οπότε χρειαζόμαστε τις τιμές της γωνίας θ . Όπως εύκολα μπορούμε να διαπιστώσουμε, η γωνία αυτή προκύπτει από τη μήτρα Di Zenzo G της \mathbf{u} . Όμως, για τον υπολογισμό αυτό, αντικαθιστούμε την \mathbf{u} με μία ελαφρώς εξομαλυμένη εκδοχή της, την \mathbf{u}_σ , σε αντιστοιχία με άλλες μεθόδους (βλ. π.χ. υποενότ. 1.3,1.7). Ο σκοπός της εξομάλυνσης αυτής είναι να αποφευχθούν οι υπερβολικά απότομες μεταβολές της \mathbf{u} και έτσι η προσέγγιση της θ να γίνει πιο αξιόπιστη. Επομένως, αντικαθιστούμε την μήτρα Di Zenzo G από τον δομικό τανυστή $J(\nabla \mathbf{u}_\sigma)$ (για τον ορισμό του, βλ. π.χ. εισαγωγή της υποενότ. 5.1).

Έτσι, το αρχικό στάδιο κάθε επανάληψης του αλγορίθμου είναι όμοιο με το αντίστοιχο του αλγορίθμου του πίνακα 5.7 (για $\rho = 0$). Έχοντας υπολογίσει τη γωνία θ , για να υπολογίσουμε την $\hat{\kappa}$ από την (5.45), απομένει η προσέγγιση των χωρικών παραγώγων της θ , για την οποία χρησιμοποιούμε κεντρικές διαφορές βήματος 2 pixels, δηλαδή:

$$\begin{aligned} \frac{\partial \theta}{\partial x}(i, j) &\approx \frac{1}{2} \delta_{2x} \theta(i, j) \\ \frac{\partial \theta}{\partial y}(i, j) &\approx \frac{1}{2} \delta_{2y} \theta(i, j) \end{aligned} \quad (5.47)$$

όπου με δ_{2x}, δ_{2y} συμβολίζουμε τις διαφορές:

$$\begin{aligned} \delta_{2x} \theta(i, j) &= \theta(i+1, j) - \theta(i-1, j) \\ \delta_{2y} \theta(i, j) &= \theta(i, j+1) - \theta(i, j-1) \end{aligned} \quad (5.48)$$

Όμως, όπως αναφέραμε στην υποενότητα 2.8.1, η συνάρτηση θ είναι πλειότιμη modulo π , δηλαδή σε κάθε σημείο αντιστοιχίζονται όλες οι τιμές $\theta_0 + \nu\pi$ με $\nu \in \mathbb{Z}$, για κάποιο βασικό $\theta_0 \in [0, \pi)$. Για τον υπολογισμό των διαφορών

(5.48), έχει σημασία ποια από αυτές τις τιμές της θ επιλέγεται τελικά σε κάθε σημείο. Στην υποενότητα 2.8.1, αναφέραμε επίσης ότι, όσον αφορά τη συνεχή περίπτωση, για να ορίζεται η $\hat{\kappa}$, πρέπει να γίνεται τέτοια επιλογή στην πλειοτιμη συνάρτηση $\theta(\mathbf{x})$, ώστε αυτή να είναι συνεχής στο εκάστοτε σημείο. Η διακριτή εκδοχή της παραπάνω διαδικασίας είναι να περιορίσουμε τις διαφορές δ_{2x}, δ_{2y} στο διάστημα $(-\frac{\pi}{2}, \frac{\pi}{2}]$, λόγω του ότι η συνάρτηση $\theta(i, j)$ μπορεί να περιγραφεί πλήρως μέσω ενός διαστήματος τιμών μήκους π (στην υλοποίηση έχουμε διαλέξει το διάστημα $[0, \pi)$). Με τον τρόπο αυτό, πετυχαίνουμε να μην παρουσιάζονται υψηλές τιμές χωρίς νόημα για τις διαφορές δ_{2x}, δ_{2y} . Επομένως, στην υλοποίηση δεν χρησιμοποιούμε τελικά τις σχέσεις (5.48), αλλά την εξής τροποποίησή τους:

$$\begin{aligned} \delta_{2x}\theta(i, j) &= \theta(i+1, j) - \theta(i-1, j) + \nu(\theta(i+1, j) - \theta(i-1, j)) \cdot \pi \\ \delta_{2y}\theta(i, j) &= \theta(i, j+1) - \theta(i, j-1) + \nu(\theta(i, j+1) - \theta(i, j-1)) \cdot \pi \end{aligned} \quad (5.49)$$

όπου ο ακέραιος $\nu(\delta\theta)$ είναι τέτοιος, ώστε το τελικό αποτέλεσμα να βρεθεί στο επιθυμητό διάστημα $(-\frac{\pi}{2}, \frac{\pi}{2}]$, δηλαδή:

$$\nu(\delta\theta) = \begin{cases} 1, & \delta\theta \in (-\pi, -\frac{\pi}{2}] \\ 0, & \delta\theta \in (-\frac{\pi}{2}, \frac{\pi}{2}] \\ -1, & \delta\theta \in (\frac{\pi}{2}, \pi) \end{cases} \quad (5.50)$$

Σημειώνουμε ότι οι παραπάνω περιπτώσεις είναι οι μόνες δυνατές, αφού στον αριθμητικό υπολογισμό της $\theta(i, j)$ αναθέτουμε τέτοια τιμή, ώστε: $\theta(i, j) \in [0, \pi)$, για όλα τα σημεία (i, j) .

Επομένως, από τις (5.45), (5.47), καταλήγουμε στην εξής αριθμητική προσέγγιση της γενικευμένης καμπυλότητας $\hat{\kappa}$:

$$\hat{\kappa} = \frac{1}{2} \{-\sin \theta \cdot \delta_{2x}\theta + \cos \theta \cdot \delta_{2y}\theta\}(i, j) \quad (5.51)$$

με τις $\delta_{2x}\theta, \delta_{2y}\theta$ να προσδιορίζονται από τις (5.49) και (5.50).

Έχοντας υπολογίσει την γενικευμένη καμπυλότητα $\hat{\kappa}$, και υπολογίζοντας τη νόρμα $\|\nabla \mathbf{u}\|$, ο συντελεστής διάχυσης υπολογίζεται άμεσα από τη σχέση (5.46). Πρέπει να σημειώσουμε ότι υπολογίζουμε τη νόρμα $\|\nabla \mathbf{u}\|$ κατευθείαν από την αρχική εικόνα \mathbf{u} (όχι από την εξομαλυσμένη \mathbf{u}_σ), χρησιμοποιώντας όμως ϵ -εξομάλυνση, ώστε στο σημείο αυτό να παρουσιάζεται αντιστοιχία με τη μέθοδο Διανυσματικής TV (βλ. υποενότ. 1.6 και 5.2.1). Όσον αφορά τη συνάρτηση $h(|\hat{\kappa}|)$, όπως είπαμε και στην υποενότητα 2.8.2, κάνουμε την εξής επιλογή (αντίστοιχη με την βαθμωτή CDD - βλ. ενότ. (2.4)):

$$h(|\hat{\kappa}|) = a |\hat{\kappa}|^p, \quad (5.52)$$

για διάφορες θετικές τιμές της σταθεράς p . Η σταθερά a επιλέγεται έτσι, ώστε το εύρος των τιμών που μπορεί να πάρει η $h(|\hat{\kappa}|)$ να είναι σταθερό και να μην επηρεάζεται από την επιλογή του εκθέτη p (όπως θα δούμε παρακάτω, αυτό μπορεί να επιτευχθεί μόνο στην υλοποίηση, λόγω της διακριτής προσέγγισης). Ελέγχοντας το εύρος αυτό, μπορούμε αξιόπιστα να συγκρίνουμε τα αποτελέσματα της μεθόδου για τις διάφορες επιλογές του p . Απαιτώντας η μέγιστη δυνατή τιμή της $h(|\hat{\kappa}|)$ να είναι 1 (και έτσι να προκύπτει το εύρος τιμών $[0, 1]$), η σχέση (5.52) γίνεται:

$$h(|\hat{\kappa}|) = \left(\frac{|\hat{\kappa}|}{|\hat{\kappa}|_{\max}} \right)^p \quad (5.53)$$

όπου $|\hat{\kappa}|_{\max}$ η μέγιστη δυνατή τιμή της υπολογιζόμενης $|k|$ (βέβαια, στο συνεχές μοντέλο, δεν υπάρχει φράγμα για την k , άρα ούτε για την $h(|\hat{\kappa}|)$). Επειδή χρησιμοποιούμε την προσέγγιση (5.51), η επίτευξη μέγιστης τιμής απαιτεί καταρχήν μέγιστη τιμή για τα $|\delta_{2x}\theta|$, $|\delta_{2y}\theta|$, η οποία είναι $\pi/2$, αφού όπως είδαμε, οι διαφορές της γωνίας θ κειμένονονται στο διάστημα $(-\frac{\pi}{2}, \frac{\pi}{2})$. Έτσι, στην περίπτωση αυτή, έχουμε:

$$|\hat{\kappa}| = \frac{\pi}{4} \{ \sin \theta \pm \cos \theta \}, \theta \in [0, \pi] \quad (5.54)$$

η οποία, όπως εύκολα μπορούμε να διαπιστώσουμε, μεγιστοποιείται για $\theta = \frac{\pi}{4}$ ή $\frac{3\pi}{4}$, με τιμή:

$$|\hat{\kappa}|_{\max} = \frac{\pi}{2\sqrt{2}} \approx 1.11 \quad (5.55)$$

Σημειώνουμε ότι το παραπάνω εκφράζει μία ακραία περίπτωση, η οποία όμως μπορεί όντως να συμβεί, όπως για παράδειγμα όταν η θ παίρνει στην γειτονιά του σημείου (i, j) τις τιμές που φαίνονται στον πίνακα:

$$\begin{bmatrix} 0 & & \\ 0 & \frac{3\pi}{4} & \frac{\pi}{2} \end{bmatrix}$$

Για τους παραπάνω λόγους, η συνάρτηση $h(|\hat{\kappa}|)$ που χρησιμοποιήσαμε είναι η εξής:

$$h(|\hat{\kappa}|) = \left(\frac{|\hat{\kappa}|}{1.11} \right)^p \quad (5.56)$$

Εκτός από τη διαδικασία υπολογισμού του συντελεστή διάχυσης $g(i, j)$ που περιγράψαμε πιο πάνω, τα υπόλοιπα βήματα του αλγορίθμου που υλοποιήσαμε είναι όμοια με τον αλγόριθμο της υλοποίησης της VTV PDE (βλ. πίν. 5.5 και επεξηγηματικό πίν. 5.6). Όμως, και στην μέθοδο αυτή, όπως και στις προηγούμενες μεθόδους για Inpainting εικόνων, το επιθυμητό αποτέλεσμα προκύπτει

όταν η ΜΔΕ (5.43) προσεγγίζει την κατάσταση ισορροπίας. Έτσι, εμπνευσμένοι από το Vectorial DTV Filter (βλ. υποενότ. 5.2.2), υλοποιήσαμε και έναν δεύτερο αλγόριθμο, αντίστοιχο με τον αλγόριθμο του πίνακα 5.10, ο οποίος προσεγγίζει κατευθείαν αυτή την κατάσταση ισορροπίας (βλ. συζήτηση στην υποενότ. 5.2.2). Συγκρίνοντας πειραματικά αυτές τις δύο εναλλακτικές υλοποιήσεις, διαπιστώσαμε ότι η δεύτερη προσεγγίζει πιο γρήγορα την ισορροπία και άρα είναι προτιμητέα. Στην συνέχεια λοιπόν, παρουσιάζουμε την υλοποίηση αυτή.

Πρέπει όμως να σημειώσουμε ότι στην αναδρομική διαδικασία της δεύτερης αυτής υλοποίησης, εμφανίζεται το άθροισμα όρων $g(i, j)$ στο παρονομαστή. Έτσι, λόγω της (5.46), για να μη μπορεί ποτέ να μηδενιστεί ούτε αυτός ο παρονομαστής, πρέπει να μην μηδενίζεται ποτέ ο όρος $h(|\hat{\kappa}|)$. Επομένως, σε αντιστοιχία με την ϵ -εξομάλυνση της νόρμας $\|\nabla \mathbf{u}\|$, τροποποιούμε ελαφρώς την συνάρτηση $h(|\hat{\kappa}|)$, ώστε να μην μηδενίζεται για $\hat{\kappa} = 0$, αλλά να είναι: $h(|\hat{\kappa}|) = \epsilon_\kappa$, για μικρή θετική σταθερά ϵ_κ . Διατηρώντας την απαίτηση να έχουμε $h(|\hat{\kappa}|_{\max}) = 1$, καταλήγουμε στην εξής τροποποίηση της (5.56):

$$h(|\hat{\kappa}|) = \epsilon_\kappa + (1 - \epsilon_\kappa) \left(\frac{|\hat{\kappa}|}{1.11} \right)^p \quad (5.57)$$

Τέλος, όσον αφορά τις μάσκες inpainting, βασισμένοι σε όσα είπαμε στην υποενότητα 5.2 και ακολουθώντας αντίστοιχη ανάλυση όπως στην υλοποίηση της TV PDE για Inpainting (βλ. υποενότ. 5.2.1), καταλήγουμε ότι, για τις διάφορες μεταβλητές, έχουμε:

$$\begin{aligned} \mathbb{D}_{\|\nabla \mathbf{u}\|} &= \mathbb{D}_\kappa = \mathbb{D}_g = \mathbb{D} \oplus B \\ \mathbb{D}_{J(\nabla \mathbf{u}_\sigma)} &= \mathbb{D}_\theta = \mathbb{D}_\kappa \oplus B \end{aligned}$$

όπου με B συμβολίζουμε και πάλι το δομικό στοιχείο:

$$B = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

Με βάση λοιπόν όλα τα παραπάνω, καταλήγουμε στον αλγόριθμο που υλοποιήσαμε, ο οποίος φαίνεται στον πίνακα 5.11.

ΕΙΣΟΔΟΙ: αρχική εικόνα \mathbf{u}^0 (διανυσματική με N κανάλια), περιοχή inpainting \mathbb{D} , αριθμός επαναλήψεων N_{it} , σταθερά εξομάλυνσης σ , σταθερές $\epsilon, \epsilon_\kappa$.

Υπολόγισε τον Gaussian πυρήνα K_σ από το σ , με έκταση σύμφωνα με την (5.20).

Υπολογισμός μασκών inpainting

$$B = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix}; \mathbb{D}_g = \mathbb{D} \oplus B; \mathbb{D}_\theta = \mathbb{D}_g \oplus B$$

Αρχικοποίηση

για $(i, j) \in \Omega \setminus \mathbb{D}$, $\mathbf{u}^{old} = \mathbf{u}^0$

για $(i, j) \in \mathbb{D}$, κάνε μία απλή αρχική πρόβλεψη στην \mathbf{u}^{old} , γεμίζοντας την \mathbb{D} με ομοιόμορφο θόρυβο ή με μία σταθερή τιμή.

For $n = 1, \dots, N_{it}$

Υπολόγισε τη νόρμα $\|\nabla \mathbf{u}\| = \left\{ \sum_{m=1}^N \|\nabla u_m\|^2 \right\}^{\frac{1}{2}}$, σύμφωνα με την σχέση (5.16) ή (5.17).

Υπολόγισε τη γωνία θ του ιδιοδιαν. $\boldsymbol{\theta}_+$ του δομικού τανυστή $J(\nabla \mathbf{u}_\sigma)$ (ίδια βήματα με τον αλγόρ. του πίν. 5.7, μόνο που $\rho = 0$ και οι υπολογισμοί μόνο στα σημεία $(i, j) \in \mathbb{D}_\theta$).

Υπολογισμός της γενικευμένης καμπυλότητας $\hat{\kappa}$

$\delta_{2x}\theta(i, j) = \theta(i+1, j) - \theta(i-1, j)$; $\delta_{2y}\theta(i, j) = \theta(i, j+1) - \theta(i, j-1)$, για $(i, j) \in \mathbb{D}_\theta$.

For each $(i, j) \in \mathbb{D}_\theta$, $z \in \{ 'x', 'y' \}$

If $\delta_{2z}\theta(i, j) \leq -\pi/2$

$\delta_{2z}\theta(i, j) = \delta_{2z}\theta(i, j) + \pi$

elseif $\delta_{2z}\theta(i, j) > \pi/2$

$\delta_{2z}\theta(i, j) = \delta_{2z}\theta(i, j) - \pi$

End if

End for

$\hat{\kappa} = \{-\sin \theta \cdot \delta_{2x}\theta + \cos \theta \cdot \delta_{2y}\theta\}/2$, για $(i, j) \in \mathbb{D}_g$

Υπολογισμός του συντελεστή διάχυσης g

$g = \{\epsilon_\kappa + (1 - \epsilon_\kappa) (|\hat{\kappa}|/1.11)^p\} / \sqrt{\epsilon^2 + \|\nabla \mathbf{u}\|^2}$, για $(i, j) \in \mathbb{D}_g$.

Υπολόγισε τους τοπικούς πυρήνες συσχέτισης $k_{i,j}$ από τις τιμές του g , σύμφωνα με τον πίνακα 5.9.

Υπολογισμός όλων των καναλιών u_m^{new} της επόμενης εικόνας \mathbf{u}^{new}

For $m = 1, \dots, N$

Σε κάθε $(i, j) \in \mathbb{D}$, βρες την $u_m^{new}(i, j)$, εφαρμόζοντας τοπική συσχέτιση της $k_{i,j}$ με την u_m^{old} .

End for

Προετοιμασία επόμενης επανάληψης

$\mathbf{u}^{old} = \mathbf{u}^{new}$

End for

ΕΞΟΔΟΣ: ανακατασκευασμένη εικόνα \mathbf{u}^{new}

Πίνακας 5.11: Αλγόριθμος της Γενικευμένης CDD Μεθόδου, που προτείνουμε για το Inpainting διανυσματικών εικόνων

5.3 Πρόταση μιας νέας μεθόδου αξιολόγησης των ΜΔΕ διάχυσης: Υπολογισμός του Ολικού Φίλτρου της ΜΔΕ

5.3.1 Εισαγωγή

Στην ενότητα αυτή, προτείνουμε τον υπολογισμό της συνολικής (για όλες τις επαναλήψεις μαζί) προσαρμοζόμενης εξομάλυνσης (adaptive smoothing) που επιτελεί μία μέθοδος ΜΔΕ διάχυσης στην διανυσματική εικόνα που δέχεται σαν είσοδο. Ο υπολογισμός αυτός προσφέρει εποπτεία της πραγματικής διαδικασίας με την οποία η μέθοδος χειρίζεται τις δομές της εικόνας, κάτι το οποίο συνήθως δεν μπορεί να φανεί τόσο καθαρά από την απλή παρατήρηση της εικόνας εξόδου ή από αριθμητικά μέτρα του σφάλματος. Επίσης, προτείνουμε έναν αλγόριθμο, ο οποίος κάνει ρεαλιστικό τον υπολογισμό αυτό. Τα παραπάνω μπορούν να φανούν χρήσιμα στην αξιολόγηση της απόδοσης μίας μεθόδου καθώς και στη σύγκριση των αποτελεσμάτων της με άλλες μεθόδους.

Πιο συγκεκριμένα, όπως αναφέραμε και στην εισαγωγή του κεφαλαίου 1, οι μέθοδοι ΜΔΕ διάχυσης για την βελτίωση των εικόνων πρέπει να λαμβάνουν υπόψιν τους τις γεωμετρικές δομές μιας εικόνας και να προσαρμόζουν την εξομάλυνση σύμφωνα με αυτές. Για παράδειγμα, τα πιο σημαντικά χαρακτηριστικά που οφείλει να έχει μία τέτοια μέθοδος είναι η εξομάλυνση να γίνεται έτσι ώστε να μην καταστρέφονται οι ακμές της εικόνας αλλά παράλληλα να απομακρύνεται ο θόρυβος. Από την έκφραση της κάθε ΜΔΕ μπορεί να διαφανεί το κατά πόσο παρουσιάζει αυτή τη συμπεριφορά, αλλά μόνο σε τοπικό επίπεδο, δηλαδή μόνο για τη συγκεκριμένη χρονική στιγμή και σε σχέση με τα σημεία γύρω της. Αντίστοιχα, από τη διακριτοποίηση, μπορούμε άμεσα να παρατηρήσουμε τον τρόπο που οι τιμές της εικόνας ανανεώνονται, σε σχέση με τις τιμές της προηγούμενης μόνο επανάληψης. Καθώς όμως η επαναληπτική διαδικασία της ΜΔΕ εξελίσσεται στο χρόνο, δεν είναι δυνατό να αποσαφηνιστεί άμεσα ο τρόπος που η μέθοδος έχει μεταλλάξει την αρχική εικόνα και τα χαρακτηριστικά της για να καταλήξει τελικά στην τρέχουσα εικόνα. Το παραπάνω είναι που αποσαφηνίζεται με τον αλγόριθμο που προτείνουμε.

Όλες οι ΜΔΕ διάχυσης για Βελτίωση που είδαμε στο κεφάλαιο αυτό (βλ. υποενότ. 5.1), όπως και αρκετές άλλες της ίδιας κατηγορίας, μπορούν να προσεγγιστούν αριθμητικά με ένα επαναληπτικό βήμα, το οποίο έχει την εξής γενική μορφή:

$$u_m^{n+1}(i, j) = u_m^n(i, j) + \tau v_m(\mathbf{u}^n)(i, j), \quad m = 1, \dots, N \quad (5.58)$$

όπου $v_m(\mathbf{u}^n)(i, j)$ το m -οστό κανάλι της διακριτοποιημένης ταχύτητας της αντίστοιχης ΜΔΕ, το οποίο, όπως είδαμε (βλ. υποενότ. 5.1), μπορεί να υπολογιστεί σαν τοπική συσχέτιση της u_m^n , με μεταβαλλόμενο πυρήνα $\tilde{k}_{i,j}(\ell_1, \ell_2)$

(έκτασης 3×3 , για τις υλοποιήσεις που κάναμε). Από την (5.58) συμπεραίνουμε ότι και η u_m^{n+1} μπορεί να εκφραστεί σαν τοπική συσχέτιση της u_m^n με έναν άλλο κατάλληλο πυρήνα συσχέτισης $k_{i,j}(\ell_1, \ell_2)$.

Επομένως, αν αναπαραστήσουμε κάθε κανάλι $u_m(i, j)$ της εικόνας σαν ένα διάνυσμα \underline{u}_m διάστασης $N_x N_y$ (όπου N_x, N_y οι διαστάσεις του ορθογωνικού πλέγματος ορισμού της εικόνας), η επαναληπτική διαδικασία προσέγγισης της ΜΔΕ μπορεί να γραφεί με συμβολισμό πινάκων ως εξής:

για κάθε $m = 1, \dots, N$:

$$\begin{aligned}\underline{u}_m^{n+1} &= \underline{u}_m^n + \tau A(\mathbf{u}^n) \underline{u}_m^n \\ &= B(\mathbf{u}^n) \underline{u}_m^n\end{aligned}\quad (5.59)$$

όπου $B(\mathbf{u}^n) = I + \tau A(\mathbf{u}^n)$, πίνακας διαστάσεων $N_x N_y \times N_x N_y$, του οποίου τα μη μηδενικά στοιχεία σε κάθε γραμμή είναι οι τιμές του αντίστοιχου πυρήνα συσχέτισης $k_{i,j}(\ell_1, \ell_2)$. Επομένως, ο πίνακας $B(\mathbf{u}^n)$ περιέχει ουσιαστικά όλους τους τοπικούς πυρήνες συσχέτισης. Ο πίνακας αυτός είναι αραιός, αφού σε κάθε γραμμή του, από τα $N_x N_y$ στοιχεία, μόνο ένας μικρός αριθμός, που αντιστοιχούν στα pixels της άμεσης γειτονιάς του εκάστοτε pixel (i, j) είναι μη μηδενικά. Για παράδειγμα, στις υλοποιήσεις που κάναμε, τα μη μηδενικά στοιχεία είναι μόνο 9, αφού οι πυρήνες $\tilde{k}_{i,j}$ και $k_{i,j}$ έχουν έκταση 3×3 .

Πρέπει να σημειώσουμε ότι οι πίνακες $A(\mathbf{u}), B(\mathbf{u})$ είναι ίδιοι για όλα τα κανάλια της εικόνας, διότι αναφερόμαστε σε εκείνες τις μεθόδους ΜΔΕ διάχυσης διανυσματικών εικόνων που επιτελούν σε όλα τα κανάλια την ίδια εξομάλυνση. Στην κατηγορία αυτή ανήκουν οι περισσότερες υπάρχουσες μέθοδοι και όλες όσες υλοποιήσαμε. Ακόμα, οι $A(\mathbf{u}), B(\mathbf{u})$ εξαρτώνται από την \mathbf{u} , δηλ. από όλα τα κανάλια u_m , αφού έτσι επιτυγχάνεται η σύζευξη μεταξύ των u_m .⁴

Αξίζει να παρατηρηθεί ότι αντίστοιχες με τις παραπάνω εκφράσεις ισχύουν και για τις ΜΔΕ για το Inpainting, με κάποιες μικρές τροποποιήσεις, που αφορούν την περιοχή inpainting. Στη συνέχεια όμως, θα επικεντρωθούμε στις ΜΔΕ διάχυσης για Βελτίωση, όπου έχουμε μεταβολή σε όλο το πεδίο ορισμού της εικόνας.

Έστω λοιπόν ότι η εικόνα εξόδου προκύπτει μετά από την εφαρμογή N_{it} επαναλήψεων της (5.58) ή ισοδύναμα της (5.59). Εφαρμόζοντας αναδρομικά την (5.59) (με αντίστροφη πορεία από αυτή της χρονικής εξέλιξης), προκύπτει

⁴για αυτόν τον λόγο, χρησιμοποιούμε τον συμβολισμό $A(\mathbf{u})$ σαν σύντμηση του συμβολισμού $A(\underline{u}_1, \dots, \underline{u}_N)$, για τον πίνακα A και αντίστοιχα για τον B και τον $\mathcal{F}_{tot}^{N_{it}}$, τον οποίο θα εισάγουμε πιο κάτω.

ότι, για κάθε $m = 1, \dots, N$:

$$\begin{aligned} \underline{u}_m^{N_{it}} &= B(\mathbf{u}^{N_{it}} - 1)B(\mathbf{u}^{N_{it}-2}) \cdots B(\mathbf{u}^0) \cdot \underline{u}_m^0 \\ &= \mathcal{F}_{tot}^{N_{it}} \cdot \underline{u}_m^0, \quad \mathcal{F}_{tot}^{N_{it}} = B(\mathbf{u}^{N_{it}-1})B(\mathbf{u}^{N_{it}-2}) \cdots B(\mathbf{u}^0) \end{aligned} \quad (5.60)$$

όπου ο πίνακας $\mathcal{F}_{tot}^{N_{it}}$ εξαρτάται από τις τιμές της εικόνας σε όλα τα προηγούμενα βήματα, από 0 έως $N_{it} - 1$. Όμως, η εικόνα \mathbf{u}^n σε κάθε βήμα n εξαρτάται προφανώς μόνο από την αρχική εικόνα \mathbf{u}^0 , αφού αυτή είναι η μόνη είσοδος της μεθόδου. Επομένως, ο πίνακας $\mathcal{F}_{tot}^{N_{it}}$ εξαρτάται ουσιαστικά μόνο από την \mathbf{u}^0 και έχουμε τελικά, για κάθε $m = 1, \dots, N$:

$$\underline{u}_m^{N_{it}} = \mathcal{F}_{tot}^{N_{it}}(\mathbf{u}^0) \cdot \underline{u}_m^0 \quad (5.61)$$

Όπως στην περίπτωση της (5.59), η παραπάνω σχέση μπορεί να ειπωθεί σαν συσχέτιση της αρχικής εικόνας με έναν χωρικά μεταβαλλόμενο πυρήνα, του οποίου τα στοιχεία προκύπτουν από την αντίστοιχη γραμμή του $\mathcal{F}_{tot}^{N_{it}}$. Επομένως, η έκφραση (5.61) έχει τη μορφή μη γραμμικού, εξαρτώμενου από την είσοδο φίλτρου, το οποίο ονομάζουμε *Ολικό Φίλτρο της ΜΔΕ*. Πρέπει να σημειώσουμε ότι η συνάρτηση $\mathcal{F}_{tot}^{N_{it}}(\mathbf{u}^0)$ είναι ισχυρά μη γραμμική και η εξάρτηση από την διακριτή εικόνα εισόδου \mathbf{u}^0 υπερβολικά περίπλοκη. Η συνάρτηση αυτή προσδιορίζεται από την συγκεκριμένη μέθοδο ΜΔΕ, τον τρόπο διακριτοποίησης και τον αριθμό επαναλήψεων N_{it} . Προφανώς δεν υπάρχει κλειστός τύπος που να την περιγράφει και η έξοδος μπορεί να προσδιοριστεί μόνο μέσω της επαναληπτικής διαδικασίας της ΜΔΕ.

