



Εθνικό Μετσόβιο Πολυτεχνείο

Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών

Τομέας Σημάτων, Ελέγχου και Ρομποτικής

Εργαστήριο Όρασης Υπολογιστών, Επικοινωνίας Λόγου και Επεξεργασίας Σημάτων

Ψηφιακή Αποκατάσταση των Τοιχογραφιών του Ακρωτηρίου Θήρας

Διπλωματική Εργασία

του

Νικολάου Καριανάκη

Επιβλέπων: Πέτρος Μαραγκός
Καθηγητής Ε.Μ.Π.

Αθήνα, Νοέμβριος 2011



Εθνικό Μετσόβιο Πολυτεχνείο

Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών

Τομέας Σημάτων, Ελέγχου και Ρομποτικής

Εργαστήριο Όρασης Υπολογιστών, Επικοινωνίας Λόγου και Επεξεργασίας Σημάτων

Ψηφιακή Αποκατάσταση των Τοιχογραφιών του Ακρωτηρίου Θήρας

Διπλωματική Εργασία

του

Νικολάου Καριανάκη

Επιβλέπων: Πέτρος Μαραγκός
Καθηγητής Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 9^η Νοεμβρίου 2011.

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

.....
Γεώργιος Καραγιάννης
Καθηγητής Ε.Μ.Π.

.....
Κωνσταντίνος Τζαφέστας
Επίκουρος Καθηγητής Ε.Μ.Π.

Αθήνα, Νοέμβριος 2011

.....

Νικόλαος Καριανάκης

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

Copyright © Νικόλαος Καριανάκης, 2011.

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 τον συγγραφέα.

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

Περίληψη

Σκοπός μας είναι η επιδιόρθωση των τοιχογραφιών του Ακρωτηρίου Θήρας, με παρόμοιο τρόπο όπως θα το έκανε ένας ειδικός τεχνίτης, μέσω μίας ημι-αυτοματοποιημένης όμως διαδικασίας. Η ψηφιακή αποκατάσταση ξεκινά με την αυτόματη αναγνώριση των φθαρμένων και ελλειπουσών περιοχών και ολοκληρώνεται με την ενδοσυμπλήρωσή τους.

Προτείνεται ένας αλγόριθμος που χρησιμοποιεί στοιχεία μαθηματικής μορφολογίας για μία αδρή ανίχνευση των ελλείψεων. Στη συνέχεια ενσωματώνεται η πληροφορία των ακμών για πληρέστερη κάλυψη των ελλειπουσών περιοχών, με παράλληλη μικρή αύξηση της ακρίβειας της ανίχνευσής τους. Ο αλγόριθμος αυτός οδήγησε στην υλοποίηση μίας αρκετά αποτελεσματικής μεθόδου για την εξαγωγή της μάσκας των φθορών. Σε κάποιες περιπτώσεις σε μία προσπάθεια ακόμη ακριβέστερης προσαρμογής της μάσκας πάνω στις ελλείψεις χρησιμοποιούνται επιπλέον τομές γράφων και συγκεκριμένα ο επαναληπτικός αλγόριθμος “GrabCut”.

Παρουσιάζονται αποτελέσματα πειραμάτων τόσο ενδοσυμπλήρωσης περιεχομένου με το μοντέλο Ολικής Μεταβολής όσο και ενδοσυμπλήρωσης υψής με τον αλγόριθμο Efros-Leung. Για μεγάλες ελλείψεις ο χρήστης ανασύρει πληροφορία αντίστοιχης σημασιολογίας, γεωμετρίας και χρώματος από άλλες περιοχές των τοιχογραφιών, η οποία επικολλάται στα κενά με αλγορίθμους οπτικά βέλτιστης συνένωσης εικόνων. Μελετώνται πρόσφατοι αλγόριθμοι σύνθεσης μωσαϊκού και για την εφαρμογή μας αξιοποιείται ο Tao et al.

Λέξεις-Κλειδιά

ψηφιακή αποκατάσταση Τοιχογραφιών Θήρας, συμπλήρωση παράστασης, εξαγωγή μάσκας, μαθηματική μορφολογία, ανίχνευση ακμών, κατάτμηση εικόνας με τομές γράφων, “GrabCut”, ενδοσυμπλήρωση περιεχομένου, μερικές διαφορικές εξισώσεις, μεταβολικές μέθοδοι, ενδοσυμπλήρωση υψής, αλγόριθμος Efros-Leung, εύρωστη σε σφάλματα σύνθεση μωσαϊκού, συνένωση στο πεδίο των παραγώγων, αλγόριθμος Tao et al., Επεξεργασία Εικόνας, Όραση Υπολογιστών

Abstract

We present a semi-automatic method for high quality restoration of Theran wall paintings, which is very close to the work of a specialist. Digital restoration begins with the automatic recognition of damaged and missing areas and is followed by their inpainting.

An algorithm based on mathematical morphology elements is proposed in order to do rough detection of damaged areas. We improve upon this approach by incorporating edge information, leading to more complete detection of damage, achieving higher recall, and usually higher precision. Our results using this approach lead to excellent identification of missing areas in the wall paintings. In some cases we further improve the extracted mask by using iterated graph cuts, specifically with the “GrabCut” algorithm.

We demonstrate the results of experiments doing structure inpainting with the Total Variation model and texture inpainting with the Efros-Leung algorithm. In the case of significant information loss, the user selects an area of similar semantics, geometry and color from a different location in the wall paintings, which is then stitched into the missing area seamlessly. We present an analysis of recent image compositing algorithms and demonstrate the results of Tao et al. on our problem.

Keywords

Theran Wall Paintings digital restoration, scene completion, mask extraction, mathematical morphology, edge detection, image segmentation with graph cuts, “GrabCut”, structure inpainting, Partial Differential Equations (PDEs), variational methods, texture inpainting, Efros-Leung algorithm, error-tolerant image stitching, gradient-domain compositing, Tao et al. algorithm, Image Processing, Computer Vision

Πρόλογος

0.1 Όραση Υπολογιστών και στόχοι διπλωματικής

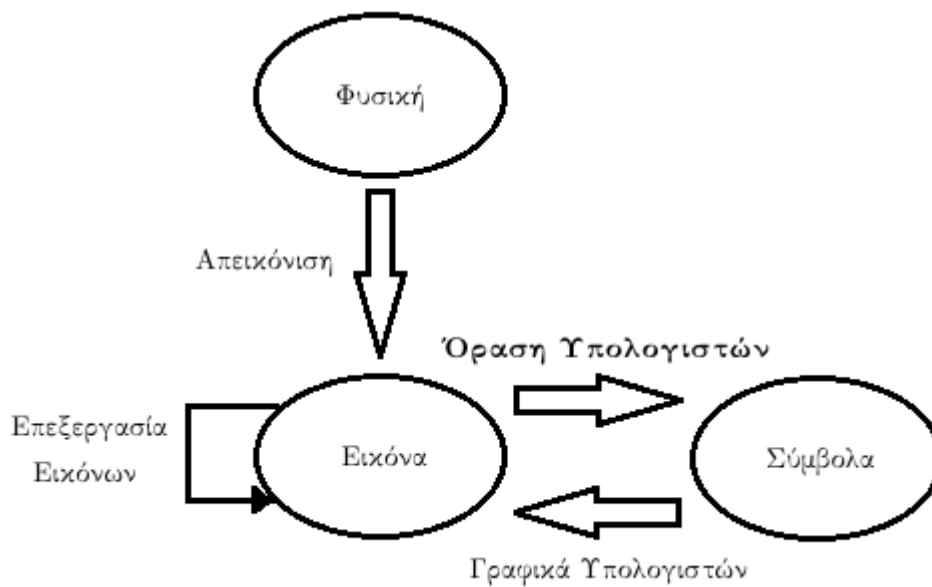
Η Όραση Υπολογιστών (ΟΥ) καθιερώθηκε ως ανεξάρτητο επιστημονικό πεδίο τη δεκαετία του 1960 λόγω της προσπάθειας ερευνητών της *Τεχνητής Νοημοσύνης* να δημιουργήσουν μηχανές που μπορούν να δουν, να αντιληφθούν και να να εξάγουν συμπεράσματα όπως και ο άνθρωπος. Τις επόμενες δύο δεκαετίες εξελίχθηκε ως κλάδος με τη συμβολή των επιστημονικών πεδίων της *Επεξεργασίας Σήματος & Εικόνας* και της *Αναγνώρισης Προτύπων*. Στην πορεία η ΟΥ έχει εξελιχθεί σε μία πολύ ευρεία διεπιστημονική περιοχή, στην οποία συμβάλλουν επιστημονικά πεδία όπως η *Νευροεπιστήμη*, η *Ψυχοφυσική* και η *Στατιστική*.

Η ΟΥ ανήκει στις *επιστήμες εικόνων* (image sciences), στις οποίες περιλαμβάνονται ακόμα τα *Γραφικά Υπολογιστών*, η *Επεξεργασία Εικόνων* και η *Απεικόνιση*. Ενώ η ΟΥ προσπαθεί από 2Δ εικόνες να εξάγει συμπεράσματα για την πιθανή ύπαρξη, τις ιδιότητες της επιφάνειας και το είδος των αντικειμένων του 3Δ κόσμου, δηλαδή επιχειρεί να επιλύσει *αντίστροφα* προβλήματα, τα Γραφικά Υπολογιστών κάνουν το ακριβώς αντίθετο, δηλαδή επιλύουν προβλήματα σύνθεσης με ευθύ τρόπο. Η Επεξεργασία Εικόνων καταπιάνεται κυρίως με προβλήματα μετασχηματισμού από μία εικόνα εισόδου σε μία εικόνα εξόδου. Το γενικό πλαίσιο λειτουργίας αυτών των κλάδων απεικονίζεται συμβολικά στο Σχήμα 1. Οι παραπάνω δύο παράγραφοι στηρίζονται στο [1].

Στόχος της διπλωματικής εργασίας είναι η διερεύνηση μεθόδων της Όρασης Υπολογιστών για την ψηφιακή αποκατάσταση των τοιχογραφιών του Προϊστορικού Οικισμού Ακρωτηρίου Θήρας. Η διαδικασία της αποκατάστασης αφορά τον αυτόματο εντοπισμό των φθορών και ελλείψεων που δημιουργήθηκαν με το πέρασμα των αιώνων και εν συνεχεία την επιδιόρθωσή τους με τεχνικές ενδοσυμπλήρωσης (inpainting). Επειδή κάποιες ελλείψεις είναι πολύ μεγάλες για την ικανοποιητική αποκατάστασή τους μόνο με μεθόδους παρεμβολής (interpolation), σε αυτές τις περιπτώσεις ανασύρεται πληροφορία παρεμφερούς γεωμετρίας και σημασιολογίας από άλλες περιοχές των τοιχογραφιών. Για την κάλυψη της έλλειψης με τη νεότερη περιοχή αξιοποιούνται πρόσφατοι αλγόριθμοι σύνθεσης μωσαϊκού, ώστε το τελικό αποτέλεσμα να είναι αισθητικά άρτιο.

0.2 Διάρθρωση και συμβολή διπλωματικής εργασίας

Το χείμενο οργανώνεται σε πέντε κεφάλαια. Στο 1^ο Κεφάλαιο παρουσιάζονται κάποιες γενικές πληροφορίες για τις τοιχογραφίες, τον τύπο των ελλείψεων που παρουσιάζουν και τη ψηφιοποίησή τους. Στο σημείο αυτό πρέπει να ευχαριστηθεί ο Γ. Παπανδρέου για τη διάθεση της βάσης ψηφιοποιημένων εικόνων των τοιχογραφιών από τον προϊστορικό οικισμό Ακρωτηρίου Θήρας που



Σχήμα 1: Σχηματικό διάγραμμα σχετικά με το γενικό περιεχόμενο των επιστημών εικόνων.

δημιούργησε για τις ανάγκες του διδακτορικού του [2]. Η ψηφιοποίησή τους έγινε στα πλαίσια του διεπιστημονικού ερευνητικού προγράμματος με τίτλο *Όραση Υπολογιστών και Εικονική Πραγματικότητα για Τρισδιάστατη Ανακατασκευή Αρχαιολογικών Ενρημάτων με Εφαρμογές σε Τοιχογραφίες και Κτίρια του Προϊστορικού Οικισμού Ακρωτηρίου Θήρας*, το οποίο εντάσσεται στην πρωτοβουλία ΠΕΝΕΔ-2003 και χρηματοδοτείται από τη Γενική Γραμματεία Έρευνας και Τεχνολογίας, με την συνδρομή του Καθ. Χρίστου Ντούμα, διευθυντή των ανασκαφών, και του εργαστηρίου τοιχογραφιών του Ακρωτηρίου Θήρας. Στο 1^ο Κεφάλαιο γίνεται ακόμη μία πλήρης εισαγωγή στο πρόβλημα της ενδοσυμπλήρωσης (inpainting). Δίνεται ο αυστηρός μαθηματικός ορισμός του και εν συνεχεία εντάσσεται στο θεμελιώδες πρόβλημα επεξεργασίας εικόνας. Παράλληλα, γίνεται μία πλήρης βιβλιογραφική επισκόπηση των μεθόδων επίλυσής του που κατηγοριοποιούνται με βάση το πώς μοντελοποιείται η εικόνα κάθε φορά.

Το 2^ο Κεφάλαιο αφορά την αυτόματη ανίχνευση των ελλειπουσών περιοχών στις τοιχογραφίες. Βρέθηκαν και υλοποιήθηκαν αλγόριθμοι που χρησιμοποιούν στοιχεία μαθηματικής μορφολογίας και ανίχνευση ακμών. Επιπλέον, παρουσιάζεται το πρόβλημα της κατάτμησης και γίνεται μία επισκόπηση της πρόσφατης βιβλιογραφίας. Το κεφάλαιο καταλήγει στη μέθοδο τομών γράφων για ακριβή κατάτμηση και συγκεκριμένα στον αλγόριθμο “GrabCut” [3]. Δίνονται αποτελέσματα εφαρμογής του στις τοιχογραφίες. Για το “GrabCut” χρησιμοποιήθηκε υλοποίηση που διανέμεται ελεύθερα στο διαδίκτυο από τους Gupta και Ramnath (Robotics Institute, CMU, Pittsburgh). Μάλιστα προτείνεται ένας αλγόριθμος που παίρνει ως είσοδο τη μάσκα που προκύπτει με μαθηματική μορφολογία και αξιοποιεί το “GrabCut” για την καλύτερη δυνατή αναγνώριση των ελλείψεων.

Το 3^ο Κεφάλαιο καταπιάνεται με το κομμάτι του inpainting. Δίνεται μία γενική περιγραφή του total variation inpainting μοντέλου, τόσο από τη θεωρητική σκοπιά και τη μαθηματική του θεμελίωση όσο και αριθμητικό σχήμα για υλοποίησή του στον υπολογιστή. Στο σημείο αυτό πρέπει να εκφραστούν οι ευχαριστίες στον Ε. Πνευματικάκη για την ευγενική παραχώρηση του πηγαίου κώδικά του για το TV inpainting μοντέλο που είχε υλοποιήσει στα πλαίσια της διπλωμα-

τικής του εργασίας [4]. Χρησιμοποιήθηκε ως βάση για την υλοποίηση του παρόντος αλγορίθμου. Μεταβλήθηκαν οι παράμετροι του μοντέλου για την προσαρμογή του στην παρούσα εφαρμογή, συστηματοποιήθηκε η διαδικασία σύγκλισής του και προστέθηκαν κάποιες συναρτήσεις για την σχηματική απόδοση των αποτελεσμάτων. Στο τέλος του κεφαλαίου γίνεται μία εισαγωγή στην υφή και παρουσιάζεται η μέθοδος σύνθεσης υφής των Efros και Leung, καθώς και η τροποποιημένη εκδοχή της για το πρόβλημα της ενδοσυμπλήρωσης. Η υλοποίηση του E. Πνευματικάκη αποτέλεσε τη βάση των πειραμάτων.

Στο 4^ο Κεφάλαιο περιγράφονται σύγχρονοι αλγόριθμοι σύνθεσης μωσαϊκού, με στοιχεία μαθηματικής θεμελίωσης και συγκριτική ανάλυση. Έμφαση δίνεται στους αλγορίθμους των Levin et al. [5] και Tao et al. [6]. Μάλιστα παρουσιάζονται αποτελέσματα από την εφαρμογή του αλγορίθμου των Tao et al., η οποία διατίθεται ελεύθερα στο διαδίκτυο από τους συγγραφείς, στην ψηφιακή επιδιόρθωση των τοιχογραφιών. Η ανάκτηση σημασιολογικά και γεωμετρικά σχετικής πληροφορίας από άλλες περιοχές των έργων έδωσε αισθητικά σπουδαία αποτελέσματα.

Τέλος, στο 5^ο Κεφάλαιο συνοψίζονται οι συνεισφορές της παρούσας εργασίας, οι άξονες στους οποίους κινήθηκε, ενώ προτείνονται και κάποιες πιθανές κατευθύνσεις για μελλοντική έρευνα.

Ευχαριστίες

Με την ευκαιρία της περάτωσης της διπλωματικής αυτής εργασίας και κατά συνέπεια των προπτυχιακών μου σπουδών στο Εθνικό Μετσόβιο Πολυτεχνείο, θα ήθελα να ευχαριστήσω όλους τους καθηγητές μου και τους ανθρώπους του ιδρύματος για τη συνεισφορά τους στη μακρά αυτή πορεία.

Θα ήθελα να ξεχωρίσω μεταξύ όλων αυτών τα μέλη του Εργαστηρίου Όρασης Υπολογιστών, Επικοινωνίας Λόγου και Επεξεργασίας Σήματος και φυσικά τον καθηγητή μου, Πέτρο Μαραγκό, στον οποίο οφείλω πολλά. Ο ενθουσιασμός του και η αγάπη του για τους φοιτητές του, μαζί με το δεδομένο επιστημονικό του κύρος, με ώθησαν να καταπιαστώ και εγώ με την ερευνητική περιοχή της Όρασης Υπολογιστών. Πάντοτε με καλή διάθεση και υπομονή, με καθοδήγησε εξαιρετικά καθ' όλη τη διάρκεια της διπλωματικής. Με τις πολύτιμες συμβουλές και τη διαρκή ενθάρρυνση από μέρους του είχα τη δυνατότητα να τη φέρω εις πέρας με επιτυχία.

Καθοριστική για την αποτελεσματική διεξαγωγή των πειραμάτων ήταν η διάθεση της βάσης ψηφιοποιημένων εικόνων των τοιχογραφιών από το Γ. Παπανδρέου που τη δημιούργησε στα πλαίσια του διδακτορικού του. Επίσης, σημαντική ήταν η συνεισφορά του Ε. Πνευματικάκη, καθώς μου διέθεσε ευγενικά μέρος του πηγαίου κώδικά του MATLAB που έγραψε στα πλαίσια της διπλωματικής του, στον οποίο βάσισα τις υλοποιήσεις μου στο κομμάτι του inpainting.