Οι πυρήνες συσχέτισης του Ολικού Φίλτρου $\mathcal{F}_{tot}^{N_{it}}$ της ΜΔΕ παρέχουν τον τρόπο με τον οποίο μπορούμε, μέσα σε ένα βήμα, εξομαλύνοντας την αρχική \mathbf{u}^0 να φτάσουμε στην τελική $\mathbf{u}^{N_{it}}$. Έτσι, εμπεριέχουν όλη την εξομάλυνση που εφάρμοσε η ΜΔΕ κατά τη χρονική ροή και τον τρόπο που χειρίστηκε την εικόνα. Μέσω της έκτασης και της κατανομής του πυρήνα που αντιστοιχεί σε κάποιο σημείο μπορούμε να διαπιστώσουμε ακριβώς ποια σημεία της αρχικής εικόνας επηρέασαν την τελική διαμόρφωση της τιμής στο σημείο αυτό και σε τι βαθμό. Επομένως, η γνώση των πυρήνων αυτών, ιδιαίτερα σε κρίσιμα σημεία της εικόνας (π.χ. κοντά σε ακμές ή γωνίες) είναι χρήσιμη για την αξιολόγηση της μεθόδου και για την εξακρίβωση του κατά πόσο η μέθοδος συμπεριφέρθηκε με τον επιθυμητό τρόπο. Για την απομάκρυνση του θορύβου, επιθυμητή συμπεριφορά θα ήταν για παράδειγμα ο πυρήνας να μην εκτείνεται έξω από τα όρια της γεωμετρικής περιοχής που βρίσκεται το αντίστοιχο σημείο ή για σημείο κοντά σε ακμές ο πυρήνας να εκτείνεται ακολουθώντας από την πλευρά της ακμής το σχήμα της.

Για τους παραπάνω λόγους, προτείνουμε τον υπολογισμό των πυρήνων συσχέτισης του Ολικού Φίλτρου της ΜΔΕ, σαν εργαλείο για την αξιολόγηση των

μεθόδων ΜΔΕ. Όπως θα δούμε, κάποιιο απλοί τρόποι υπολογισμού των πυρήνων αυτών έχουν απαγορευτική πολυπλοκότητα. Για το λόγο αυτό, όπως θα παρουσιάσουμε στη συνέχεια, κατασκευάζουμε έναν αλγόριθμο που υπολογίζει ένα τέτοιο πυρήνα συσχέτισης, για ένα μόνο σημείο. Ο αλγόριθμος αυτός μπορεί εύκολα να επεκταθεί για ένα μικρό σύνολο από σημεία. Για την σημαντική μείωση της υπολογιστικής πολυπλοκότητας, έχουμε το τίμημα του περιορισμού σε ένα ή λίγα σημεία. Βέβαια, για την σύγκριση της απόδοσης διαφόρων μεθόδων, η γνώση όλων των πυρήνων είναι περιττή, αφού αρκούν λίγα σημεία, σε αντιπροσωπευτικές περιοχές, για να καταλήξουμε στα συμπεράσματά μας.

5.3.2 Αλγόριθμος υπολογισμού ενός πυρήνα συσχέτισης του Ολικού Φίλτρου

Με βάση όσα είπαμε στην προηγούμενη υποενότητα, ο προφανής τρόπος υπολογισμού του Ολικού Φίλτρου $\mathcal{F}_{tot}^{N_{it}}$ είναι, παράλληλα με την επαναληπτική διαδικασία της ΜΔΕ, σε κάθε βήμα n , να σώζουμε τον πίνακα \mathcal{F}_{tot}^n και να υπολογίζουμε την επόμενη τιμή, βασισμένοι στην (5.60), ως εξής:

$$F_{tot}^{n+1} = B(\mathbf{u}^n)F_{tot}^n \quad (5.62)$$

Όμως, οι πίνακες $B(\mathbf{u}^n)$, \mathcal{F}_{tot}^n είναι μεγέθους $N_x N_y \times N_x N_y$. Έτσι, σε κάθε επανάληψη, απαιτούνται για τον υπολογισμό αυτό $(N_x N_y)^3$ πολλαπλασιασμοί. Επίσης, αν για την αποθήκευση κάθε στοιχείου του πίνακα (πραγματικός αριθμός) απαιτούνται 8 bytes, για όλον τον \mathcal{F}_{tot}^n , απαιτούνται $8(N_x N_y)^2$ bytes. Τα μεγέθη αυτά όμως, είναι απαγορευτικά. Για παράδειγμα, για μία συνηθισμένη περίπτωση εικόνας 300×300 pixels, ο \mathcal{F}_{tot}^n περιέχει $(N_x N_y)^2 = 8.1 \cdot 10^9$ στοιχεία και πρέπει να δεσμεύσει χώρο στη μνήμη, μεγέθους περίπου 60 GiB. Επιπρόσθετα, αν θεωρήσουμε ότι εφαρμόζονται συνολικά $N_{it} = 100$ επαναλήψεις (που είναι επίσης μία συνηθισμένη περίπτωση), απαιτούνται μόνο για τον υπολογισμό (5.62) $100(N_x N_y)^3 \approx 7 \cdot 10^{16}$ πολλαπλασιασμοί. Συνεπώς, για τα δεδομένα της σύγχρονης τεχνολογίας των υπολογιστών, ο προφανής αυτός τρόπος δεν είναι εφικτό να υλοποιηθεί.

Έτσι, κατασκευάσαμε έναν αλγόριθμο, με σημαντικά μικρότερη πολυπλοκότητα, για τον υπολογισμό αυτό. Για να γίνουν εφικτοί οι υπολογισμοί, περιοριζόμαστε στην εύρεση του πυρήνα του Ολικού Φίλτρου $\mathcal{F}_{tot}^{N_{it}}$ για ένα μόνο σημείο (i_0, j_0) , το οποίο αποτελεί και είσοδο του αλγορίθμου. Για το (i_0, j_0) θα χρησιμοποιήσουμε την ορολογία *pixel ενδιαφέροντος*. Η βασική ιδέα έγκειται στην αποθήκευση σε κάθε βήμα της ΜΔΕ μόνο εκείνων των πυρήνων συσχέτισης που θα μας χρειαστούν. Μετά από την ολοκλήρωση όλων των βημάτων της ΜΔΕ, υπολογίζουμε αναδρομικά τον ζητούμενο πυρήνα του Ολικού Φίλτρου, με μία διαδικασία οπισθοδρόμησης.

Έστω λοιπόν ότι ο συνολικός αριθμός επαναλήψεων είναι N_{it} και θέλουμε να υπολογίσουμε τον πυρήνα του Ολικού Φίλτρου για το σημείο (i_0, j_0) .

Στους αλγορίθμους υλοποίησης των ΜΔΕ για Βελτίωση, που παρουσιάσαμε στην υποενότητα 5.1, χρησιμοποιείται τοπικός πυρήνας συσχέτισης σε κάθε βήμα n και για κάθε σημείο (i, j) (τον οποίο συμβολίζουμε με $k_{i,j}^n(\ell_1, \ell_2)$), για τον υπολογισμό των καναλιών $v_m^n(i, j)$ της ταχύτητας της ΜΔΕ στο σημείο αυτό από τα αντίστοιχα κανάλια u_m^n . Έτσι, η επόμενη επανάληψη u_m^{n+1} προκύπτει ως εξής, για κάθε $m = 1, \dots, N$:

$$\begin{aligned} v_m^n(i, j) &= (\tilde{k}_{i,j}^n * u_m^n)(i, j) \\ u_m^{n+1}(i, j) &= u_m^n(i, j) + \tau v_m^n(i, j) \end{aligned} \quad (5.63)$$

όπου με “ $*$ ” συμβολίζουμε την συσχέτιση. Εδώ όμως, θέλουμε να εντάξουμε τα δύο βήματα (5.63) σε ένα, χρησιμοποιώντας ενιαίο πυρήνα, έστω $k_{i,j}^n(\ell_1, \ell_2)$, ώστε η τοπική συσχέτισή του με το κανάλι u_m^n της εικόνας να δίνει κατευθείαν την επόμενη εικόνα u_m^{n+1} . Λόγω της (5.63), ο πυρήνας $k_{i,j}^n(\ell_1, \ell_2)$ προκύπτει από τον $\tilde{k}_{i,j}^n(\ell_1, \ell_2)$, ως εξής:

$$k_{i,j}^n(\ell_1, \ell_2) = \delta(\ell_1, \ell_2) + \tau \tilde{k}_{i,j}^n(\ell_1, \ell_2) \quad (5.64)$$

όπου $\delta(\ell_1, \ell_2)$ η διακριτή 2Δ συνάρτηση Dirac.

Όμως, ο πυρήνας $k_{i,j}^n$ χρειάζεται μία μικρή τροποποίηση ώστε να μπορεί τελικά να χρησιμοποιηθεί από τον αλγόριθμο. Όπως έχουμε αναφέρει (βλ. υποενότ. 5.1.9), σε όλες τις υλοποιήσεις χρησιμοποιούμε ανάκλαση για τις τιμές των u_m^n . Έτσι, για τον υπολογισμό της συσχέτισης στα σημεία του συνόρου, ο αντίστοιχος πυρήνας εκτείνεται και εκτός ορίων της εικόνας, αλλά ο πολλαπλασιασμός που τελικά γίνεται είναι με την κατοπτρισμένη τιμή της u_m^n , εντός των ορίων. Εδώ όμως, θέλουμε όλοι οι πυρήνες $k_{i,j}^n$ να εκτείνονται μόνο εντός των ορίων της εικόνας, για αυτό τροποποιούμε τις τιμές των πυρήνων $k_{i,j}^n$ στο σύνορο, ώστε τα εσωτερικά σημεία να ενσωματώσουν αυτή την ανάκλαση και τα εξωτερικά να μηδενιστούν.

Όπως έχουμε αναφέρει, στις διακριτοποιήσεις που κάναμε, ο πυρήνας $k_{i,j}^n(\ell_1, \ell_2)$ έχει σε όλες τις περιπτώσεις έκταση 3×3 (βλ. υποενότ. 5.1.7). Αυτή είναι η περίπτωση που θα θεωρήσουμε στη συνέχεια για την ανάπτυξη του αλγορίθμου, αλλά ο αλγόριθμος μπορεί άμεσα να τροποποιηθεί και για πυρήνες άλλου μεγέθους (μερικές φορές, για μεγαλύτερη ακρίβεια, χρησιμοποιούνται διακριτοποιήσεις με μεγαλύτερη έκταση από την τυπική 3×3 , βλ. π.χ. [64]). Έτσι, στην περίπτωση αυτή, οι τιμές ορισμού του $\tilde{k}_{i,j}^n(\ell_1, \ell_2)$ κειμώνται γύρω από το $(0, 0)$ ως εξής: $(\ell_1, \ell_2) \in \{-1, 0, 1\}^2$. Σύμφωνα με τα παραπάνω, η επαναληπτική διαδικασία της ΜΔΕ μπορεί λοιπόν να εκφραστεί ως εξής, για κάθε

$m = 1, \dots, N$:

$$\begin{aligned} u_m^n(i, j) &= (k_{i,j}^n \tilde{*} u_m^n)(i, j) \\ &= \sum_{\ell_1=-1}^1 \sum_{\ell_2=-1}^1 k_{i,j}^{n-1}(\ell_1, \ell_2) \cdot u_m^{n-1}(i + \ell_1, j + \ell_2) \end{aligned} \quad (5.65)$$

Θεωρούμε επίσης τον πυρήνα συσχέτισης $h^n(i, j)$, μέσω του οποίου η τιμή $u_m^{N_{it}}(i_0, j_0)$ της τελικής εικόνας στο pixel ενδιαφέροντος (i_0, j_0) προκύπτει κατευθείαν από το κανάλι u_m^n της εικόνας, την χρονική στιγμή n .⁵ Λόγω του ότι οι πυρήνες $k_{i,j}^m(\ell_1, \ell_2)$, για κάθε m , περιορίζονται εντός των ορίων της εικόνας, ο πυρήνας $h^n(\ell_1, \ell_2)$ θα καταλαμβάνει το πολύ την έκταση της εικόνας. Επειδή μας ενδιαφέρει η εφαρμογή της συνέλιξης του πυρήνα αυτού μόνο στο σημείο (i_0, j_0) , δεν χρησιμοποιούμε την κλασσική αρίθμηση των πυρήνων για το πεδίο ορισμού του (i, j) , αλλά τέτοια, που να συμβαδίζει με την εικόνα. Έτσι, η συχέτιση με την u_m^n γράφεται:

$$u_m^{N_{it}}(i_0, j_0) = \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} h^n(i, j) \cdot u_m^n(i, j) \quad (5.66)$$

Ο πυρήνας $h^n(i, j)$ μπορεί να ειπωθεί και σαν ο αντίστοιχος πυρήνας του Ολικού Φίλτρου, αν θεωρήσουμε την u^n σαν αρχική εικόνα. Επομένως, ο πυρήνας που ψάχνουμε είναι ουσιαστικά ο $h^0(i, j)$.

Λόγω των παραπάνω, βασικό στάδιο του αλγορίθμου είναι ο υπολογισμός του $h^{n-1}(i, j)$ από τον $h^n(i, j)$, με αναδρομή προς τα πίσω (οπισθοδρόμηση). Συγκεκριμένα, αν στην (5.66) αντικαταστήσουμε το $u_m^n(i, j)$ με την έκφραση της (5.65), προκύπτει:

$$u_m^{N_{it}}(i_0, j_0) = \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} \sum_{\ell_1=-1}^1 \sum_{\ell_2=-1}^1 h^n(i, j) k_{i,j}^{n-1}(\ell_1, \ell_2) \cdot u_m^{n-1}(i + \ell_1, j + \ell_2) \quad (5.67)$$

από την άλλη όμως, αν εφαρμόσουμε την (5.66) αντικαθιστώντας το n με $n - 1$, παίρνουμε:

$$u_m^{N_{it}}(i_0, j_0) = \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} h^{n-1}(i, j) \cdot u_m^{n-1}(i, j) \quad (5.68)$$

Από τις σχέσεις (5.67) και (5.68) διαφαίνεται η σύνδεση των $h^{n-1}(i, j)$ και $h^n(i, j)$. Έτσι, βασισμένοι στις σχέσεις αυτές, υπολογίζουμε στον αλγόριθμο

⁵ο πυρήνας h^n σχετίζεται άμεσα με το pixel ενδιαφέροντος (i_0, j_0) και την στιγμή N_{it} , αλλά για απλότητα, αυτό δεν φαίνεται στον συμβολισμό που χρησιμοποιούμε, αφού δεν ενδιαφερόμαστε για κάποιο άλλο pixel ή άλλη στιγμή (ο πλήρης συμβολισμός θα ήταν $h_{i_0, j_0}^{N_{it}, n}$).

τις τιμές του $h^{n-1}(i, j)$ από τις τιμές του $h^n(i, j)$, με επαναληπτικό τρόπο. Ο υπολογισμός αυτός βασίζεται στην παρατήρηση ότι για κάθε σημείο (i, j) , η τιμή $h^n(i, j)$, όταν δεν είναι μηδενική, συνεισφέρει 9 προσθετικούς όρους στον πυρήνα h^{n-1} της μορφής $h^n(i, j)k_{i,j}^{n-1}(\ell_1, \ell_2)$, στα σημεία $(i + \ell_1, j + \ell_2)$ γύρω από το (i, j) (βλ. σχέση (5.67)).

Επίσης, σαν αρχικοποίηση στον προς τα πίσω υπολογισμό του πυρήνα $h^n(i, j)$, έχουμε την τιμή για τη στιγμή $n = N_{it}$. Τότε, αν στην (5.66) θέσουμε $n = N_{it}$, λόγω της βασικής ιδιότητας της 2Δ συνάρτησης Dirac $\delta(i, j)$, παίρνουμε την απλή έκφραση:

$$h_{it}^N(i, j) = \delta(i - i_0, j - j_0) \quad (5.69)$$

Τέλος, από τα παραπάνω μπορούμε να συμπεράνουμε ποιούς πυρήνες $k_{i,j}^n$ χρειαζόμαστε, ώστε στην αρχική φάση εξέλιξης της ΜΔΕ, να αποθηκεύουμε μόνο αυτούς: Σε κάποιο βήμα n , χρειαζόμαστε τους πυρήνες σε ένα τετράγωνο γύρω από το pixel ενδιαφέροντος (i_0, j_0) και, προχωρώντας προς τα πίσω, στο βήμα $n - 1$ βλέπουμε ότι το τετράγωνο που χρειαζόμαστε μεγαλώνει, προσθέτοντας 1 pixel σε κάθε διεύθυνση. Η διαδικασία αυτή σταματάει σε μία διεύθυνση, όταν φτάσουμε στο αντίστοιχο όριο της εικόνας. Επίσης, όταν $n = N_{it} - 1$, το τετράγωνο αυτό είναι μόνο το pixel (i_0, j_0) . Από τα παραπάνω προκύπτει ότι στο n -οστό βήμα, αρκεί να αποθηκεύσουμε τους πυρήνες $k_{i,j}^n$ για (i, j) , τέτοια ώστε:

$$\max\{1, i_0 - N_{it} + n\} \leq i \leq \min\{N_x, i_0 + N_{it} - n\} \quad (5.70)$$

$$\max\{1, j_0 - N_{it} + n\} \leq j \leq \min\{N_y, j_0 + N_{it} - n\} \quad (5.71)$$

Με βάση όλη τη συζήτηση που προηγήθηκε, παραθέτουμε τον αλγόριθμο που προτείνουμε: βλ. πίνακα 5.12.

5.4 Υλοποίηση ΜΔΕ για την Παρεμβολή εικόνων

Στην ενότητα 3.2, για την επίλυση του προβλήματος της Παρεμβολής διανυσματικών εικόνων, χρησιμοποιήσαμε τελικά την υπόθεση ότι η αρχική, χαμηλής χωρικής ευκρίνειας εικόνα $\mathbf{z}(n_1, n_2)$ έχει προέλθει από την τψηλότερης ευκρίνειας εικόνα \mathbf{u} που ψάχνουμε, μέσω βαθυπερατού φιλτραρίσματος και δειγματοληψίας. Η εικόνα \mathbf{u} ικανοποιεί, δηλαδή, την συνθήκη αναστρεψιμότητας:

$$z_m = Q(s * u_m), \quad m = 1, \dots, N \quad (5.72)$$

όπου s κατάλληλος πυρήνας συνέλιξης (χρυστική απόκριση του φίλτρου). Η συνθήκη (5.72) ορίζει τον υποχώρο $\mathbf{U}_{\mathbf{z},s}$.

ΕΙΣΟΔΟΙ: αρχική διανυσματική εικόνα \mathbf{u}^0 (έκτασης $N_x \times N_y$), αριθμός επαναλήψεων N_{it} , pixel ενδιαφέροντος (i_0, j_0) και άλλες παράμετροι, ανάλογα με τη συγκεκριμένη μέθοδο ΜΔΕ.

ΕΦΑΡΜΟΓΗ ΜΔΕ ΔΙΑΧΥΣΗΣ

Αρχικοποίηση: $\mathbf{u} = \mathbf{u}^0$

For $n = 1, \dots, N_{it}$

Εφαρμοσε τα βήματα μίας επανάληψης, όπως στον αλγόριθμο της αντίστοιχης ΜΔΕ (βλ. πίν. 5.5, 5.7) και βρες την επόμενη εικόνα, έχοντας υπολογίσει τους τοπικούς πυρήνες συσχέτισης $\tilde{k}_{i,j}$ για την εύρεση της ταχύτητας.

Αποθήκευση μόνο των πυρήνων συσχέτισης που θα χρειαστούν

For $i = \max(1, i_0 - N_{it} + n), \dots, \min(N_x, i_0 + N_{it} - n)$

For $j = \max(1, j_0 - N_{it} + n), \dots, \min(N_y, j_0 + N_{it} - n)$

αποθήκευση πυρήνα που δίνει κατευθείαν την επόμενη εικόνα

$$k_{i,j}^n(\ell_1, \ell_2) = \delta(\ell_1, \ell_2) + \tau \tilde{k}_{i,j}(\ell_1, \ell_2), \quad (\ell_1, \ell_2) \in \{-1, 0, 1\}^2$$

προσαρμογή του $k_{i,j}^n$, ώστε να μένει εντός της εικόνας

If $i == 1$

$$k_{i,j}^n(0, \ell_2) = k_{i,j}^n(0, \ell_2) + k_{i,j}^n(-1, \ell_2), \quad \ell_2 \in \{-1, 0, 1\}$$

$$k_{i,j}^n(-1, \ell_2) = 0, \quad \ell_2 \in \{-1, 0, 1\}$$

End if

If $i == N_x$

$$k_{i,j}^n(0, \ell_2) = k_{i,j}^n(0, \ell_2) + k_{i,j}^n(1, \ell_2), \quad \ell_2 \in \{-1, 0, 1\}$$

$$k_{i,j}^n(1, \ell_2) = 0, \quad \ell_2 \in \{-1, 0, 1\}$$

End if

{Εντελώς παρόμοιες εντολές, για τις περιπτώσεις $j == 1$ και $j == N_y$ }

End for

End for

End for

ΥΠΟΛΟΓΙΣΜΟΣ ΠΡΟΣ ΤΑ ΠΙΣΩ ΤΟΥ ΠΥΡΗΝΑ ΤΟΥ $\mathcal{F}_{tot}^{N_{it}}$

αρχικοποίηση: $h^{cur}(i, j) = \delta(i - i_0, j - j_0)$

For $n = N_{it}, N_{it} - 1, \dots, 1$

αρχικοποίηση: $h^{bef}(i, j) = 0$, για κάθε pixel (i, j)

For all $(i, j) : h^{cur}(i, j) \neq 0$ do

For $(\ell_1, \ell_2) \in \{-1, 0, 1\}^2$

$$h^{bef}(i + \ell_1, j + \ell_2) = h^{bef}(i + \ell_1, j + \ell_2) + h^{cur}(i, j) \cdot k_{i,j}^{n-1}(\ell_1, \ell_2)$$

End for

End for

ανανέωση της τρέχουσας τιμής του πυρήνα $h^n(i, j)$

$$h^{cur}(i, j) = h^{bef}(i, j), \quad \text{για κάθε } (i, j)$$

End for

ΕΞΟΔΟΣ: $h^{cur}(i, j)$, πυρήνας Ολικού Φίλτρου $\mathcal{F}_{tot}^{N_{it}}$ για το (i_0, j_0)

Πίνακας 5.12: Ο Αλγόριθμος που προτείνουμε, για τον Υπολογισμό του Πυρήνα του Ολικού Φίλτρου της ΜΔΕ

Για την τελική κατάληξη σε μια εικόνα ανάμεσα σε αυτές που ικανοποιούν την παραπάνω συνθήκη, προτείναμε την χρήση της εξής γενικής ΜΔΕ:

$$\frac{\partial \mathbf{u}}{\partial t} = P_{\mathbf{u}_{0,s}}(\mathbf{v}(\mathbf{u})) \quad (5.73)$$

όπου σαν αρχική συνθήκη \mathbf{u}^0 της ΜΔΕ αυτής επιλέξαμε και πάλι την παρεμβολή zero-padding για κάθε κανάλιχωριστά. Με $\mathbf{v}(\mathbf{u})$ συμβολίζουμε τη διακριτοποίηση της ταχύτητας (2^ο μέλος) οποιασδήποτε από τις ΜΔΕ διάχυσης για Βελτίωση ή Inpainting (βλ. ενότητες 1 και 2). Επίσης, το $P_{\mathbf{u}_{0,s}}(\cdot)$ συμβολίζει την προβολή στον υποχώρο $\mathbf{U}_{0,s}$, ο οποίος είναι ο χώρος των συναρτήσεων \mathbf{v} που ικανοποιούν τη συνθήκη αναστρεψιμότητας για μηδενική αρχική εικόνα, δηλαδή:

$$Q(s * v_m) = 0, \quad m = 1, \dots, N \quad (5.74)$$

Έστω ότι η ζητούμενη εικόνα \mathbf{u} έχει μεγαλύτερο μέγεθος, λόγω πυκνότερου πλέγματος, από την γνωστή εικόνα $\mathbf{z}(n_1, n_2)$ κατά τον ακέραιο παράγοντα d (d φορές μεγαλύτερη σε κάθε διάσταση). Για το d , θα χρησιμοποιήσουμε στα επόμενα την ορολογία παράγοντας μεγέθυνσης. Επίσης, για το $d \times d$ μπλοκ από pixel του πυκνότερου πλέγματος, που αντιστοιχεί σε ένα αρχικό μεγάλο pixel του αραιότερου πλέγματος, θα χρησιμοποιήσουμε την ονομασία μπλοκ μεγέθυνσης.

Στην υλοποίηση που πραγματοποιήσαμε, θα θεωρούμε ότι ο διακριτός πυρήνας s έχει μη μηδενικές τιμές μόνο στα σημεία που αντιστοιχούν σε ένα μπλοκ μεγέθυνσης, έχει δηλαδή μέγεθος $d \times d$. Αυτή η υπόθεση, όπως είδαμε στην ενότητα 3.1, οδηγεί σε αξιόπιστα αποτελέσματα καθώς επίσης είναι αντιπροσωπευτική της πραγματικής διαδικασίας της ψηφιοποίησης. Η πιο απλή επιλογή είναι ο πυρήνας αυτός να έχει σταθερή τιμή για όλα τα σημεία του μπλοκ μεγέθυνσης (διακριτή κυβική συνάρτηση).

Υλοποιήσαμε τη μέθοδο αυτή για τις συγκεκριμένες ΜΔΕ διάχυσης που υλοποιήσαμε για το πρόβλημα Βελτίωσης (βλ. υποενότητα 5.1). Έτσι, οι αλγόριθμοι για την υλοποίηση είναι παρόμοιοι με τους αντίστοιχους αλγορίθμους της υποενότητας 5.1.7 (βλ. πίνακες 5.5, 5.7), με μόνες διαφορές ότι εδώ, η αρχική εικόνα δεν παρέχεται σαν είσοδος, αλλά υπολογίζεται με τη μέθοδο zero-padding και ότι σε κάθε επανάληψη, ακριβώς μετά τον υπολογισμό της ταχύτητας $\mathbf{v}(n_1, n_2)$, προστίθεται το βήμα της προβολής της στον χώρο $\mathbf{U}_{0,s}$. Στην συνέχεια λοιπόν, θα παρουσιάσουμε την υλοποίηση των δύο αυτών επιπλέον βημάτων:

5.4.1 Υλοποίηση της μεθόδου Zero-Padding

Υλοποιήσαμε τη μέθοδο zero-padding, με τον απλό τρόπο, ο οποίος προκύπτει από τον ορισμό της και χρησιμοποιεί Μετασχηματισμούς Fourier. Η

υλοποίηση αυτή δεν είναι υπολογιστικά αποδοτική, αλλά το αποτέλεσμα που προκύπτει έχει μεγάλη ακρίβεια.

Αναφερόμενοι στην προηγούμενη υποενοότητα, θεωρούμε ότι η γνωστή, χαμηλής ευκρίνειας, εικόνα $\mathbf{z}(n_1, n_2)$ έχει έκταση $M_x \times M_y$ και ο παράγοντας μεγέθυνσης είναι d , οπότε η ζητούμενη εικόνα $\mathbf{u}(n_1, n_2)$ έχει έκταση $N_x \times N_y$, με $N_x = dM_x$, $N_y = dM_y$. Στην ακόλουθη ανάλυση και στον αλγόριθμο που παρουσιάζουμε, θα υποθέσουμε, για απλότητα στην παρουσίαση, ότι οι ακέραιοι M_x, M_y, d είναι όλοι περιττοί. Οι υπόλοιπες περιπτώσεις είναι πιο περίπλοκες, αλλά μπορούν να προκύψουν σχετικά εύκολα, με κάποιες τροποποιήσεις στις σχέσεις (που χρειάζονται εξαιτίας κυρίως της διαίρεσης με το 2). Λόγω της υπόθεσης αυτής, οι N_x, N_y είναι προφανώς επίσης περιττοί.