Τέλος, θα ήθελα να ευχαριστήσω τους φίλους μου, οι οποίοι με έχουν στηρίξει πολύ όλα αυτά τα χρόνια και έχουν προσθέσει το δικό τους λιθαράκι για να καταφέρω να φτάσω έως εδώ. Μα πάνω απ' όλα θα ήθελα να εκφράσω την ευγνωμοσύνη μου στην οικογένειά μου, τους γονείς μου και την αδερφή μου, οι οποίοι μου έχουν προσφέρει τα πάντα.

Περιεχόμενα

0.1	Όραση Υπολογιστών και στόχοι διπλωματικής	9
0.2	Διάρθρωση και συμβολή διπλωματικής εργασίας	9
1	Εισαγωγή	23
1.1	Τοιχογραφίες Ακρωτηρίου Θήρας	23
1.1.1	Γενική περιγραφή	23
1.1.2	Κατηγορίες ελλείψεων	25
1.1.3	Στοιχεία για το διαθέσιμο φωτογραφικό υλικό από τις τοιχογραφίες	27
1.2	Το πρόβλημα της ενδοσυμπλήρωσης (inpainting)	29
1.2.1	Γενική περιγραφή του προβλήματος και εφαρμογές	29
1.2.2	Μαθηματικός ορισμός του προβλήματος	30
1.2.3	Bayesian προσέγγιση για τη δημιουργία inpainting μοντέλων	31
1.2.4	Ένταξη του inpainting στο θεμελιώδες πρόβλημα της επεξεργασίας εικόνας	31
2	Αυτόματη εύρεση ελλειπόντων τμημάτων στις τοιχογραφίες	39
2.1	Εισαγωγή	39
2.2	Το πρόβλημα της κατάτμησης εικόνας	40
2.2.1	Γενικά στοιχεία για το πρόβλημα	40
2.2.2	Κάποιες πρόσφατες μέθοδοι	41
2.3	Εύρεση ελλειπόντων τμημάτων με χρήση μαθηματικής μορφολογίας	42
2.3.1	Στοιχεία μαθηματικής μορφολογίας	42
2.3.2	Μερική ανίχνευση ελλείψεων με μορφολογικούς τελεστές	44
2.3.3	Βελτίωση εξαχθείσας μάσκας με την προσάρτηση των ακμών	46
2.3.4	Εφαρμογή αλγορίθμου μορφολογίας και ανίχνευσης ακμών στις τοιχογραφίες	47
2.4	Κατάτμηση εικόνας με τομές γράφων	50
2.4.1	Διαδραστική κατάτμηση με χρήση επαναληπτικών γραφοτομών ("GrabCut")	51
2.4.2	Αξιοποίηση "GrabCut" για την περαιτέρω βελτίωση του μηχανισμού εύρεσης των ελλειπόντων τμημάτων	54
2.4.3	Εφαρμογή του "GrabCut" στις τοιχογραφίες	55
3	Αποκατάσταση των τοιχογραφιών με ενδοσυμπλήρωση (inpainting)	59
3.1	Εισαγωγή	59
3.2	Οφέλη χρήσης PDEs στην Όραση Υπολογιστών	60
3.3	Ενδοσυμπλήρωση περιεχομένου (structure inpainting) με το TV μοντέλο	60
3.3.1	Αποθρομβοποίηση εικόνας και μοντέλο ολικής μεταβολής	60
3.3.2	Βασικές αρχές που πρέπει να ικανοποιεί ένα inpainting μοντέλο	62
3.3.3	Το μοντέλο ενδοσυμπλήρωσης ολικής μεταβολής	63
3.3.4	Μαθηματική θεμελίωση της μεθόδου	65

3.3.5	Αριθμητικό σχήμα υλοποίησης της μεθόδου	66
3.3.6	Εφαρμογή του TV inpainting μοντέλου στις τοιχογραφίες	68
3.4	Ενδοσυμπλήρωση με σύνθεση υφής (texture inpainting)	71
3.4.1	Υφή	71
3.4.2	Σύνθεση υφής (texture synthesis)	72
3.4.3	Ο αλγόριθμος σύνθεσης υφής Efros-Leung	73
3.4.4	Ενδοσυμπλήρωση υφής βάση του αλγορίθμου Efros-Leung	73
3.4.5	Εφαρμογή του αλγορίθμου Efros-Leung στις τοιχογραφίες	74
4	Αξιοποίηση πληροφορίας όλης της επιφάνειας των τοιχογραφιών	77
4.1	Εισαγωγή	77
4.2	Οπτικά βέλτιστη συνένωση εικόνων (seamless image stitching)	79
4.2.1	Εργασία στο πεδίο των παραγώγων	79
4.2.2	Η μέθοδος των Levin, Zomet, Peleg, Weiss	79
4.2.3	Η μέθοδος των Tao, Johnson, Paris	83
4.3	Εφαρμογή οπτικά βέλτιστης συνένωσης εικόνων στις τοιχογραφίες	88
5	Συμπεράσματα	95
5.1	Συμβολή της διπλωματικής εργασίας	95
5.2	Κατευθύνσεις για μελλοντική έρευνα	96
	Βιβλιογραφία	97

Κατάλογος σχημάτων

1	Σχηματικό διάγραμμα σχετικά με το γενικό περιεχόμενο των επιστημών εικόνων.	10
1.1	Τμήμα από το τοπογραφικό διάγραμμα του αρχαιολογικού χώρου στο Ακρωτήρι Θήρας. Σχέδιο των Αξιώτη-Σάλη και Κριτσωτάκη. Έχει ληφθεί από το [2].	24
1.2	Η τοιχογραφία <i>Κροκοσυλλέκτρια & Πότνια</i> (από το [2]).	26
1.3	Η τοιχογραφία <i>Ακολουθία Ανδρών</i> (από το [2]).	27
1.4	Η τοιχογραφία <i>Ακολουθία Γυναικών</i> (από το [2]).	28
1.5	Το πρόβλημα της ενδοσυμπλήρωσης.	30
1.6	Αλλοιωμένη εικόνα και Navier-Stokes Inpainting [7].	32
1.7	Αναπαραγωγή βιολογικών προτύπων με χρήση του RD μοντέλου [8].	33
1.8	Η φτέρη του Barnsley, ένα κλασικό IFS fractal.	34
1.9	Παραδείγματα εικόνων fractal flame κατασκευασμένων με μη γραμμικά IFSs [9].	35
1.10	Inpainting φραγμένης μεταβολής [10]. Στην πάνω (θολωμένη) εικόνα λείπουν 60 τυχαία επιλεγμένα πακέτα που χάθηκαν κατά τη φάση της μετάδοσης. Η κάτω εικόνα έχει προκύψει μετά από αποθρομβοποίηση και inpainting φραγμένης μεταβολής.	36
1.11	Mumford-Shah inpainting [10]. Αριστερά είναι η θορυβημένη εικόνα με την περιοχή G να είναι ο κυκλικός δακτύλιος που ενώνει τους δύο δίσκους. Δεξιά είναι η αποθρομβοποιημένη εικόνα, έχοντας εφαρμοστεί inpainting στον κυκλικό δακτύλιο.	37
2.1	Σύγκριση μερικών κλασικών μεθόδων για κατάτμηση (εξαγωγή προσκηνίου). Στην πάνω σειρά παρουσιάζεται η απαιτούμενη εργασία από το χρήστη προτού εκτελεστεί ο αλγόριθμος για τον προσδιορισμό είτε του προσκηνίου (λευκό χρώμα) και/ή του παρασκηνίου (κόκκινο χρώμα) είτε του συνόρου (κίτρινα σημάδια) του εξαγόμενου αντικειμένου. Στην κάτω σειρά απεικονίζεται το αποτέλεσμα της κατάτμησης για την κάθε μέθοδο. Το “GrabCut” καταφανώς πλεονεχτεί τόσο σε ποιότητα αποτελέσματος όσο και στην απαιτούμενη εργασία από τη μεριά του χρήστη. Έχει αντληθεί από το [3], στο οποίο έχει βασιστεί και η παρούσα υποενοότητα.	43
2.2	Εξαγωγή δυαδικής μάσκας για μικρό τμήμα με ρωγμές της τοιχογραφίας <i>Ακολουθία Ανδρών</i> και υπέρθεση του περιγράμματος της μάσκας πάνω στο αυθεντικό έργο.	45
2.3	Αυτόματη εύρεση της μάσκας ενός άλλου τμήματος της τοιχογραφίας <i>Ακολουθία Ανδρών</i> και υπέρθεση της πάνω σε αυτό.	46
2.4	Εξαγωγή μάσκας σε περιοχή της πλάτης του μεγαλύτερου άνδρα στην τοιχογραφία <i>Ακολουθία Ανδρών</i> και υπέρθεση του περιγράμματος στο αυθεντικό έργο. Στη βελτιωμένη μάσκα το περίγραμμα έχει εξαλειφθεί, ενώ έχουν εντοπιστεί περισσότερες φθορές.	48

- 2.5 Οι εξαχθείσες μάσκες σε αυτή την περίπτωση αφορούν τμήμα της τοιχογραφίας *Κροκοσυλλέκτρια & Πότνια*, όπου απεικονίζεται το σκουλαρίκι της Πότνιας. Στη μάσκα που αξιοποιείται και η πληροφορία των ακμών αποσοβείται η εμφάνιση περιγράμματος και γενικότερα ο εντοπισμός των ελλείψεων είναι πληρέστερος. Αναπόφευκτα ανιχνεύονται και οι έντονες ακμές που είναι μέρος της παράστασης της τοιχογραφίας, κάτι που θα οδηγήσει σε όχι ιδανικό inpainting, όπως θα δειχθεί στο 3^ο Κεφάλαιο με το Σχήμα 3.5. 49
- 2.6 Εξαγωγή μάσκας σε περιοχή του σώματος του μεγαλύτερου άνδρα στην τοιχογραφία *Ακολουθία Ανδρών* και υπέρθεση του περιγράμματος στο αυθεντικό έργο. Όπως διακρίνεται στη βελτιωμένη μάσκα το περίγραμμα έχει εξαλειφθεί, οι ρωγμές που δε βρέθηκαν αρχικά έχουν εντοπιστεί, ενώ τέλος οι ελλείψεις περικλείονται πιο «ανοικτά» από τη λευκή περιοχή της μάσκας, κάτι που αν και αφελώς μπορεί να ερμηνευτεί ως λιγότερο καλή προσαρμογή, στην πραγματικότητα επιτρέπει πολύ πιο ασφαλές και «καθαρό» inpainting αργότερα. 50
- 2.7 Ο χρήστης παρεμβαίνει και ορίζει ρητά (explicitly) τη θέση κάποιων pixel που έχουν ταξινομηθεί σε λάθος περιοχή. Από το [3], στο οποίο έχει βασιστεί η Υποενοότητα 53
- 2.8 Επ' ακριβώς εύρεση έλλειψης με το "GrabCut" σε περιοχή της τοιχογραφίας *Ακολουθία Ανδρών* που απεικονίζει το κεφάλι του μεγαλύτερου άνδρα. Με μία πρόχειρη περιγραφή του background και του foreground, ως είσοδο του αλγορίθμου, αυτός επιστρέφει μία πολύ ικανοποιητική ταυτοποίηση του ελλείποντος τμήματος. Σε συνδυασμό με την αξιοποίηση του αλγορίθμου μαθηματικής μορφολογίας η αναγνώριση μπορεί να γίνει με ακρίβεια. 55
- 2.9 Ανίχνευση της σημαντικής έλλειψης στην κεφαλή της γυναίκας στην τοιχογραφία *Ακολουθία Γυναικών*. 56
- 2.10 Ανίχνευση των ευμεγεθών ελλείψεων στην κροκοσυλλέκτρια από την τοιχογραφία *Κροκοσυλλέκτρια & Πότνια*. 57
- 3.1 Το μοντέλο ενδοσυμπλήρωσης ολικής μεταβολής υπολογίζει την ελλείπουσα περιοχή $u|_D$ βάση της TV νόρμας στο πεδίο $E \cup D$ και του περιορισμού θορύβου στο E [11]. 63
- 3.2 Τα γειτονικά σημεία ενός pixel επί του διακριτού πλέγματος της εικόνας [12]. 66
- 3.3 Εφαρμογή inpainting σε τμήμα της τοιχογραφίας *Ακολουθία Ανδρών* από την πλάτη του μεγαλύτερου άνδρα. Το τμήμα αυτό παρουσιάζει μικρές ελλείψεις, δηλαδή ρωγμές μικρού ή μέσου πλάτους, οι οποίες αντιμετωπίζονται αποτελεσματικά από το TV inpainting μοντέλο. 68
- 3.4 Εφαρμογή inpainting σε τμήμα της ίδιας τοιχογραφίας που αφορά μέρος του σώματος του μεγαλύτερου άνδρα. Η περιοχή εμφανίζει πολλές ρωγμές, αλλά και κάποιες ελλείψεις μέσης έκτασης ως προς το μέγεθος της εικόνας, οι οποίες δεν αντιμετωπίζονται εύκολα. Παρόλ' αυτά, σε σχετικά μικρό πλήθος επαναλήψεων η παράσταση επιδιορθώνεται αρκούτως ικανοποιητικά. Ένας αλγόριθμος που αξιοποιεί το συναρτησιακό *elastica* του Euler θα μπορούσε να αποφέρει σαφώς βελτιωμένη αποκατάσταση. 69

3.5	Εφαρμογή inpainting σε τμήμα της τοιχογραφίας <i>Κροκοσυλλέκτρια & Πότνια</i> που απεικονίζει το σκουλαρίκι της Πότνιας. Αν και το περίγραμμα της μάσκας δείχνει ικανοποιητικό, παρόλ' αυτά αρχικοποιείται μεγαλύτερη επιφάνεια με τυχαίες τιμές από την ελάχιστη δυνατή και έτσι χάνεται σημαντική για την αποκατάσταση της τοιχογραφίας πληροφορία. Βασική αιτία είναι ότι αρκετές ακμές της παράστασης αναγνωρίζονται ως ελλείψεις. Η εκπαίδευση ενός μοντέλου, ώστε να αναγνωρίζει ποιες ακμές δεν αφορούν ελλείψεις, θα μπορούσε να προσφέρει σημαντική βελτίωση στο αποτέλεσμα.	70
3.6	Μία εικόνα με πολλούς τύπους υφής. Η ταξινόμηση υφής αφορά την ταυτοποίηση της υφής καθεμιάς περιοχής, ενώ στην κατάτμηση υφής διαχωρίζονται οι περιοχές υφής σε μία εικόνα, μέσω του προσδιορισμού των συνόρων τους, χωρίς την ακριβή αναγνώριση του τύπου υφής ανά περιοχή. Έχει αντληθεί από το [13], στο οποίο βασίζεται η Υποενότητα 3.4.1.	71
3.7	Σύνθεση υφής σε περιοχή του σώματος του μεγαλύτερου άνδρα στην τοιχογραφία <i>Ακολουθία Ανδρών</i> . Είναι το ίδιο τμήμα με το Σχήμα 3.4, όπου είχε εφαρμοστεί το TV inpainting model. Όπως βλέπουμε πλέον δεν εμφανίζεται θόλωμα και το αποτέλεσμα είναι αισθητικά καλύτερο. Με διαδοχικές εφαρμογές σύνθεσης υφής σε όλες τις περιοχές της εικόνας μπορεί να γίνει πλήρης αποκατάσταση.	74
3.8	Σύνθεση υφής στο λαιμό του μεγαλύτερου άνδρα στην τοιχογραφία <i>Ακολουθία Ανδρών</i> . Η γραμμή του λαιμού επεκτείνεται και η κόκκινη περιοχή συμπληρώνεται σταδιακά. Ο αλγόριθμος των Efros-Leung μπορεί να αποφέρει αισθητικά άρτια αποκατάσταση.	75
4.1	Δίνονται δύο φωτογραφικές λήψεις I_1 και I_2 του ίδιου τοπίου που απεικονίζουν γειτονικές ως προς τον ορίζοντα περιοχές του $\tau_1 \cup \omega$ και $\omega \cup \tau_2$. Η επικόλλησή τους δίπλα-δίπλα φυσικά δεν έχει το επιθυμητό αποτέλεσμα και δημιουργείται μία «ραφή» (π.χ. λόγω μη ομοιόμορφης φωτομετρικής απόκρισης της κάμερας). Ο αλγόριθμος των Levin et al. αποδίδει ένα φυσικό οπτικά αποτέλεσμα [5].	80
4.2	Για την αποφυγή των οπτικών ατελειών που ανακύπτουν στην απλή σύνθεση Poisson, ο αλγόριθμος των Tao, Johnson & Paris υλοποιείται σε δύο φάσεις. Στην πρώτη φάση προσαρμόζονται οι παράγωγοι της έντασης στο σύνορο μέσω της ελαχιστοποίησης του στροβιλισμού και στη δεύτερη φάση σταθμίζεται η διαδικασία ανακατασκευής έτσι ώστε τα υπόλοιπα να συγκεντρώνονται κυρίως στις περιοχές υφής. Ο αλγόριθμος των Tao et al. αποδίδει ένα άψογο οπτικά αποτέλεσμα [6].	87
4.3	Προσάρτηση τμήματος της κεφαλής του νεότερου άνδρα, μετά από αναστροφή, κατάλληλη περιστροφή και μετατόπιση, στην ελλείπουσα περιοχή της κεφαλής του μεγαλύτερου άνδρα (τοιχογραφία <i>Ακολουθία Ανδρών</i>).	89
4.4	Ψηφιακή αποκατάσταση του λαιμού της ανδρικής φιγούρας με πληροφορία που αντλείται από τη μέση και την πλάτη (τοιχογραφία <i>Ακολουθία Ανδρών</i>).	91
4.5	Ψηφιακή επιδιόρθωση τμήματος του υφάσματος που κρατάει ο μεγαλύτερος άνδρας με τη λήψη υποπεριοχής της εικόνας αντίστοιχης πληροφορίας (τοιχογραφία <i>Ακολουθία Ανδρών</i>).	92
4.6	Αποκατάσταση ενδεικτικού τμήματος της κορδέλας στα μαλλιά της Πότνιας (τοιχογραφία <i>Κροκοσυλλέκτρια & Πότνια</i>).	93

Λίστα Αλγορίθμων

1	Αυτόματη εύρεση ελλείψεων με χρήση στοιχείων <i>Μαθηματικής Μορφολογίας</i> και ανίχνευση ακμών (region and boundary based approach).	47
2	Κατάτμηση εικόνας με τον αλγόριθμο “GrabCut”.	53
3	Αξιοποίηση “GrabCut” ώστε η μάσκα που έχει εξαχθεί με μαθηματική μορφολογία και ανίχνευση ακμών να βρίσκει ακριβώς τις ελλείψεις.	54
4	Ο 1ος αλγόριθμος για σύνθεση μωσαϊκού των Levin et al.	81
5	Ο 2ος αλγόριθμος για σύνθεση μωσαϊκού των Levin et al.	81
6	Οι κύριες συνεισφορές του αλγορίθμου των Tao et al. για εύρωστη σύνθεση μωσαϊκού	85

Κεφάλαιο 1

Εισαγωγή

Στην παρούσα διπλωματική εργασία αναπτύσσονται και μελετώνται τεχνικές Όρασης Υπολογιστών με σκοπό τη ψηφιακή αποκατάσταση των τοιχογραφιών του Προϊστορικού Οικισμού του Ακρωτηρίου Θήρας. Στον *Πρόλογο* περιγράφονται αναλυτικότερα οι στόχοι της εργασίας και επεξηγείται η διάρθρωση των κεφαλαίων του κειμένου.

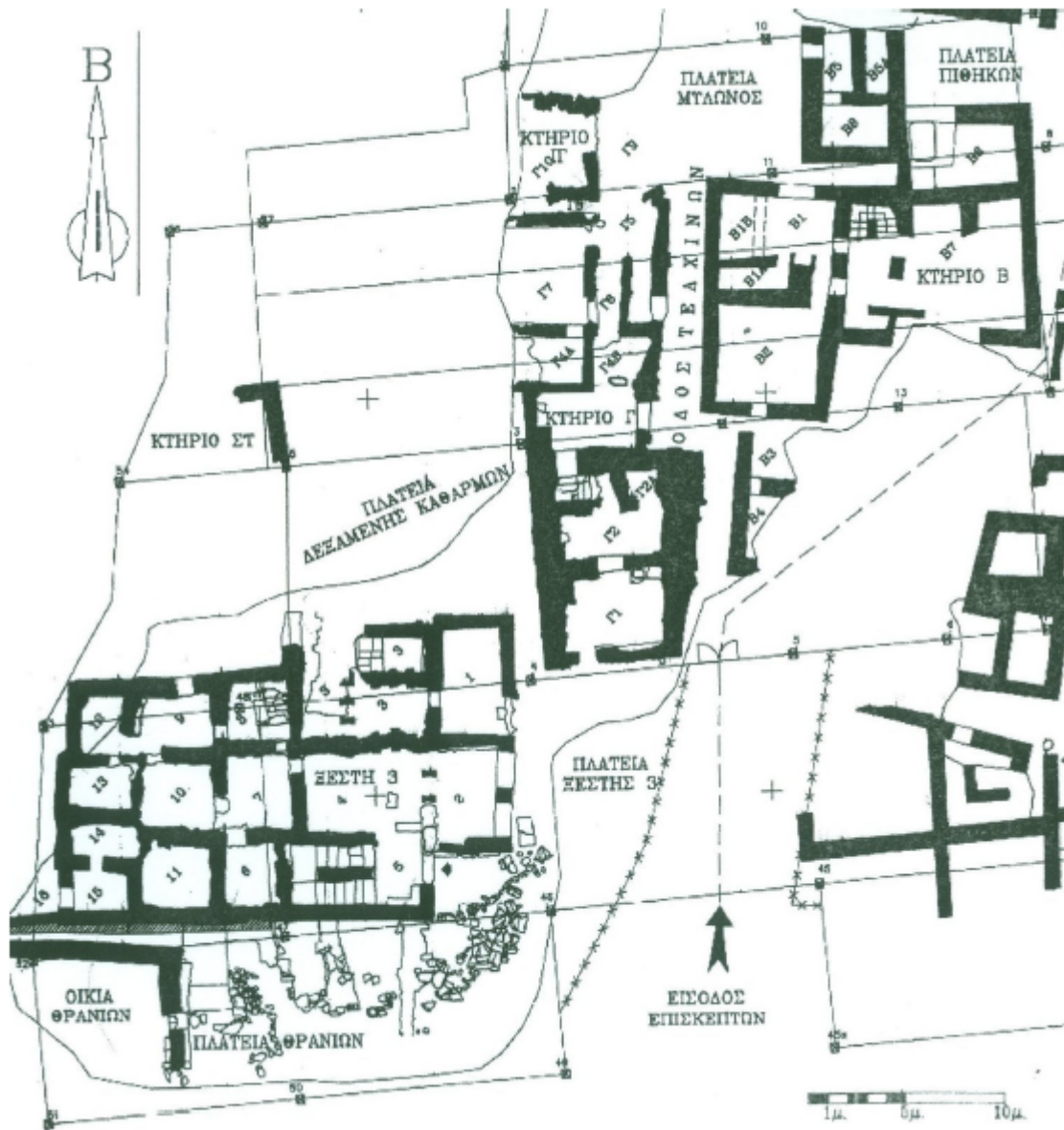
Στην ακόλουθη ενότητα, η οποία έχει βασιστεί στο [2], παρουσιάζονται κάποιες πληροφορίες για τις τοιχογραφίες. Πιο συγκεκριμένα, επεξηγείται η σημασία τους και περιγράφεται το ερευνητικό ενδιαφέρον που έχουν προσελκύσει, μαζί με κάποιες πρόσφατες εργασίες που τις αφορούν. Στη συνέχεια, κατηγοριοποιούνται τα ελλείποντα τμήματά τους και δίνονται κάποια εισαγωγικά στοιχεία σχετικά με την αποκατάστασή τους. Τέλος, περιγράφεται συνοπτικά η διαδικασία ψηφιοποίησής τους από το G. Papandreou.

1.1 Τοιχογραφίες Ακρωτηρίου Θήρας

1.1.1 Γενική περιγραφή

Οι τοιχογραφίες που ανακαλύφθηκαν στον Προϊστορικό Οικισμό του Ακρωτηρίου Θήρας είναι διάσημες για τη θεματογραφία τους, την καλλιτεχνική τους αξία και για τον πλούτο πληροφορίας που παρέχουν για τον Αιγιακό Πολιτισμό πριν από περίπου 3500 χρόνια. Οφείλουν τη διατήρησή τους ως τις μέρες μας στην ισχυρότατη έκρηξη του ηφαιστείου της Θήρας κατά τον 17^ο αιώνα π.Χ. που είχε ως συνέπεια να καταπλακωθεί ολόκληρος ο προϊστορικός οικισμός από παχύ στρώμα ηφαιστειακής τέφρας. Ο οικισμός ανακαλύφθηκε και άρχισε να ανασκάπτεται το 1967 υπό την καθοδήγηση του καθηγητή Μαρινάτου, τον οποίο διαδέχθηκε το 1974 ο καθηγητής Ντούμας, ως και σήμερα διευθυντής της ανασκαφής. Στο Σχήμα 1.1 απεικονίζεται κάτοψη τμήματος του αρχαιολογικού χώρου που έχει σχεδιαστεί από τους Αξιώτη-Σάλη και Κριτσωτάκη.

Η αποκατάσταση αυτών των ιστορικών τοιχογραφιών είναι ιδιαίτερα δύσκολη και χρονοβόρα. Η διαδικασία συντήρησης περιλαμβάνει μεταξύ άλλων συλλογή και αποκόλληση των τμημάτων της τοιχογραφίας από τον Αρχαιολογικό χώρο, καθαρισμό και στερέωση των σπαραγμάτων στο εργαστήριο, καθώς και ανασύνθεση κάθε τοιχογραφίας από τα ευρεθέντα τμήματά της. Πληροφορίες σχετικά με την ιστορία και τη σημασία της ανασκαφής, με ιδιαίτερη έμφαση τις τοιχογραφίες, περιέχεται στο [14]. Ενδιαφέρον αρχαιολογικό υλικό από την ανασκαφή μπορεί να αναζητηθεί στο [15].



Σχήμα 1.1: Τμήμα από το τοπογραφικό διάγραμμα του αρχαιολογικού χώρου στο Ακρωτήρι Θήρας. Σχέδιο των Αξιώτη-Σάλη και Κριτσωτάκη. Έχει ληφθεί από το [2].

Πλούσια έρευνα έχει πραγματοποιηθεί για την ανασκαφή κατά το παρελθόν. Οι Rousopoulos et al. [16] διερευνούν τις πιθανές διαδικασίες που ακολουθήθηκαν για τη δημιουργία των μοτίβων σπείρας στις τοιχογραφίες. Κάνουν κάποιες υποθέσεις για την πιθανή ύπαρξη και τη μορφή ενός μηχανισμού για το χειρισμό των πινέλων, για το υπό ποια γωνία τοποθετούνταν αυτά στον τοίχο, την πιθανή ύπαρξη ενός πρόχειρου σχεδίου αναφοράς και ούτω καθεξής. Το [17] αποτελεί μία μελέτη των ευρημάτων στο Ακρωτήρι Θήρας που αποδίδονται στην Κυκλαδική Εποχή του Χαλκού, σχετικά με τις πρώτες ύλες και τα εργαλεία για την προετοιμασία και τη χρήση των μολυβιών για το χρωματισμό των τοιχογραφιών. Για να ταυτοποιήσουν τη χημική σύνθεση των χρωματισμών

και να διερευνήσουν την προέλευσή τους από γεωλογική πλευρά, οι Sotiropoulou et al. διενήργησαν ποσοτική ανάλυση με *περίθλαση ακτίνων Χ*. Μέσω μίας μικροσκοπικής και μακροσκοπικής εξέτασης της μορφολογίας των υλών έγινε δυνατός ο καθορισμός των φυσικών γνωρισμάτων και η εξαγωγή συμπερασμάτων για τη φύση τους και ο βαθμός που έχουν κατεργαστεί.

Οι τοιχογραφίες της Θήρας έχουν γενικότερα βρεθεί στο επίκεντρο πολλών διεθνών συμποσίων και συνεδρίων [18]. Στο [19] έχει υλοποιηθεί ένα σύστημα προσδιορισμού της γεωμετρίας και των μετρικών των σπαραγμάτων από τις τοιχογραφίες, το οποίο είναι σχετικά φθηνό και προορισμένο για χρήση από μη ειδικούς. Με επίπεδο σαρωτή λαμβάνονται υψηλής ανάλυσης εικόνες και τα μετρικά των τοιχογραφιών, ενώ με πολλαπλούς σαρωτές εξάγεται η 3Δ γεωμετρία τους, και έτσι είναι δυνατή η καταγραφή, η ανάλυση και η εύρεση των ταιριασμάτων μεταξύ των σπαραγμάτων. Εξαιρετική είναι και η δουλειά στο [20] όπου χρησιμοποιούνται αλγόριθμοι γεωμετρικής ανάλυσης για τον ακριβή υπολογισμό της γεωμετρίας και των πιθανών ταιριασμάτων των συνόρων των σπαραγμάτων και εν συνεχεία τη δημιουργία ενός στατιστικού μοντέλου που αφορά τα μήκη, τις γωνίες, τις επιφάνειες και τις γειτονίες στα συναρμολογημένα μωσαϊκά. Με βάση το μοντέλο αυτό προτείνεται ένα ιεραρχικό πρότυπο για τη σύνθεση των σπαραγμάτων, όπου κάθε κομμάτι επί των ρωγμών σπάει αναδρομικά σε μικρότερα κομμάτια σχεδόν ορθογώνια ως προς το αρχικό, ένας μηχανισμός που μπορεί να αποδειχθεί βοηθητικός στην εκτίμηση προτύπων σπαραγμάτων σε άλλες τοιχογραφίες είτε σε μελλοντικούς αυτόματους αλγόριθμους ανακατασκευής.

Τις τελευταίες δεκαετίες υπάρχει έντονο ερευνητικό ενδιαφέρον σχετικά με την ανάπτυξη αυτοματοποιημένων τεχνικών και λογισμικού που θα δρουν επικουρικά στο έργο των αρχαιολόγων και των συντηρητών τέχνης που αφορά την ευρύτερη πολιτισμική κληρονομιά του τόπου. Το εργαστήριο του καθηγητή του Εθνικού Μετσόβιου Πολυτεχνείου Πέτρου Μαραγκού είναι ερευνητικά ενεργό στην περιοχή την τελευταία δεκαετία [2]. Με την ολοκλήρωση της διαδικασίας συντήρησης από τους αρχαιολόγους, κάποια τμήματα κάθε τοιχογραφίας παραμένουν κενά είτε γιατί αποτελούνται από κομμάτια που δεν έχουν ακόμη ταυτοποιηθεί είτε γιατί τα αντίστοιχα κομμάτια έχουν καταστραφεί εντελώς με το πέρασμα των αιώνων. Σκοπός του εργαστηρίου είναι η αυτοματοποιημένη ψηφιακή αποκατάσταση των περιοχών αυτών τόσο για τη δυνατότητα έκθεσης των έργων στο ευρύ κοινό όσο για την υποβοήθηση του έργου των ειδικών αρχαιολόγων στη διαδικασία ερμηνείας των παραστάσεών τους. Στο Σχήμα 1.2 παρουσιάζεται η τοιχογραφία “Κροκοσυλλέκτρια & Πότνια”, όπου διακρίνονται αρκετά στοιχεία της παράστασης αλλά η ψηφιακή αποκατάσταση θα μπορούσε να δώσει πολύ ανώτερο οπτικό αποτέλεσμα.

1.1.2 Κατηγορίες ελλείψεων

Οι ελλείψεις που παρατηρούνται στις τοιχογραφίες διακρίνονται σε διάφορες κατηγορίες ανάλογα με την έκταση και τη μορφή τους. Η συντριπτική πλειοψηφία τους είναι αλλοιώσεις μικρής έκτασης, όπως για παράδειγμα ρωγμές. Συναντώνται στις επαφές συγκολλημένων σπαραγμάτων ή οφείλονται σε σκάσιμο του χρώματος. Αυτές επιδιορθώνονται σχετικά εύκολα, με ντετερμινιστικό τρόπο μάλιστα. Αρκετά είναι και τα κενά μεσαίας σχετικά έκτασης στις τοιχογραφίες. Οφείλονται στην απώλεια ή αδυναμία αντιστοίχισης μικρού αριθμού διάσπαρτων σπαραγμάτων από τμήματα της τοιχογραφίας. Για το ανθρώπινο μάτι, ακόμα και για μη ειδικούς, η συμπλήρωση των κενών αυτών και η σχιαγράφηση των ελλειπόντων μοτίβων δεν είναι ιδιαίτερα δύσκολη, αν και σε ορισμένες περιπτώσεις περισσότερες από μία παρεμφερείς εκδοχές είναι αποδεκτές. Σίγουρα οι αυτοματοποιημένες ψηφιακές μέθοδοι συνεισφέρουν για μία πιο συστηματική επιδιόρθωση αυτών των ελλείψεων. Τέλος, δυστυχώς υπάρχουν και τα μεγάλα ελλείποντα τμήματα. Οφείλονται στην



Σχήμα 1.2: Η τοιχογραφία *Κροκοσυλλέκτρια & Πότνια* (από το [2]).

απώλεια ή αδυναμία αντιστοίχισης μεγάλου αριθμού ή μεγάλου μεγέθους σπαραγμάτων σε μία περιοχή της τοιχογραφίας. Καθιστούν την αποκατάσταση του συγκεκριμένου τμήματος ιδιαίτερα προβληματική, με την έννοια ότι απαιτείται διεξοδική μελέτη και ερμηνεία της όλης σύνθεσης από αρχαιολόγους και ζωγράφους για να γίνει απόπειρα αποκατάστασης. Σε κάποιες περιπτώσεις είναι τέτοια η έκταση της αλλοίωσης που η προτεινόμενη αποκατάσταση είναι απλώς ενδεικτική.

Η ψηφιακή αποκατάσταση των τοιχογραφιών απαιτεί διαφορετικές τεχνικές ανάλογα με τον τύπο της έλλειψης. Οι ρωγμές μπορούν να ανιχνευθούν και να αντιμετωπιστούν με απλές μη γραμμικές τεχνικές επεξεργασίας εικόνas, όπως για παράδειγμα αυτές που συναντώνται στο [21]. Οι Giakoumis et al. χρησιμοποιούν κατωφλιοποίηση (thresholding) σε ένα μορφολογικό *top-hat* μετασχηματισμό για την ανίχνευση των ρωγμών και στη συνέχεια οι λεπτές πινελιές που εσφαλμένα εντοπίστηκαν ως ρωγμές εξαλείφονται μέσω Median RBF νευρωνικού δικτύου. Τα κενά μεσαίου μεγέθους απαιτούν την εφαρμογή σύγχρονων μεθόδων της Όρασης Υπολογιστών, τις οποίες πραγματεύεται η παρούσα εργασία. Τέλος, οι μεγάλες βλάβες είναι εξαιρετικά πιο δύσκολο να αντιμετωπιστούν με πλήρως αυτόματες ψηφιακές τεχνικές, καθώς απαιτούν την εκπαίδευση του ψηφιακού συστήματος αποκατάστασης. Μπορούν να αντιμετωπιστούν αρκετά αποτελεσματικά με ανάκτηση πληροφορίας από άλλες περιοχές των τοιχογραφιών, αλλά επί του παρόντος η εύρεση του κατάλληλου κομματιού γίνεται με παρέμβαση του χρήστη και όχι αυτόματα.



Σχήμα 1.3: Η τοιχογραφία *Ακολουθία Ανδρών* (από το [2]).

1.1.3 Στοιχεία για το διαθέσιμο φωτογραφικό υλικό από τις τοιχογραφίες

Η ύπαρξη λεπτομερούς φωτογραφικού υλικού από τις τοιχογραφίες της ανασκαφής Ακρωτηρίου Θήρας ήταν απαραίτητη προϋπόθεση για την περάτωση της παρούσας εργασίας. Αν και φωτογραφικό υλικό βρισκόταν στη διάθεση της αρχαιολογικής ομάδας από δεκαετίες, δεν ήταν κατάλληλο για τους σκοπούς της εφαρμογής, γιατί μέρος του είχε ληφθεί για συγκεκριμένες αρχαιολογικές μελέτες και ήταν ανομοιογενές ως προς την κάλυψη των τοιχογραφιών και ένα άλλο μέρος του προοριζόταν για την προβολή τους στο ευρύ κοινό και χαρακτηριζόταν από περιορισμένη λεπτομέρεια, ακόμη και ύστερη ψηφιακή παρέμβαση στους χρωματισμούς σε κάποιες περιπτώσεις. Σημαντική διευκόλυνση παρείχε το γεγονός ότι ο G. Papandreou, στα πλαίσια της διδακτορικής του διατριβής, κατέφυγε σε συστηματική λήψη και συλλογή φωτογραφικού υλικού. Η προσέγγιση που ακολούθησε για αποτύπωση των τοιχογραφιών σε υπερυψηλή ανάλυση με χρήση απλού φωτογραφικού εξοπλισμού και ύστερη συρραφή των στιγμιotypών σε ένα ενιαίο

μωσαϊκό αποτελεί εναλλακτική μεθόδων που απαιτούν ειδικευμένο εξοπλισμό υψηλού κόστους. Στο [2] περιγράφεται με μεγάλη λεπτομέρεια η μέθοδος που υιοθετήθηκε.