Επίσης, συμβολίζουμε με $Z_m(m_1, m_2)$ και $U_m(m_1, m_2)$ τους Διακριτούς Μετασχηματισμούς Fourier (DFT) των καναλιών $z_m(n_1, n_2)$ και $u_m(n_1, n_2)$, οι οποίοι έχουν το ίδιο μέγεθος με τα αντίστοιχα 2Δ σήματα, δηλαδή $M_x \times M_y$ και $N_x \times N_y$, αντίστοιχα. Κατά τα γνωστά της Θεωρίας Σημάτων, μπορούμε (λόγω της χωρικής διακριτοποίησης) να θεωρήσουμε ότι οι DFT επεκτείνονται περιοδικά σε όλο το υπόλοιπο διακριτό επίπεδο συχνοτήτων. Έτσι, αν το 2Δ σήμα έχει μέγεθος $L_x \times L_y$ (όπου $(L_x, L_y) = (M_x, M_y)$ ή (N_x, N_y)), για τον DFT του σήματος αυτού, μπορούμε αυθαίρετα να θεωρήσουμε ότι οι χωρικές συχνότητες m_1 και m_2 περιορίζονται σε οποιοδήποτε διάστημα διαδοχικών ακεραίων με μήκος L_x και L_y σημεία αντίστοιχα. Στην παρουσίαση αυτή λοιπόν, επιλέγουμε τα διαστήματα συχνοτήτων με κέντρο το 0. Έτσι, επειδή οι L_x, L_y είναι περιττοί, το διάστημα για το L_x είναι $\{-\frac{L_x-1}{2}, \dots, \frac{L_x-1}{2}\}$ και αντίστοιχα για το L_y . Συμβολίζουμε με B_Z και B_U τα σύνολα ορισμού των DFT $Z_m(m_1, m_2)$ και $U_m(m_1, m_2)$ αντίστοιχα, οπότε, σύμφωνα με την επιλογή που κάναμε, τα σύνολα αυτά δίνονται από τις σχέσεις:

$$\begin{aligned} B_Z &= \left\{ -\frac{M_x-1}{2}, \dots, \frac{M_x-1}{2} \right\} \times \left\{ -\frac{M_y-1}{2}, \dots, \frac{M_y-1}{2} \right\} \\ B_U &= \left\{ -\frac{N_x-1}{2}, \dots, \frac{N_x-1}{2} \right\} \times \left\{ -\frac{N_y-1}{2}, \dots, \frac{N_y-1}{2} \right\} \end{aligned} \quad (5.75)$$

Για την συνέχεια, αναφέρουμε καταρχήν το Θεώρημα Δειγματοληψίας, όπως εκφράζεται στην τελείως διακριτή του μορφή, δηλαδή όταν και η αρχική εικόνα είναι διακριτή. Έτσι, έστω $w(n_1, n_2)$ μία διακριτή εικόνα περιορισμένης έκτασης $N_x \times N_y$ και $\tilde{w}(n_1, n_2) = Q(w)(n_1, n_2)$ η υποδειγματοληψία της κατά τον παράγοντα d , δηλαδή η εικόνα που προκύπτει κρατώντας ένα σημείο κάθε d σημεία, για κάθε άξονα. Θεωρούμε ότι οι ακέραιοι N_x, N_y είναι πολλαπλάσια του d , οπότε η $\tilde{w}(n_1, n_2)$ έχει έκταση $M_x \times M_y$, με $M_x = N_x/d$, $M_y = N_y/d$. Τότε, οι Διακριτοί Μετασχηματισμοί Fourier \tilde{W} και W των $\tilde{w}(n_1, n_2)$ και $w(n_1, n_2)$ αντίστοιχα, συνδέονται μέσω της σχέσης (στην περίπτωση που οι M_x, M_y, d

είναι περιττοί):

$$\tilde{W}(m_1, m_2) = \frac{1}{d^2} \sum_{\ell_1=-(d-1)/2}^{(d-1)/2} \sum_{\ell_2=-(d-1)/2}^{(d-1)/2} W(m_1 + \ell_1 M_x, m_2 + \ell_2 M_y) \quad (5.76)$$

Θα χρησιμοποιήσουμε την παραπάνω ιδιότητα για την έκφραση της σχέσης (5.72) στη μορφή που μας χρειάζεται εδώ. Πριν από αυτό όμως, πρέπει να ασχοληθούμε με τον όρο $s * u_m$. Θεωρούμε ότι η συνέλιξη “*” είναι κυκλική συνέλιξη, που ισοδυναμεί με το να θεωρήσουμε ότι ο πυρήνας s έχει ίδια έκταση $N_x \times N_y$ με την u καθώς επίσης ότι οι s και u επεκτείνονται περιοδικά εκτός του $N_x \times N_y$ πλέγματος. Επειδή, όπως έχουμε θεωρήσει, ο s έχει μη μηδενικές τιμές μόνο στο τετράγωνο $d \times d$, γειμίζουμε όλες τις υπόλοιπες τιμές του με 0, ώστε να αποκτήσει έκταση $N_x \times N_y$.

Επιπρόσθετα, πρέπει να προσέξουμε ότι, για να μπορέσει να εφαρμοστεί σωστά η (5.76), για κάθε μπλοκ μεγέθυνσης, το αποτέλεσμα της συνέλιξης που κρατάμε μετά την διακριτοποίηση πρέπει να βρίσκεται στο 1^ο στοιχείο του μπλοκ ($i = 0, j = 0$, στην τοπική αρίθμηση του μπλοκ). Αν λάβουμε υπόψιν και την περιστροφή που εφαρμόζεται στον πυρήνα s πριν την συνέλιξη, το παραπάνω συνεπάγεται ότι τα μη μηδενικά στοιχεία του s που αντιστοιχούν στο μπλοκ πρέπει να βρίσκονται στις πρώτες $d \times d$ τιμές με $i \leq 0, j \leq 0$. Έτσι, για να απλοποιηθεί η έκφραση του πυρήνα, χρησιμοποιούμε συσχέτιση αντί για συνέλιξη, με τον αντίστοιχο πυρήνα συσχέτισης, έστω \tilde{s} , να προκύπτει από περιστροφή του πυρήνα συνέλιξης s , δηλαδή $\tilde{s}(n_1, n_2) = s(-n_1, -n_2)$. Άρα, ο \tilde{s} έχει τα μη μηδενικά του στοιχεία στις πρώτες $d \times d$ τιμές με $i \geq 0, j \geq 0$, δηλαδή στην αρχή του πλέγματος της εικόνας u . Ο πίνακας του πυρήνα συσχέτισης προσφέρει εκτός των άλλων και καλύτερη εποπτεία, αφού στην συσχέτιση, κάθε τιμή του πολλαπλασιάζεται με την τιμή του εκάστοτε μπλοκ μεγέθυνσης που βρίσκεται στην ίδια ακριβώς θέση (σε αντίθεση με την συσχέτιση, όπου χρειάζεται και περιστροφή).

Σημειώνουμε ότι, όπως προκύπτει από τα παραπάνω, ο $d \times d$ πίνακας $\tilde{\mathbf{S}}$ που περιγράφει τον πυρήνα συσχέτισης \tilde{s} , αφού περιέχει τα μη μηδενικά του στοιχεία, είναι ουσιαστικά ο ανάστροφος του αντίστοιχου πίνακα \mathbf{S} του πυρήνα συνέλιξης s , δηλαδή $\tilde{\mathbf{S}} = \mathbf{S}^T$. Όπως είναι προφανές, για να είναι αξιόπιστα τα αποτελέσματα οποιασδήποτε μεθόδου Παρεμβολής, επιθυμούμε η εξομάλυνση της εικόνας u να είναι συμμετρική, άρα σε όλες σχεδόν τις περιπτώσεις, ο πυρήνας συνέλιξης s είναι συμμετρικός, κάτι το οποίο συνεπάγεται ότι οι $\tilde{\mathbf{S}} = \mathbf{S}$. Έτσι, για να απλοποιηθούν οι υπολογισμοί, η μέθοδος που υλοποιήσαμε δέχεται σαν είσοδο τον πίνακα συσχέτισης, ο οποίος προκύπτει άμεσα, από αναστροφή του πίνακα συνέλιξης και συνήθως ταυτίζεται με αυτόν (λόγω συμμετρίας).

Με τη χρήση λοιπόν του πυρήνα συσχέτισης, η συνθήκη αναστροφής

(5.72) μπορεί να γραφεί και ως εξής:

$$z_m = Q(\tilde{s} * u_m), \quad m = 1, \dots, N \quad (5.77)$$

όπου με “ $\tilde{*}$ ” συμβολίζουμε την πράξη της συσχέτισης.

Λόγω της σχέσης (5.76), η συνθήκη (5.77) γράφεται στο πεδίο των συχνοτήτων ως εξής, για κάθε $m = 1, \dots, N$:

$$Z_m(m_1, m_2) = \frac{1}{d^2} \sum_{-(d-1)/2}^{(d-1)/2} \sum_{-(d-1)/2}^{(d-1)/2} (\tilde{S}(m_1 + \ell_1 M_x, m_2 + \ell_2 M_y))^* U_m(m_1 + \ell_1 M_x, m_2 + \ell_2 M_y) \quad (5.78)$$

όπου με a^* , για a μιγαδικό αριθμό, συμβολίζουμε το συζυγή μιγαδικό του a .

Σημειώνουμε ότι, επειδή $\tilde{s}(n_1, n_2) = s(-n_1, -n_2)$, οι \tilde{S} και S συνδέονται μέσω της σχέσης: $\tilde{S}(m_1, m_2) = (S(m_1, m_2))^*$.

Μέχρι το σημείο αυτό, όσα παρουσιάστηκαν ισχύουν γενικά για τη συνθήκη αναστρεψιμότητας. Όσον αφορά τώρα τη μέθοδο Zero-Padding, κάθε κανάλι u_m της λύσης της μεθόδου αυτής έχει φασματικό περιεχόμενο $U_m(m_1, m_2)$ που μηδενίζεται στο υπόλοιπο πεδίο ορισμού του B_U , εκτός της 2Δ ζώνης συχνοτήτων B_Z , που αντιστοιχεί στο πεδίο ορισμού του $Z_m(m_1, m_2)$. Αν αυτό ισχύει και επιπλέον θεωρήσουμε ότι, για $(m_1, m_2) \in B_Z$, ισχύει $S(m_1, m_2) \neq 0$ τότε και $\tilde{S}(m_1, m_2) \neq 0$, τότε από τη σχέση (5.78) προκύπτει:

$$U_m(m_1, m_2) = \begin{cases} d^2 \frac{Z_m(m_1, m_2)}{(\tilde{S}(m_1, m_2))^*} & , (m_1, m_2) \in B_Z \\ 0 & , (m_1, m_2) \in B_U \setminus B_Z \end{cases}, \quad m = 1, \dots, N \quad (5.79)$$

όπου τα σύνολα B_Z, B_U έχουν οριστεί στην (5.75).

Επομένως, μέσω της σχέσης (5.79), καταλήγουμε άμεσα στη λύση της μεθόδου Zero-Padding:

$$u_m(n_1, n_2) = \mathcal{DF}^{-1} \{U_m(m_1, m_2)\} \quad (5.80)$$

όπου $\mathcal{DF}^{-1}(\cdot)$ ο αντίστροφος Διακριτός Μετασχηματισμός Fourier.

5.4.2 Υπολογισμός της προβολής $P_{u_{0,s}}(\cdot)$

Εφόσον η ταχύτητα \mathbf{v} είναι διακριτή, στην σχέση (5.74), που ορίζει τον χώρο $\mathbf{u}_{0,s}$, έχουμε διακριτή συνέλιξη της u_m με τον πυρήνα που προκύπτει από απλή δειγματοληψία του συνεχούς πυρήνα s στο πυκνότερο πλέγμα. Έτσι, για κάθε μπλοκ μεγέθυνσης (n_1, n_2) και για κάθε κανάλι m , έχουμε:

$$(s * v_m)(n_1 d, n_2 d) = 0, \quad (n_1, n_2) \in \{0, \dots, M_x - 1\} \times \{0, \dots, M_y - 1\}, \quad m = 1, \dots, N$$

Αν χρησιμοποιήσουμε και πάλι τον πυρήνα συσχέτισης \tilde{s} , που παρουσιάσαμε στην προηγούμενη υποενότητα, η παραπάνω σχέση γράφεται και σε μορφή συσχέτισης ως εξής:

$$(\tilde{s} * v_m)(n_1 d, n_2 d) = 0, (n_1, n_2) \in \{0, \dots, M_x - 1\} \times \{0, \dots, M_y - 1\}, m = 1, \dots, N \quad (5.81)$$

Χάρη στην υπόθεση που έχουμε κάνει, ότι ο πυρήνας s περιορίζεται στην έκταση ενός μπλοκ μεγέθυνσης, οι εξισώσεις (5.81) για κάθε μπλοκ (n_1, n_2) αποσυζεύγγονται, δηλαδή κάθε τιμή $v_m(n_1, n_2)$ της ταχύτητας \mathbf{v} συμμετέχει σε μία μόνο εξίσωση. Έτσι, κάθε μπλοκ τιμών της v_m πρέπει να ικανοποιεί τη σχέση (5.81) ανεξάρτητα από τα άλλα και οι υπολογισμοί απλοποιούνται.

Η σχέση (5.81) μπορεί να γραφεί χρησιμοποιώντας μόνο διανύσματα. Για τον σκοπό αυτό, θεωρούμε την αναπαράσταση κάθε καναλιού v_m της ταχύτητας \mathbf{v} σαν ένα διάνυσμα \underline{v}_m , διάστασης $N_x N_y$, που αποτελείται από υποδιανύσματα $\underline{v}_m^{n_1, n_2}$ διάστασης d^2 . Κάθε υποδιάνυσμα $\underline{v}_m^{n_1, n_2}$ αντιστοιχεί στο αντίστοιχο pixel (n_1, n_2) της αρχικής εικόνας. Για το σχηματισμό κάθε τέτοιου υποδιανύσματος $\underline{v}_m^{n_1, n_2}$ χρησιμοποιούμε μία μέθοδο σάρωσης των στοιχείων του μπλοκ μεγέθυνσης. Αντίστοιχα, για το σχηματισμό του διανύσματος \underline{v}_m από τα υποδιανύσματα $\underline{v}_m^{n_1, n_2}$, σαρώνουμε τα μπλοκ μεγέθυνσης με την ίδια μέθοδο σάρωσης. Για παράδειγμα, μπορούμε να χρησιμοποιήσουμε σάρωση γραμμής, δηλαδή από αριστερά προς τα δεξιά και ύστερα από πάνω προς τα κάτω.

Επίσης, αναπαριστούμε σαν διάνυσμα και τον πυρήνα συσχέτισης \tilde{s} , που όπως υποθέσαμε, περιορίζεται στην έκταση ενός μπλοκ μεγέθυνσης. Έτσι, το αντίστοιχο διάνυσμα $\underline{\tilde{s}}$ έχει διάσταση d^2 και θεωρούμε ότι προκύπτει από τα στοιχεία του πυρήνα, με την ίδια σάρωση όπως και τα διανύσματα $\underline{v}_m^{n_1, n_2}$.

Χρησιμοποιώντας λοιπόν τις παραπάνω αναπαραστάσεις, από τη σχέση (5.81) προκύπτει άμεσα ότι, για κάθε μπλοκ μεγέθυνσης, ισχύει:

$$\underline{\tilde{s}}^T \underline{v}_m^{n_1, n_2} = 0, (n_1, n_2) \in \{0, \dots, M_x\} \times \{0, \dots, M_y\}, m = 1, \dots, N \quad (5.82)$$

Οι παραπάνω σχέσεις είναι γραμμικές και σημαίνουν ότι κάθε υποδιάνυσμα $\underline{v}_m^{n_1, n_2}$ πρέπει να ανήκει στο υπερεπίπεδο του \mathbb{R}^{d^2} , που διέρχεται από την αρχή των αξόνων και είναι κάθετο στο διάνυσμα $\underline{\tilde{s}}$ του πυρήνα. Συμβολίζουμε με $\mathcal{U}_{0,s}^b$ το υπερεπίπεδο αυτό, δηλαδή τον χώρο των διανυσμάτων του \mathbb{R}^{d^2} (που αντιστοιχούν στις τιμές σε ένα μπλοκ μεγέθυνσης) που ικανοποιούν τη σχέση:

$$\underline{\tilde{s}}^T \underline{x} = 0 \quad (5.83)$$

Έτσι, για να ανήκει η ταχύτητα \mathbf{v} στον $\mathcal{U}_{0,s}$, πρέπει όλα τα υποδιανύσματα $\underline{v}_m^{n_1, n_2}$, για κάθε κανάλι της, m και κάθε μπλοκ μεγέθυνσης (n_1, n_2) , να ανήκουν στο $\mathcal{U}_{0,s}^b$. Αυτό συνεπάγεται ότι η προβολή της \mathbf{v} στον $\mathcal{U}_{0,s}$ γίνεται μέσω της προβολής κάθε υποδιανύσματος $\underline{v}_m^{n_1, n_2}$ στον $\mathcal{U}_{0,s}^b$.

Όπως είναι γνωστό από την αναλυτική γεωμετρία, η προβολή του \underline{x} στο υπερεπίπεδο $\mathcal{U}_{0,s}^b$, που ορίζεται από την (5.83), γίνεται μέσω της αφαίρεσης της συνιστώσας του \underline{x} που είναι κάθετη στο υπερεπίπεδο αυτό. Άρα, η $P_{\mathcal{U}_{0,s}^b}(\cdot)$ δίνεται από τη σχέση:

$$P_{\mathcal{U}_{0,s}^b}(\underline{x}) = \underline{x} - \frac{\tilde{\underline{s}}^T \underline{x}}{\tilde{\underline{s}}^T \tilde{\underline{s}}} \tilde{\underline{s}} \quad (5.84)$$

Με βάση τη τελευταία σχέση, η προβολή $P_{\mathcal{U}_{0,s}^b}(\underline{v})$ προκύπτει τελικά εφαρμόζοντας την προβολή $P_{\mathcal{U}_{0,s}^b}(\underline{v}_m^{n_1, n_2})$ για κάθε $m = 1, \dots, N$ και $(n_1, n_2) \in \{0, \dots, M_x\} \times \{0, \dots, M_y\}$.

Σύμφωνα με όλα τα παραπάνω, καταλήγουμε τελικά στον αλγόριθμο που υλοποιήσαμε: βλ. πίνακα 5.13.

ΕΙΣΟΔΟΙ: αρχική εικόνα \mathbf{z} (έκτασης $M_x \times M_y$, με N κανάλια), παράγοντας μεγέθυνσης d , πίνακας συσχέτισης $\tilde{\mathbf{S}}$, αριθμός επαναλήψεων N_{it} και άλλες παράμετροι για κάθε συγκεκριμένη μέθοδο ΜΔΕ (πρέπει M_x, M_y, d περιττοί).

Εύρεση διαστάσεων ζητούμενης εικόνας: $N_x = dM_x$; $N_y = dM_y$

Εύρεση των πεδίων ορισμού συχνοτήτων B_U, B_Z από τις (5.75)

ΑΡΧΙΚΟΠΟΙΗΣΗ ΜΕ ΠΑΡΕΜΒΟΛΗ ZERO PADDING

Βρες το σήμα $\tilde{s}(n_1, n_2)$, γειμίζοντας με 0 τον πίνακα $\tilde{\mathbf{S}}$ ώστε να αποκτήσει μέγεθος $N_x \times N_y$.

$\tilde{\mathbf{S}}(m_1, m_2) = \mathcal{DF}\{\tilde{s}(n_1, n_2)\}$

For $m = 1, \dots, N$

$Z_m(m_1, m_2) = \mathcal{DF}\{z_m(n_1, n_2)\}$

For each $(m_1, m_2) \in B_U$

If $(m_1, m_2) \in B_Z$

$U_m^0(m_1, m_2) = d^2 Z_m(m_1, m_2) / \left(\tilde{\mathbf{S}}(m_1, m_2)\right)^*$

else

$U_m^0(m_1, m_2) = 0$

End if

End for

$u_m^0(n_1, n_2) = \mathcal{DF}^{-1}\{U_m^0(m_1, m_2)\}$

End for

$\mathbf{u} = \mathbf{u}^0$

ΕΦΑΡΜΟΓΗ ΜΔΕ ΔΙΑΧΥΣΗΣ ΜΕ ΠΡΟΒΟΛΗ

Βρες το διάνυσμα $\tilde{\mathbf{x}}$ που προκύπτει από τον πίνακα $\tilde{\mathbf{S}}$, με κάποια σάρωση.

For $n = 1, \dots, N_{it}$

Υπολογισμός της ταχύτητας \mathbf{v} με τα ίδια ακριβώς βήματα, όπως στον αλγόριθμο της αντίστοιχης ΜΔΕ (βλ. πίνακες 5.5, 5.7)

Προβολή της \mathbf{v} στον υποχώρο $\mathbf{U}_{0,s}$

For $m = 1, \dots, N$

For each block $(i, j) \in \{0, \dots, M_x - 1\} \times \{0, \dots, M_y - 1\}$

Πάρε τις τιμές του block $v_m(id + i', jd + j')$, $(i', j') \in \{0, \dots, d - 1\}^2$ και γράψτε σαν διάνυσμα \underline{v}_m^{m,m_2} , με την ίδια σάρωση όπως για το $\tilde{\mathbf{x}}$

Ανανέωσε τις τιμές της v_m στο block αυτό, ως εξής:

$$\underline{v}_m^b = \underline{v}_m^b - \frac{\tilde{\mathbf{S}}^T \underline{v}_m^b}{\tilde{\mathbf{S}}^T \tilde{\mathbf{x}}}$$

End for

End for

Υπολογισμός της επόμενης εικόνας

$\mathbf{u} = \mathbf{u} + \tau \mathbf{v}$

End for

ΕΞΟΔΟΣ: μεγεθυμένη εικόνα \mathbf{u}

Πίνακας 5.13: Αλγόριθμος της Γενικής Μεθόδου με ΜΔΕ διάχυσης, που προτείνουμε για την Παρεμβολή διανυσματικών εικόνων

Κεφάλαιο 6

Πειράματα και Συγκρίσεις

Στο κεφάλαιο αυτό, παρουσιάζουμε και αναλύουμε διάφορα παραδείγματα, που προέκυψαν από την εφαρμογή των αλγορίθμων που παρουσιάσαμε στο κεφάλαιο 5 σε διάφορες εικόνες, βαθμωτές και κυρίως έγχρωμες. Αναφέρουμε ότι για την υλοποίηση των αλγορίθμων και την διεξαγωγή των πειραμάτων, χρησιμοποιήσαμε την γλώσσα Matlab (version 6.5).

Σημειώνουμε επίσης ότι, σε όλα τα παραδείγματα με έγχρωμες εικόνες, χρησιμοποιούμε για την επεξεργασία τους, την αναπαράστασή τους στον χώρο χρώματος (R, G, B) , με τιμές μεταξύ 0 και 255 .

6.1 Πειράματα για το πρόβλημα Βελτίωσης Εικόνων

6.1.1 Παράδειγμα 1

Στο παράδειγμα αυτό, χρησιμοποιούμε μία θορυβώδη εκδοχή της έγχρωμης εικόνας του σχήματος 6.1a (που προέρχεται από τη βάση εικόνων του Πανεπιστημίου του Berkeley). Ο θόρυβος αυτός προκύπτει ως εξής: Σε κάθε κανάλι ($m = 1(\text{red}), 2(\text{green})$ ή $3(\text{blue})$), ανεξάρτητα από τα υπόλοιπα, προσθέτουμε θόρυβο Gaussian κατανομής, τυπικής απόκλισης $\sigma = 20$ (όπου, όπως είπαμε, οι τιμές της εικόνας κειμένονται στο διάστημα $[0, 255]$). Εφαρμόσαμε τις διάφορες μεθόδους ΜΔΕ Βελτίωσης σε αυτήν θορυβώδη εικόνα και τα αποτελέσματα φαίνονται στο σχήμα 6.2.

Η διαδικασία που ακολουθήσαμε για την εξαγωγή των αποτελεσμάτων αυτών είναι η ακόλουθη: Για κάθε μέθοδο, τρέξαμε τον αλγόριθμο που υλοποιήσαμε για διάφορες τιμές των παραμέτρων και για επαρκή αριθμό επαναλήψεων, σώζοντας τα ενδιαμέσα βήματα. Συγκρίνοντας όλες τις εξόδους, επιλέξαμε εκείνες τις παραμέτρους που αντιστοιχούσαν στην “καλύτερη” εικόνα (σύμφω-



(a)

(b)

Σχήμα 6.1: Η εικόνα που χρησιμοποιήθηκε στο Παράδειγμα 1. (a) Αρχική, καθαρή εικόνα (481×321 pixels). (b) Αρχική εικόνα, με προσθετικό (σε κάθε κανάλι χωριστά) Gaussian θόρυβο, τυπικής απόκλισης $\sigma = 20$.



(a)



(b)

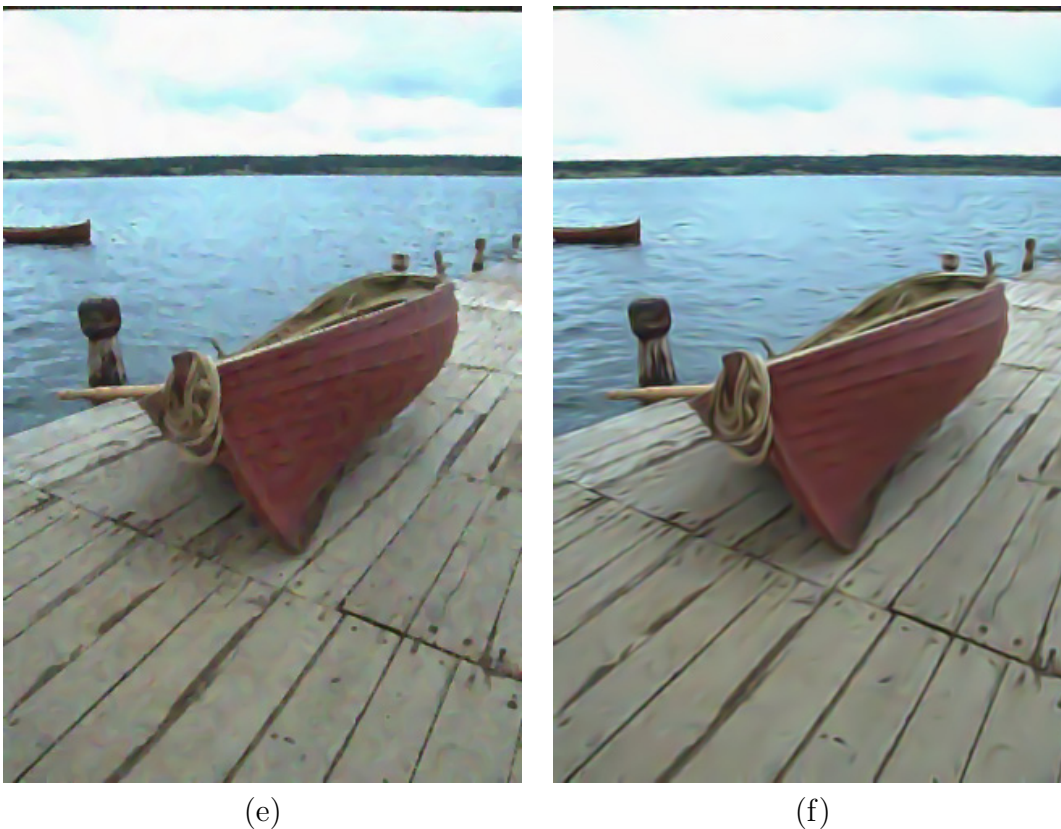


(c)



(d)

(συνέχεια στην επόμενη σελίδα)



(e)

(f)

Σχήμα 6.2: Σύγκριση των μεθόδων βελτίωσης, με εφαρμογή τους στην έγχρωμη εικόνα του σχήματος 6.1b (a) Ομαλοποιημένη $P-M$ μέθοδος, γενικευμένη για έγχρωμες εικόνες (10 επαναλήψεις, $\tau = 0.2$, $\sigma = 0.5$, $K = 20$, $g(s)$ από σχέση (5.3)). (b) *Vectorial TV PDE*, με ϵ -εξομάλυνση (100 επαναλήψεις, $\tau = 0.2$, $\epsilon = 1$). (c) Διάχυση Βελτίωσης της Συνοχής (70 επαναλήψεις, $\tau = 0.2$, $\sigma = 0$, $\rho = 1.5$). (d) *Beltrami Flow* (500 επαναλήψεις, $\tau = 0.2$). (e) Διάχυση στην Κατεύθυνση θ_- (200 επαναλήψεις, $\tau = 0.2$). (f) Ανισοτρ. Εξομάλυνση των *Tschump.-Deriche* (190 επαναλήψεις, $\beta_{\max} = 5$, $\rho = 1.5$).

να με την υποκειμενική ανθρώπινη αντίληψη). Κριτήριο για την επιλογή του αριθμού των επαναλήψεων ήταν ο θόρυβος να έχει απομακρυνθεί όσο το δυνατόν περισσότερο και η εικόνα να έχει θολώσει όσο το δυνατόν λιγότερο.

Παρατηρώντας, λοιπόν τα αποτελέσματα του σχήματος 6.2, καταλήγουμε στα ακόλουθα συμπεράσματα για κάθε μέθοδο:

1. *Ομαλοποιημένη μέθοδος των Perona-Malik, γενικευμένη για διανυσματικές εικόνες* (βλ. ενότ. 1.2,1.3):

Παρατηρούμε ότι η μέθοδος αυτή δίνει ένα αρκετά ικανοποιητικό αποτέλεσμα και μάλιστα πολύ πιο γρήγορα σε σχέση με όλες τις υπόλοιπες μεθόδους (μέσα σε 10 μόνο επαναλήψεις, με ένα τυπικό βήμα $\tau = 0.2$). Αυτό όμως έχει ως αποτέλεσμα την απομάκρυνση του θορύβου με κάπως “απότομο” τρόπο, αφού κάποιες λεπτές ακμές θολώνονται (βλ. π.χ. τις οριζόντιες ακμές στην επιφάνεια της βάρκας), ενώ οι ομογενείς περιοχές εξομαλύνονται υπερβολικά, δίνοντας στο αποτέλεσμα μία αρκετά αφύσικη όψη (το οποίο γίνεται ακόμα πιο έντονο μετά από λίγο μεγαλύτερο αριθμό επαναλήψεων από αυτόν που παρουσιάζουμε).