Σχήμα 1.4: Η τοιχογραφία *Ακολουθία Γυναικών* (από το [2]).

Τα Σχήματα 1.2, 1.3 και 1.4 απεικονίζουν τις τρεις τοιχογραφίες που τα μωσαϊκά τους αξιοποιήθηκαν στα πειράματα της παρούσας εργασίας. Στο Σχήμα 1.2 παρουσιάζεται η *Πότνια Θηρών* (δέσποινα των ζώων), η οποία καθισμένη σε θρόνο δέχεται την προσφορά των κρίνων από κροκοσυλλέτρια που αδειάζει τα καλάθι της σε ένα πανέρι που βρίσκεται μπροστά της. Δίπλα στην Πότνια θηρών βρίσκεται ένας γρύπας, ενώ μπροστά της ένας κυανοπίθηκος. Η θεά έχει εντυπωσιακή κόμμωση με χρυσό διάδημα, ενώ στα μαλλιά της υπάρχει και κορδόνι με χρυσούς κόμπους. Στο λαιμό της φοράει ένα περιδέραιο από χάντρες και ένα περιδέραιο, ενώ στο αυτί έχει χρυσό

σκουλαρίκι. Το φόντο της τοιχογραφίας είναι ένα λιβάδι με κρόκους.

Στο Σχήμα 1.3 απεικονίζονται δύο γυμνοί νεαροί άνδρες. Έχουν και οι δύο ξυρισμένα κεφάλια, τα οποία απεικονίζονται με μπλε χρώμα. Ο μεγαλύτερος άνδρας κρατάει με τα δύο του χέρια ένα ύφασμα και στρέφει το κεφάλι του προς τα πίσω. Ο νεαρότερος, που τον ακολουθεί, κοιτάζει προς τα εμπρός και έχει στο εσωτερικό του αριστερού του χεριού ένα μικρό αγγείο, το οποίο σκεπάζει προστατευτικά ή τελετουργικά με το δεξί του χέρι [22]. Τέλος, το Σχήμα 1.4 αφορά τη λιγότερο καλοδιατηρημένη από τις τρεις τοιχογραφίες στο πέρασμα των αιώνων. Απεικονίζονται δύο γυναικείες φιγούρες με μακριά εντυπωσιακά φορέματα.

1.2 Το πρόβλημα της ενδοσυμπλήρωσης (inpainting)

1.2.1 Γενική περιγραφή του προβλήματος και εφαρμογές

Η ενδοσυμπλήρωση (*inpainting*) είναι η διαδικασία ανακατασκευής ελλειπόντων ή φθαρμένων κομματιών σε εικόνες και βίντεο με έναν τρόπο που ο παρατηρητής, δεδομένου ότι δε γνωρίζει το αυθεντικό έργο, να μη μπορεί να αντιληφθεί ότι έχει υποστεί επεξεργασία. Το *inpainting* έχει μεγάλο εύρος εφαρμογών όπως για παράδειγμα στη συμπίεση εικόνας, στο ψηφιακό *zoom in*, στην επιδιόρθωση σφαλμάτων κατά την κωδικοποίηση και τη μετάδοση εικόνας, στην επεξεργασία φωτογραφιών ή και ταινιών του κινηματογράφου με σκοπό τη διόρθωση λαθών (παραδείγματος χάρη απομάκρυνση παρεμβαλλόμενου αντικειμένου ή κόκκινων ματιών στις φωτογραφίες, επιδιόρθωση των περιοχών σκόνης ή φθορών στο κινηματογραφικό φιλμ), στην απομάκρυνση κειμένου ή ημερομηνιών από εικόνες αλλά και στη συντήρηση έργων τέχνης.

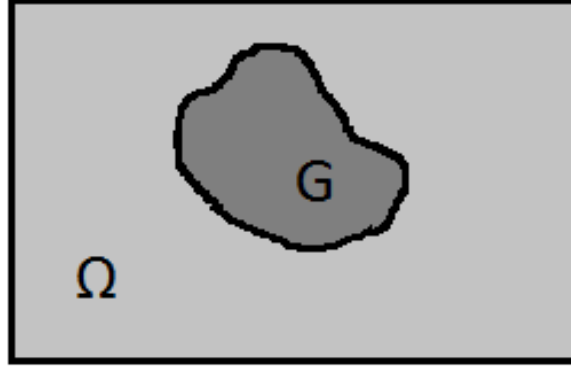
Στην περίπτωση των πολύτιμων τοιχογραφιών η ενδοσυμπλήρωση μπορεί σε κάποιο βαθμό να πραγματοποιηθεί από ειδικούς στην αποκατάσταση έργων τέχνης. Με μία ιδιαίτερα χρονοβόρα και απαιτητική διεργασία απομακρύνουν σχολαστικά πιθανή σκόνη, διορθώνουν «σκασίματα» στο χρώμα, αποκαθιστούν κομμάτια που απουσιάζουν. Επιπλέον, η αποκατάσταση ενέχει συχνά το υποκειμενικό στοιχείο και μπορεί να διαφοροποιηθεί σημαντικά ανάλογα τον τεχνίτη. Στον ψηφιακό κόσμο όμως η ενδοσυμπλήρωση αφορά την εφαρμογή ειδικών αλγορίθμων για τη συστηματική αντικατάσταση απολεσθέντων ή κατεστραμμένων δεδομένων της εικόνας με έναν αυτόματο ή ημιαυτόματο τρόπο. Βέβαια, συχνά είναι απαραίτητη η επέμβαση και του χρήστη, είτε για να ορίσει την ακριβή περιοχή που θα εφαρμοστεί το *inpainting* είτε για το άμεσο προσδιορισμό (*explicitly annotated*) του προσκηνίου ή/και του παρασκηνίου της εικόνας.

Ο όρος *inpainting* έχει υιοθετηθεί από τον καλλιτεχνικό κόσμο και αφορά αυτό που οι ερευνητές στην επεξεργασία εικόνας αποκαλούν «παρεμβολή» (*interpolation*). Στο [23] οι Bertalmio, Sapiro, Caselles και Ballester έκανα για πρώτη φορά εκτενή καταγραφή των κινήτρων για να ερευνηθεί το πρόβλημα αυτό και ουσιαστικά ήταν αυτοί που εισήγαγαν τον όρο *inpainting* στην Επεξεργασία Εικόνας.

Στην Ενότητα 1.2.2 δίνεται ο τυπικός μαθηματικός ορισμός του προβλήματος. Στην Ενότητα 1.2.3 παρουσιάζεται η Bayesian προσέγγιση για τη δημιουργία *inpainting* μοντέλων, ενώ στην 1.2.4 εντάσσεται η ενδοσυμπλήρωση στο θεμελιώδες πρόβλημα επεξεργασίας εικόνας. Βάση αυτού, πρωταρχικό ζήτημα σε κάθε πρόβλημα επεξεργασίας εικόνας, άρα και στο *inpainting*, είναι πώς μοντελοποιείται η εικόνα. Έτσι, μέσω της διερεύνησης των πιθανών μοντέλων της εικόνας παρουσιάζεται μία πληθώρα από κατηγορίες αλγορίθμων ενδοσυμπλήρωσης.

1.2.2 Μαθηματικός ορισμός του προβλήματος

Αν και στην περίπτωση των τοιχογραφιών συναντώνται κατηγορίες ελλείψεων με μεγάλη ποικιλία σε μορφολογία, μέγεθος, άρα και μεθοδολογία αντιμετώπισης για την αποκατάστασή τους, παρόλ' αυτά μπορεί να θεμελιωθεί ένα πιο αφηρημένο μαθηματικό πλαίσιο για την ενιαία θεωρητική αντιμετώπιση και φορμαλισμό του προβλήματος της αποκατάστασης των ελλειπόντων τμημάτων των τοιχογραφιών.



Σχήμα 1.5: Το πρόβλημα της ενδοσυμπλήρωσης.

Με αναφορά στο Σχήμα 1.5, έστω Ω ένα πλήρες πεδίο εικόνας, συχνά μία ορθογώνια περιοχή ή ένα έργο τέχνης, είτε πιο αυστηρά μαθηματικά ένα πεπερασμένο πεδίο *Lipschitz*¹ στο R^2 . Συχνά, λείπουν ή είναι μη προσβάσιμα δεδομένα της εικόνας σε μία περιοχή της G . Στόχος της ενδοσυμπλήρωσης είναι η ανάκτηση των δεδομένων της εικόνας u σε ολόκληρο το πεδίο Ω , με βάση μόνο τη μερική και συχνά παραμορφωμένη παρατήρηση $u_0|_{\Omega \setminus G}$.

Στα πλαίσια της εργασίας αυτής το χωρίο G εξάγεται αυτόματα ως δυαδική μάσκα που οριοθετεί τις ελλείπουσες περιοχές. Η χρωματική ομοιότητα μεταξύ του κονιάματος που έχει τοποθετηθεί από τους συντηρητές στις ελλείπουσες περιοχές και του ασβεστοκονιάματος στις αρχικές τοιχογραφίες δυσκολεύει τη διάκριση των κενών με χρήση αποκλειστικά χρωματικής πληροφορίας. Επίσης η ποικιλομορφία στη γεωμετρία των ελλειπουσών περιοχών δυσχεραίνει περαιτέρω την αυτόματη ανίχνευσή τους. Παρόλ' αυτά με χρήση είτε μαθηματικής μορφολογίας είτε κατάτμησης μέσω γραφοτομών ανά περίπτωση ανιχνεύονται αρκούντως ικανοποιητικά.

Μία δημοφιλής προσέγγιση για την ανάπτυξη inpainting μοντέλων είναι η *bayesian*, όπου κρίσιμο ρόλο παίζει η αναγνώριση κατάλληλων *a priori* πιθανοτήτων της εικόνας. Η προσέγγιση σταχυολογείται ακολουθώντας με βάση το [10], στο οποίο έχει στηριχθεί ο κύριος άξονας της παρούσας ενότητας.

¹Έστω ένα ανοικτό, φραγμένο σύνολο $\Omega \subseteq R^2$ και $\partial\Omega$ το σύνορό του. Τότε το Ω έχει Lipschitz σύνορο ή είναι Lipschitz πεδίο, αν $\forall p \in \partial\Omega$ υπάρχει μία ακτίνα $r > 0$ και μία απεικόνιση $h_p : B_r(p) \rightarrow Q$ έτσι ώστε:

- h_p αμφιμονοσήμαντη
- h_p και h_p^{-1} Lipschitz συνεχείς συναρτήσεις
- $h_p(\partial\Omega \cap B_r(p)) = Q_0$
- $h_p(\Omega \cap B_r(p)) = Q_+$

όπου $B_r(p) := \{x \in R^2 \mid \|x - p\| < r\}$ προσδιορίζει έναν κυκλικό δίσκο ακτίνας r γύρω από το p , Q προσδιορίζει το μοναδιαίο κυκλικό δίσκο $B_1(0)$ και $Q_0 := \{(x_1, x_2) \in Q \mid x_2 = 0\}$, $Q_+ := \{(x_1, x_2) \in Q \mid x_2 > 0\}$. [24]

1.2.3 Bayesian προσέγγιση για τη δημιουργία inpainting μοντέλων

Τα περισσότερα αντικείμενα στο φυσικό κόσμο δεν είναι διαφανή και ο άνθρωπος συνεχώς συναντά το φαινόμενο της παρεμβολής αντικειμένων στο οπτικό του πεδίο. Επιπλέον, το ανθρώπινο μάτι βλέπει τον κόσμο ως τέλεια διατεταγμένο και όχι σαν ένα σύνολο ανεξάρτητων διακριτών αντικειμένων. Οι επιστήμονες που μελετούν την ανθρώπινη όραση και νόηση πιστεύουν ότι ο άνθρωπος, χωρίς να το συνειδητοποιεί, εφαρμόζει συνεχώς και έξυπνα τους κανόνες Bayesian συμπερασμού και λήψης απόφασης.

Σε Bayesian όρους το inpainting ορίζεται ως η μεγιστοποίηση της ύστερης πιθανότητας $p(u|u_0, G)$. Ο τύπος του Bayes γράφεται:

$$p(u|u_0, G) = \frac{p(u_0|u, G)p(u|G)}{p(u_0|G)} \quad (1.1)$$

Στις περισσότερες εφαρμογές η απώλεια πληροφορίας είναι ανεξάρτητη του περιεχομένου της εικόνας και άρα μπορεί να γίνει η υπόθεση ότι $p(u|G) = p(u)$. Αν τα u_0 και G είναι γνωστά, το $p(u_0|G)$ είναι απλώς μία σταθερά κανονικοποίησης. Συνεπώς, επιζητείται η μεγιστοποίηση του γινομένου του μοντέλου δεδομένων $p(u_0|u, G)$ επί του πρότερου μοντέλου της εικόνας $p(u)$.

Εισάγοντας μεγέθη εντροπίας, όπως η log-likelihood συνάρτηση $E = -\ln p/\beta$, επιθυμούμε να ελαχιστοποιήσουμε την ποσότητα:

$$E(u|u_0, G) = E(u_0|u, G) + E[u] \quad (1.2)$$

σε μία προσθετική σταθερά. Αν $\beta = 1/(kT)$, όπου k είναι η σταθερά του Boltzmann και T η απόλυτη θερμοκρασία, τότε λαμβάνεται $p = \exp(-\beta E)/Z$ που είναι ο τύπος του Gibbs από τη στατιστική μηχανική και συνδέει την ενέργεια μίας πιθανής κατάστασης του συστήματος με την πιθανότητα εμφάνισής της.

Με την προσθήκη Gaussian θορύβου, το μοντέλο δεδομένων γίνεται:

$$E(u_0|u, G) = \frac{1}{2\sigma^2|\Omega \setminus G|} \int_{\Omega \setminus G} (u_0 - Ku)^2 dx \quad (1.3)$$

όπου $dx = dx_1 dx_2$ είναι το στοιχειώδες τμήμα μίας περιοχής Lebesgue, σ^2 η διασπορά και $|\Omega \setminus G|$ το μέτρο Lebesgue. Ουσιαστικά το κλειδί για το inpainting πρόβλημα είναι η εφαρμογή του κατάλληλου πρότερου μεγέθους $p(u)$ ή $E[u]$, κάτι που συνιστά και το θεμελιώδες πρόβλημα στην επεξεργασία εικόνας.

1.2.4 Ένταξη του inpainting στο θεμελιώδες πρόβλημα της επεξεργασίας εικόνας

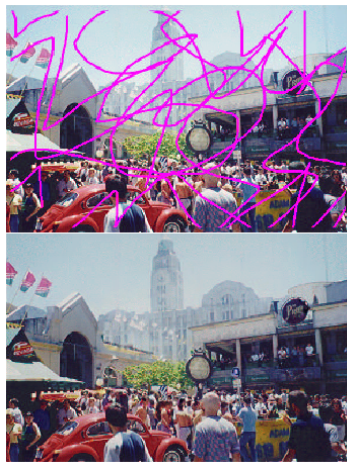
Θεμελιώδες ερώτημα στην επεξεργασία εικόνας αποτελεί το τι ακριβώς είναι οι εικόνες και για να θεωρείται κλάδος των μαθηματικών το ερώτημα αυτό πρέπει να απαντάται αρκούντως ικανοποιητικά. Γενικά, ένα πρόβλημα στην επεξεργασία εικόνας μπορεί να μοντελοποιηθεί ως ένα πρόβλημα εισόδου-εξόδου $U = \tau(U_0)$, όπου U_0 είναι είτε μία μεμονωμένη εικόνα είτε μία παρατηρούμενη ακολουθία εικόνων, τ είναι ένας γραμμικός ή μη γραμμικός τελεστής και U είναι το χαρακτηριστικό που αναζητείται στο εκάστοτε πρόβλημα επεξεργασίας εικόνας. Για παράδειγμα

στο inpainting ισχύουν $U_0 = (u_0, G)$ και $U = u$. Έχοντας γνώση των κλάσεων αντικειμένων U_0 και U (ή του ορισμού και των πεδίων του τ) μπορεί να γίνει αποτελεσματική μαθηματική μοντελοποίηση και ανάλυση της εικόνας. Οι τρεις πιο σημαντικές προσεγγίσεις στο θεμελιώδες πρόβλημα της επεξεργασίας εικόνας μαζί με σχετιζόμενες inpainting μεθόδους περιγράφονται στις ακόλουθες παραγράφους.

Φυσική προσομοίωση

Οι εικόνες παράγονται μέσω προσομοιώσεων φυσικών, χημικών ή βιολογικών διεργασιών. Η τεχνική της φυσικής προσομοίωσης χρησιμοποιείται ευρέως στα *Γραφικά Υπολογιστών*, ώστε η κίνηση να μοιάζει φυσική και ρεαλιστική (αφορά τον τομέα του computer animation στον κινηματογράφο και στα βιντεοπαιχνίδια), αλλά και σε άλλες εφαρμογές όπως στην αυτοκινητοβιομηχανία για την προσομοίωση συγκρούσεων.