2. *Μέθοδος Vectorial TV με ϵ -εξομάλυνση* (βλ. ενότ. 1.6):

Στα πειράματα, παρατηρήσαμε ότι η μέθοδος αυτή καταφέρνει να συντηρήσει τις ακμές της εικόνας για αρκετή ώρα, κατά τη χρονική ροή. Παράλληλα, όμως διατηρεί και αρκετό μέρος του θορύβου, ακόμα και σε περιοχές μακριά από ακμές. Αυτό οφείλεται στο ότι δεν γίνεται κάποια εξομάλυνση στην εικόνα (όπως π.χ. στην Ομαλοποιημένη Μέθοδο των Perona-Malik) και η διάχυση επιβραδύνεται σημαντικά από τον υψίσυχνο θόρυβο. Έτσι, για να προκύψει μία εικόνα στην οποία ο θόρυβος να έχει απομακρυνθεί ικανοποιητικά, αναγκάζομαστε να συνεχίσουμε τη διάχυση για σχετικά μεγάλο διάστημα (σε σύγκριση, για παράδειγμα, με την προηγούμενη μέθοδο των Perona-Malik, που για το ίδιο χρονικό βήμα $\tau = 0.2$, χρειάζονται μόνο 10 επαναλήψεις έναντι 100, στην περίπτωση αυτή). Αυτό όμως έχει ως αποτέλεσμα οι διανυσματικές ακμές να θολώνουν και σε κάποια σημεία να καταστρέφονται (βλ. π.χ. στην εικόνα τις ξύλινες σανίδες του δαπέδου ή τις οριζόντιες ακμές στην επιφάνεια της βάρκας). Τέλος, όσον αφορά την περιοχή της θάλασσας, όπου υπάρχει υφή, η επιβράδυνση από τις απότομες μεταβολές, που είναι αρνητική για την απομάκρυνση του θορύβου, έχει εδώ το θετικό αποτέλεσμα της διατήρησης ενός σημαντικού μέρους της υφής αυτής.

3. *Διάχυση Βελτίωσης της Συνοχής (Coherence Enhancing Diffusion) διανυσματικών εικόνων* (βλ. ενότ. 1.7.1):

Όπως έχουμε πει, η μέθοδος αυτή έχει σχεδιαστεί για τη βελτίωση της συνοχής εικόνων (δηλαδή τον τονισμό των μονοδιάστατων δομών), αλλά την εφαρμόσαμε και εδώ, για την απομάκρυνση του θορύβου, με σκοπό τη σύγκριση των διαδικασιών διάχυσης (για μια πιο κατάλληλη εφαρμογή της μεθόδου αυτής, βλ. το παράδειγμα 3). Παρατηρούμε λοιπόν ότι η μέθοδος καταφέρνει να απομακρύνει τελείως τον θόρυβο, αλλά παράλληλα αλλοιώνει σημαντικά την εικόνα, δίνοντάς της μια όψη “ιμπρεσιονιστικού πίνακα”. Πιο συγκεκριμένα, συντηρεί όλες τις ακμές, ακόμα και τις πιο λεπτές (βλ. π.χ. τις οριζόντιες ακμές στην επιφάνεια της βάρκας), και αρκετές από αυτές τις επεκτείνει (βλ. π.χ. τη θάλασσα και τις περιοχές στο εσωτερικό των σανίδων του δαπέδου), δίνοντας έτσι την αφύσικη αυτή όψη. Σε αυτή την απομάκρυνση του θορύβου συντελεί η εξομάλυνση του δομικού τανυστή που επιτελεί η μέθοδος. Επιπλέον, η διάχυση είναι έντονα ανισοτροπική, ακόμα και σε περιοχές με μικρές μεταβολές (π.χ. στο εσωτερικό των ξύλινων σανίδων του δαπέδου) και κατευθύνεται από τις τοπικές μονοδιάστατες δομές.

4. *Poř Beltrami (Beltrami Flow)* (βλ. ενότ. 1.8):

Η μέθοδος αυτή καταφέρνει να συντηρήσει τις ακμές της εικόνας για την περισσότερη ώρα, σε σχέση με όλες τις υπόλοιπες μεθόδους, (παρουσιάζοντας μια συμπεριφορά όπως αυτή που αναφέραμε για την μέθοδο Vectorial TV, αλλά σε πολύ πιο έντονο βαθμό). Αυτό οφείλεται στον πολλαπλασιαστικό όρο $1/\sqrt{\det(I + G)}$ της ταχύτητας, ο οποίος ελαττώνει πολύ γρήγορα τη διάχυση κοντά στις ακμές. Παράλληλα, όμως διατηρείται και υπερβολικά μεγάλο μέρος του θορύβου για μεγάλο διάστημα, λόγω του ότι ούτε στη μέθοδο αυτή γίνεται κάποια εξομάλυνση (στην εικόνα ή στον τανυστή διάχυσης). Έτσι, αναγκάζομαστε να συνεχίσουμε τη διάχυση για τον περισσότερο χρόνο, σε σχέση με όλες τις υπόλοιπες μεθόδους (500 επαναλήψεις, με το τυπικό βήμα $\tau = 0.2$). Μετά από τόσο χρόνο όμως, οι δομές της εικόνας έχουν εξομαλυνθεί υπερβολικά, ενώ ο θόρυβος δεν έχει εξαφανιστεί τελείως (βλ. π.χ. τις περιοχές στο εσωτερικό των σανίδων του δαπέδου). Η μέθοδος αυτή δίνει το χειρότερο αποτέλεσμα στο παράδειγμα αυτό, κι αυτό διότι η μέθοδος καταφέρνει να λειτουργήσει ικανοποιητικά μόνο σε περιπτώσεις θορύβου μικρότερης ισχύος.

5. *Διάχυση Διανυσματικών Εικόνων στην Κατεύθυνση θ_-* (βλ. ενότ. 1.9):

Παρατηρούμε ότι στην συγκεκριμένη εικόνα, το αποτέλεσμα της μεθόδου αυτής έχει παρόμοια ποιότητα με αυτό της μεθόδου Vectorial TV. Η μέθοδος καταφέρνει να συντηρήσει τις πιο έντονες ακμές, αλλά καταστρέφει

τοπικά κάποιες λεπτές ακμές (βλ. π.χ. κάποια σημεία στις οριζόντιες ακμές στην επιφάνεια της βάρκας), ενώ δεν μπορεί να απομακρύνει τελείως τον θόρυβο (βλ. π.χ. το ξύλινο δάπεδο). Τα παραπάνω οφείλονται στο ότι η εξομάλυνση γίνεται μόνο κατά την διεύθυνση της ελάχιστης μεταβολής, οπότε παρουσιάζεται επιθυμητή συμπεριφορά κοντά στις ισχυρές ακμές, αλλά σε περιοχές όπου υπάρχουν λεπτές ακμές ή δεν υπάρχουν καθόλου, η διάχυση κατευθύνεται από τον θόρυβο.

6. *Ανισοτροπική Εξομάλυνση των Tschumperlé-Derliche* (βλ. ενότητα 1.10):

Όπως παρατηρούμε, το αποτέλεσμα της μεθόδου αυτής είναι το πιο ικανοποιητικό (για την ανθρώπινη αντίληψη τουλάχιστον) αποτέλεσμα σε σχέση με όλα τα υπόλοιπα. Ο θόρυβος έχει τελείως απομακρυνθεί, ενώ οι περισσότερες ακμές της εικόνας έχουν συντηρηθεί. Αυτό επιτυγχάνεται χάρη στην εξομάλυνση της μεθόδου, η οποία είναι ανισοτροπική κοντά στις ακμές και ισοτροπική στις ομογενείς περιοχές, λαμβάνοντας υπόψιν τις μεταβολές της εικόνας με έναν συνεκτικό τρόπο (λόγω της εξομάλυνσης στον δομικό τανυστή). Επίσης, παρατηρούμε ότι η μέθοδος εντοπίζει επιτυχώς και τις διανυσματικές γωνίες και καταφέρνει να τις συντηρήσει ενώ παράλληλα να απομακρύνει τον θόρυβο στην κοντινή τους περιοχή (βλ. π.χ. τις γωνίες στις ξύλινες σανίδες του δαπέδου).

Πρέπει όμως να παρατηρήσουμε επίσης ότι, ακόμα και αυτό το αποτέλεσμα παρουσιάζει κάποια προβλήματα. Πρώτα από όλα, η μέθοδος έχει καταστρέψει το μεγαλύτερο μέρος της υφής της αρχικής εικόνας, όπως στη θάλασσα και στις ξύλινες σανίδες του δαπέδου. Αυτό το πρόβλημα παρουσιάζεται σε όλες τις μεθόδους ΜΔΕ διάχυσης της κατηγορίας που μελετάμε στην εργασία αυτή και οφείλεται στο ότι δεν μπορούν να ξεχωρίσουν την υφή από τον θόρυβο (αφού και τα δύο παρουσιάζουν μεγάλη τοπική μεταβολή). Επίσης, η μέθοδος αυτή έχει καταστρέψει τις πιο λεπτές από τις οριζόντιες ακμές στην επιφάνεια της βάρκας και έχει αλλοιώσει τις μικρές δομές, στην περιοχή πίσω από τη βάρκα. Αυτό οφείλεται στην εξομάλυνση του δομικού τανυστή, που ενώ από τη μία δίνει όπως είπαμε επιθυμητή συμπεριφορά για τις πιο σημαντικές δομές της εικόνας, από την άλλη κάνει τη διάχυση να μην λαμβάνει υπόψιν της μικρές δομές (μεγέθους της τάξης της σταθεράς εξομάλυνσης ρ), με αποτέλεσμα να τις καταστρέφει ή να τις αλλοιώνει με αφύσικο τρόπο.

Αξιολόγηση των αποτελεσμάτων του Παραδείγματος 1 με Σηματοθορυβικούς λόγους (SNR)

Για την αξιολόγηση των αποτελεσμάτων, υπολογίσαμε επιπρόσθετα τους Σηματοθορυβικούς λόγους (SNR). Έτσι, αν με \mathbf{u}^0 συμβολίσουμε την αρχική

καθαρή εικόνα (βλ. σχήμα 6.1a) και με \mathbf{u} το αποτέλεσμα κάποιας μεθόδου, υπολογίσαμε τον SNR μέσω της σχέσης:

$$\text{SNR} = 10 \log \left(\frac{\sum_{i=1}^{N_x} \sum_{j=1}^{N_x} \|\mathbf{u}^0(i, j) - \overline{\mathbf{u}^0}\|^2}{\sum_{i=1}^{N_x} \sum_{j=1}^{N_x} \|\mathbf{u}(i, j) - \mathbf{u}^0(i, j)\|^2} \right) \quad (6.1)$$

όπου N_x, N_y είναι οι χωρικές διαστάσεις των εικόνων και $\log(\cdot)$ ο λογάριθμος με βάση το 10. Επίσης, με $\overline{\mathbf{u}^0}$ συμβολίζουμε την μέση τιμή της εικόνας \mathbf{u}^0 σε όλη την έκταση της εικόνας.

Χρησιμοποιήσαμε επίσης τον Average Local SNR, ο οποίος προκύπτει ως εξής: Χωρίζουμε τα pixels της εικόνας σε μπλοκ έκτασης $b_x \times b_y$ ¹ και υπολογίζουμε χωριστά σε κάθε τέτοιο μπλοκ τον αντίστοιχο SNR (έστω SNR_{blk}). Τότε, ο Average Local SNR προκύπτει σαν ο μέσος όρος των παραπάνω, δηλαδή:

$$(\text{Average Local SNR}) = \overline{\text{SNR}_{\text{blk}}} \quad (6.2)$$

Διευκρινίζουμε ότι όσο μεγαλύτερος είναι ο SNR (ή ο Average Local SNR), τόσο λιγότερο (με την έννοια της ισχύος) θόρυβο περιέχει η εικόνα \mathbf{u} . Σημειώνουμε επίσης ότι ο Average Local SNR παρέχει μία μέτρηση, η οποία είναι πιο κοντά, σε σχέση με τον απλό SNR, στην υποκειμενική αντίληψη του ανθρώπου για την σχέση σήματος προς θόρυβο σε μία εικόνα.

Οι παραπάνω μετρήσεις για τα αποτελέσματα του σχήματος 6.2 φαίνονται στον πίνακα 6.2.

Από τον πίνακα 6.2, φαίνεται ότι όλες οι μέθοδοι έχουν βελτιώσει τους SNR της αρχικής θορυβώδους εικόνας. Ο μεγαλύτερος SNR αλλά και Average Local SNR επιτυγχάνεται με τη μέθοδο *Vectorial TV*. Βλέπουμε λοιπόν ότι από τις μετρήσεις αυτές δεν εξάγονται τα ίδια συμπεράσματα, όσον αφορά τη σύγκριση των αποτελεσμάτων των μεθόδων, που εξάγονται από την παρατήρηση των εικόνων. Αυτό οφείλεται στο ότι οι σηματοθορυβικοί λόγοι αυτοί δεν καταφέρνουν να αντικατοπτρίσουν την “ποσότητα” της βελτίωσης που ψυχολογικά γίνεται αντιληπτή από τον ανθρώπινο παρατηρητή.

Για παράδειγμα, η μέθοδος της *Ανισοτροπικής Εξομάλυνσης των Tschumperlé-Derliche* (βλ. σχήμα 6.2f), που δίνει όπως είπαμε το πιο ικανοποιητικό για την ανθρώπινη αντίληψη αποτέλεσμα, καταφέρνει να απομακρύνει σε πολύ μεγαλύτερο βαθμό τις διαταραχές του θορύβου, σε σχέση με την μέθοδο

¹ στην περίπτωση που τα μπλοκ αυτά δεν μπορούν να καλύψουν όλη την έκταση της εικόνας (υπάρχει κάποιο υπόλοιπο), θεωρούμε μία ελαφρώς μικρότερη εκδοχή της εικόνας, που να την καλύπτουν πλήρως (έκταση $(N_x - \text{mod}(N_x, b_x)) \times (N_y - \text{mod}(N_y, b_y))$).

| Μέθοδος | Σχήμα | SNR (dB) | Aver.Loc.SNR, με 40×40 μπλοκ (dB) |
|-----------------------------------|-------|----------|--------------------------------------|
| (Αρχική με θόρυβο) | 6.1b | 10.45 | 2.65 |
| Perona-Malik | 6.2a | 16.41 | 8.93 |
| Vectorial TV | 6.2b | 16.50 | 9.00 |
| Coherence Enhancing | 6.2c | 14.32 | 7.39 |
| Beltrami Flow | 6.2d | 14.51 | 7.26 |
| Διάχυση στην κατεύθ. θ_{-} | 6.2e | 15.49 | 8.33 |
| Ανισοτρ. Εξομάλ. Tsch.-Der. | 6.2f | 16.13 | 8.75 |

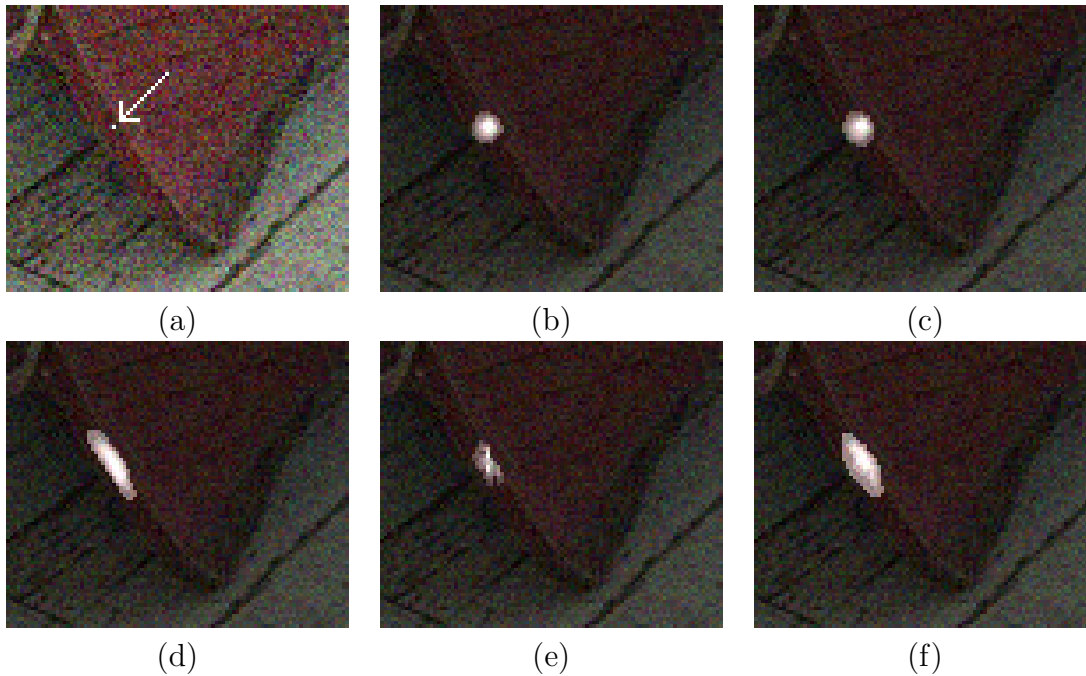
Πίνακας 6.1: Σηματοθορυβικοί λόγοι (SNR) για τα αποτελέσματα του σχήματος 6.2

Vectorial TV (βλ. σχήμα 6.2b), που δίνει τους καλύτερους SNR. Για να το επιτύχει αυτό όμως, εξομαλύνει την εικόνα σε περισσότερα σημεία (ακόμα και κοντά στις ακμές, χάρη στην ανισοτροπική συμπεριφορά της), με συνέπεια το τελικό αποτέλεσμα να έχει απομακρυνθεί περισσότερο από την αρχική εικόνα. Η αλλοίωση αυτή αυξάνει αριθμητικά την ισχύ του θορύβου, αλλά για τον ανθρώπινο παρατηρητή δεν είναι σημαντική, αντίθετα οδηγεί σε ένα αποτέλεσμα πολύ πιο ικανοποιητικό για το μάτι.

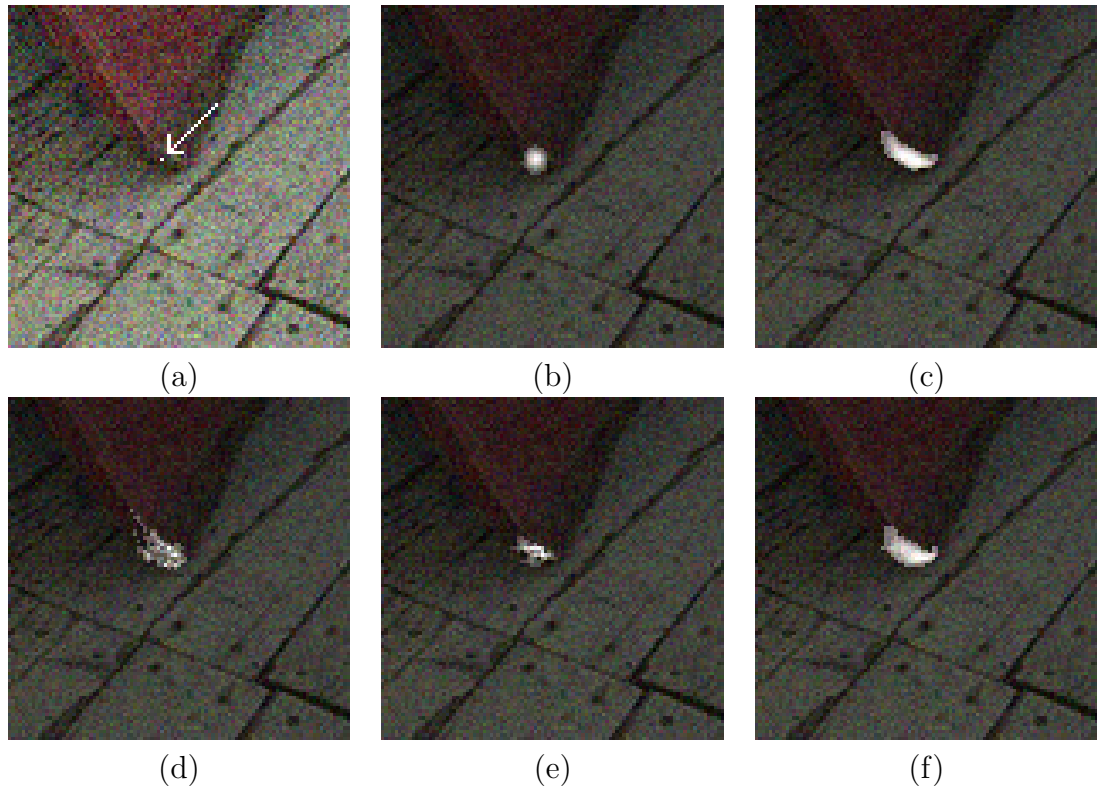
Αξιολόγηση της απόδοσης των ΜΔΕ στο Παράδειγμα 1, με τη νέα μέθοδο του Ολικού Φίλτρου που προτείνουμε

Επιπρόσθετα, για τη σύγκριση της απόδοσης των μεθόδων στο παραπάνω παράδειγμα, εφαρμόσαμε την μέθοδο αξιολόγησης που προτείναμε στην υποενότητα 5.3 : Υπολογίσαμε για δύο pixel ενδιαφέροντος, τους τοπικούς πυρήνες συσχέτισης του Ολικού Φίλτρου της κάθε ΜΔΕ: βλ. σχήματα 6.3 και 6.4 (όπου έχουμε μεγεθύνει μία μικρή περιοχή της αρχικής εικόνας, γύρω από το εκάστοτε pixel ενδιαφέροντος). Όπως παρουσιάσαμε στην υποενότητα 5.3, οι πυρήνες αυτοί παρέχουν εποπτεία της πραγματικής διαδικασίας μέσω της οποίας η κάθε μέθοδος βελτιώνει την αρχική εικόνα. Πιο συγκεκριμένα, είναι οι τοπικοί πυρήνες ενός ισοδύναμου μη γραμμικού, προσαρμοζόμενου φίλτρου εξομάλυνσης, το οποίο αν εφαρμοζόταν στην αρχική εικόνα θα έδινε κατευθείαν (σε ένα βήμα) την τελική εικόνα. Έτσι, οι πυρήνες αυτοί φανερώσουν ποια pixels της αρχικής εικόνας και σε ποιο βαθμό επηρέασαν τελικά την διαμόρφωση της τιμής στο συγκεκριμένο pixel ενδιαφέροντος.

Επιλέξαμε τα δύο αυτά pixels για την εφαρμογή της παραπάνω μεθόδου, επειδή βρίσκονται σε δύο κρίσιμα μέρη στην εικόνα του σχήματος 6.1, κοντά στα όρια της βάρκας. Το πρώτο σημείο (βλ. σχήμα 6.3a) βρίσκεται ανάμεσα σε δύο παράλληλες, λεπτές ακμές, οι οποίες είναι δυσδιάκριτες, λόγω της έντονης



Σχήμα 6.3: Εφαρμογή της νέας μεθόδου του Ολικού Φίλτρου που προτείνουμε, για την σύγκριση των μεθόδων στο παράδειγμα 1. (a) Περιοχή της θορυβώδους εικόνας του σχήμα 6.1b, με το βέλος να δείχνει το pixel ενδιαφέροντος. (b-f) Τοπικός πυρήνας (για το pixel αυτό) του Ολικού Φίλτρου της αντίστοιχης ΜΔΕ, για τα αποτελέσματα του σχήμ. 6.2 (με υπέρθεση της αρχικής εικόνας (a)). (b) Ομαλοποιημένη P-M μέθοδος, γενικευμένη για έγχρωμες εικόνες. (c) *Vectorial TV PDE*, με ϵ -εξομάλυνση. (d) Διάχυση Βελτίωσης της Συνοχής. (e) Διάχυση στην Κατεύθυνση θ_- . (f) Ανισοτρ. Εξομάλυνση των *Tschump.-Deriche*.



Σχήμα 6.4: Άλλη μία εφαρμογή της μεθόδου του Ολικού Φίλτρου, για το ίδιο παράδειγμα με το προηγούμενο σχήμα 6.3, αλλά για διαφορετικό pixel ενδιαφέροντος. (a) Περιοχή της θορυβώδους εικόνας του σχήμ. 6.1b, με το βέλος να δείχνει το νέο pixel ενδιαφέροντος. (b-f) Τοπικοί πυρήνες των Ολικών Φίλτρων των μεθόδων (αντίστοιχα με το προηγούμενο σχήμα 6.3). (b) *Vectorial TV PDE*, με ϵ -εξομάλυνση. (c) *Διάχυση Βελτίωσης της Συνοχής*. (d) *Beltrami Flow*. (e) *Διάχυση στην Κατεύθυνση θ_** . (f) *Ανισοτρ. Εξομάλυνση των Tschump.-Deriche*.

παρουσίας του θόρυβου. Το δεύτερο σημείο (βλ. σχήμα 6.4a) βρίσκεται και αυτό ανάμεσα στις ακμές αυτές, αλλά πιο κάτω, πολύ κοντά στη συνάντηση με μία τρίτη ακμή (που θα μπορούσαμε να πούμε ότι αποτελεί μία “καμπυλωμένη” διανυσματική γωνία).

Από τα σχήματα 6.3 και 6.4, παρατηρούμε πρώτα από όλα, ότι όλες οι μέθοδοι σε γενικές γραμμές εμφανίζουν την επιθυμητή συμπεριφορά οι πυρήνες να μην εκτείνονται πέρα από τα όρια της περιοχής που βρίσκονται τα σημεία ενδιαφέροντος. Επιπρόσθετα, διαφαίνεται καθαρά μία σημαντική διαφορά ανάμεσα στις μεθόδους μη γραμμικής αλλά ιστροπικής διάχυσης (με βαθμωτό συντελεστή διάχυσης - σχήματα 6.2a, 6.2b) και στις μεθόδους μη γραμμικής ανιστροπικής διάχυσης (με τανυστή διάχυσης - σχήματα 6.2c-6.2f): Ενώ στις πρώτες, οι πυρήνες εκτείνονται σε περιοχές με σχεδόν κυκλικά σχήματα, στις δεύτερες εκτείνονται σε περιοχές που ακολουθούν τις γεωμετρικές δομές της εικόνας. Έτσι, γίνεται αισθητό το πλεονέκτημα της χρήσης ανιστροπικού τανυστή διάχυσης αντί βαθμωτού συντελεστή.

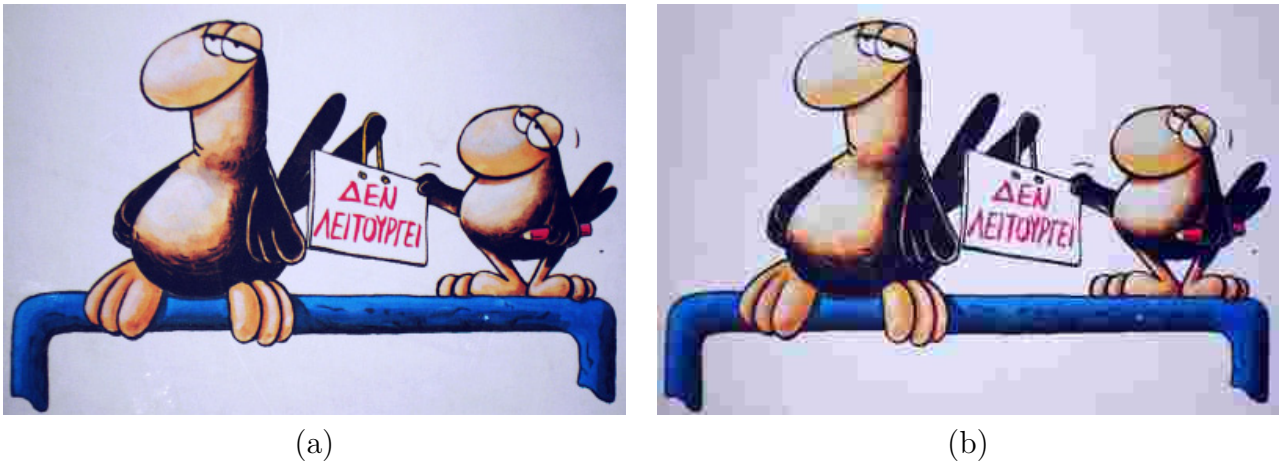
Πιο συγκεκριμένα, οι πρώτες εντοπίζουν τις ακμές και επιβραδύνουν τη διάχυση, με αποτέλεσμα οι πυρήνες να μην επεκτείνονται κάθετα στις ακμές. Λόγω όμως του ιστροπικού συντελεστή, οι πυρήνες δεν μπορούν εύκολα να επεκταθούν ούτε παράλληλα στις ακμές, για αυτό το σχήμα τους παραμένει σχεδόν κυκλικό. Αντίθετα, οι δεύτερες, χάρη στον τανυστή διάχυσης, κατευθύνουν τη διάχυση παράλληλα στις ακμές, με αποτέλεσμα οι πυρήνες να επεκτείνονται στη διεύθυνση αυτή.