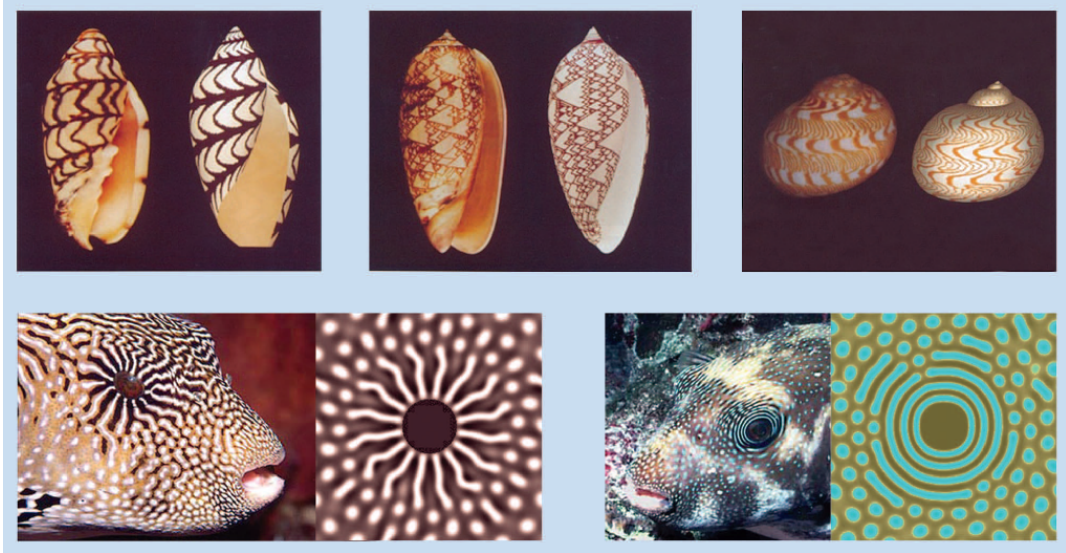
Κλασική εφαρμογή της φυσικής προσομοίωσης είναι η σύνθεση εικόνας από ροές ρευστών, οι οποίες προσδιορίζονται από την επίλυση των εξισώσεων *Navier-Stokes*. Συγκεκριμένα, η ένταση της εικόνας έχει το ρόλο της συνάρτησης ροής για ένα δισδιάστατο ασυμπίεστο ρευστό και η laplacian της παίζει το ρόλο της *στροβιλότητας* του ρευστού. Για την επίλυση του inpainting προβλήματος οι ροές διαδίδονται εντός της ελλειπούσας περιοχής, όπως ακριβώς απαιτείται η συνέχιση των *ισοφάτων* εντός του inpainting χωρίου, ενώ παράλληλα αντιστοιχίζονται τα διανύσματα παραγώγων στο σύνορό του. Η μέθοδος βασίζεται στις *Navier-Stokes* εξισώσεις από τη ρευστοδυναμική, κάτι που συνεπάγεται καλά θεμελιωμένο θεωρητικό υπόβαθρο και ήδη γνωστά αριθμητικά σχήματα επίλυσης των εξισώσεων. Είναι μία προσέγγιση που εισάγει τις ιδέες της Υπολογιστικής Ρευστοδυναμικής στα προβλήματα της Όρασης Υπολογιστών. Το Σχήμα 1.6 είναι ένα παράδειγμα της εφαρμογής της μεθόδου των Bertalmio et al [7].



Σχήμα 1.6: Αλλοιωμένη εικόνα και Navier-Stokes Inpainting [7].

Επίσης, το περίφημο θεωρητικό μοντέλο διάχυσης-ανάδρασης του Turing (*Reaction-Diffusion model*, *RD model*) χρησιμοποιείται για τη δημιουργία προτύπων δέρματος. Μάλιστα πρόσφατα στο [8] αξιοποιήθηκε για την ερμηνεία του αυτορρυθμιζόμενου σχηματισμού προτύπων στο αναπτυσσόμενο ζωικό έμβρυο. Αν και στο παρελθόν είχε αμφισβητηθεί η χρησιμότητά του σε εφαρμογές

του πραγματικού κόσμου, οι Kondo και Miura διέλυσαν την όποια δυσπιστία απέναντί του μέσω της ανάδειξης εντυπωσιακών εφαρμογών, όπως στην κατασκευή χωρικών προτύπων σε όστρακα και ψάρια, και έδωσαν στο μοντέλο αυτό τη δυναμική εφαρμογής του ως πειραματική υπόθεση εργασίας σε μία ευρεία ποικιλία μορφολογικών φαινομένων, ειδικά από την επιστήμη της βιολογίας.



Σχήμα 1.7: Αναπαραγωγή βιολογικών προτύπων με χρήση του RD μοντέλου [8].

Ένα ακόμη παράδειγμα φυσικής προσομοίωσης αφορά τις κατασκευές που παρουσιάζουν την ιδιότητα της *αυτοομοιότητας* (fractals), όπως για παράδειγμα φύλλα, βράχοι ή φυσικά τοπία, και που προσομοιώνονται μέσω επαναληπτικών συναρτησιακών συστημάτων (Iterated Function Systems, IFSs). Τα *IFS fractals*, όπως συχνά συναντώνται στη βιβλιογραφία [25], μπορούν να είναι κάθε διάστασης αλλά συνήθως υπολογίζονται και απεικονίζονται στο επίπεδο (2D). Ένα fractal αποτελείται από την ένωση πολλών αντιγράφων του εαυτού του που μετασχηματίζονται από μία *συστολική*² συνάρτηση. Αντίστοιχα, τα αντίγραφα αυτά αποτελούνται από αντίγραφα του εαυτού τους και ούτω καθεξής (αρχή της αυτοομοιότητας). Τυπικά το IFS είναι ένα πεπερασμένο σύνολο συστολικών απεικονίσεων σε ένα πλήρη μετρικό χώρο. Συμβολικά, το

$$\{f_i : X \rightarrow X | i = 1, 2, \dots, N\}, N \in \mathbb{N} \quad (1.4)$$

είναι ένα IFS αν κάθε f_i είναι μία συστολή στον πλήρη μετρικό χώρο X . Μερικές φορές απαιτείται κάθε συνάρτηση f_i να είναι γραμμική ή γενικότερα ένας αφινικός μετασχηματισμός και έτσι αναπαριστάται με πίνακα. Παρόλ' αυτά, τα IFSs μπορούν να κατασκευαστούν και από μη γραμμικές συναρτήσεις, συμπεριλαμβανομένων προβολικών και Möbius μετασχηματισμών. Παράδειγμα IFS μη γραμμικής συνάρτησης είναι η fractal φλόγα που δημιουργήθηκε από τον Scott Draves [9].

²Σε ένα μετρικό χώρο (M, d) μία συνάρτηση $f : M \rightarrow M$ ονομάζεται συστολική, αν έχει την ιδιότητα για κάποιο μη αρνητικό πραγματικό $k < 1$, $\forall x \in M, \forall y \in M: d(f(x), f(y)) \leq kd(x, y)$.



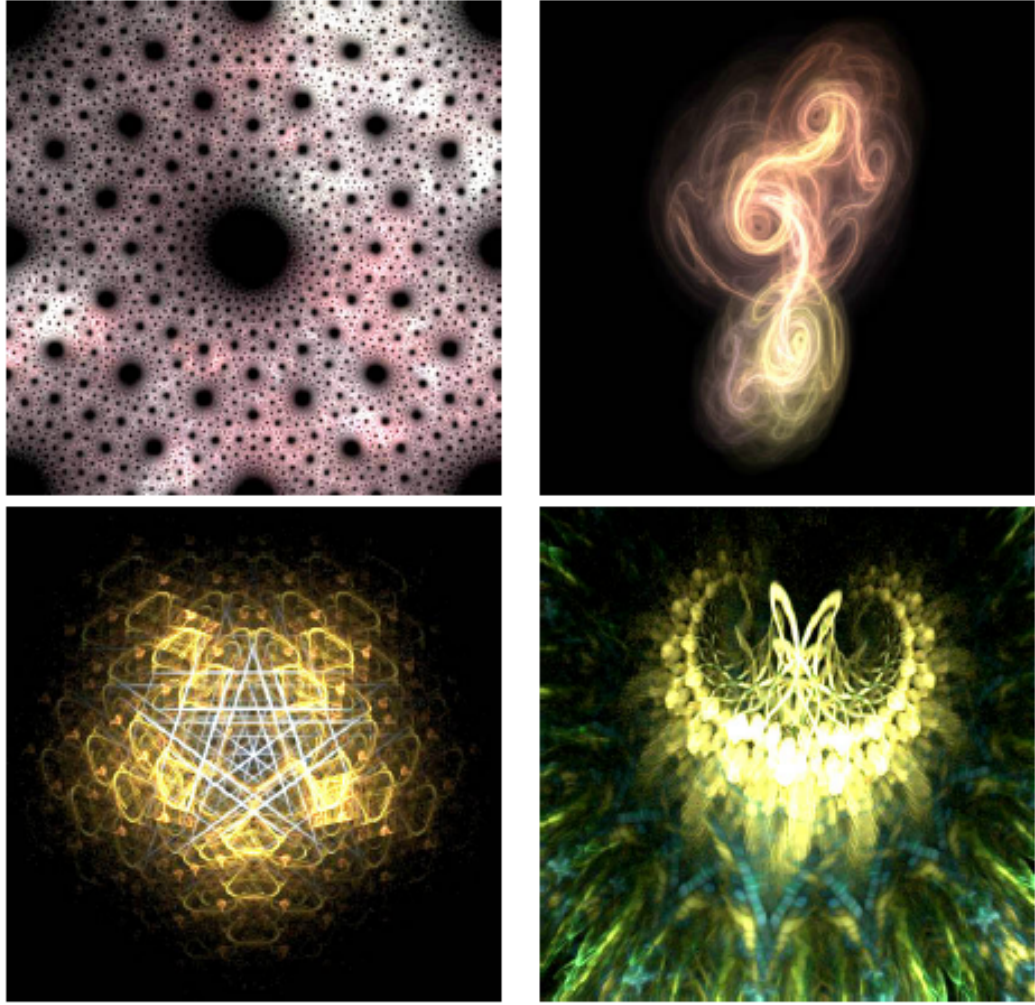
Σχήμα 1.8: Η φτέρη του Barnsley, ένα κλασικό IFS fractal.

Τυχαία πεδία

Οι εικόνες μοντελοποιούνται ως δείγματα που έχουν ληφθεί από συγκεκριμένα τυχαία πεδία. Πρωταρχικός στόχος αυτής της μοντελοποίησης είναι η κατανόηση της συνάρτησης κατανομής πιθανότητας $p(u)$, ειδικά όταν το Ω είναι διάταξη από ψηφιακά pixels. Κλασικά μοντέλα έχουν κυρίως εμπνευστεί από τα πεδία Gibbs στη *Στατιστική Μηχανική*, όπου οι τοπικοί περιορισμοί ενέργειας επιβάλλονται με τον ίδιο τρόπο όπως στους κρυστάλλους Ising [26].

Στα πλαίσια της αναλογίας μεταξύ των εικόνων και των στατιστικών μηχανικών συστημάτων, τα επίπεδα έντασης φωτεινότητας των pixels και η παρουσία και ο προσανατολισμός των ακμών θεωρούνται ως οι καταστάσεις ατόμων και μορίων σε ένα φυσικό σύστημα με δομή πλέγματος. Η ανάθεση μιας συνάρτησης ενέργειας σε ένα φυσικό σύστημα καθορίζει την κατανομή Gibbs του, κάτι που επακόλουθα προσδιορίζει ένα MRF μοντέλο. Για μια σειρά μηχανισμών αλλοίωσης της εικόνας όπως θόλωμα, μη γραμμικές παραμορφώσεις και θόρυβος, η ύστερη κατανομή είναι ένα MRF που αντίστοιχα εκφράζει ένα διαφορετικό του αρχικού φυσικό σύστημα. Για παράδειγμα, η μείωση της θερμοκρασίας σε ένα φυσικό σύστημα οδηγεί στην απομόνωση καταστάσεων χαμηλής ενέργειας που ισοδύναμα είναι οι πιο πιθανές καταστάσεις υπό την κατανομή Gibbs. Η ανάλογη διαδικασία υπό την ύστερη κατανομή αποφέρει τη βέλτιστη εκτίμηση (maximum a posteriori estimate) της εικόνας, δεδομένων των αλλοιωμένων εκδοχών της.

Τα τυχαία πεδία μπορούν επίσης να εκπαιδευτούν από μία βάση δεδομένων με εικόνες μέσω τεχνικών όπως το φιλτράρισμα και η μη-παραμετρική εκτίμηση με την *αρχή της μέγιστης εντροπίας* [27]. Στη bayesian θεωρία η αρχή της μέγιστης εντροπίας είναι ένα αξίωμα που δηλώνει ότι η κατανομή πιθανότητας που καλύτερα αντιπροσωπεύει την τρέχουσα γνώση, βάση γνωστών περιορισμών, είναι εκείνη με τη μεγαλύτερη εντροπία. Εάν δίνονται οι περιορισμοί για μια συνάρτηση κατανομής πιθανότητας, δεδομένου του συνόλου των πιθανών κατανομών πιθανότητας που κωδικοποιούν την πληροφορία, η κατανομή πιθανότητας που μεγιστοποιεί την εντροπία της πληροφορίας είναι η αληθής γι' αυτούς τους περιορισμούς.



Σχήμα 1.9: Παραδείγματα εικόνων fractal flame κατασκευασμένων με μη γραμμικά IFSs [9].

Χώροι συναρτήσεων

Είναι μία ντετερμινιστική προσέγγιση όπου χώροι συναρτήσεων χρησιμοποιούνται για να εκφραστούν οι κανονικότητες της εικόνας μετρημένες με μία ενέργεια $E[u]$. Οι κλασικές φασματικοί μέθοδοι, όπως ο Fourier, υποθέτουν ότι οι εικόνες ανήκουν στον L^2 και ότι $E[u] = \|u\|_{L^2} = \|\hat{u}\|_{L^2}$, όπου $\hat{u}$ είναι ο μετασχηματισμός Fourier. Η θεωρία γραμμικού φιλτραρίσματος υποθέτει ότι οι εικόνες ανήκουν στο Sobolev χώρο $W^{1,2}(\Omega)$ και το οπτικό τους περιεχόμενο μετράται ως $E[u] = \|\nabla u\|_{L^2}$.

Σαν μια προσπάθεια να καταδείξουν τη μεγάλη σημασία των ακμών στην ανθρώπινη οπτική αντίληψη [28], οι Rudin, Osher και Fatemi πρότειναν το μοντέλο εικόνας φραγμένης μεταβολής (*Bounded Variation image model* [29]). Το $E[u] = \int_{\Omega} \|Du\|$ είναι το μέτρο Radon ολικής μεταβολής. Οι εικόνες φραγμένης μεταβολής και γενικότερα οι Besov εικόνες έχουν μελετηθεί επίσης εκτενώς στα πλαίσια της θεωρίας κυματιδίων. Για την επίλυση του inpainting προβλήματος

με χρήση του μοντέλου φραγμένης μεταβολής [11] ελαχιστοποιείται η έκφραση:

$$E_{tv}[u|u_0, G] = \alpha \int_{\Omega} \|Du\| + \frac{1}{2\sigma^2 \|\Omega \setminus G\|} \int_{\Omega \setminus G} (Ku - u_0)^2 dx \quad (1.5)$$

όπου η διασπορά του θορύβου σ^2 μπορεί να εκτιμηθεί στατιστικά και η μόνη παράμετρος είναι το α . Το μοντέλο αυτό είναι παρόμοιο με το μοντέλο αποκατάστασης των Rudin, Osher & Fatemi, διαφοροποιούμενο μόνο στο ότι ο πολλαπλασιαστής Lagrange λ είναι ο όρος:

$$\frac{1}{2\sigma^2 \|\Omega \setminus G\|} 1_{\Omega \setminus G}(x) \quad (1.6)$$

Το Σχήμα 1.10 είναι μία τυπική έξοδος του μοντέλου.



Σχήμα 1.10: Inpainting φραγμένης μεταβολής [10]. Στην πάνω (θολωμένη) εικόνα λείπουν 60 τυχαία επιλεγμένα πακέτα που χάθηκαν κατά τη φάση της μετάδοσης. Η κάτω εικόνα έχει προκύψει μετά από αποθορυβοποίηση και inpainting φραγμένης μεταβολής.

Για την επίλυση του παραπάνω μοντέλου γίνεται χρήση υπολογιστικών PDEs. Μέσω μεταβολικών μεθόδων προκύπτει η ακόλουθη μη γραμμική εξίσωση ελλειπτικού τύπου:

$$-\alpha \nabla \cdot \left[\frac{\nabla u}{|\nabla u|} \right] + K^* \lambda_G (Ku - u_0) = 0 \quad (1.7)$$

όπου K^* είναι το συμπληρωματικό. Η ανωτέρω εξίσωση επιλύεται με τεχνικές γραμμικοποίησης.

Για τον ακριβή διαχωρισμό των ακμών, οι Mumford και Shah [30] πρότειναν το περίφημο μοντέλο ελεύθερου συνόρου (*free boundary model*) για κατά τμήματα λείες εικόνες:

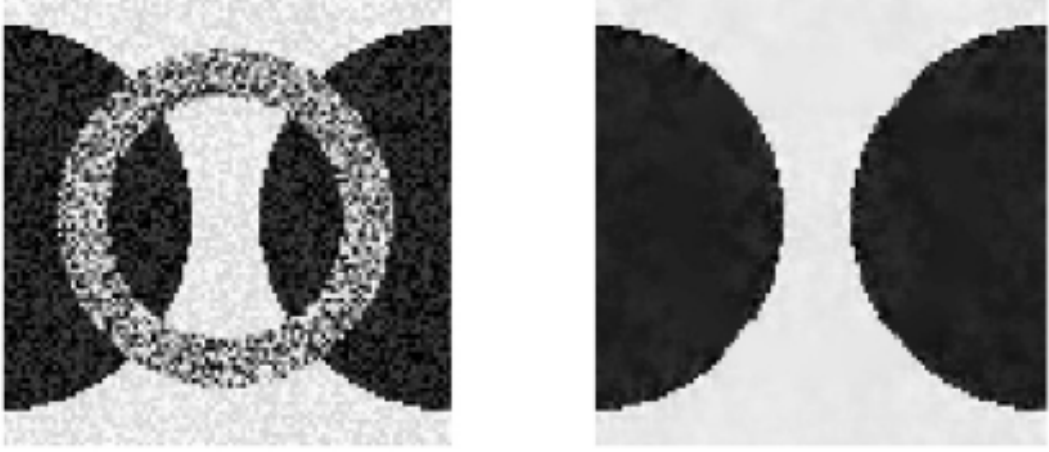
$$E[u, \Gamma] = E[u|\Gamma] + E[\Gamma] = \frac{\mu}{2} \int_{\Omega \setminus \Gamma} |\nabla u|^2 dx + \alpha H^1(\Gamma) \quad (1.8)$$

όπου το Γ προσδιορίζει το σύνολο αλμάτων και H^1 τη μονοδιάστατη Hausdorff απόσταση που συχνά αντικαθίσταται από το μήκος Γ στους υπολογισμούς.

Το inpainting που βασίζεται στο μοντέλο των Mumford-Shah ελαχιστοποιεί την ενέργεια:

$$E_{ms}[u, \Gamma | u_0, G] = E[u | \Gamma] + E[\Gamma] + E[u_0 | u, G] \quad (1.9)$$

Το Σχήμα 1.11 είναι αποτέλεσμα της μεθόδου.



Σχήμα 1.11: Mumford-Shah inpainting [10]. Αριστερά είναι η θορυβημένη εικόνα με την περιοχή G να είναι ο κυκλικός δακτύλιος που ενώνει τους δύο δίσκους. Δεξιά είναι η αποθορυβοποιημένη εικόνα, έχοντας εφαρμοστεί inpainting στον κυκλικό δακτύλιο.

Τόσο το μοντέλο φραγμένης μεταβολής όσο και το Mumford-Shah μοντέλο αξιοποιούν μόνο την πρώτη τάξης γεωμετρία των επιπεδοσυνόλων και των ακμών. Συχνά είναι επαρκή για κλασικές εφαρμογές όπως κατάτμηση, αλλά για το inpainting συνήθως απαιτείται και πληροφορία υψηλότερης τάξης, όπως καμπυλότητα, για καλύτερο αποτέλεσμα. Θεμελιώδες εργαλείο στην περιοχή είναι το μοντέλο καμπυλότητας *elastica* του Euler:

$$e[\gamma] = \int_{\lambda} (\alpha + b\kappa^2) dx \quad (1.10)$$

όπου το κ προσδιορίζει την καμπυλότητα της καμπύλης γ . Το μοντέλο αυτό εισήχθη στην Όραση Υπολογιστών για πρώτη φορά από τον Mumford. Επιβάλλοντας το $e[\gamma]$ σε όλα τα επιπεδοσύνολα, λαμβάνουμε το μοντέλο εικόνας *elastica* [12]:

$$E[u] = \int_{\Omega} (\alpha + b\kappa^2) |Du| \quad (1.11)$$

$$\kappa = \nabla \cdot \left[\frac{\nabla u}{|\nabla u|} \right] \quad (1.12)$$

Αντικαθιστώντας την ενέργεια μήκους στο μοντέλο εικόνας Mumford-Shah με το $e[\Gamma]$, λαμβάνεται το Mumford-Shah-Euler μοντέλο εικόνας [30].

Κεφάλαιο 2

Αυτόματη εύρεση ελλειπόντων τμημάτων στις τοιχογραφίες

2.1 Εισαγωγή

Βασικός στόχος της παρούσας εργασίας είναι η ψηφιακή επιδιόρθωση των τοιχογραφιών. Στην κατεύθυνση αυτή το πρώτο στάδιο είναι ο εντοπισμός των ρωγμών/κενών και γενικότερα των ελλειπουσών περιοχών στις τοιχογραφίες στο καλύτερο δυνατό βαθμό. Για το σκοπό αυτό χρησιμοποιούνται δύο μέθοδοι που το ποια επιλέγεται εξαρτάται από την περιοχή των τοιχογραφιών όπου επιχειρείται αποκατάσταση. Η πρώτη προσέγγιση αξιοποιεί στοιχεία μαθηματικής μορφολογίας, με περαιτέρω αξιοποίηση της πληροφορίας των ακμών, για την καλύτερη δυνατή αποτελεσματικότητα στην ανίχνευση των ελλειπόντων τμημάτων, με το μικρότερο δυνατό ποσοστό απώλειας σημαντικής πληροφορίας έξω από την εξαγόμενη μάσκα. Η δεύτερη μέθοδος στηρίζεται σε κατάτμηση με γραφοτομές. Για τη συστηματική αξιολόγηση της εξαχθείσας μάσκας ανά περίπτωση χρησιμοποιήθηκαν κατάλληλοι δείκτες απόδοσης.

Γενικά, η αυτόματη εξαγωγή ικανοποιητικής μάσκας ήταν ένας δυσχερής και σε κάποιες περιπτώσεις ανέφικτος στόχος. Η χρωματική ομοιότητα μεταξύ του κονιάματος που έχει τοποθετηθεί από τους αρχαιολόγους στα κενά και του ασβεστοκονιάματος στις αρχικές τοιχογραφίες καθιστά πολύ δύσκολη την εύρεση των ελλειπουσών περιοχών βάση της χρωματικής πληροφορίας μόνο. Επίσης, η ευρεία ποικιλία στη γεωμετρία των περιοχών αυτών είναι ένα ακόμη εμπόδιο για την αυτοματοποίηση της διαδικασίας ανίχνευσής τους.

Από την άλλη μεριά, υπάρχει μία κατηγορία μεθόδων που θα μπορούσε να φέρει ελπιδοφόρα αποτελέσματα, όπως προτείνεται και στο [2], διότι οι τοιχογραφίες έχουν μία ιδιαιτερότητα που θα μπορούσε να αξιοποιηθεί. Συγκεκριμένα, κατά το τελικό στάδιο της συντήρησής τους, όταν τα σπαράγματα εναποτίθενται πάνω σε φορητό πινάκα, οι ειδικοί φροντίζουν ώστε τα ευρεθέντα τμήματα να προεξέχουν μερικά χιλιοστά του κονιάματος που τοποθετείται στα κενά. Έτσι, η ανίχνευση των ασυνεχειών βάθους με τεχνικές 3Δ Υπολογιστικής Όρασης σε συνδυασμό φυσικά με αξιοποίηση της πληροφορίας του χρώματος θα μπορούσε να φέρει καλύτερα αποτελέσματα. Τεχνικές στερεοσκοπίας θα μπορούσαν να χρησιμοποιηθούν, αν και πιο αποδοτικές για το συγκεκριμένο πρόβλημα πιθανότατα θα ήταν τεχνικές ανίχνευσης ασυνεχειών βάθους μέσω ελεγχόμενου φωτισμού. Πάντως, η εφαρμογή τέτοιων μεθόδων 3Δ Υπολογιστικής Όρασης προ-

υποθέτει τη λήψη νέου φωτογραφικού υλικού υπό καλά ελεγχόμενες συνθήκες φωτισμού, οπότε η υλοποίησή τους αφήνεται για μελλοντική εργασία.

Συχνά ο μηχανισμός αλλοίωσης των τοιχογραφιών περιλαμβάνει πολύπλοκες διεργασίες όπως ξεθώριασμα του χρώματος ή μερική απόξεση της ζωγραφικής επιφάνειας. Μοντελοποιώντας τέτοιες αλλοιώσεις ως προσθήκη θορύβου αποτελεί φτωχή προσέγγιση του μηχανισμού αλλοίωσης. Η ανάπτυξη εξειδικευμένων μοντέλων παρατήρησης για τη μοντελοποίηση της φυσικής διαδικασίας φθοράς του τοιχογραφικού υποστρώματος και της χημικής αλλοίωσης των χρωστικών ουσιών της τοιχογραφίας είναι ένας σπουδαίος στόχος, αλλά με βάση τις έως τώρα ερευνητικές χημικές μελέτες (π.χ. [31]) είναι δύσκολο να κατασκευαστούν συγκεκριμένα υπολογιστικά μοντέλα της χημικής αλλοίωσης των χρωμάτων της εικόνας.

Σε κάθε περίπτωση η αυτοματοποίηση της διαδικασίας εύρεσης των ελλειπόντων τμημάτων είναι ένας καλός στόχος γιατί οι τοιχογραφίες έχουν πολλές τέτοιες περιοχές και η διεκπεραίωση της ανίχνευσής τους από αρχαιολόγους και ειδικευμένους τεχνίτες είναι ιδιαίτερα χρονοβόρα. Στην παρούσα εργασία με χρήση μαθηματικής μορφολογίας εντοπίζονται αρκετά καλά οι ρωγμές αλλά και μεγαλύτερα κενά πληροφορίας σε κάποιες περιπτώσεις. Ο αλγόριθμος συναντά πρόβλημα σε περιοχές με γειτονικές ρωγμές που κάποιες φορές τις ενώνει, όπως επίσης και σε σημεία των παραστάσεων με έντονες ακμές, καθώς αναγνωρίζονται ως σύνορα ελλείψεων. Η εκπαίδευση ενός δικτύου στις συγκεκριμένες παραστάσεις και τεχνοτροπία θα μπορούσε να οδηγήσει στην εξάλειψη αυτού του φαινομένου, όπως γίνεται στο [21] όπου χρησιμοποιείται Median RBF νευρωνικό δίκτυο για το σκοπό αυτό.

Σε μια προσπάθεια περαιτέρω βελτιστοποίησης της εξαχθείσας μάσκας ή και ορθότερης ανίχνευσης σε περιπτώσεις που ο αλγόριθμος μαθηματικής μορφολογίας δε δίνει ικανοποιητικά αποτελέσματα, χρησιμοποιείται ο αλγόριθμος “GrabCut” των Rother, Kolmogorov και Blake [3]. Η ακόλουθη ενότητα αφορά τον αλγόριθμο εξαγωγής μάσκας που βασίζεται σε στοιχεία μαθηματικής μορφολογίας και η αμέσως επόμενη σχετίζεται με το “GrabCut”. Αρχικά παρουσιάζεται μία εισαγωγή στο πρόβλημα της κατάτμησης με βάση το [1] και ακολουθούν κάποιες κλασικές και μερικές πιο πρόσφατες προσεγγίσεις για την επίλυση του.

2.2 Το πρόβλημα της κατάτμησης εικόνας

2.2.1 Γενικά στοιχεία για το πρόβλημα

Η κατάτμηση εικόνας είναι ένα από τα δυσκολότερα προβλήματα στην Όραση Υπολογιστών. Είναι η διαδικασία εύρεσης αντικειμένων σε μία εικόνα και διαχωρισμού από το φόντο τους, μέσω διαμέρισης της εικόνας σε ξεχωριστές περιοχές, καθεμία εκ των οποίων είναι συνεκτική και ομογενής με βάση κάποιο κριτήριο, όπως η ένταση της φωτεινότητας, η υφή, το χρώμα, η κίνηση, το βάθος. Επειδή το κριτήριο ομογένειας δεν είναι συγκεκριμένο, η κατάτμηση ως πρόβλημα εξαρτάται από την εφαρμογή και το εκάστοτε κριτήριο.

Στην κατάτμηση εμφανίζεται μία δυικότητα μεταξύ των *boundary-based* προσεγγίσεων που περιλαμβάνουν μεθόδους ανίχνευσης ακμών και ενεργών χαμπυλών (active contour methods) και των *region-based* προσεγγίσεων που μπορούν να ταξινομηθούν σε στατιστικές μεθόδους βασισμένες στα Markov τυχαία πεδία (Markov Random Fields, *MRFs*), σε *variational* μεθόδους όπου ελαχιστοποιούνται συναρτησιακά ενέργειας και σε μεθόδους επέκτασης περιοχών (region

growing), όπως είναι τα *watersheds*. Η προσέγγιση με MRFs έχει το μειονέκτημα ότι έχει μεγάλες χρονικές απαιτήσεις. Στις region growing μεθόδους ισχυρή και υπολογιστικά αποδοτική είναι η προσέγγιση με μορφολογικά watershed μαζί με μη γραμμικούς, πολυκλιμακωτούς μορφολογικούς τελεστές για την εξαγωγή των κατάλληλων δεικτών με βάση το περιεχόμενο της εικόνας για την τροφοδότηση της κατάτμησης.

Τελευταία υπάρχει μια προσπάθεια για συγκερασμό των region-based με τις boundary-based μεθόδους για τη δημιουργία πιο ισχυρών αλγορίθμων. Αυτό έχει να κάνει με το γεγονός ότι οι δυο αυτές συμπληρωματικές μέθοδοι έχουν διαφορετικούς στόχους: οι region-based μέθοδοι εκμεταλλεύονται τις ιδιότητες ομογένειας, ενώ οι boundary-based προσπαθούν να αξιοποιήσουν τις πιθανές ανομοιογένειες στα ίδια δεδομένα. Μια τέτοια προσπάθεια γίνεται και στην παρούσα εργασία στον αλγόριθμο εύρεσης των ελλείψεων στις τοιχογραφίες που περιγράφεται στην επόμενη ενότητα.

2.2.2 Κάποιες πρόσφατες μέθοδοι

Στο παρελθόν έχουν χρησιμοποιηθεί διάφορες προσεγγίσεις για την επίλυση του προβλήματος της κατάτμησης. Το σύνολο των μεθόδων αυτών απαιτεί κάποια εργασία από το χρήστη συνήθως στην αρχή για έναν αρχικό προσδιορισμό του προσκηνίου ή/και του παρασκηνίου, αλλά κάποιες φορές και κατά τη διάρκεια για τυχόν συμπλήρωση της απαιτούμενης πληροφορίας για την επιτυχή εκτέλεση του αλγορίθμου ή και στο τέλος για πιθανή διόρθωση του αποτελέσματος της κατάτμησης (interactive tools).

Οι Caselles et al. το 1995 [32] χρησιμοποίησαν επιπεδοσύνολα (*levelsets*) για κατάτμηση εικόνας. Η τεχνική τους βασίζεται σε ενεργές καμπύλες που εξελίσσονται στο χρόνο με βάση συγκεκριμένα εσωτερικά γεωμετρικά χαρακτηριστικά της εικόνας. Οι καμπύλες αυτές χωρίζονται και συγχωνεύονται κατά τη μεταβολή του χρόνου επιτρέποντας την παράλληλη ανίχνευση πολλών αντικειμένων, καθώς και τόσο των εσωτερικών όσο και των εξωτερικών συνόρων τους. Η προσέγγιση αυτή επιτρέπει τη σύνδεση των κλασικών *snakes*¹ που βασίζονται στην ελαχιστοποίηση ενέργειας και των γεωμετρικών ενεργών καμπυλών που στηρίζονται στη θεωρία της εξέλιξης των καμπυλών. Αν και ως μέθοδος πλεονεκτεί στο ότι σχεδόν οποιαδήποτε ενέργεια μπορεί να χρησιμοποιηθεί, έχει το μειονέκτημα ότι υπολογίζει τοπικό ελάχιστο και άρα η αρχικοποίηση παίζει ρόλο στο αποτέλεσμα. Έτσι, στις περιπτώσεις που η συνάρτηση ενέργειας μπορεί να ελαχιστοποιηθεί ακριβώς μέσω τομών γράφων είναι προτιμότερη η χρήση των τελευταίων.

Αυτή τη σκέψη αξιοποίησαν οι Boykov και Kolmogorov το 2003 στο [34] για τον υπολογισμό γεωδαιτικών καμπυλών και ελαχιστικών (minimal) επιφανειών σε χώρους Riemann. Μέσω της παρατήρησής τους ότι μία τομή γράφου ενσωματωμένη σε ένα συνεχή χώρο μπορεί να ερμηνευτεί ως μία καμπύλη (στο 2Δ χώρο) ή ως μία επιφάνεια (στο 3Δ χώρο) συσχέτισαν γεωδαιτικές ενεργές καμπύλες και τομές γράφων για μία νέα μέθοδο κατάτμησης που συνδυάζει κάποια από τα οφέλη των δυο μεθόδων. Χτίζουν ένα γράφημα-πλέγμα με κατάλληλα βάρη στις ακμές, έτσι ώστε το κόστος των τομών να είναι αυθαίρετα κοντά στο μήκος (εμβαδόν) της αντίστοιχης καμπύλης (επιφάνειας) για ένα ανισοτροπικό μετρικό Riemann. Μέσω αυτής της ισοδυναμίας αφενός υπολογίζουν ολικά ελάχιστα γεωδαιτικές καμπύλες (minimal επιφάνειες στο 3Δ χώρο) και αφετέρου

¹Το 1988 οι Kass, Witkin και Terzopoulos δημοσίευσαν την πρωτοποριακή εργασία τους [33] για τα snakes, τα οποία ήταν splines ελαχιστοποίησης ενέργειας οδηγούμενα από εξωτερικές δυνάμεις και το περιεχόμενο της εικόνας, και τα οποία μπορούσαν να προσαρμόζονται και έτσι να εντοπίζουν χαρακτηριστικά της εικόνας όπως ακμές και καμπύλες.

ελαχιστοποιούν τις ατέλειες μετρολογίας σε διάφορες μεθόδους της Όρασης Υπολογιστών με τομές γράφων μέσω συσχέτισης διάφορων όρων στα μοντέλα αυτά με συγκεκριμένα γεωμετρικά μετρικά.

Στα πλαίσια κλειστού λογισμικού έχουν χρησιμοποιηθεί αρκετές ακόμη ενδιαφέρουσες μέθοδοι. Οι Mortensen & Barrett το 1995 πρότειναν μία μεθοδολογία (*Intelligent Scissors*) που επιτρέπει στο χρήστη να επιλέγει μία καμπύλη ελάχιστου κόστους ζητώντας του ένα πρόχειρο προσδιορισμό του συνόρου δίνοντας μερικά σημεία με το «ποντίκι». Καθώς ο χρήστης τροφοδοτεί τον αλγόριθμο με σημεία χαράσσεται το μονοπάτι ελάχιστου κόστους. Αν κάπου αυτό αποκλίνει από την επιθυμητή πορεία, απαιτούνται επιπλέον σημεία από το χρήστη. Η μέθοδος υστερεί σε περιοχές έντονης ή καθόλου υψής όταν υπάρχουν περισσότερα από ένα ελάχιστα μονοπάτια. Επίσης, απαιτούνται πολλά σημεία από το χρήστη.

Ο αλγόριθμος *Magic Wand* που χρησιμοποιείται από το Adobe Photoshop 7 (2002) απαιτεί επίσης αρκετή παρέμβαση από το χρήστη που δίνει κάποια σημεία του αντικειμένου που θέλει να εξάγει, όπως επίσης και ένα περιθώριο του πόσο μπορεί να αποκλίνουν τα στατιστικά του χρώματος των pixel που ανήκουν στην ίδια συνεκτική συνιστώσα. Δεν είναι πάντα εφικτός ο προσδιορισμός του κατάλληλου περιθωρίου, ενώ στο Σχήμα 2.1, όπου χρησιμοποιούνται οι παραπάνω αλγόριθμοι σε συγκεκριμένη εφαρμογή, όπως είναι φανερό η ομοιότητα στην κατανομή του χρώματος μεταξύ φόντου και αντικειμένου δυσκολεύει τη σωστή κατάτμηση με το Magic Wand.