Αν τώρα συγκρίνουμε τις μεθόδους ΜΔΕ με τανυστή διάχυσης μεταξύ τους, παρατηρούμε ότι οι μέθοδοι *Διάχυσης Βελτίωσης της Συνοχής και Ανιστροπική Εξομάλυνση των Tschumperlé-Deriche* παρουσιάζουν πυρήνες με την πιο μακρόστενη έκταση από όλες, ενώ για το 2^ο pixel ενδιαφέροντος (βλ. σχήμα 6.4), οι πυρήνες αυτοί καταφέρνουν να ακολουθήσουν ακόμα και την καμπύλωση της ακμής. Ακόμα, για τη μέθοδο *Διάχυση Διανυσματικών Εικόνων στην Κατεύθυνση θ* παρατηρούμε ότι οι πυρήνες της, παρόλο που καταφέρνουν να περιορίζονται στην επιθυμητή περιοχή, παρουσιάζουν μία ακανόνιστη έκταση, κάτι που αντικατοπτρίζει την ανεπιθύμητη επιρροή της διεύθυνσης διάχυσης από τον θόρυβο. Παρόμοιο φαινόμενο παρουσιάζεται και στη μέθοδο *Beltrami Flow*, όπου η έκταση του πυρήνα δεν είναι συμπαγής, λόγω της επιρροής από τον θόρυβο.

6.1.2 Παράδειγμα 2

Στο παράδειγμα αυτό, χρησιμοποιούμε την έγχρωμη cartoon εικόνα του σχήματος 6.5a (ζωγραφιά του σκιτσογράφου Αρχά), στην οποία εφαρμόζουμε συμπίεση JPEG με συντελεστή ποιότητας $Q = 10\%$. Έτσι, προκύπτει η χαμηλής ποιότητας συμπίεσμένη εικόνα του σχήματος 6.5b, η οποία όπως παρατη-

ρούμε είναι έντονα αλλοιωμένη. Έτσι, με σκοπό τη βελτίωσή της, εφαρμόζουμε πάλι τις ίδιες μεθόδους ΜΔΕ και τα αποτελέσματα φαίνονται στο σχήμα 6.6. Η διαδικασία που ακολουθήσαμε για την εξαγωγή των αποτελεσμάτων αυτών είναι ίδια με αυτή του παραδείγματος 1.



Σχήμα 6.5: Η εικόνα που χρησιμοποιήθηκε στο Παράδειγμα 2. (a) Αρχική, έγχρωμη cartoon εικόνα (262×397 pixels). (b) Αρχική, εικόνα, μετά από JPEG συμπίεση, με συντελεστή ποιότητας $Q = 10\%$.

Από τα αποτελέσματα του σχήματος 6.6, καταλήγουμε ξανά σε αντίστοιχα συμπεράσματα για την απόδοση των διαφόρων μεθόδων. Παρατηρούμε και πάλι λοιπόν ότι η μέθοδος *Ανισοτροπικής Εξομάλυνσης των Tschumperlé-Deriche* καταλήγει στο πιο ικανοποιητικό αποτέλεσμα. Επιπρόσθετα, στην περίπτωση αυτή, παρατηρούμε την υπερβολική αλλοίωση που έχει επιφέρει στα γράμματα η μέθοδος *Διάχυσης Βελτίωσης της Συνοχής*, για τους λόγους που αναφέραμε και προηγουμένως.

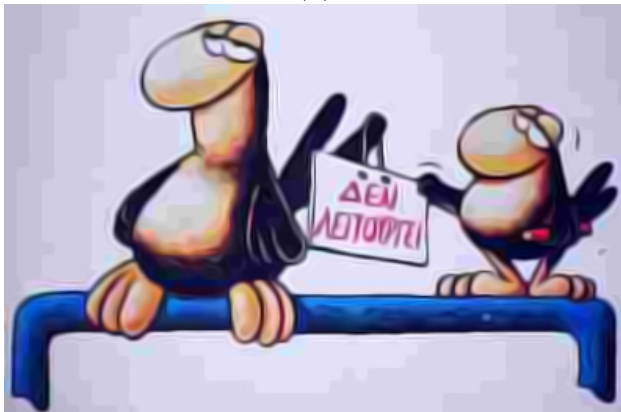
Ακόμα, πρέπει να σημειώσουμε ότι η εικόνα προς βελτίωση περιέχει, όπως παρατηρούμε, κάποιες ψευδοκαμπύλες (*false contours*) στο φόντο της, οι οποίες έχουν δημιουργηθεί από τη συμπίεση JPEG. Πιο συγκεκριμένα, η αρχική εικόνα, περιέχει μία ανεπαίσθητη διαδοχική μεταβολή της φωτεινότητας στο φόντο (που οφείλεται στο φλας που χρησιμοποιήθηκε, κατά τη φωτογράφιση της ζωγραφιάς αυτής). Κατά τη συμπίεση όμως, λόγω της χαμηλής ποιότητας, αυτή η μεταβολή έγινε απότομη, με αποτέλεσμα τη δημιουργία των ψευδοκαμπύλων. Θα ήταν επιθυμητό λοιπόν μία μέθοδος βελτίωσης να μπορεί να απομακρύνει αυτές τις καμπύλες. Παρατηρούμε ότι η *Ομαλοποιημένη μέθοδος των Perona-Malik* και η μέθοδος *Ανισοτροπικής Εξομάλυνσης των Tschumperlé-Deriche* παρουσιάζουν την καλύτερη συμπεριφορά, καταφέροντας να εξαφανίσουν εντελώς τις καμπύλες αυτές. Αυτό γίνεται χάρη στις εξομαλύνσεις με Gaussian



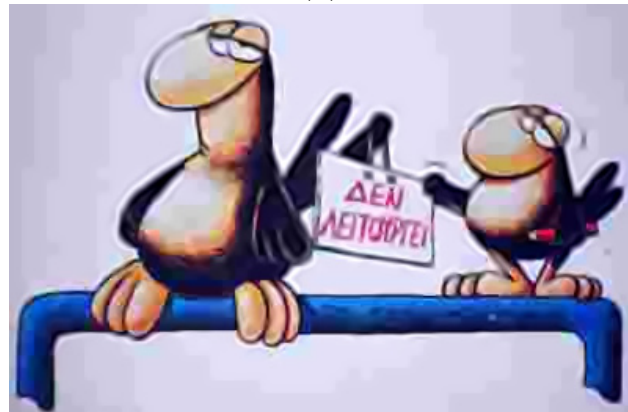
(a)



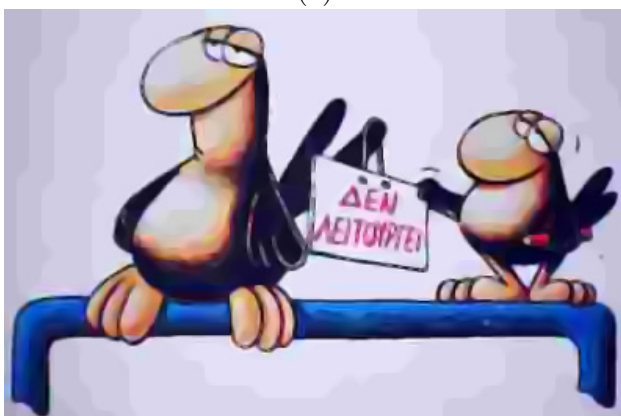
(b)



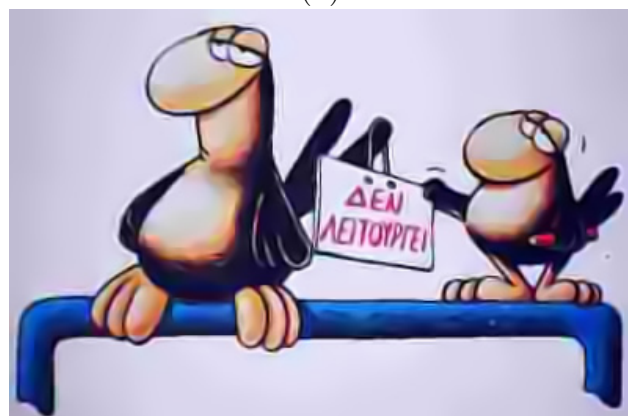
(c)



(d)



(e)



(f)

Σχήμα 6.6: Σύγκριση των μεθόδων βελτίωσης, με εφαρμογή τους στην εικόνα του σχήματος 6.5b (a) Ομαλοποιημένη P-M μέθοδος, γενικευμένη για έγχρωμες εικόνες (40 επαναλήψεις, $\tau = 0.2$, $\sigma = 0.6$, $K = 10$, $g(s)$ από σχέση (5.3)). (b) Vectorial TV PDE, με ϵ -εξομάλυνση (100 επαναλήψεις, $\tau = 0.2$, $\epsilon = 1$). (c) Διάχυση Βελτίωσης της Συνοχής (30 επαναλήψεις, $\tau = 0.2$, $\sigma = 0$, $\rho = 1$). (d) Beltrami Flow (200 επαναλήψεις, $\tau = 0.1$). (e) Διάχυση στην Κατεύθυνση θ_- (50 επαναλήψεις, $\tau = 0.2$). (f) Ανισοτρ. Εξομάλυνση των Tschump.-Deriche (250 επαναλήψεις, $\beta_{\max} = 2$, $\rho = 1$).

πυρήνες που η μία κάνει στην εικόνα και η άλλη στον δομικό τανυστή, με αποτέλεσμα να μην λαμβάνονται υπόψη στη διάχυση οι καμπύλες αυτές, διότι έχουν πολύ μικρό πάχος.

Παρατηρούμε επίσης ότι οι ακμές συντηρούνται αλλά και ανακατασκευάζονται (μέσω της εξομάλυνσης κατά μήκος τους) με τον καλύτερο συγκριτικά τρόπο στη μέθοδο *Ανισοτροπικής Εξομάλυνσης των Tschumperlé-Deriche*. Η *Ομαλοποιημένη μέθοδος των Perona-Malik* και η *Vectorial TV* καταφέρνουν κι αυτές να εξομαλύνουν με την ίδια επιτυχία τις ομογενείς περιοχές, εξαφανίζοντας τα τετράγωνα της JPEG συμπίεσης, αλλά δεν παρουσιάζουν τόσο ικανοποιητική απόδοση στις ακμές. Τέλος, το χειρότερο αποτέλεσμα προκύπτει και στην περίπτωση αυτή από τη μέθοδο *Beltrami Flow*, η οποία για να απομακρύνει τα τετράγωνα της συμπίεσης, αναγκάζεται να θολώσει υπερβολικά τις ακμές.

6.1.3 Παράδειγμα 3

Στο παράδειγμα αυτό, παρουσιάζουμε μία εφαρμογή της μεθόδου *Διάχυσης Βελτίωσης της Συνοχής διανυσματικών εικόνων* στην δημιουργία ειδικών εφέ. Χρησιμοποιούμε μία έγχρωμη φωτογραφία πεύκου (βλ. σχήμα 6.7a) και εφαρμόζουμε τη μέθοδο αυτή: βλ. σχήματα 6.7b1,b2. Παρατηρούμε ότι η μέθοδος τονίζει και επεκτείνει τις μονοδιάστατες δομές της εικόνας (πευκοβελόνες) δημιουργώντας μία αφαιρετική εικόνα, ιμπρεσιονιστικού στυλ. Στα σχήματα 6.7b1,b2 παρατίθενται δύο χρονικές στιγμές και έτσι διαφαίνεται η χρονική εξέλιξη: όσο η διάχυση προχωράει, τονίζονται οι πιο σημαντικές δομές, ενώ εξαφανίζονται οι λεπτομέρειες.

Επίσης, στο σχήματα 6.7c1,c2 βλέπουμε το αποτέλεσμα της μεθόδου *Ανισοτροπικής Εξομάλυνσης των Tschumperlé-Deriche*, η οποία έχει σχεδιαστεί για την απομάκρυνση του θορύβου, αλλά την εφαρμόσαμε και στην εικόνα αυτή, με σκοπό τη σύγκριση των διαδικασιών διάχυσης. Παρατηρούμε ότι και η μέθοδος αυτή παρουσιάζει μία κάπως παρόμοια συμπεριφορά, λόγω του ανισοτροπικού τανυστή διάχυσης και της μεγάλης τιμής στην σταθερά εξομάλυνσης ρ , αλλά τα αποτελέσματά της είναι λιγότερο ικανοποιητικά στην εφαρμογή αυτή. Ιδιαίτερα μετά από αρκετές επαναλήψεις, η μέθοδος αυτή καταστρέφει κάποιες λεπτές μονοδιάστατες δομές (βλ. σχήμα 6.7c2), σε αντίθεση με τη *Διάχυση Βελτίωσης της Συνοχής* όπου το αποτέλεσμα έχει περισσότερο ινώδη μορφή, λόγω του ότι ο τανυστής της *Διάχυσης Βελτίωσης της Συνοχής* είναι πιο έντονα ανισοτροπικός.



(a)



(b1)



(b2)



(c1)



(c2)

Σχήμα 6.7: Παράδειγμα Βελτίωσης της Συνοχής, για δημιουργία ειδικών εφέ. (a) Αρχική εικόνα πεύκου, με πολυάριθμες μονοδιάστατες δομές (πευκοβελόνες). (b) Διάχυση Βελτίωσης της Συνοχής ($\tau = 0.2, \sigma = 0, \rho = 3$). (b1) 24 επαναλήψεις. (b2) 100 επαναλήψεις. (c) Ανισοτρ. Εξομάλυνση των Tschump.-Deriche ($\beta_{\max} = 6, \rho = 3$). (c1) 36 επαναλήψεις. (c2) 80 επαναλήψεις.

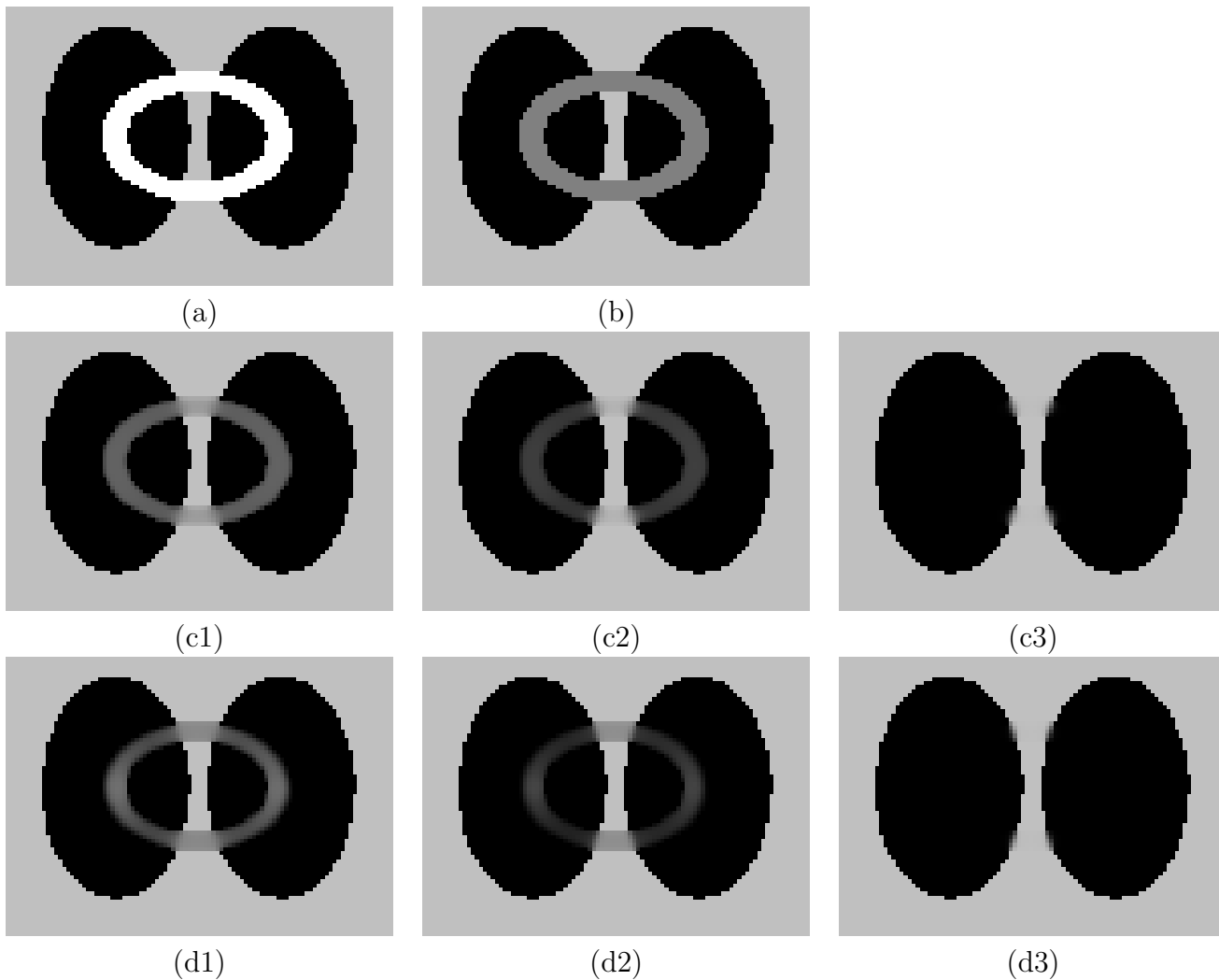
6.2 Πειράματα για το πρόβλημα Inpainting Εικόνων

6.2.1 Παράδειγμα 1

Το παράδειγμα αυτό έχει ως σκοπό τη σύγκριση της απόδοσης των μεθόδων *TV PDE* και *Κίνησης Μέσης Καμπυλότητας (Mean Curvature Motion - MCM)*, όταν αυτές εφαρμόζονται στο Inpainting. Χρησιμοποιούμε την συνθετική εικόνα και την περιοχή inpainting που φαίνονται στο σχήμα 6.8a (παρόμοιες εικόνες, με την ίδια μορφή σχημάτων και περιοχής inpainting χρησιμοποιούνται συχνά στα παραδείγματα, βλ. π.χ. [14, 3]). Εφαρμόζοντας τις δύο αυτές μεθόδους, παίρνουμε τα αποτελέσματα του σχήματος 6.8. Παρατηρούμε ότι οι μέθοδοι αυτές δίνουν, όπως περιμέναμε το ίδιο τελικό αποτέλεσμα, αφού η εξίσωση της συνθήκης ισορροπίας είναι ίδια και για τις δύο ΜΔΕ. Βλέπουμε όμως ότι η *MCM* καταφέρνει να προσεγγίσει την κατάσταση ισορροπίας πιο γρήγορα από την *TV PDE*, κάτι το οποίο, οφείλεται στον πολλαπλασιαστικό παράγοντα $\|\nabla u\|$ της ταχύτητας, που περιέχει επιπλέον η ΜΔΕ αυτή. Παρατηρούμε τέλος ότι οι δύο ΜΔΕ καταλήγουν στην ίδια τελική κατάσταση ακολουθώντας σημαντικά διαφορετική ροή. Από τη μία, η *MCM* μικραίνει συνεχώς το πάχος του δακτυλίου, μεταβάλλοντας λιγότερο τις τιμές στο εσωτερικό του (δηλ. το γκρι χρώμα), ενώ από την άλλη η *TV PDE* επιτελεί εξαρχής διάχυση, σε όλη την επιφάνεια του δακτυλίου, μεταβάλλοντας σχεδόν παράλληλα όλες τις τιμές της περιοχής inpainting. Η διαφορά αυτή έγκειται στο ότι η *MCM* επιτελεί κίνηση ισοϋψών καμπύλων, ενώ η *TV PDE* συμπεριφέρεται περισσότερο σαν μη γραμμική εξομάλυνση. Με την κίνηση όμως των ισοϋψών, η ανακατασκευή διαδίδεται πιο γρήγορα και καταλήγουμε πιο εύκολα στην κατάσταση ισορροπίας.

6.2.2 Παράδειγμα 2

Το παράδειγμα αυτό αποτελεί μία εφαρμογή του Inpainting στην απομάκρυνση κειμένου από εικόνες. Στο σχήμα 6.9a φαίνεται η εικόνα που χρησιμοποιήσαμε καθώς και το κείμενο, που αποτελεί την περιοχή inpainting. Στα υπόλοιπα υποσχήματα του σχήματος 6.9, φαίνονται τα αποτελέσματα των μεθόδων *Vectorial Digital TV Filter (Vectorial DTVF)* (που όπως έχουμε πει δίνει τα ίδια αποτελέσματα με την μέθοδο *Vectorial TV*), *Ανισοτροπικής Εξομάλυνσης των Tschumperlé-Deriche* και της νέας μεθόδου *Γενικευμένης Διάχυσης Οδηγούμενης από την Καμπυλότητα (Gen. CDD)*, που προτείναμε στην ενότητα 2.8.2. Παρατηρούμε ότι και οι τρεις μέθοδοι καταφέρνουν να αποκαταστήσουν την εικόνα με αρκετά ικανοποιητικό τρόπο. Σημειώνουμε εδώ ότι γενικά οι μέθοδοι Inpainting με ΜΔΕ παρουσιάζουν το πλεονέκτημα ότι μπορούν με



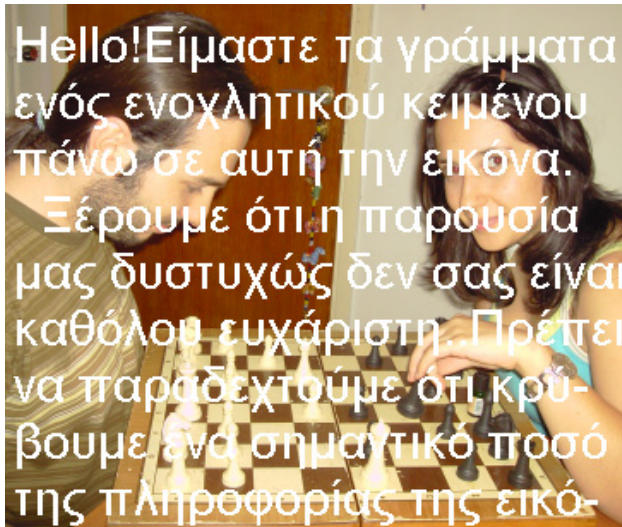
Σχήμα 6.8: Παράδειγμα Inpainting. (a) Αρχική εικόνα, με την περιοχή που λείπει (άσπρο). (b) Αρχική πρόβλεψη των μεθόδων. (c) *Total Variation PDE*, με ϵ -regulation ($\tau = 0.2, \epsilon = 1$). (c1) 300 επαναλήψεις. (c2) 600 επαναλήψεις. (c3) 1500 επαναλήψεις. (d) *Mean Curvature Motion*, με ϵ -regulation ($\tau = 0.2, \epsilon = 1$). (d1) 40 επαναλήψεις. (d2) 80 επαναλήψεις. (d3) 800 επαναλήψεις.

ευκολία να αντιμετωπίσουν προβλήματα με inpainting περιοχές περίπλοκης τοπολογίας, όπως εδώ, λόγω του ότι αρκεί απλά να οριστεί το σύνολο σημείων στο οποίο συντελείται διάχυση.

Συγκρίνοντας τα αποτελέσματα του σχήματος 6.9, παρατηρούμε ότι οι τρεις μέθοδοι παρουσιάζουν παρόμοια επίδοση και οι διαφορές τους συνίστανται σε κάποιες λεπτομέρειες. Πιο συγκεκριμένα, η μέθοδος *Gen. CDD* που προτείναμε, καταφέρνει σε αρκετά σημεία να ανακατασκευάσει καλύτερα τις ακμές της εικόνας σε σχέση με την *Vectorial DTVF* (βλ. π.χ. τα λευκά πιόνια στο βάθος και τα περισσότερα τετράγωνα της σκακιέρας). Επίσης, η μέθοδος *Ανισοτροπικής Εξομάλυνσης των Tschumperlé-Deriche*, σε πολλά σημεία, καταφέρνει ακόμα καλύτερα να ενώσει τις σπασμένες ακμές, αλλά από την άλλη δεν καταφέρνει να εξαφανίσει τελείως κάποια γράμματα, παρά τις πολύ περισσότερες επαναλήψεις που απαιτεί (1200 επαναλήψεις, σε σχέση με 200 για την *Gen. CDD* και 225 για την *Vectorial DTVF*).

6.2.3 Παράδειγμα 3

Στο παράδειγμα αυτό συγκρίνουμε τις μεθόδους *Vectorial TV* και *Vectorial DTVF* για διανυσματικές εικόνες. Χρησιμοποιούμε ένα μικρό κομμάτι της εικόνας και της περιοχής inpainting του προηγούμενου παραδείγματος και εφαρμόζουμε τις μεθόδους αυτές: βλ. σχήμα 6.10. Παρατηρούμε ότι και οι μέθοδοι αυτές (όπως στο παράδειγμα 1) καταλήγουν στο ίδιο αποτέλεσμα, κάτι το οποίο περιμέναμε, αφού το φίλτρο *Vectorial DTVF* συγκλίνει στην κατάσταση ισορροπίας της ΜΔΕ της *Vectorial TV*. Όμως, οι μέθοδοι αυτές ακολουθούν διαφορετική πορεία, όπως φαίνεται στο σχήμα 6.10. Επίσης, η μέθοδος *Vectorial DTVF* παρουσιάζει το σημαντικό πλεονέκτημα ότι προσεγγίζει την κατάσταση ισορροπίας πολύ πιο γρήγορα και με πολύ καλύτερη συμπεριφορά. Το τελευταίο φαίνεται και στο σχήμα 6.11, όπου έχει σχεδιαστεί μία νόρμα της μεταβολής της εικόνας σε μια επανάληψη (μέγιστο μέτρο του διανύσματος της μεταβολής), κατά την εξέλιξη των βημάτων των μεθόδων αυτών. Παρατηρούμε ότι στη μέθοδο *Vectorial DTVF*, η εικόνα μεταβάλλεται αρχικά πολύ πιο έντονα, κάτι που τη βοηθάει να προσεγγίσει πιο γρήγορα τη κατάσταση ισορροπίας, ενώ γρήγορα η μεταβολή αυτή γίνεται ασήμαντη, κάτι που φανερώνει την πλήρη ισορροπία της διαδικασίας. Αντίθετα, στη μέθοδο *Vectorial TV*, η μεταβολή της εικόνας είναι από την αρχή μικρή και έτσι η κατάσταση ισορροπίας προσεγγίζεται με αργό ρυθμό. Ακόμα όμως και όταν η εικόνα φτάσει κοντά στην κατάσταση ισορροπίας, η διαδικασία δεν καταφέρνει να ισορροπήσει πλήρως, αφού οι τιμές σε κάποια pixels ταλαντώνονται. Αυτός είναι ο λόγος που παρουσιάζεται η απότομη αύξηση στην αντίστοιχη γραφική παράσταση του σχήματος 6.11 (σημειώνουμε ότι το φαινόμενο αυτό δεν θα φαινόταν τόσο έντονα, αν είχαμε χρησιμοποιήσει κάποια άλλη νόρμα, όπως την L^2 , που λαμβάνει



(a)



(b)



(c)



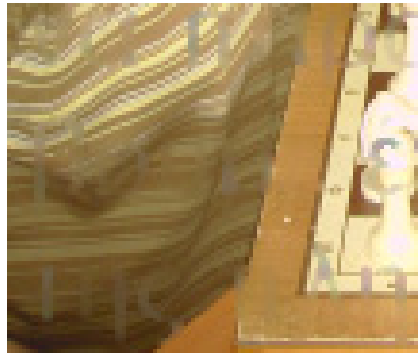
(d)

Σχήμα 6.9: Παράδειγμα Inpainting για απομάκρυνση κειμένου από έγχρωμη εικόνα. (a) Αρχική εικόνα, με κείμενο. (b) *Vectorial DTVF* (225 επαναλήψεις, $\epsilon = 0.01$). (c) *Ανισοτρ. Εξομάλυνση των Tschump.-Deriche* (1200 επαναλήψεις, $\beta_{\max} = 20$, $\rho = 3$). (d) *Γενικευμένη CDD* (νέα μέθοδος που προτείνουμε) (200 επαναλήψεις, $p = 2$, $\sigma = 1$, $\epsilon = 0.5$, $\epsilon_{\kappa} = 0.1$).

υπόψιν της όλες τις τιμές στην περιοχή inpainting και όχι μόνο τη μέγιστη τιμή).