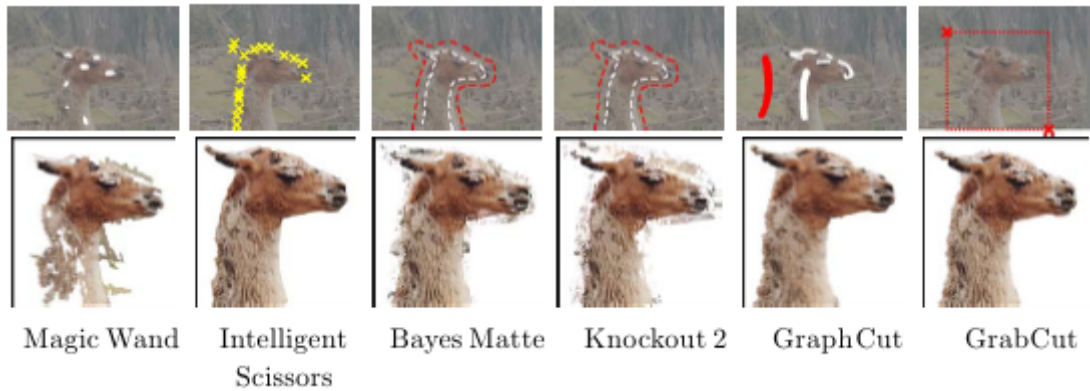
Επίσης τα μοντέλα *Bayes matting* μοντελοποιούν τις κατανομές χρώματος πιθανοτικά για την εκτίμηση των παραμέτρων α της κατάτμησης. Ο χρήστης αρχικοποιεί το χάρτη (*trimap*) $T = \{T_B, T_U, T_F\}$ μέσω προσδιορισμού του φόντου T_B και του προσκηνίου T_F και οι παράμετροι α υπολογίζονται στην ενδιάμεση περιοχή T_U . Εξαιρετική κατάτμηση λαμβάνει χώρα όταν η περιοχή T_U έχει σχετικά μικρή έκταση και οι κατανομές χρώματος προσκηνίου-παρασκηνίου διαφοροποιούνται επαρκώς. Βέβαια απαιτείται αρκετή εργασία από το χρήστη για την κατασκευή του εσωτερικού και του εξωτερικού μονοπατιού. Ο αλγόριθμος *Knockout 2* [Corel Corporation 2002] περιλαμβάνει την αρχικοποίηση του χάρτη T από το χρήστη και πάλι και παρουσιάζει παρόμοια αποτελέσματα με το Bayes matting.

Τέλος, οι Τομές Γράφων [35] είναι μία ισχυρή τεχνική βελτιστοποίησης που αξιώνει την αρχικοποίηση του χάρτη T και χρησιμοποιεί πιθανοτικά μοντέλα χρώματος, όπως το Bayes matting, και επιτυγχάνει εύρωστη κατάτμηση, ακόμη και όταν οι κατανομές χρώματος προσκηνίου-παρασκηνίου δε διαφοροποιούνται σημαντικά (*camouflage*). Ο αλγοριθμός με τομές γράφων αποτελεί τον πυρήνα του αλγορίθμου “GrabCut” που παρουσιάζεται στην Ενότητα 2.4.

2.3 Εύρεση ελλειπόντων τμημάτων με χρήση μαθηματικής μορφολογίας

2.3.1 Στοιχεία μαθηματικής μορφολογίας

Πολλές μέθοδοι στην Όραση Υπολογιστών απαιτούν την επίλυση ενός γραμμικού συστήματος. Συχνά όμως ένα πρόβλημα δεν επιλύεται ικανοποιητικά με την κατάστρωση ενός γραμμικού συστήματος, κάτι που είναι συχνό φαινόμενο σε προβλήματα που σχετίζονται με τη γεωμετρία της εικόνας, όπως είναι η κατάτμηση. Έτσι, στα πλαίσια της αναζήτησης μη γραμμικών γεωμετρικών προσεγγίσεων, μία αποτελεσματική μη γραμμική μεθοδολογία είναι η μαθηματική μορφολογία.



Σχήμα 2.1: Σύγκριση μερικών κλασικών μεθόδων για κατάτμηση (εξαγωγή προσκηνίου). Στην πάνω σειρά παρουσιάζεται η απαιτούμενη εργασία από το χρήστη προτού εκτελεστεί ο αλγόριθμος για τον προσδιορισμό είτε του προσκηνίου (λευκό χρώμα) και/ή του παρασκηνίου (κόκκινο χρώμα) είτε του συνόρου (κίτρινα σημάδια) του εξαγόμενου αντικειμένου. Στην κάτω σειρά απεικονίζεται το αποτέλεσμα της κατάτμησης για την κάθε μέθοδο. Το “GrabCut” καταφανώς πλεονεχτεί τόσο σε ποιότητα αποτελέσματος όσο και στην απαιτούμενη εργασία από τη μεριά του χρήστη. Έχει αντληθεί από το [3], στο οποίο έχει βασιστεί και η παρούσα υποενότητα.

Η μαθηματική μορφολογία είναι «μία θεωρητική μεθοδολογία συνόλων και πλεγμάτων (lattices) για την επεξεργασία εικόνας, η οποία στοχεύει στην ποσοτική περιγραφή της γεωμετρικής δομής των αντικειμένων της εικόνας» [36]. Τη δεκαετία του '60 χρησιμοποιήθηκε για την ανάλυση δυαδικών εικόνων από γεωλογικά και βιοϊατρικά δεδομένα, όπως επίσης για το φορμαλισμό εργασιών στη δυαδική αναγνώριση προτύπων, βασισμένων στα κυτταρικά αυτόματα και στη λογική Boole. Στα τέλη της δεκαετίας του '70 επεκτάθηκε στις grayscale εικόνες, ενώ στα μέσα της επόμενης δεκαετίας καθιερώθηκε ως κλασική μέθοδος, στενά συνυφασμένη με άλλες μη γραμμικές προσεγγίσεις φιλτραρίσματος. Στη δεκαετία του '90 γενικεύτηκε σε αυθαίρετα πλέγματα. Η εξέλιξη αυτή των ιδεών έχει διαμορφώσει την περιοχή που αποκαλείται σήμερα ως *μορφολογική επεξεργασία εικόνας* με πλούσιο θεωρητικό πλαίσιο, πολλούς αποδοτικούς αλγορίθμους και υλοποιήσεις, σε μία ευρεία ποικιλία προβλημάτων στην Επεξεργασία Εικόνας, στην Όραση Υπολογιστών και στην Αναγνώριση Προτύπων.

Οι έννοιες και οι ορισμοί στη μαθηματική μορφολογία είναι γενικοί και μπορούν να διατυπωθούν για όλα τα δυνατά μοντέλα μίας εικόνας, δηλαδή μπορεί να είναι δυαδική, grayscale ή έγχρωμη. Παρακάτω θα διατυπωθούν κάποια βασικά στοιχεία για grayscale εικόνες, επειδή σε αυτή τη μορφή έγινε η εφαρμογή των μορφολογικών τελεστών στην παρούσα εργασία, αλλά ισοδύναμοι ορισμοί ισχύουν και για τα άλλα μοντέλα εικόνας. Η ακόλουθη παράγραφος έχει βασιστεί στο [36].

Βασικοί ορισμοί για grayscale εικόνες

Έστω μία εικόνα $f(x)$ ορισμένη στο συνεχές ή διακριτό επίπεδο $\mathbb{E} = \mathbb{R}^2$ or $\mathbb{Z}^2$. Η κατωφλίωση της f σε όλα τα επίπεδα πλάτους v παράγει ένα σύνολο (ensemble) από δυαδικές εικόνες που αντιπροσωπεύονται από τα άνω επιπεδοσύνολα (upper level sets):

$$X_v(f) \triangleq \{x \in \mathbb{E} : f(x) \geq v\}, \quad -\infty < v < +\infty. \quad (2.1)$$

Η εικόνα μπορεί να ανακατασκευαστεί ακριβώς από όλα τα επιπεδοσύνολά της ως εξής:

$$f(x) = \sup\{v \in \mathbb{R} : x \in X_v(f)\}. \quad (2.2)$$

Μετασχηματίζοντας τα επιπεδοσύνολα της f μέσω ενός τελεστή συνόλων Ψ και θεωρώντας τα μετασχηματισμένα επιπεδοσύνολα ως μία νέα εικόνα, ορίζεται ένας flat τελεστής εικόνας ψ που η έξοδος του είναι:

$$\psi(f)(x) = \sup\{v \in \mathbb{R} : x \in \psi[X_v(f)]\}. \quad (2.3)$$

Έτσι, όταν το Ψ είναι η διαστολή (*dilation*) και η συστολή (*erosion*) με το σύνολο B , η παραπάνω διαδικασία ορίζει τους ακόλουθους θεμελιώδεις στη μαθηματική μορφολογία τελεστές:

$$(f \oplus B)(x) \triangleq \bigvee_{y \in B} f(x - y) \quad (2.4)$$

$$(f \ominus B)(x) \triangleq \bigwedge_{y \in B} f(x + y) \quad (2.5)$$

όπου το $\bigvee$ συμβολίζει το supremum και το $\bigwedge$ το infimum. Εν γένει, η διαστολή μίας συνάρτησης f με ένα μικρό κυρτό σύνολο B αυξάνει τις «κοιλιάδες» και επεκτείνει τα μέγιστα της συνάρτησης, ενώ η συστολή μειώνει τις «κορυφές» και επεκτείνει τα ελάχιστα.

Ο flat τελεστής ανοίγματος (*opening*) της συνάρτησης f από το σύνολο B ορίζεται ως

$$f \circ B = (f \ominus B) \oplus B \quad (2.6)$$

και ο ρόλος του είναι να ομαλοποιήσει το γράφημα της f «από κάτω», αποκόβοντας τις κορυφές (peaks).

Ο flat τελεστής κλεισίματος (*closing*) αντίθετα ορίζεται ως

$$f \bullet B = (f \oplus B) \ominus B \quad (2.7)$$

και ο ρόλος του είναι να ομαλοποιήσει το γράφημα «από πάνω», γεμίζοντας τις κοιλιάδες (valleys).

2.3.2 Μερική ανίχνευση ελλείψεων με μορφολογικούς τελεστές

Στην παρούσα ενότητα γίνεται φανερό το πώς οι παραπάνω έννοιες αξιοποιούνται στην υπερεσία ενός αλγορίθμου αυτόματης ανίχνευσης των κενών. Ακολουθώντας, περιγράφονται το πρώτο μέρος του αλγορίθμου, ενώ στην επόμενη υποενότητα παρουσιάζεται το δεύτερο μέρος του που αφορά μία προσπάθεια βελτίωσης της εξαχθείσας μάρσας.

Το πρώτο βήμα είναι το διάβασμα της έγχρωμης εικόνας, τμήμα κάποιας τοιχογραφίας, όπου επιχειρείται να βρεθούν οι ελλείψεις. Για την επεξεργασία της μετασχηματίζεται σε grayscale μορφή $I(i, j)$, $1 \leq i \leq N, 1 \leq j \leq M$ και εκτελείται κατωφλιοποίηση στο πεδίο των παραγώγων της (image gradient domain). Για το σκοπό αυτό υπολογίζονται οι διακριτές παράγωγοι της έντασης της φωτεινότητας τόσο στην οριζόντια όσο και στην κάθετη διεύθυνση της εικόνας και αναζητούνται οι περιοχές που αυτές υπερβαίνουν ένα κατώφλι (threshold). Το κατώφλι υπολογίζεται με βάση την πληροφορία της εικόνας ως εξής:

$$t = k \frac{\sum_i \sum_j I(i, j)}{\sqrt{N * M}} \quad (2.8)$$

Μέσω πειραμάτων προέκυψε για τη σταθερά η τιμή $k = 6$ ως κατάλληλη.

Στη συνέχεια, λαμβάνεται η λογική διάζευξη των δύο δυαδικών εικόνων για τον έλεγχο ύπαρξης έντονων μεταβολών της έντασης της εικόνας σε τουλάχιστον μία από τις κύριες διευθύνσεις της. Στη δυαδική εικόνα που προκύπτει εφαρμόζεται μορφολογικό closing για την ομαδοποίηση πολύ γειτονικών περιοχών που πιθανότατα ανήκουν στα κενά και την ομαλοποίηση των συνόρων των περιοχών υψηλής μεταβολής της έντασης της φωτεινότητας. Κατ' αυτό τον τρόπο εξάγεται μία πρώτη εκδοχή της δυαδικής μάσκας που ταυτοποιεί τις ελλείπουσες περιοχές. Η μάσκα αυτή ορίζει τις περιοχές όπου θα εφαρμοστεί το inpainting κατά το επόμενο στάδιο της αποκατάστασης των τοιχογραφιών.

Ο απλός αυτός αλγόριθμος συνήθως δουλεύει ικανοποιητικά σε εικόνες με μικρά κενά ή περιοχές μεγάλου εμβαδού αλλά στενής ζώνης. Στα Σχήματα 2.2 και 2.3 παρουσιάζονται δύο παραδείγματα επιτυχούς εξαγωγής της μάσκας με βάση το πρώτο μόνο μέρος του αλγορίθμου.



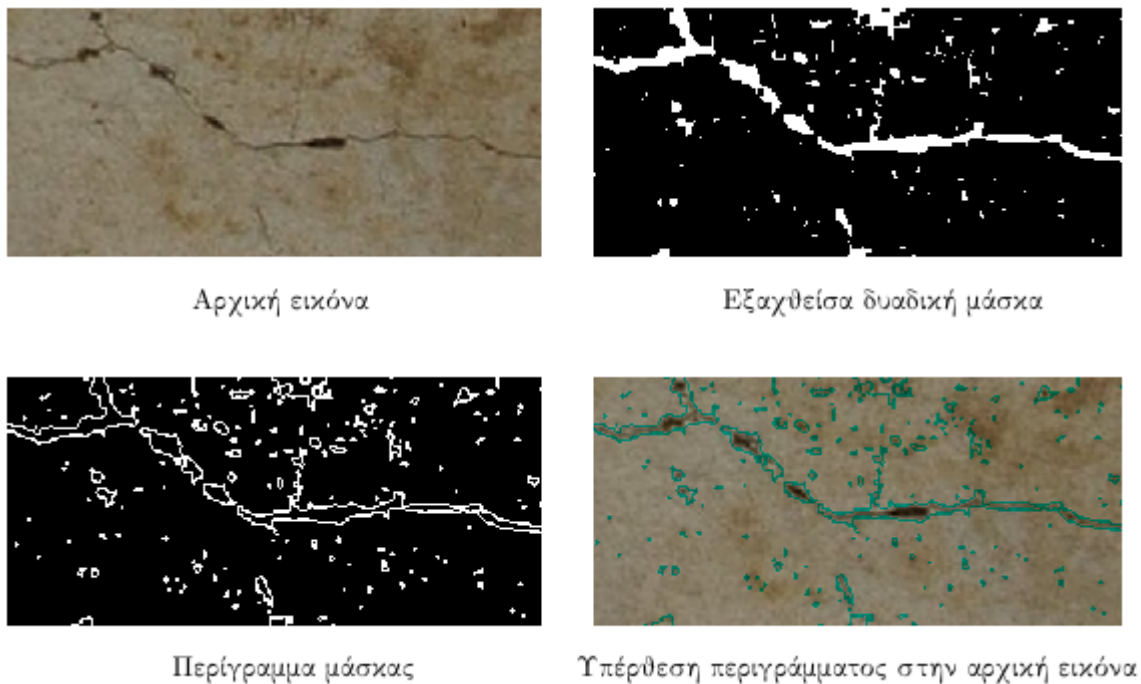
Σχήμα 2.2: Εξαγωγή δυαδικής μάσκας για μικρό τμήμα με ρωγμές της τοιχογραφίας *Ακολουθία Ανδρών* και υπέρθεση του περιγράμματος της μάσκας πάνω στο αυθεντικό έργο.

Στις εικόνες παρατίθεται και το περίγραμμα της δυαδικής μάσκας μετά την υπέρθεσή του πάνω στην αυθεντική εικόνα, ώστε να αξιολογείται ευκολότερα οπτικά το αποτέλεσμα. Το περίγραμμα υπολογίζεται και πάλι με χρήση μορφολογικών τελεστών ως η διαφορά της συστολής από τη διαστολή της δυαδικής μάσκας M με ένα σύνολο B που είναι ένας κυκλικός δίσκος:

$$outline(M) = (M \oplus B) - (M \ominus B) \quad (2.9)$$

Η ακτίνα του κυκλικού δίσκου εξαρτάται από τη μορφή και το μέγεθος των κενών και κυμαίνεται από $3 - 7 \text{ pixel}$ στα πειράματα. Στα παραδείγματα των Σχημάτων 2.2 και 2.3 λήφθηκε $r = 3$.

Ο *region-based* αυτός αλγόριθμος σε μερικές περιπτώσεις «χάνει» τμήματα των ακμών που εμφανίζονται κατά μήκος των ελλειπουσών περιοχών, όταν αυτές δεν είναι αρκετά έντονες σε σχέση με τη γειτονιά τους. Αυτό έχει ως συνέπεια την όχι ιδανική προσαρμογή της μάσκας στην ελλείπουσα περιοχή, άρα και αποτυχία της καλύτερης δυνατής ενδοσυμπλήρωσης αργότερα. Αυτό συμβαίνει γιατί pixel από τα κενά των τοιχογραφιών μένουν εκτός του περιγράμματος της μάσκας και άρα κατά το inpainting στέλνουν την πληροφορία τους εντός του περιγράμματος. Συνεπώς, διαδίδεται ένα θόλωμα μέσα στην inpainting περιοχή, κάτι που δρα αρνητικά στο αισθητικό αποτέλεσμα της ενδοσυμπληρωμένης εικόνας.



Σχήμα 2.3: Αυτόματη εύρεση της μάσκας ενός άλλου τμήματος της τοιχογραφίας *Ακολουθία Ανδρών* και υπέρθεση της πάνω σε αυτό.

Για να αποφευχθεί αυτό το φαινόμενο είναι επιθυμητό η συντριπτική πλειοψηφία των pixel ακμών να βρίσκεται εντός της περιοχής ενδοσυμπλήρωσης. Έτσι, καθίσταται επιτακτική η ανάγκη για το *boundary based* κομμάτι του αλγορίθμου. Στην επόμενη υποενότητα περιγράφεται με ποιο τρόπο βελτιώνεται ο μηχανισμός εξαγωγής της μάσκας.

2.3.3 Βελτίωση εξαχθείσας μάσκας με την προσάρτηση των ακμών

Για την ανίχνευση των ακμών της εικόνας χρησιμοποιείται η μέθοδος του *Canny* [37]. Ακολουθώντας με την εφαρμογή ενός τελεστή μορφολογικής διαστολής (*dilation*) επιτυγχάνεται η ομογενοποίηση των περιοχών της μάσκας που περιλαμβάνουν πολύ γειτονικές υποπεριοχές ακμών και η εξάλειψη των πολύ μικρών χωρίων μεταξύ των ακμών που δεν περιέχουν «καθαρή» την πληροφορία της εικόνας και άρα περισσότερο βλάπτουν παρά υποβοηθούν στο στάδιο της ενδοσυμπλήρωσης. Ως δομικό στοιχείο για τη διαστολή χρησιμοποιείται κυκλικός δίσκος μεταβλητής ακτίνας ανάλογα την εικόνα. Συνηθισμένες τιμές στα πειράματα ανήκουν στο διάστημα 3 – 5 pixel. Στη συνέχεια, λαμβάνεται η λογική διάζευξη της δυαδικής εικόνας με τις πεπλατυσμένες ακμές και της δυαδικής μάσκας που έχει προκύψει από το πρώτο στάδιο του αλγορίθμου, ώστε προκύπτει μία «πεπλατυσμένη» εκδοχή της τελικής μάσκας.