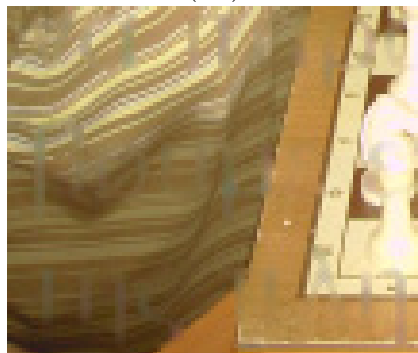
(a)



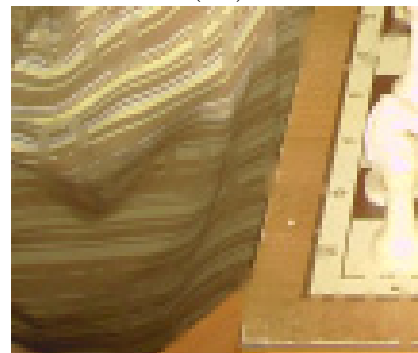
(b1)



(b2)

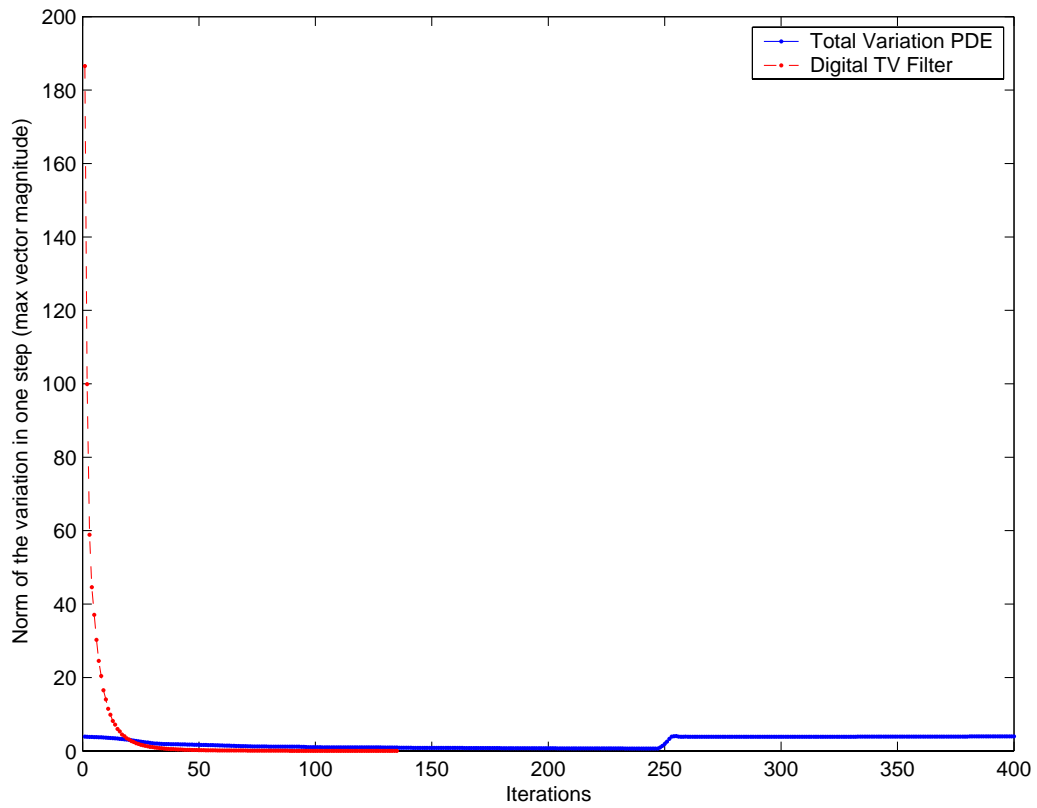


(c1)



(c2)

Σχήμα 6.10: Παράδειγμα για τη σύγκριση της σύγκλισης στο πρόβλημα Inpainting. (a) Αρχική εικόνα, με κείμενο. (b) *Vectorial TV PDE*, με ϵ -regulation ($\tau = 0.5, \epsilon = 1$). (b1) 170 επαναλήψεις. (b2) 300 επαναλήψεις. (c) *Vectorial DTVF* ($\epsilon = 0.01$). (c1) 10 επαναλήψεις. (c2) 80 επαναλήψεις.



Σχήμα 6.11: Σύγκριση της σύγκλισης των μεθόδων *Vectorial TV* και *Vectorial DTVF*.

6.2.4 Παράδειγμα 4

Στο σχήμα 6.12 φαίνεται άλλο ένα παράδειγμα απομάκρυνσης κειμένου. Αυτή τη φορά, χρησιμοποιήσαμε την εικόνα που φαίνεται στο σχήμα 6.10a (εικόνα “peppers”, που διατίθεται σαν demo στο πρόγραμμα Matlab 6), με περιοχή inpainting το κείμενο που φαίνεται επίσης στο σχήμα 6.10a. Το κείμενο αυτό αποτελείται από γράμματα, που είναι αρκετά μεγαλύτερα σε σχέση με αυτά του προηγούμενου παραδείγματος. Παρατηρούμε ότι η νέα μέθοδος *Gen. CDD* που προτείναμε καταφέρνει με επιτυχία να εξαφανίσει τα γράμματα. Παρόμοια επιτυχία παρουσιάζει και η μέθοδος *Vectorial DTVF*, αλλά σε κάποια σημεία, όπου η μέθοδος αυτή έχει αφήσει κάποια υπολείμματα από τα γράμματα, η νέα μέθοδος *Gen. CDD* έχει καταφέρει να τα εξαφανίσει σχεδόν τελείως.

6.2.5 Παράδειγμα 5

Παραθέτουμε ένα τελευταίο παράδειγμα, στο οποίο διαφαίνεται πιο καθαρά η βελτιωμένη συμπεριφορά της νέας μεθόδου *Gen. CDD* που προτείναμε, σε σχέση με τη μέθοδο *Vectorial DTVF*. Χρησιμοποιήσαμε εδώ μία λεπτομέρεια της εικόνας “peppers” του προηγούμενου παραδείγματος και μία περιοχή inpainting με μεγάλο πάχος, βλ. σχήμα 6.13a. Η περίπτωση αυτή είναι δυσκολότερη από τις προηγούμενες περιπτώσεις των γραμμάτων, τα οποία είχαν σημαντικά μικρότερο πάχος.

Τα αποτελέσματα παρουσιάζονται στα σχήματα 6.13b, 6.13c. Σαν αρχική πρόβλεψη της μεθόδου *Gen. CDD* που προτείνουμε χρησιμοποιήσαμε το αποτέλεσμα της μεθόδου *Vectorial DTVF*. Πρέπει να διευκρινίσουμε ότι η αρχικοποίηση αυτή βοηθάει μόνο στην γρηγορότερη σύγκλιση στην κατάσταση ισορροπίας της *Gen. CDD*. Το ίδιο τελικό αποτέλεσμα θα προέκυπτε εάν χρησιμοποιούσαμε κάποια απλή αρχικοποίηση, όπως το γέμισμα της περιοχής inpainting με μία σταθερή τιμή. Η μόνη διαφορά είναι ότι θα απαιτούνταν αρκετά περισσότερες επαναλήψεις για την κατάληξη στην κατάσταση ισορροπίας.

Από το σχήμα 6.13, παρατηρούμε ότι, ενώ η μέθοδος *Vectorial DTVF* αδυνατεί να επανασυνδέσει τις κομμένες ακμές, καταλήγωντας σε ένα θολό αποτέλεσμα, η νέα μέθοδος *Gen. CDD* καταφέρνει να συνδέσει τις ακμές αυτές, παρόλο το μεγάλο κομμάτι που λείπει. Η βελτίωση αυτή οφείλεται στον όρο της γενικευμένης καμπυλότητας που εισάγαμε. Πιο συγκεκριμένα, η έξοδος της μεθόδου *Vectorial DTVF* (που χρησιμοποιούμε όπως είπαμε σαν είσοδο για τη μέθοδό μας) παρουσιάζει υψηλή γενικευμένη καμπυλότητα στα σημεία που οι ακμές συναντούν την περιοχή inpainting, πράγμα το οποίο κάνει τη διάχυση στη μέθοδό μας να συνεχίζεται, έως ότου επανενωθούν οι κομμένες ακμές.



(a)

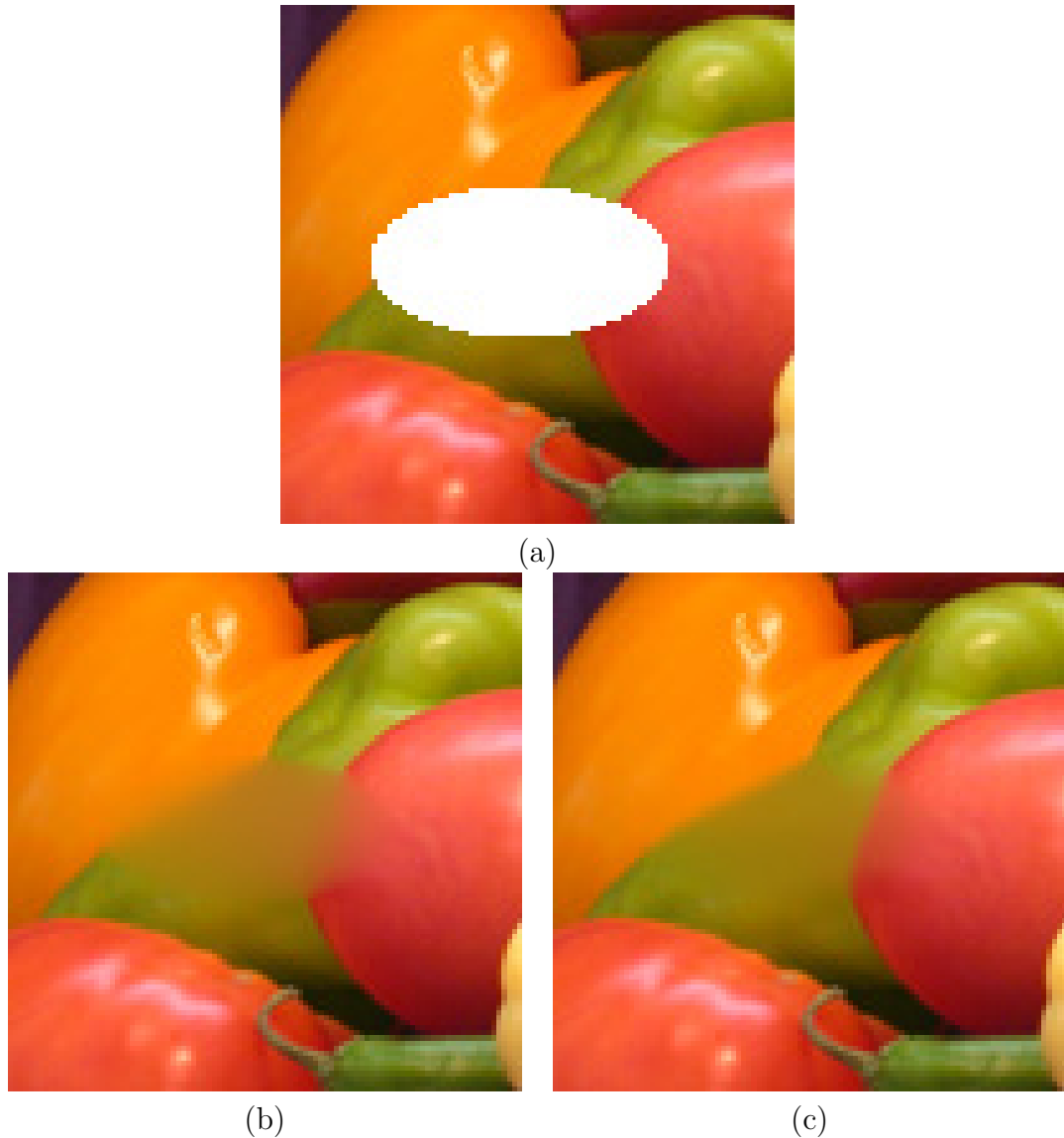


(b)



(c)

Σχήμα 6.12: Παράδειγμα Inpainting για απομάκρυνση κειμένου από έγχρωμη εικόνα. (a) Αρχική εικόνα, με κείμενο. (b) *Vectorial DTVF* (545 επαναλήψεις, $\epsilon = 0.01$). (c) *Γενικευμένη CDD* (νέα μέθοδος που προτείνουμε) (1000 επαναλήψεις, $p = 2, \sigma = 1, \epsilon = 0.5, \epsilon_{\kappa} = 0.1$).



Σχήμα 6.13: Παράδειγμα Inpainting έγχρωμης εικόνας, με περιοχή inpainting μεγάλου πάχους. (a) Αρχική εικόνα, με την περιοχή που λείπει (άσπρο). (b) *Vectorial DTVF* (2024 επαναλήψεις, $\epsilon = 0.01$). (c) *Γενικευμένη CDD* (νέα μέθοδος που προτείνουμε) (με αρχική εικόνα την (b), 4000 επαναλήψεις, $p = 1, \sigma = 1, \epsilon = 0.5, \epsilon_k = 0.1$).

6.3 Πειράματα για το πρόβλημα Παρεμβολής Εικόνων

6.3.1 Παράδειγμα 1

Στο παράδειγμα αυτό, συγκρίνουμε την απόδοση διαφόρων μεθόδων στο πρόβλημα της Παρεμβολής. Για τον σκοπό αυτό, χρησιμοποιούμε την έγχρωμη εικόνα του σχήματος 6.14a1, η οποία αποτελεί ένα κομμάτι της δημοφιλούς εικόνας ενός πιθήκου (“madrill”), που χρησιμοποιείται συχνά στην βιβλιογραφία για τα παραδείγματα. Η εικόνα αυτή έχει μία αρκετά ικανοποιητική χωρική ευκρίνεια (255×225 pixels).

Στο σχήμα 6.14a2, φαίνεται η εικόνα που χρησιμοποιήθηκε σαν είσοδος για τις διάφορες μεθόδους παρεμβολής, η οποία προήλθε από ελάττωση της χωρικής ευκρίνειας της εικόνας 6.14a1 κατά τον παράγοντα $(1/5) \times (1/5)$. Η ελάττωση αυτή έγινε με τρόπο σύμφωνο με την συνθήκη αναστρεψιμότητας (βλ. υποενότητα 3.1.1), εφαρμόζοντας δηλαδή πρώτα βαθυπερατή εξομάλυνση με έναν πυρήνα συσχέτισης s (για την αποφυγή του aliasing) και στη συνέχεια υποδειγματοληψία. Για αυτό τον πυρήνα συσχέτισης, κάναμε την απλή επιλογή του διακριτού 2Δ παλμού, έκτασης 5×5 pixels (που ισοδυναμεί με τον υπολογισμό της τιμής σε κάθε pixel της εικόνας 6.14a2 σαν τον μέσο όρο των τιμών στο αντίστοιχο 5×5 μπλοκ από pixels της εικόνας 6.14a1).

Έτσι, εφαρμόσαμε τους διάφορους αλγόριθμους των μεθόδων παρεμβολής στην εικόνα 6.14a2, με σκοπό την αύξηση της χωρικής της ευκρίνειας κατά τον παράγοντα 5×5 , καταλήγοντας σε μία εικόνα ίδιας χωρικής ευκρίνειας με την εικόνα 6.14a1 (χωρίς βέβαια οι αλγόριθμοι αυτοί να “γνωρίζουν” την εικόνα 6.14a1, από την οποία προήλθε η 6.14a2). Με τον τρόπο αυτό, γίνεται εφικτή η σύγκριση των αποτελεσμάτων με την εικόνα 6.14a1 (ιδανικό αποτέλεσμα της παρεμβολής) και η εξαγωγή κάποιων μέτρων απόδοσης (σηματοθορυβικοί λόγοι, κ.τ.λ.).

Τα αποτελέσματα αυτά των διαφόρων μεθόδων παρεμβολής φαίνονται στα σχήματα 6.14 και 6.15.

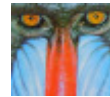
Διευκρινίσεις

Πριν την σύγκριση των μεθόδων, πρέπει να διευκρινίσουμε τα εξής, όσον αφορά τις διάφορες μεθόδους που εφαρμόσαμε:

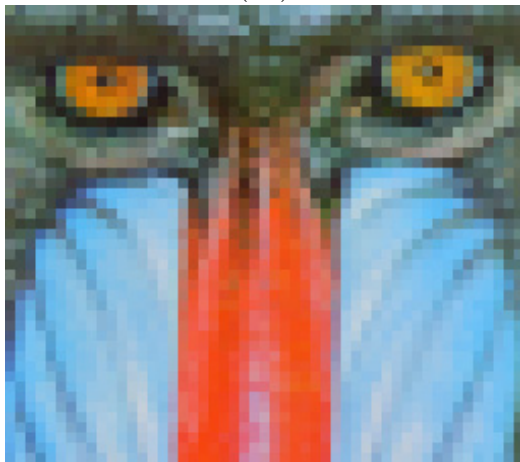
- το αποτέλεσμα του σχήματος 6.14f έχει προκύψει με μία απλή και άμεση γενίκευση στις διανυσματικές εικόνες της μεθόδου Παρεμβολής βασισμένης στην Ολική Μεταβολή (TV) των Malgouyres και Guichard [26, 35, 36] (προτείναμε την γενίκευση αυτή στην υπενότητα 3.2 - βλ.



(a1)



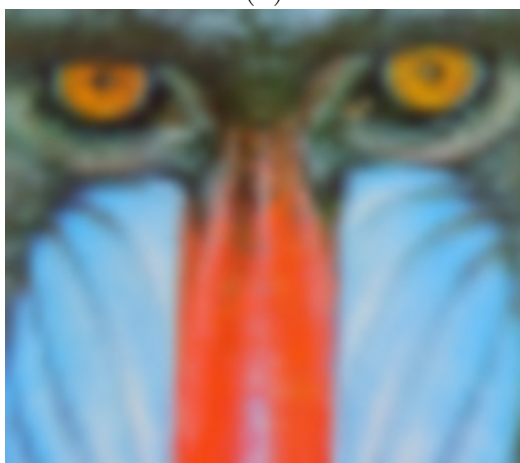
(a2)



(b)



(c)

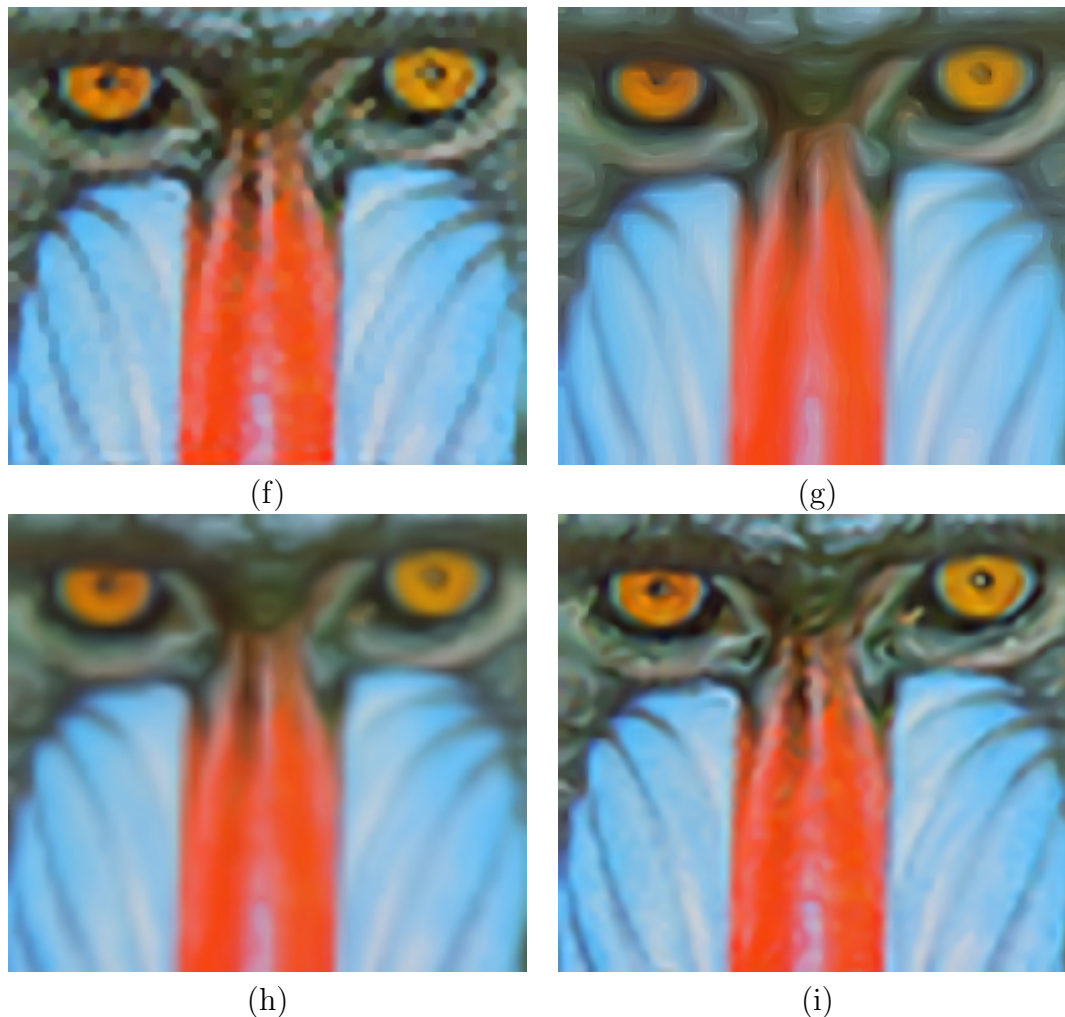


(d)



(e)

(συνέχεια στην επόμενη σελίδα)



Σχήμα 6.14: 1^ο παράδειγμα σύγκρισης διάφορων μεθόδων παρεμβολής.

(a1) Κομμάτι της έγχρωμης εικόνας “madrill” (μέγεθος 255×225 pixels).

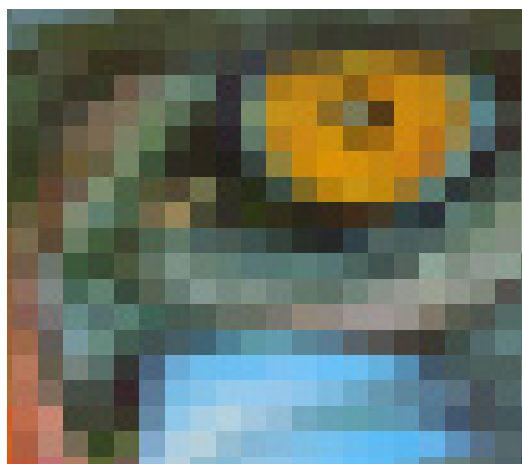
(a2) Ελάττωση της ευκρίνειας της (a1) κατά τον παράγοντα $(1/5) \times (1/5)$ (μέγεθος 51×45 pixels).

(b-i) Αποτελέσματα μεθόδων παρεμβολής της εικόνας (b) για αύξηση της ευκρίνειας κατά 5×5 (τελικό μέγεθος 255×225 pixels):

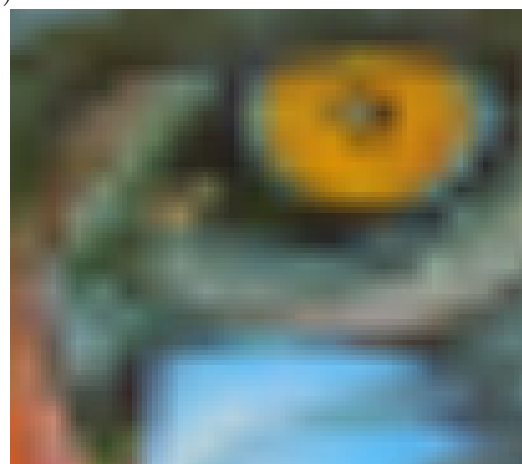
(b) Απλή επανάληψη των pixels (*Duplication*). (c) *Bilinear* (αρχική πρόβλεψη για την (g)). (d) *B-spline*. (e) *Zero-padding*, με πυρήνα συνθήκης αναστρέψ.: 5×5 2Δ παλμό (αρχική πρόβλεψη για τις (f),(i)). (f) Άμεση γενίκευση στις διανυσματικές εικόνες της Παρεμβολής βασισμένης στην TV, των *Malgouyres-Guichard* [26, 35, 36] (400 επαναλήψεις, $\tau = 0.1$, $\epsilon = 0.5$, με πυρήνα συνθήκης αναστρέψ.: 5×5 2Δ παλμό). (g) Παρεμβολή με την *Ανισοτρ. Εξομάλυνση των Tschump.-Deriche* (όπως προτείνεται στα [52, 55]) (200 επαναλήψεις, $\beta_{\max} = 5$, $\rho = 1.5$). (h) Παρεμβολή βασισμένη σε LICs, του *Tschump.* (1 επανάληψη). (i) Νέα μέθοδος που προτείνουμε (600 επαναλήψεις, $\beta_{\max} = 5$, $\rho = 2$, με πυρήνα συνθήκης αναστρέψ. 5×5 2Δ παλμό).



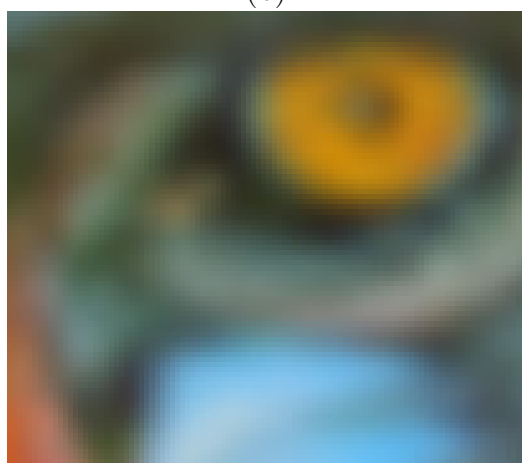
(a)



(b)



(c)

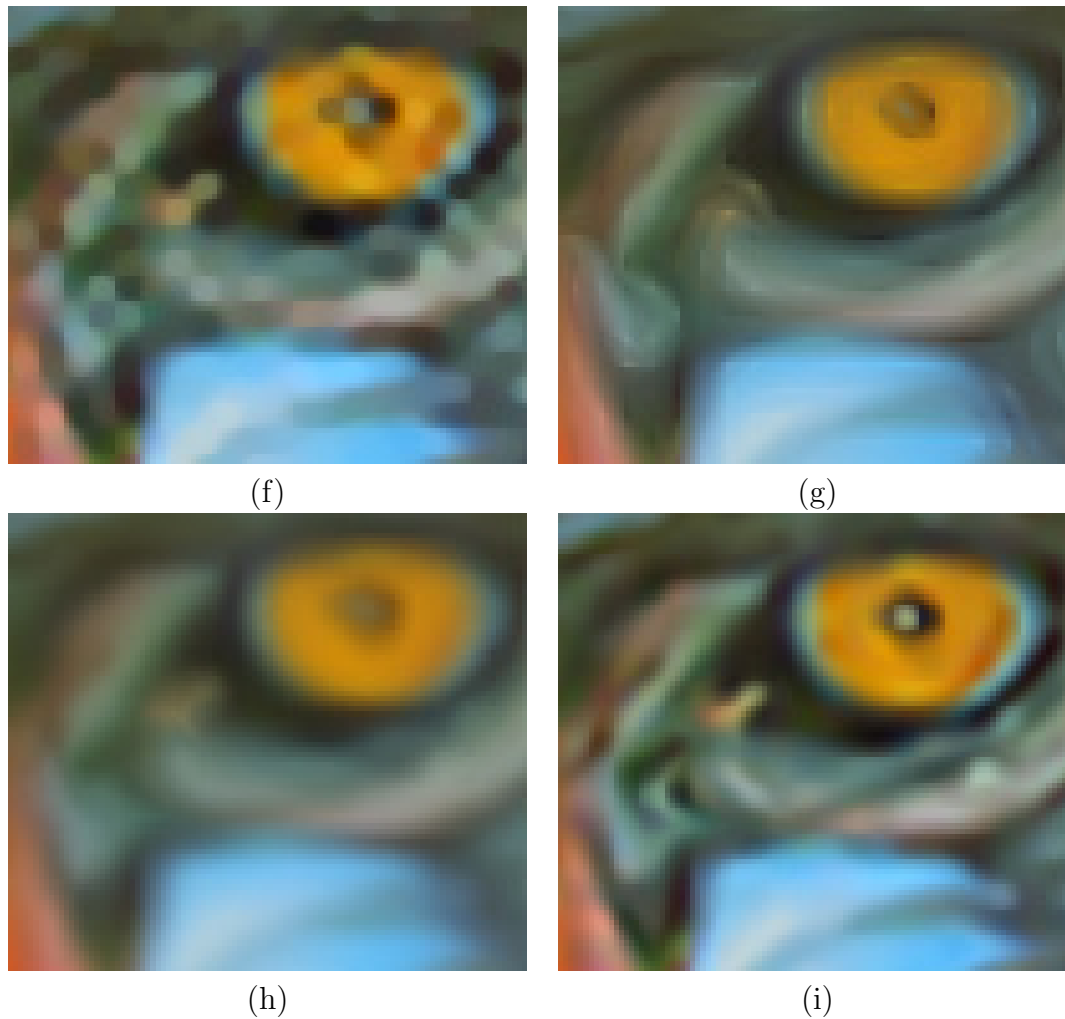


(d)



(e)

(συνέχεια στην επόμενη σελίδα)



Σχήμα 6.15: Αντίστοιχες λεπτομέρειες των αποτελεσμάτων του σχήματος 6.14 (μεγέθους 102×90 pixels).

(a) Εικόνα “madrill”. (b) *Duplication*. (c) *Bilinear*. (d) *B-spline*. (e) *Zero-padding*. (f) *Total Variation (Malg.-Guich.)*. (g) *Ανισοτρ. Εξομάλ. Tschump.-Der.* (h) *LIC (Tschump.)*. (i) *Νέα προτεινόμενη μέθοδος*.

σχέση (3.7)). Όπως είναι φυσικό, η γενίκευση αυτή διατηρεί τις ιδιότητες και τη συμπεριφορά της μεθόδου των Malgouyres και Guichard. Χάρη στη γενίκευση αυτή, μάς δίνεται η δυνατότητα της σύγκρισης της μεθόδου αυτής με άλλες μεθόδους, ακόμα και σε παραδείγματα με έγχρωμες εικόνες (όπως είναι το τρέχον και το επόμενο παράδειγμα). Σαν πυρήνα συνέλιξης της μεθόδου, επιλέξαμε τον διακριτό 2Δ παλμό, έκτασης 5×5 pixels. Η υλοποίηση της μεθόδου αυτής περιγράφεται από τον γενικό αλγόριθμο του πίνακα 5.13 (κάνοντας τις επιλογές που αναφέραμε προηγουμένως).

- Επίσης, εξάγαμε την εικόνα 6.14g, εφαρμόζοντας την μέθοδο παρεμβολής των Tschumperlé-Deriche, όπως ακριβώς περιγράφεται και χρησιμοποιείται στα [52, 55] (βλ. και υπενόηση 3.1.3, όπου παρουσιάσαμε τα βασικά στοιχεία της μεθόδου αυτής). Έκτός όσων αναφέραμε στην 3.1.3, σημειώνουμε ότι η μέθοδος αυτή χρησιμοποιεί σαν αρχική πρόβλεψη την παρεμβολή bilinear, την οποία εξελίσσει μέσω της ΜΔΕ (2.18) (βλ. ενότητα 2.7), που πρότειναν οι Tschumperlé-Deriche για το Inpainting (με την εξομάλυνση δηλαδή να γίνεται μόνο κατά την διεύθυνση των ακμών θ_*). Υλοποιήσαμε την μέθοδο αυτή, εφαρμόζοντας τον γενικό αλγόριθμο που προτείνουμε στον πίνακα 5.13, αφού η μέθοδος αυτή είναι ισοδύναμη με την επιλογή στον αλγόριθμο αυτό της διακριτής συνάρτησης dirac σαν πυρήνα συνέλιξης, της ταχύτητας που προκύπτει από την ΜΔΕ (2.18) και με μόνη τροποποίηση ότι η αρχικοποίηση πρέπει να γίνει με παρεμβολή bilinear αντί του zero-padding.
- Ακόμα, η εικόνα 6.14h προέκυψε με εφαρμογή της πολύ πρόσφατης, νέας μεθόδου του Tschumperlé, που βασίζεται σε *Συνελίξεις Επικαμπύλιας Ολοκλήρωσης (Line Integral Convolution - LICs)*. Για λεπτομέρειες πάνω στη μέθοδο αυτή, παραπέμπουμε στα [53, 54].

Εξάγαμε το αποτέλεσμα της εικόνας 6.14h εφαρμόζοντας την υλοποίηση της μεθόδου αυτής που είναι δημόσια διαθέσιμη από τον Tschumperlé στην ιστοσελίδα:

<http://www.greyc.ensicaen.fr/~dtschump/greycstoration/>

Πιο συγκεκριμένα, τρέξαμε (με τις κατάλληλες παραμέτρους) τον αλγόριθμο “greycstoration” της έκδοσης 1.0.8 του πακέτου “CImg Library”.

- Η εικόνα 6.14i προέκυψε με εφαρμογή της νέας μεθόδου που προτείνουμε στην ενότητα 3.2 (βλ. γενική ΜΔΕ (3.8)). Σαν ταχύτητα που προβάλλεται στον υποχώρο $\mathbf{U}_{0,s}$, επιλέξαμε την ταχύτητα της ΜΔΕ της *Ανισοτροπικής Εξομάλυνσης των Tschumperlé-Deriche* για την Βελτίωση εικόνων (βλ. ενότητα 1.10, με την εξομάλυνση δηλαδή να γίνεται και

στις δύο κάθετες κατευθύνσεις θ_-^* και θ_+^* , σε αντίθεση με τη μέθοδο που προτείνουν οι Tschumperlé, Deriche στα [52, 55] και είδαμε πιο πάνω).

Όπως είπαμε στην ενότητα 3.2, η μέθοδος που προτείνουμε είναι γενική και μπορεί, για την εξαγωγή της τελικής ΜΔΕ να χρησιμοποιήσει οποιαδήποτε ΜΔΕ διάχυσης. Ο λόγος που κάναμε την συγκεκριμένη επιλογή της μεθόδου της Ανισοτροπικής Εξομάλυνσης των Tschumperlé-Deriche είναι ότι, όπως είδαμε και στα παραδείγματα 1 και 2 της ενότητας 6.1, η μέθοδος αυτή παρουσιάζει στο πρόβλημα της Βελτίωσης τα πιο ικανοποιητικά (για την ανθρώπινη αντίληψη τουλάχιστον) αποτελέσματα σε σχέση με όλες τις υπόλοιπες μεθόδους που υλοποιήσαμε. Λόγω των αντιστοιχιών του προβλήματος της Βελτίωσης με το πρόβλημα της Παρεμβολής, όπως το τελευταίο αντιμετωπίζεται με τη μεθόδό μας (βλ. συζήτηση στην ενότητα 3.2), συμπεραίνουμε ότι η παραπάνω ΜΔΕ αποτελεί την πιο κατάλληλη επιλογή για τη μεθόδό μας. Σημειώνουμε τέλος ότι, όπως και στη μέθοδο παρεμβολής βασισμένης στην TV, επιλέξαμε σαν πυρήνα συνελίξης, τον διακριτό 2Δ παλμό, έκτασης 5×5 pixels.

Σχολιασμός των αποτελεσμάτων και συγκρίσεις

Από τα σχήματα 6.14 και 6.15, προκύπτουν οι εξής παρατηρήσεις για τα αποτελέσματα των διαφόρων μεθόδων:

- Η bilinear παρεμβολή (σχήματα 6.14c, 6.15c) δίνει ένα σημαντικά θολωμένο αποτέλεσμα, το οποίο παρουσιάζει επιπλέον σε ορισμένα σημεία (βλ. εικόνα 6.15c) το φαινόμενο των σκαλοπατιών (staircase effect). Η μέθοδος b-spline (σχήματα 6.14d, 6.15d) καταφέρνει να αποφύγει το φαινόμενο των σκαλοπατιών, αλλά το αποτέλεσμά της είναι υπερβολικά θολωμένο. Βλέπουμε λοιπόν ότι, σε περιπτώσεις σαν αυτή όπου η αρχική εικόνα είναι πολύ χαμηλής ευκρίνειας, οι μέθοδοι που βασίζονται σε γραμμικούς τελεστές δίνουν αποτελέσματα, τα οποία δεν είναι καθόλου ικανοποιητικά. Διαφαίνεται λοιπόν η ανάγκη χρήσης μη γραμμικών μεθόδων παρεμβολής.
- Το αποτέλεσμα της μεθόδου zero-padding (σχήματα 6.14e, 6.15e) παρουσιάζει υπερβολικές ταλαντώσεις (Gibbs effects, λόγω του απότομου μηδενισμού στο φασματικό περιεχόμενο - βλ. υποενότητα 3.1.1) αλλά καταφέρνει να ανακατασκευάσει με ικανοποιητικό τρόπο τις γεωμετρικές δομές της εικόνας (το αποτέλεσμα είναι πολύ λιγότερο θολωμένο σε σχέση με την εικόνα 6.14c αλλά και την 6.14d). Βέβαια, λόγω των Gibbs effects, το αποτέλεσμα του zero-padding δεν μπορεί σε καμία περίπτωση να γίνει αποδεκτό σαν τελικό, αλλά αποτελεί μία πολύ καλή αρχικοποίηση για τις μεθόδους που βασίζονται στην Συνθήκη Αναστρεψιμότητας (βλ. εικόνες 6.14f και 6.14i).

- Η μέθοδος Total Variation των Malgoyres-Guichard (σχήματα 6.14f, 6.15f) καταφέρνει στα περισσότερα σημεία να εξαφανίσει τα Gibbs effects του zero-padding. Από την άλλη όμως, σε περιοχές με έντονες μεταβολές (βλ. π.χ. την κόρη του ματιού στην εικόνα 6.15(f)), εμφανίζονται ανεπιθύμητα μπλοκ με σχεδόν σταθερή τιμή (bloc effect), που κάνουν το αποτέλεσμα να μοιάζει με την απλή επανάληψη των pixels (duplication) της αρχικής εικόνας (βλ. εικόνα 6.15(b)). Το τελευταίο οφείλεται στο ότι, κατά την χρονική εξέλιξη της αντίστοιχης ΜΔΕ, η προβολή της ταχύτητας στον χώρο $\mathcal{U}_{0,s}$ (βλ. ενότητα 3.2) δημιουργεί κάποιες μικρές μεταβολές στις τιμές ανάμεσα σε γειτονικά 5×5 μπλοκ μεγέθυνσης. Η μέθοδος TV θεωρεί εσφαλμένα τις μεταβολές αυτές σαν ακμές και έτσι τις ενισχύει αντί να τις εξομαλύνει.
- Η μέθοδος Ανισοτροπικής Εξομάλυνσης των Tschump.-Der. (σχήματα 6.14g, 6.15g) δίνει ένα αποτέλεσμα που φαίνεται έντονα τεχνητό, στο οποίο οι γεωμετρικές δομές της εικόνας έχουν αλλοιωθεί και έχουν δημιουργηθεί νέες που δεν υπήρχαν. Αυτό οφείλεται στο ότι η ΜΔΕ που χρησιμοποιείται συντελεί διάχυση αποκλειστικά στην διεύθυνση των αξόνων θ^* . Επίσης, παρατηρούμε ότι σε “δύσκολα” σημεία, με λεπτές γεωμετρικές δομές (βλ. π.χ. το περίγραμμα και τη κόρη του ματιού στην εικόνα 6.15g), η μέθοδος αυτή ανακατασκευάζει τις ακμές δείχνοντας μία ανεπιθύμητη προτίμηση στην οριζόντια και κάθετη διεύθυνση. Το τελευταίο οφείλεται στο ότι η μέθοδος αυτή αφήνει αμετάβλητες τις τιμές στα pixels που αντιστοιχούν στο αρχικό αραιότερο πλέγμα, με συνέπεια να επέρχεται ισορροπία στη χρονική ροή της ΜΔΕ μόνον όταν η εξομάλυνση τείνει να αφήσει ανέπαφες τις τιμές αυτές.
- Η καινούργια μέθοδος LIC του Tschumperlé (σχήματα 6.14h, 6.15h), δεν εμφανίζει πλέον τα προβλήματα της προηγούμενης μεθόδου, δίνοντας ένα αποτέλεσμα αρκετά πιο φυσικό. Όμως, το αποτέλεσμα αυτό είναι σημαντικά θολωμένο, με πολύ μικρή βελτίωση σε σχέση με την bilinear παρεμβολή. Παρατηρούμε ότι η μέθοδος χρειάζεται μόνο 1 επανάληψη για να καταλήξει στο τελικό αποτέλεσμα.
- Τέλος, η νέα μέθοδος που προτείνουμε (σχήματα 6.14i, 6.15i) δίνει το πιο ικανοποιητικό αποτέλεσμα (για την ανθρώπινη αντίληψη τουλάχιστον), σε σχέση με όλες τις υπόλοιπες μεθόδους που εφαρμόσαμε. Βλέπουμε ότι η μέθοδος αυτή καταφέρνει να εξαφανίσει τελείως τα Gibbs effects της παρεμβολής zero-padding και παράλληλα να ανακατασκευάσει επιτυχώς τις διανυσματικές ακμές, δίνοντας το λιγότερο θολωμένο αποτέλεσμα.

Πιο αναλυτικά, συγκρίνοντας τις εικόνες 6.14i, 6.15i με τις 6.14f, 6.15f

παρατηρούμε ότι στη μέθοδό μας, σε αντίθεση με τη μέθοδο Total Variation, δεν διακρίνονται τα pixels της αρχικής εικόνας χαμηλής ευκρίνειας (ξεπερνιέται δηλ. το bloc effect). Επίσης, οι ακμές ανακατασκευάζονται με πιο φυσικό και αποτελεσματικό τρόπο (η διαφορά φαίνεται έντονα στις ακμές υψηλής καμπυλότητας - συνέκρινε π.χ. το μάτι στις εικόνες 6.15i και 6.15f). Τα παραπάνω οφείλονται στο ότι (παρόλο που και οι δύο μέθοδοι αρχίζουν από την εικόνα zero-padding και χρησιμοποιούν προβολή στον χώρο $U_{0,s}$), η ΜΔΕ που χρησιμοποιεί η συγκεκριμένη εφαρμογή της μεθόδου μας (βλ. σχέση (1.67) ενότητας 1.10) είναι πιο αποτελεσματική από την TV PDE (βλ. συζήτηση στο τέλος της ενότητας 3.2). Μία από τις κύριες διαφορές τους είναι ότι η ΜΔΕ που χρησιμοποιούμε χρησιμοποιεί τανυστή διάχυση αντί του απλού συντελεστή καθώς επίσης ότι συντελεί εξομάλυνση στον δομικό τανυστή $J_p(\nabla \mathbf{u})$ (βλ. σχέση (1.68)), με συνέπεια η διάχυση να ακολουθεί τις πραγματικές γεωμετρικές δομές της εικόνας με συνεκτικό τρόπο, χωρίς να επηρεάζεται από ψεύτικες μικρές ακμές στα όρια των μπλοκ μεγέθυνσης.

Συγκρίνοντας τώρα τις εικόνες 6.14i, 6.15i με τις 6.14g, 6.15g, βλέπουμε ότι το αποτέλεσμα της μεθόδου που προτείνουμε δίνει ένα αποτέλεσμα, που φαίνεται σημαντικά πιο φυσικό και με τις γεωμετρικές δομές να έχουν ανακατασκευαστεί με μεγαλύτερη επιτυχία, σε σχέση με την μέθοδο Ανισοτροπικής Εξομάλυνσης των Tschump.-Der.. Αυτό οφείλεται στο ότι, παρόλο που και οι δύο μέθοδοι βασίζονται στην ίδια γενική ΜΔΕ (1.67), την χρησιμοποιούν με σημαντικά διαφορετικό τρόπο (βλ. όσα διευκρίνισαμε για τις μεθόδους αυτές πιο πάνω, στην υποενότητα 6.3.1).

Τέλος, από την σύγκριση των εικόνων 6.14i, 6.15i με τις 6.14h, 6.15h, βλέπουμε ότι το αποτέλεσμα της μεθόδου μας είναι πιο ικανοποιητικό και από αυτό της πολύ πρόσφατης μεθόδου LIC του Tschumperlé, μιας και είναι αρκετά λιγότερο θολό και με καλύτερη ανακατασκευή των διανυσματικών ακμών. Από την άλλη, όμως, διαπιστώνουμε ότι η μέθοδος LIC παρουσιάζει πολύ μικρότερη υπολογιστική πολυπλοκότητα σε σχέση με τη μέθοδό μας (αλλά και με τις υπόλοιπες μη γραμμικές μεθόδους), μιας και χρειάζεται μόλις 1 επανάληψη (σε σύγκριση με 300 ή 200 των άλλων μεθόδων) και αρκετά λιγότερο χρόνο.

Αξιολόγηση των αποτελεσμάτων της Παρεμβολής με την εύρεση των Σηματοθορυβικών λόγων (SNR)

Επιπρόσθετα, για τα αποτελέσματα του σχήματος 6.14, υπολογίσαμε τους SNRs και Average Local SNRs (βλ. σχέσεις (6.1) και (6.2) αντίστοιχα), θεωρώντας σαν αρχική καθαρή εικόνα \mathbf{u}^0 την εικόνα 6.14a1, από την οποία, όπως

είπαμε προήλθε η αρχική, χαμηλής ευκρίνειας εικόνα 6.14a2. Οι μετρήσεις αυτές παρουσιάζονται στον πίνακα 6.2.

| Μέθοδος Παρεμβολής | Σχήμα | SNR (dB) | Aver.Loc.SNR, με 25×25 μπλοκ (dB) |
|--------------------------------|-------|----------|--------------------------------------|
| Duplication | 6.14b | 09.11 | 4.25 |
| Bilinear | 6.14c | 09.58 | 4.80 |
| B-spline | 6.14d | 08.74 | 3.85 |
| Zero-padding | 6.14e | 09.56 | 4.56 |
| Total Variation (Malg.-Guich.) | 6.14f | 10.04 | 5.26 |
| Ανισοτρ. Εξομάλ. Tschump.-Der. | 6.14g | 09.39 | 4.64 |
| LIC (Tschump.) | 6.14h | 08.98 | 4.09 |
| Νέα προτεινόμενη μέθοδος | 6.14i | 10.24 | 5.57 |

Πίνακας 6.2: Σηματοθορυβικοί λόγοι (SNR) για τα αποτελέσματα του σχήματος 6.14

Από τον πίνακα αυτό, φαίνεται ότι ο καλύτερος SNR αλλά και Average Local SNR επιτυγχάνεται με την νέα μέθοδο που προτείναμε, κάτι που συμβαδίζει με αυτά που συμπεράναμε παρατηρώντας και συγκρίνοντας τις αντίστοιχες εικόνες. Βέβαια, πρέπει να σημειώσουμε ότι οι μετρήσεις αυτές δεν είναι πάντα αντιπροσωπευτικές του “μεγέθους της ομοιότητας” με την αρχική εικόνα, όπως ψυχολογικά την αντιλαμβάνεται ο ανθρώπινος παρατηρητής (βλ. και συζήτηση για τις μετρήσεις αυτές στην υποενότητα 6.1.1).

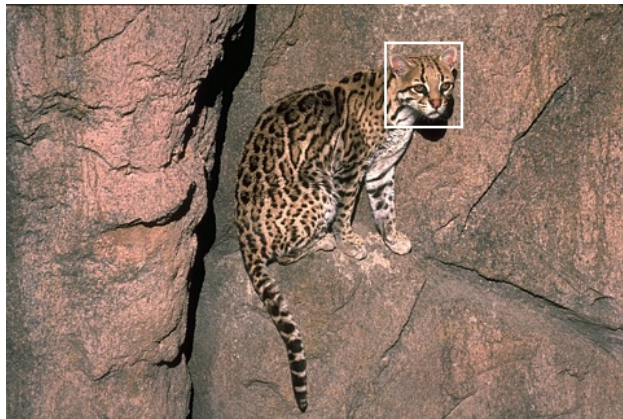
6.3.2 Παράδειγμα 2

Στο παράδειγμα αυτό, χρησιμοποιούμε την έγχρωμη εικόνα του σχήματος 6.16a0 (που προέρχεται από τη βάση εικόνων του Πανεπιστημίου του Berkeley). Σκοπός μας είναι η απομόνωση και μεγέθυνση (κατά τον παράγοντα 5×5) του μικρού κομματιού που φαίνεται στο σχήμα 6.16a0. Εφαρμόσαμε τις διάφορες μεθόδους παρεμβολής και καταλήξαμε στα αποτελέσματα που φαίνονται στα σχήματα 6.16, 6.17. Βέβαια στο παράδειγμα αυτό, σε αντίθεση με το προηγούμενο, δεν έχουμε εξαρχής στη διάθεσή μας κάποια εικόνα υψηλότερης ευκρίνειας.

Όσον αφορά τον τρόπο εφαρμογής των διαφόρων μεθόδων που εφαρμόσαμε, ισχύουν ξανά όσα διευκρινίσαμε στο προηγούμενο παράδειγμα.

Παρατηρώντας τα σχήματα 6.16, 6.17, διαπιστώνουμε ότι τα συμπεράσματα που μπορούν να εξαχθούν για τις μεθόδους είναι εντελώς αντίστοιχα με αυτά που παρουσιάσαμε αναλυτικά στο παράδειγμα 1.

Έτσι, η μέθοδος που προτείνουμε (βλ. εικόνες 6.16h, 6.17h) παρουσιάζει και στην περίπτωση αυτή τα καλύτερα αποτελέσματα σε σχέση με τις άλλες



(a0)



(a)



(b)

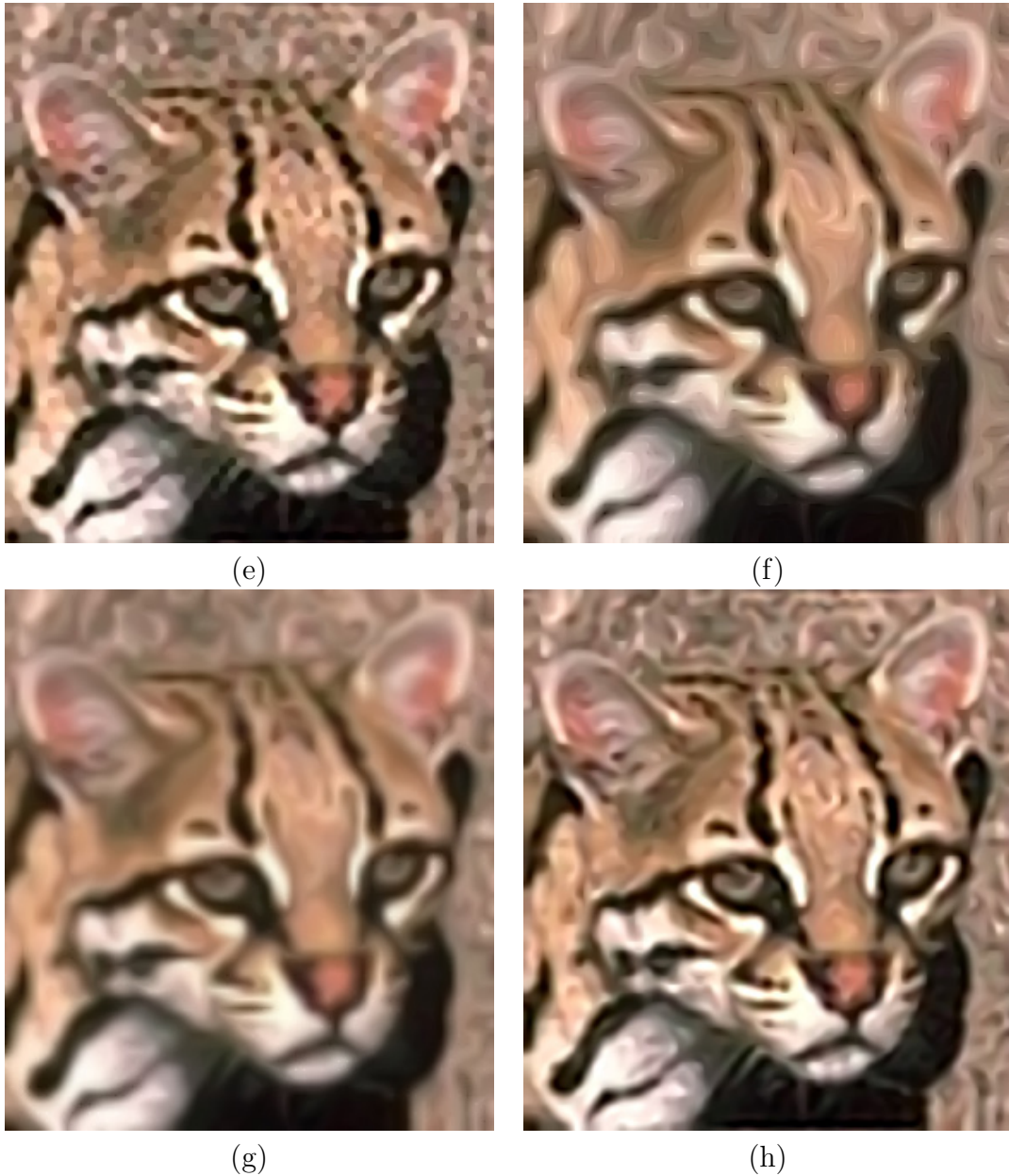


(c)



(d)

(συνέχεια στην επόμενη σελίδα)

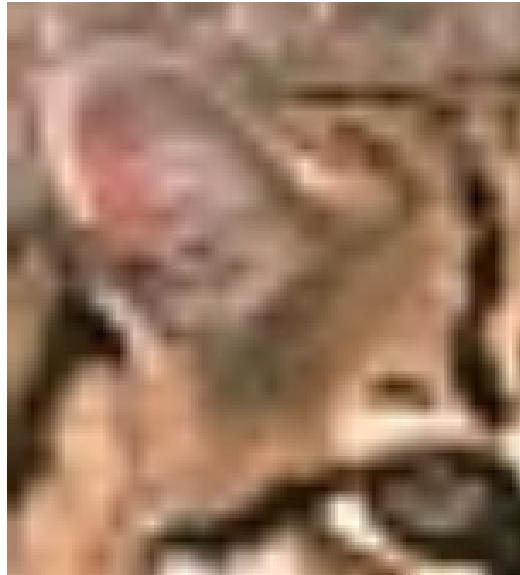


Σχήμα 6.16: 2^ο παράδειγμα σύγκρισης διάφορων μεθόδων παρεμβολής. (a0) Έγχρωμη εικόνα λεοπάρδαλης, όπου δείχνουμε το τμήμα προς μεγέθυνση (εντός του άσπρου πλαισίου, με μέγεθος 57×63 pixels). (a-h) Αποτελέσματα μεθόδων παρεμβολής για μεγέθυνση του τμήματος της (a0) με παράγοντα 5×5 (τελικό μέγεθος 285×315 pixels):

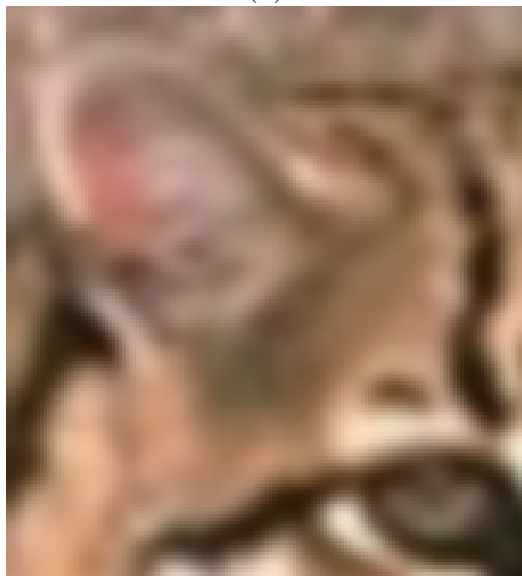
(a) *Duplication*. (b) *Bilinear*. (c) *B-spline*. (d) *Zero-padding*, με πυρήνα 5×5 2Δ παλμό (αρχική πρόβλεψη για τις (e),(h)). (e) Διανυσματική γενίκευση της *Παρεμβολής βασισμένης στην TV*, των *Malg.-Guich*. (250 επαναλήψεις, $\tau = 0.2$, $\epsilon = 1$, με πυρήνα 5×5 2Δ παλμό). (f) *Παρεμβολή με την Ανισοτρ. Εξομάλυνση των Tschump.-Deriche* (όπως προτείνεται στα [52, 55]) (300 επαναλήψεις, $\beta_{\max} = 5$, $\rho = 2$). (g) *Παρεμβολή βασισμένη σε LICs*, του *Tschump.* (1 επανάληψη). (h) *Νέα μέθοδος που προτείνουμε* (300 επαναλήψεις, $\beta_{\max} = 5$, $\rho = 2$, με πυρήνα 5×5 2Δ παλμό).



(a)



(b)

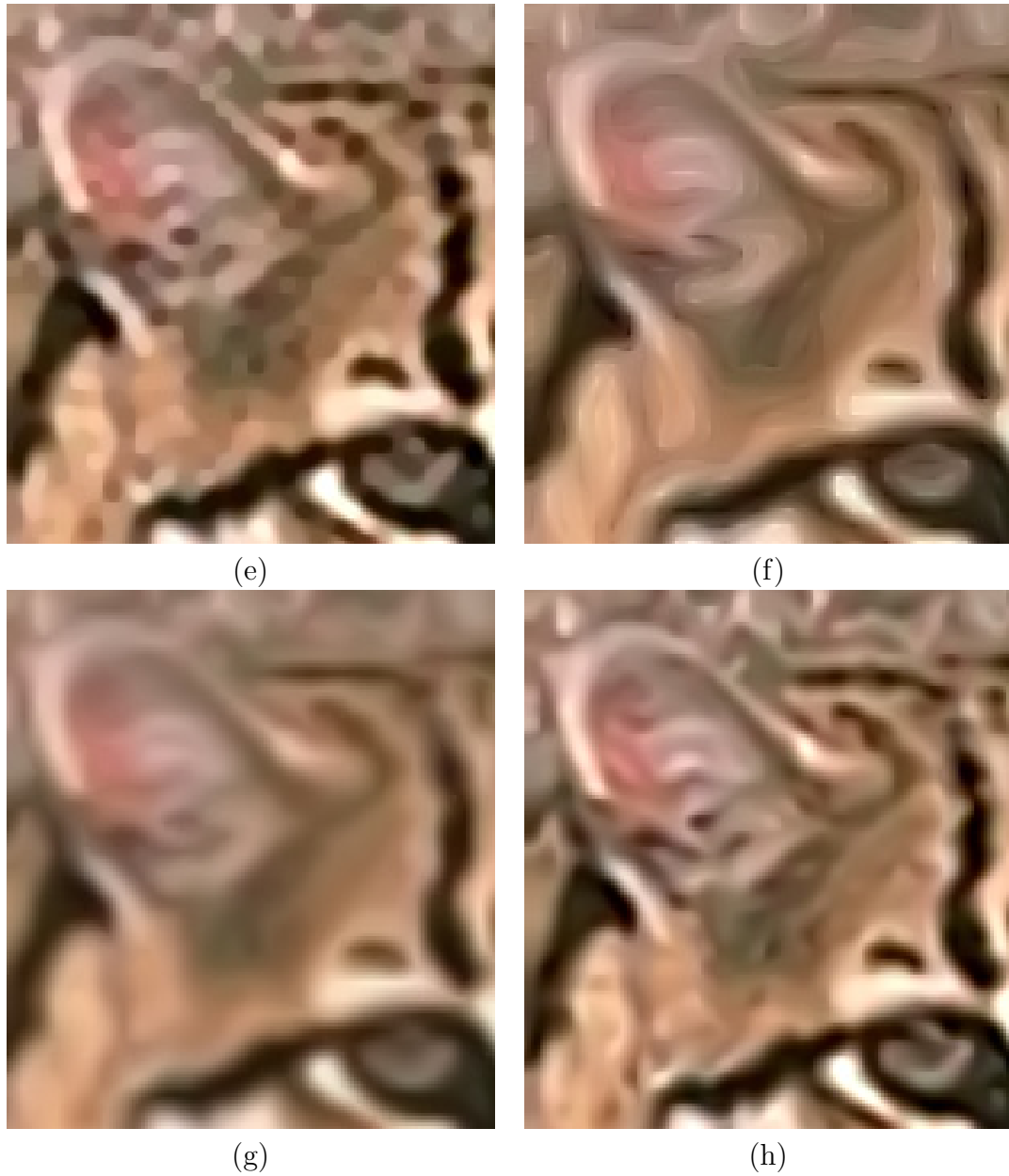


(c)



(d)

(συνέχεια στην επόμενη σελίδα)



Σχήμα 6.17: Αντίστοιχες λεπτομέρειες των αποτελεσμάτων του σχήματος 6.16 (μεγέθους 152×168 pixels).

(a) *Duplication*. (b) *Bilinear*. (c) *B-spline*. (d) *Zero-padding*. (e) *Total Variation (Malg.-Guich.)*. (f) *Ανισοτρ. Εξομάλ. Tschump.-Der.* (g) *LIC (Tschump.)*. (h) *Νέα προτεινόμενη μέθοδος*.

μεθόδους που εφαρμόσαμε. Σε σύγκριση με την μέθοδο Total Variation, η μέθοδός μας πλεονεκτεί διότι έχει καταφέρει με μεγαλύτερη επιτυχία να απομακρύνει τα Gibbs effects, δεν εμφανίζει το bloc effect και έχει καταλήξει σε ακμές που φαίνονται πιο ομαλές. Επίσης, το αποτέλεσμα της μεθόδου μας είναι πολύ πιο φυσικό από αυτό της μεθόδου Ανισοτροπικής Εξομάλυνσης των Tschump.-Der., η οποία έχει αλλοιώσει υπερβολικά τις δομές της εικόνας. Τέλος, οι ακμές είναι σημαντικά πιο εντωπισμένες χωρικά (λιγότερο θόλωμα) στο αποτέλεσμα της μεθόδου μας, σε σχέση με τη μέθοδο LIC του Tschumperlé (η οποία όμως παρουσιάζει πολύ μικρότερη υπολογιστική πολυπλοκότητα).

Όμως, από το σχήμα 6.16h παρατηρούμε και ένα μειονέκτημα της μεθόδου που προτείνουμε. Βλέπουμε ότι σε ορισμένα σημεία της περιοχής που αντιστοιχεί στον βράχο πίσω από το κεφάλι της λεοπάρδαλης (περιοχή με υφή), η μέθοδός μας έχει δημιουργήσει μονοδιάστατες δομές, οι οποίες φαίνονται τεχνητές. Βέβαια, το πρόβλημα αυτό εμφανίζεται επίσης στη μέθοδο LIC (Tschump.) και πολύ πιο έντονα στη μέθοδο Ανισοτροπικής Εξομάλυνσης (Tschump.-Der.). Τα παραπάνω οφείλονται στο ότι και οι τρεις αυτές μέθοδοι χρησιμοποιούν μία εξομάλυνση του δομικού τανυστή, η οποία, παρά τα θετικά της, οδηγεί σε αφύσικη αλλοίωση της υφής. Τέλος, σημειώνουμε ότι ούτε η μέθοδος Total Variation (βλ. εικόνα 6.16e) καταφέρει να χειριστεί με ικανοποιητικό τρόπο την υφή. Το παραπάνω αποτελεί ένα γενικό πρόβλημα όλων των μεθόδων ΜΔΕ διάχυσης που εξετάσαμε στην εργασία αυτή (βλ. και σχετική συζήτηση για την υφή στην υποενότητα 6.1.1).

Κεφάλαιο 7

Συμπεράσματα

7.1 Συμβολή της διπλωματικής εργασίας

Η συμβολή της παρούσας διπλωματικής εργασίας έγκειται κυρίως στα ακόλουθα σημεία:

- Αναλυτική παρουσίαση και σύγκριση πολλών από τις υπάρχουσες μεθόδους ΜΔΕ διάχυσης για τρία σημαντικά προβλήματα της Όρασης Υπολογιστών, την Βελτίωση και Απομάκρυνση του Θορύβου (Enhancement and Denoising), το Inpainting και την Παρεμβολή (Interpolation).
- Πρόταση δύο νέων ποσοτήτων για τον χειρισμό των γεωμετρικών δομών των διανυσματικών εικόνων, των Καμπυλών Ελάχιστης Μεταβολής και της Γενικευμένης Καμπυλότητας. Επίσης, χρησιμοποίηση της Γενικευμένης Καμπυλότητας για την πρόταση μίας νέας μεθόδου Inpainting διανυσματικών εικόνων, της Γενικευμένης CDD.
- Πρόταση μιας νέας γενικής μη γραμμικής μεθόδου Παρεμβολής διανυσματικών εικόνων, που βασίζεται σε ΜΔΕ διάχυσης. Η μέθοδος αυτή καταφέρνει σε αρκετές περιπτώσεις να καταλήξει σε σημαντικά βελτιωμένα αποτελέσματα, σε σχέση με αρκετές από τις υπάρχουσες γραμμικές αλλά και μη γραμμικές μεθόδους.
- Απόδειξη κάποιων νέων θεωρημάτων, που αφορούν την ακρίβεια και την σύγκλιση της προσέγγισης των ΜΔΕ διάχυσης (που μπορούν να γραφούν με Κατευθυνόμενες Λαπλασιανές) με επαναληπτική τοπική συνέλιξη.
- Απόδειξη ενός νέου θεωρήματος, το οποίο αποσαφηνίζει τη σχέση ανάμεσα στην συνεχή και στην εξαρχής διακριτή θεώρηση των ϕ -συναρτησιακών, τα οποία περιγράφουν με γενικό τρόπο πολλές από τις μεθόδους μεταβολών (variational methods) της ΟΥ.

- Πρόταση μίας νέας μεθόδου αξιολόγησης και σύγκρισης των διαφόρων ΜΔΕ διάχυσης μέσω της εισαγωγής της έννοιας του Ολικού Φίλτρου της ΜΔΕ. Επίσης, πρόταση ενός αποδοτικού αλγορίθμου για την επίτευξη του παραπάνω σκοπού.
- Ανάπτυξη και διεξοδική παρουσίαση, με όσο το δυνατόν πιο ενοποιητικό τρόπο, της αριθμητικής υλοποίησης πολλών από τις ΜΔΕ διάχυσης που εφαρμόζονται στην Όραση Υπολογιστών.

7.2 Κατευθύνσεις για μελλοντική έρευνα

Από τα θέματα με τα οποία ασχοληθήκαμε στην παρούσα διπλωματική εργασία και κυρίως από τα σημεία πρωτοτυπίας της, μπορούν να προκύψουν ορισμένες κατευθύνσεις για μελλοντική έρευνα. Οι πιο κύριες από αυτές είναι οι ακόλουθες:

- *Χρήση της Γενικευμένης Καμπυλότητας για την αποτελεσματική γενίκευση στις διανυσματικές εικόνες και άλλων μεθόδων επεξεργασίας βαθμωτών εικόνων.*

Η Γενικευμένη Καμπυλότητα που προτείναμε καταφέρει να μεταφέρει στις διανυσματικές εικόνες τις χρήσιμες ιδότητες της απλής καμπυλότητας των ισοϋψών γραμμών (που ορίζεται μόνο για τις βαθμωτές εικόνες). Στην εργασία αυτή, τη χρησιμοποιήσαμε μόνο για τη γενίκευση της υπάρχουσας μεθόδου Inpainting CDD. Θα μπορούσε όμως να χρησιμεύσει στην αποτελεσματική γενίκευση και άλλων μεθόδων επεξεργασίας βαθμωτών εικόνων που χρησιμοποιούν την καμπυλότητα.

Για παράδειγμα, η μέθοδος Euler's Elastica Inpainting (βλ. ενότητα 2.5) που προτάθηκε πρόσφατα βελτιώνει σημαντικά αρκετές από τις προηγούμενες μεθόδους Inpainting (συμπεριλαμβανομένης και της CDD). Μία γενίκευσή της στις διανυσματικές εικόνες \mathbf{u} με χρήση της Γενικευμένης Καμπυλότητας μπορεί να προκύψει ελαχιστοποιώντας το συναρτησιακό (αποτελεί γενίκευση του συναρτησιακού (2.13)):

$$J[\mathbf{u}] = \iint_{\mathbb{D}} \phi(\hat{\kappa}) \|\nabla \mathbf{u}\| \, d\mathbf{x} \quad (7.1)$$

όπου $\hat{\kappa}$ είναι η Γενικευμένη Καμπυλότητα της εικόνας \mathbf{u} στο εκάστοτε σημείο.

Με σκοπό την ελαχιστοποίηση του συναρτησιακού αυτού, από τις Euler-Lagrange Εξισώσεις και με ροή ταχύτερης καθόδου, θα μπορούσαμε να

καταλήξουμε σε μία ΜΔΕ (που θα αποτελεί γενίκευση της αντίστοιχης (2.15)).

Πιστεύουμε ότι η μέθοδος που θα προκύψει από την ΜΔΕ αυτή θα είναι ιδιαίτερα χρήσιμη, μιας και θα είναι το ίδιο αποτελεσματική με την αντίστοιχή της για τις βαθμωτές εικόνες.

- *Περαιτέρω μελέτη των ζητημάτων που αφορούν την ακρίβεια και την σύγκλιση της προσέγγισης της Γενικής Ανισοτροπικής Διάχυσης Κατευθυνόμενων Λαπλασιανών με επαναληπτική τοπική συνέλιξη.*

Στο θεώρημα 4.5, για την απόδειξη της σύγκλισης, υποθέσαμε ότι ο τανυστής T είναι ανισοτροπικός αλλά σταθερός. Μένει να διερευνηθεί αν το θεώρημα αυτό συνεχίζει να ισχύει όταν ο τανυστής αυτός μεταβάλλεται και εξαρτάται από την ίδια την εικόνα \mathbf{u} . Εάν κάτι τέτοιο αποδειχθεί, τότε η χρήση επαναληπτικής τοπικής συνέλιξης για την προσέγγιση της Γενικής Ανισοτροπικής Διάχυσης Κατευθυνόμενων Λαπλασιανών θα έχει θεμελιωθεί πλήρως θεωρητικά.

Επίσης, από το φράγμα του σφάλματος που εξάγαμε στο θεώρημα 4.4, προκύπτει ένας παράγοντας $A[k(\mathbf{x})]$, χαρακτηριστικός του πυρήνα συνέλιξης $k(\mathbf{x})$ που χρησιμοποιείται στην προσέγγιση αυτή. Η ελάττωση αυτού του “μέτρου” $A[k(\mathbf{x})]$, ελαττώνει το φράγμα σφάλματος που βρήκαμε. Μένει να διερευνηθεί εάν όντως η επιλογή ενός πυρήνα με κριτήριο την ελάττωση του $A[k(\mathbf{x})]$ μπορεί πρακτικά να οδηγήσει σε μία καλύτερη προσέγγιση.

- *Σχέση ανάμεσα στην συνεχή και στην εξαρχής διακριτή θεώρηση των ϕ -συναρτησιακών για την περίπτωση άλλων πιο αξιόπιστων διακριτοποιήσεων.*

Στην ενότητα 4.2, διερευνήσαμε τη σχέση ανάμεσα στην συνεχή και στην εξαρχής διακριτή θεώρηση των ϕ -συναρτησιακών, μόνο για την περίπτωση κάποιων απλών χωρικών διακριτοποιήσεων. Πιστεύουμε ότι παρόμοια αποτελέσματα θα μπορούσαν να εξαχθούν και στην περίπτωση άλλων διακριτοποιήσεων, πιο αξιόπιστων και με μεγαλύτερη πρακτική αξία, κάτι όμως το οποίο μένει να αποδειχθεί.

- *Περαιτέρω χρήση της μεθόδου του Ολικού Φίλτρου της ΜΔΕ*

Στην εργασία αυτή, εφαρμόσαμε τη μέθοδο του Ολικού Φίλτρου που προτείναμε μόνο στην αξιολόγηση κάποιων μεθόδων στο πρόβλημα της Βελτίωσης. Θα ήταν ιδιαίτερα ενδιαφέρον αυτή η μέθοδος αξιολόγησης να εφαρμοστεί πιο διεξοδικά, με σκοπό την αξιολόγηση αλλά και την σχεδία-

ση μεθόδων, τόσο στο πρόβλημα της Βελτίωσης όσο και στα προβλήματα του Inpainting και της Παρεμβολής.

Για παράδειγμα, η μέθοδος του Ολικού Φίλτρου θα μπορούσε να συνδυαστεί με την γενική μέθοδο Παρεμβολής που προτείναμε, δοκιμάζοντας επιλογές διαφόρων ΜΔΕ διάχυσης, ώστε τελικά να επιλεγεί η ΜΔΕ αυτή που πραγματικά παρουσιάζει την καλύτερη απόδοση.

- *Πιο αποδοτική υλοποίηση των αλγορίθμων που προτείναμε*

Η υλοποίηση των ΜΔΕ διάχυσης που αναπτύξαμε στην εργασία αυτή ήταν σχετικά απλή. Θα ήταν ιδιαίτερα χρήσιμο από πρακτικής σκοπιάς, ειδικά για τις νέες μεθόδους που προτείνουμε, να αναπτυχθούν πιο περίπλοκες υλοποιήσεις, με διακριτοποίηση μεγαλύτερης ακρίβειας.

Για παράδειγμα, αντί του απλού explicit σχήματος που χρησιμοποιήσαμε, θα μπορούσε να χρησιμοποιηθεί το AOS σχήμα (Additive Operator Splitting - βλ. [65]), το οποίο οδηγεί σε πολύ γρηγορότερη υλοποίηση. Έτσι, η χρονική πολυπλοκότητα των αλγορίθμων μπορεί να μειωθεί σημαντικά.

Βιβλιογραφία

- [1] A. Albiol και J. Serra. Morphological image enlargements. *J. Visual Comm. and Image Representation*, 8:367–383, 1997.
- [2] J. Allebach και P. W. Wong. Edge-directed interpolation. Στο *International Conference on Image Processing*, τόμος 3, σελίδες 707–710, Los Alamitos, CA, 1996. IEEE Press.
- [3] C. Ballester, M. Bertalmio, V. Caselles, G. Sapiro και J. Verdera. Filling-in by joint interpolation of vector fields and gray levels. *IEEE Trans. on Image Processing*, 10(8):1200–1211, 2001.
- [4] C. Ballester, V. Caselles και J. Verdera. Dissocclusion by joint interpolation of vector fields and gray levels. *Multiscale Model. Simul.*, 2(1):80–123, 2003.
- [5] D. Barash. Bilateral filtering and anisotropic diffusion : Towards a unified viewpoint. Τεχνική Αναφορά υπ. αριθμ., HP Laboratories Israel, 2000.
- [6] A. Belahmidi και F. Guichard. A partial differential equation approach to image zoom. *Int. Conf. on Image Processing*, σελίδες 649–652, 2004.
- [7] M. Bertalmio, G. Sapiro, V. Caselles και C. Ballester. Image inpainting. *Computer Graphics*, SIGGRAPH 2000, Ιούλιος 2000.
- [8] M. Black, G. Sapiro, D. Marimont και D. Heeger. Robust anisotropic diffusion. *IEEE Trans. on Image Processing*, 7(3):421–432, 1998.
- [9] P. Blomgren και T.F. Chan. Color TV: Total Variation methods for restoration of vector-valued images. *IEEE Trans. on Image Processing*, 7(3):304–309, 1998. Special Issue on Partial Differential Equations and Geometry-Driven Diffusion in Image Processing and Analysis.

- [10] F. Catte, P.L. Lions, J.M. Morel και T. Coll. Image selective smoothing and edge detection by nonlinear diffusion. *SIAM Journal Numer. Anal.*, 29(1):182–193, 1992.
- [11] A. Chambolle και P. Lions. Image recovery via total variation minimization and related problems. *Numer. Math.*, 76:167–188, 1997.
- [12] T. F. Chan, S. H. Kang και J. Shen. Euler’s elastica and curvature based inpaintings. Τεχνική Αναφορά, CAM 01-12, UCLA Dept. of Math., Απρίλιος 2001.
- [13] T. F. Chan, S. Osher και J. Shen. The digital TV filter and nonlinear denoising. *IEEE Trans. on Image Processing*, 10(2):231–241, 2001.
- [14] T. F. Chan και J. Shen. Mathematical models for local non-textured inpaintings. Τεχνική Αναφορά υπ. αριθμ., UCLA Dept. of Math., Μάρτιος 2000.
- [15] T. F. Chan και J. Shen. Mathematical models for local deterministic inpaintings. *SIAM Journal of Applied Math.*, 62(3):1019–1043, 2001.
- [16] T. F. Chan και J. Shen. Morphologically invariant PDE inpaintings. Τεχνική Αναφορά, CAM 01-15, UCLA Dept. of Math., 2001.
- [17] T. F. Chan και J. Shen. Non-texture inpainting by curvature-driven diffusion (CDD). *Journal of Visual Comm. and Image Rep.*, 12(4):436–449, 2001.
- [18] P. Charbonnier, G. Aubert, L. Blanc-Féraud και M. Barlaud. Two deterministic half-quadratic regularization algorithms for computed imaging. Στο *Proceedings of the International Conference on Image Processing*, τόμος II, σελίδες 168–172, 1994.
- [19] P. Charbonnier, L. Blanc-Féraud, G. Aubert και M. Barlaud. Deterministic edgepreserving regularization in computed imaging. *IEEE Trans. on Image Processing*, 6(2):298–311, 1997.
- [20] R. Deriche. Recursively implementing the gaussian and its derivatives. Τεχνική Αναφορά υπ. αριθμ. 1893, INRIA, Unité de Recherche Sophia-Antipolis, 1993.
- [21] R. Deriche και O. Faugeras. Les EDP en traitement des images et vision par ordinateur. Rapport de Recherche, INRIA, Νοέμβριος 1995.

- [22] S. Di Zeno. A note on the gradient of a multi-image. *Computer Vision, Graphics, and Image Processing*, 33:116–125, 1986.
- [23] N. A. Dodgson. Quadratic interpolation for image resampling. *IEEE Trans. on Image Processing*, 6(9):1322–1326, Σεπτέμβριος 1997.
- [24] G. Gilboa. *Super-resolution Algorithms Based on Inverse Diffusion-type Processes*. Διδακτορική Διατριβή, Department of Electrical Engineering, Technion - Israel Institute of Technology, 2004.
- [25] G. Gilboa, Y.Y. Zeevi και N. Sochen. Resolution enhancement by forward-and-backward nonlinear diffusion processes. *Nonlinear Signal and Image Processing, Baltimore, Maryland*, Ιούνιος 2001.
- [26] F. Guichard και F. Malgouyres. Total variation based interpolation. *EUSIPCO III*, σελίδες 1741–1744, 1998.
- [27] F. Guichard και J.M. Morel. *Image Analysis and P.D.E.'s*, κεφάλαιο 3. IPAM GBM Tutorials, 2001.
- [28] S. H. Kang. *Mathematical approaches to color denoising and image inpainting problems*. Διδακτορική Διατριβή, UCLA CAM 02-25, 2002.
- [29] R. G. Keys. Cubic convolution interpolation for digital image processing. *IEEE Trans. on Acoustics, Speech and Signal Processing*, 29:1153–1160, 1981.
- [30] R. Kimmel, R. Malladi και N. Sochen. Images as embedded maps and minimal surfaces: movies, color, texture, and volumetric medical images. *Int. Journal of Comp. Vision*, 39(2):111–129, Σεπτέμβριος 2000.
- [31] J.J. Koenderink. The structure of images. *Biological Cybernetics*, 50:363–370, 1984.
- [32] P. Kornprobst. *Contributions à la Restauration d'Images et à l'Analyse de Séquences: Approches Variationnelles et Solutions de Viscosité*. Διδακτορική Διατριβή, Université de Nice-Sophia Antipolis, 1998.
- [33] P. Kornprobst, R. Deriche και G. Aubert. Image restoration via PDE's. Στο *First Annual Symposium on Enabling Technol. for Law Enforc. and Security - SPIE Conference 2942 : Invest. Image Processing*, Boston, Massachusetts, USA, Νοέμβριος 1996.

- [34] P. Kornprobst, R. Deriche και G. Aubert. Nonlinear operators in image restoration. Στο *Proceedings of the Int. Conf. on Comp. Vision and Pattern Recogn.*, σελίδες 325–331, Puerto Rico, Ιούλιος 1997. IEEE Comp. Society, IEEE.
- [35] F. Malgouyres. *Increase in the Resolution of Digital Images: Variational Theory and Applications*. Διδακτορική Διατριβή, Ecole Normale Supérieure de Cachan, Cachan, France, 2000.
- [36] F. Malgouyres και F. Guichard. Edge direction preserving image zooming: a mathematical and numerical analysis. *SIAM Journal Numer. Anal.*, 39(1):1–37, 2002.
- [37] S. Masnou και J. M. Morel. Level lines based disocclusion. *5th IEEE Int. Conf. on Image Processing, Chicago*, 3:259–263, 1998.
- [38] D. Mumford. Elastica and computer vision. Στο *Algebraic Geometry and its Applications*, επιμελητής: C. L. Bajaj, σελίδες 491–506. Springer-Verlang, New York, 1994.
- [39] J. P. Oackley και M. J. Cunningham. A function space model for digital image sampling and its application in image reconstruction. *Computer Vision, Graphics, and Image Processing*, 49:171–197, 1990.
- [40] P. Perona και J. Malik. Scale space and edge detection using anisotropic diffusion. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 12(7):629–639, Ιούλιος 1990.
- [41] L. Rudin, S. Osher και E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60:259–268, 1992.
- [42] G. Sapiro και D.L. Ringach. Anisotropic diffusion of multivalued images with applications to color filtering. *IEEE Trans. on Image Processing*, 5(11):1582–1585, 1996.
- [43] C. E. Shannon. Communication in the presence of noise. *Proceedings of the IRE*, 37:10–21, 1949.
- [44] N. Sochen, R. Kimmel και A. M. Bruckstein. Diffusions and confusions in signal and image processing. *Journal of Mathematical Imaging and Vision*, 14(3):195–209, 2001.
- [45] N. Sochen, R. Kimmel και R. Malladi. From high energy physics to low level vision. Τεχνική Αναφορά υπ. αριθμ. 39243, UC Berkeley, 1996.

- [46] N. Sochen, R. Kimmel και R. Malladi. A geometrical framework for low level vision. *IEEE Transaction on Image Processing, Special Issue on PDE based Image Processing*, 7(3):310–338, 1998.
- [47] Π. Μαραγκός. *Ανάλυση Εικόνων και Όραση Υπολογιστών*. Ε.Μ.Π., Αθήνα, 2005.
- [48] Α. Μπακόπουλος και Ι. Χρυσοβέργης. *Εισαγωγή στην Αριθμητική Ανάλυση*. Εκδόσεις Συμεών, 1999.
- [49] Ν. Καδιανάκης, Σ. Καρανάσιος και Α. Φελλούρης. *Ανάλυση ΙΙ, Συναρτήσεις Πολλών Μεταβλητών*. Αθήνα, 5η έκδοση, 2000.
- [50] S. Thurnhofer και S. K. Mitra. Edge-enhanced image zooming. *Optical Engrg.*, 35:1862–1870, 1996.
- [51] C. Tomasi και R. Manduchi. Bilateral filtering for gray and color images. Στο *Proceedings of the IEEE International Conference on Computer Vision*, σελίδες 839–846, Ιανουάριος 1998.
- [52] D. Tschumperlé. *PDE's Based Regularization of Multivalued Images and Applications*. Διδακτορική Διατριβή, Université de Nice-Sophia Antipolis, Δεκέμβριος 2002.
- [53] D. Tschumperlé. Fast anisotropic smoothing of multi-valued images using curvature-preserving PDE's. Research Report “Les Cahiers du GREYC”, No 05/01, Equipe IMAGE/GREYC (CNRS UMR 6072), Φεβρουάριος 2005.
- [54] D. Tschumperlé. LIC-based regularization of multi-valued images. *IEEE Int. Conf. on Image Processing*, Σεπτέμβριος 2005.
- [55] D. Tschumperlé και R. Deriche. Vector-valued image regularization with PDE's : A common framework for different applications. Στο *IEEE Conf. on Comp. Vision and Pattern Recogn.*, Madison, Wisconsin (United States), Ιούνιος 2003.
- [56] A. Unser, M. Aldroubi και M. Eden. Fast b-spline transforms for continuous image representation and interpolation. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 13(3):277–285, 1991.
- [57] M. Unser, A. Aldroubi και M. Eden. Enlargement or reduction of digital image with minimum loss of information. *IEEE Trans. on Image Processing*, 4:247–258, 1995.

- [58] L. Vese. *Problèmes variationnelles et EDP pour l'analyse d'images et l'évolution de courbes*. Διδακτορική Διατριβή, Université de Nice-Sophia Antipolis, Νοέμβριος 1996.
- [59] J. Weickert. Multiscale texture enhancement. Στο *Computer Analysis of Images and Patterns*, επιμελητές: V. Hlavác και R. Sára, τόμος 970, σελίδες 230–237. *Lecture Notes in Computer Science*, Springer, Berlin, 1995.
- [60] J. Weickert. *Anisotropic Diffusion in Image Processing*. Διδακτορική Διατριβή, University of Kaiserslautern, Germany, Laboratory of Technomathematics, Ιανουάριος 1996.
- [61] J. Weickert. Coherence-enhancing diffusion of colour images. *7th National Symposium on Pattern Recognition and Image Analysis*, Απρίλιος 1997.
- [62] J. Weickert. A review of nonlinear diffusion filtering. *Scale-Space Theory in Computer Vision, Lecture Notes in Comp. Science (Springer, Berlin)*, 1252:3–28, 1997. Invited Paper.
- [63] J. Weickert. Nonlinear diffusion filtering. Στο *Handbook on Computer Vision and Applications, Vol. 2: Signal Processing and Pattern Recognition*, επιμελητές: B. Jähne, H. Haußecker και P. Geißler, σελίδες 423–450. Academic Press, San Diego, 1999.
- [64] J. Weickert και H. Scharr. A scheme for coherence-enhancing diffusion filtering with optimized rotation invariance. *Journal of Visual Communication and Image Representation*, 13:103–118, 2002.
- [65] J. Weickert, B.ter Haar Romeny και M. Viergever. Efficient and reliable schemes for nonlinear diffusion filtering. *IEEE Trans. on Image Processing*, 7(3):398–410, Μάρτιος 1998.
- [66] A.P. Witkin. Scale-space filtering. Στο *Int. Joint Conf. on Artificial Intel.*, σελίδες 1019–1022, 1983.
- [67] L. Yaroslavsky. Efficient algorithm for discrete sinc interpolation. *Appl. Optics*, 36:460–463, 1997.