Το τελευταίο βήμα αφορά την εφαρμογή ενός τελεστή μορφολογικής συστολής (*erosion*²), κάτι που συνεισφέρει στην εξάλειψη των περιοχών χωρίς αξιόλογη μεταβολή της έντασης της

²Χρησιμοποιείται μοναδιαίος κυκλικός δίσκος ως δομικό στοιχείο (*structural element*) για την όσο δυνατόν πιο λεπτομερή προσαρμογή της μάσκας πάνω στις ακμές. Εξάλλου, το συγκεκριμένο βήμα εκτελείται επαναληπτικά έως ότου «χαθεί» το πολύ το 0.5% της πληροφορίας των ακμών.

φωτεινότητας και στη συρρίκνωση της μάσκας, ώστε να προσαρμόζεται στον καλύτερο δυνατο βαθμό στις ελλείπουσες περιοχές, για την αποδοτικότερη μετέπειτα εφαρμογή του inpainting. Το βήμα αυτό εκτελείται επαναληπτικά έως ότου η μάσκα απολέσει όχι περισσότερο από το 0.5% της πληροφορίας των ανιχνευθείσων ακμών. Κατά αυτό τον τρόπο προκύπτει η τελική, αρκετά βελτιωμένη εκδοχή της μάσκας. Ακολουθεί η σχηματική περιγραφή των κύριων σημείων του αλγορίθμου.

Αλγόριθμος 1: Αυτόματη εύρεση ελλείψεων με χρήση στοιχείων *Μαθηματικής Μορφολογίας* και ανίχνευση ακμών (region and boundary based approach).

Αξιοποίηση πληροφορίας παραγώγων (region based)

- Έστω η εικόνα $I(i, j)$, $1 \leq i \leq N$, $1 \leq j \leq M$.
- Υπολογίζονται οι διακριτές παράγωγοι στους δύο κύριους άξονές της και βρίσκονται οι περιοχές με παράγωγο πάνω από το κατώφλι:

$$t = k \frac{\sum_i \sum_j I(i, j)}{\sqrt{N * M}}$$

- Εφαρμόζεται closing στη λογική διάζευξη των δύο δυαδικών εικόνων.

Αξιοποίηση πληροφορίας ακμών (boundary based)

- Με τη μέθοδο του Canny ανιχνεύονται οι κύριες ακμές της εικόνας.
 - Εφαρμόζεται dilation στη δυαδική εικόνα των ακμών με κυκλικό δίσκο ακτίνας 5 pixel.
 - Γίνεται διάζευξη των δυαδικών εικόνων των ακμών και της αρχικής μάσκας.
 - Εφαρμόζεται erosion με μοναδιαίο δίσκο επαναληπτικά και διακόπτεται πριν η μάσκα απολέσει πάνω από το 0.5% της επιφάνειας έντονων ακμών της.
-

2.3.4 Εφαρμογή αλγορίθμου μορφολογίας και ανίχνευσης ακμών στις τοιχογραφίες

Στα Σχήματα 2.4 - 2.6 που ακολουθούν παρουσιάζονται αποτελέσματα εφαρμογής του αλγορίθμου στην εξαγωγή μάσκας σε διάφορες περιοχές των τοιχογραφιών με κενά σχετικά μικρού εμβαδού και χωρίς πολλές «παραπλανητικές» για τον αλγόριθμο ακμές.

Όπως είναι φανερό η μάσκα μετά τη βελτιστοποίηση έχει τρία κύρια πλεονεκτήματα σε σχέση με την αρχική μάσκα. Καταρχάς, περιέχει τις ακμές ολόκληρες και δεν είναι οι τελευταίες μέρος του συνόρου της. Συνεπώς, δεν υπάρχει κίνδυνος κάποια pixel των ακμών να μείνουν εκτός περιγράμματος μάσκας και άρα να «μολύνουν» το κενό κατά το inpainting αργότερα με ένταση άσχετη με την παράσταση της τοιχογραφίας. Κατά δεύτερον, με μία γρήγορη επισκόπηση διακρίνεται ότι η νέα μάσκα εντοπίζει περισσότερα φθαρμένα σημεία στην τοιχογραφία. Τέλος, μετά το δεύτερο στάδιο του αλγορίθμου αποβάλλεται το περίγραμμα που συχνά κατά το πρώτο στάδιο εκλαμβάνεται ως ρωγμή. Αυτό συμβαίνει γιατί η μέθοδος του Canny αποφεύγει να εκλάβει το περίγραμμα της εικόνας ως ακμή και άρα κατά το στάδιο του επαναληπτικού erosion, καθώς το περίγραμμα δεν έχει τομή με pixel ακμών στη γενική περίπτωση, δεν επιβαρύνει τη συνθήκη του να μη απολεσθεί πάνω από το 0.5% της πληροφορίας των ακμών και άρα εξαλείφεται.

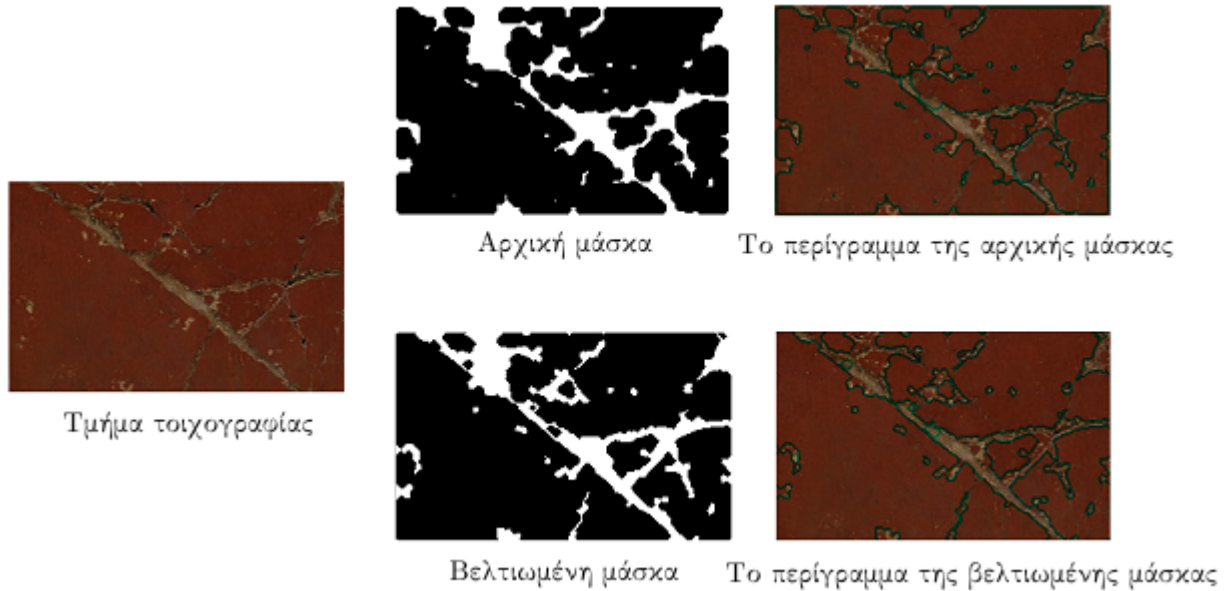
Για την ποσοτική αξιολόγηση της εξαχθείσας μάσκας με τον κάθε αλγόριθμο χρησιμοποιούνται ως δείκτες απόδοσης το *precision* που είναι το ποσοστό των ανιχνευμένων στιγμιοτύπων που είναι σχετικά και το *recall* που είναι το ποσοστό των σχετικών στιγμιοτύπων που ανιχνεύονται. Ως γενικό μέτρο ποιότητας χρησιμοποιείται το ημίθροισμα των δύο δεικτών q :

$$precision = \frac{hit}{detected} \quad (2.10)$$

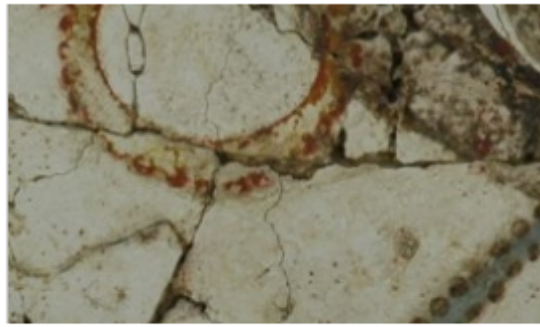
$$recall = \frac{hit}{right} \quad (2.11)$$

$$q = \frac{precision + recall}{2} \quad (2.12)$$

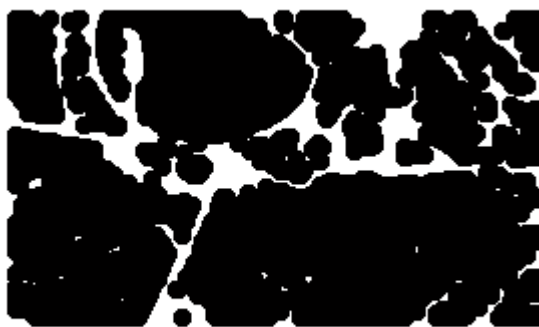
Για παράδειγμα, στο Σχήμα 2.4 η μάσκα μαθηματικής μορφολογίας έχει τιμές $precision = 0.5332$, $recall = 0.7517$ και $q = 0.6425$, ενώ η μάσκα που συνδυάζει και την πληροφορία των ακμών εμφανίζει τις βελτιωμένες τιμές $precision_{opt} = 0.6040$, $recall_{opt} = 0.8210$ και $q_{opt} = 0.7125$.



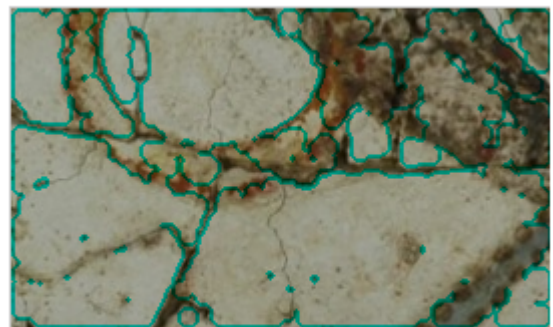
Σχήμα 2.4: Εξαγωγή μάσκας σε περιοχή της πλάτης του μεγαλύτερου άνδρα στην τοιχογραφία *Ακολουθία Ανδρών* και υπέρθεση του περιγράμματος στο αυθεντικό έργο. Στη βελτιωμένη μάσκα το περίγραμμα έχει εξαλειφθεί, ενώ έχουν εντοπιστεί περισσότερες φθορές.



Τμήμα της τοιχογραφίας που απεικονίζει το σκουλαρίκι της Πότνιας



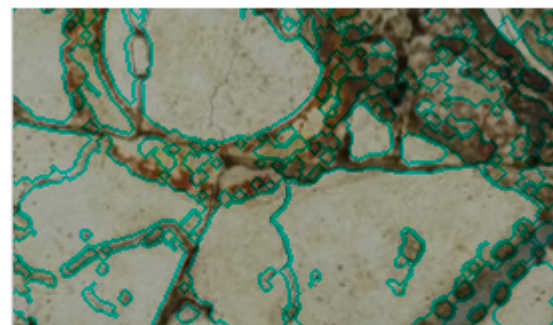
Αρχική μάσκα



Το περίγραμμα της αρχικής μάσκας

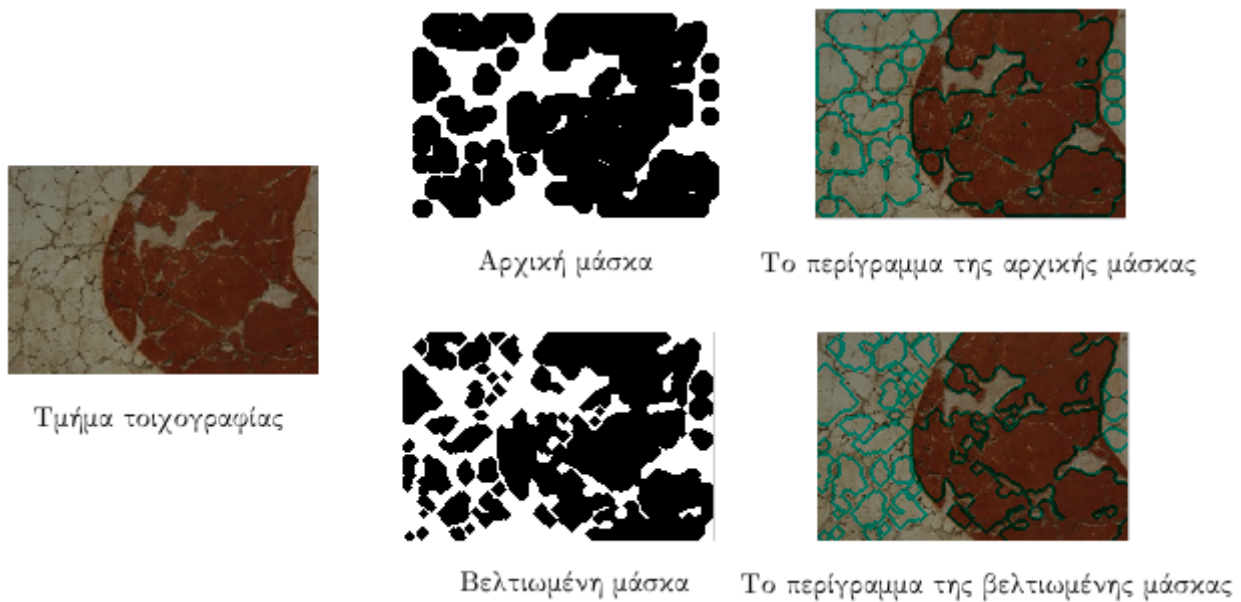


Βελτιωμένη μάσκα (μετά τη βέβαιη προσάρτηση τουλάχιστον του 99.5% της πληροφορίας των ακμών)



Το περίγραμμα της βελτιωμένης μάσκας

Σχήμα 2.5: Οι εξαχθείσες μάσκες σε αυτή την περίπτωση αφορούν τμήμα της τοιχογραφίας *Κροκοσυλλέκτρια & Πότνια*, όπου απεικονίζεται το σκουλαρίκι της Πότνιας. Στη μάσκα που αξιοποιείται και η πληροφορία των ακμών αποσοβείται η εμφάνιση περιγράμματος και γενικότερα ο εντοπισμός των ελλείψεων είναι πληρέστερος. Αναπόφευκτα ανιχνεύονται και οι έντονες ακμές που είναι μέρος της παράστασης της τοιχογραφίας, κάτι που θα οδηγήσει σε όχι ιδανικό inpainting, όπως θα δειχθεί στο 3^ο Κεφάλαιο με το Σχήμα 3.5.



Σχήμα 2.6: Εξαγωγή μάσκας σε περιοχή του σώματος του μεγαλύτερου άνδρα στην τοιχογραφία *Ακολουθία Ανδρών* και υπέρθεση του περιγράμματος στο αυθεντικό έργο. Όπως διακρίνεται στη βελτιωμένη μάσκα το περίγραμμα έχει εξαλειφθεί, οι ρωγμές που δε βρέθηκαν αρχικά έχουν εντοπιστεί, ενώ τέλος οι ελλείψεις περικλείονται πιο «ανοικτά» από τη λευκή περιοχή της μάσκας, κάτι που αν και αφελώς μπορεί να ερμηνευτεί ως λιγότερο καλή προσαρμογή, στην πραγματικότητα επιτρέπει πολύ πιο ασφαλές και «καθαρό» inpainting αργότερα.

2.4 Κατάτμηση εικόνας με τομές γράφων

Σε αρκετές περιοχές των τοιχογραφιών τα κενά είναι σημαντικής έκτασης. Σε αυτές τις περιπτώσεις η μαθηματική μορφολογία δε μπορεί να βοηθήσει στην ανίχνευσή τους, γιατί ακόμη και πολύ μεγάλο δομικό στοιχείο να χρησιμοποιηθεί ώστε να ενσωματωθεί όλο το κενό στην εξαχθείσα μάσκα, η αναγνώριση της inpainting περιοχής γίνεται σε αδρές γραμμές, χωρίς μεγάλη ακρίβεια. Το μέγεθος των κενών, μαζί με άλλες δυσκολίες που έχουν αναφερθεί στην εισαγωγή, όπως η μεγάλη ποικιλία στη γεωμετρία των ελλειπουσών περιοχών και η χρωματική ομοιότητα κάποιες φορές μεταξύ των κενών και της αυθεντικής τοιχογραφίας, καθιστά κάποιες φορές απαραίτητη την παρέμβαση από τη μεριά του χρήστη στη διαδικασία της αποκατάστασης.

Παρόλ' αυτά, ο ακριβής προσδιορισμός της περιοχής των κενών είναι χρονοβόρα διαδικασία ακόμη και για τον εξειδικευμένο τεχνίτη και η ανάγκη αυτοματοποίησής της είναι επιτακτική. Σε αυτό το πρόβλημα τη λύση έρχεται να δώσει η μεθοδολογία της κατάτμησης με τομές γράφων. Μέσω αυτών είναι δυνατός ο προσδιορισμός των ελλείψεων με μεγάλη ακρίβεια, ενώ από τη μεριά του χρήστη απαιτείται μόνο ένας προσεγγιστικός προσδιορισμός του παρασκηνίου και του προσκηνίου των εξαχθεισών περιοχών. Μάλιστα ο επαναληπτικός αλγόριθμος "GrabCut" που πρότειναν οι Rother, Kolmogorov και Blake το 2004 [3], ο οποίος έχει στον πυρήνα του την επεξεργασία με τομές γράφων, είναι ακόμη πιο αποτελεσματικός και με λιγότερη απαιτούμενη εργασία από το χρήστη. Το μόνο που χρειάζεται από την πλευρά του είναι δύο σημεία στο επίπεδο, τα οποία ορίζουν ένα ορθογώνιο που ταυτοποιεί το φόντο του αντικειμένου (εν προκειμένω το

ελλείπον τμήμα) που επιθυμεί ο χρήστης να εξάγει. Ακολουθεί περιγραφή του αλγορίθμου και στη συνέχεια παραδείγματα εφαρμογής του στις τοιχογραφίες. Τέλος, προτείνεται και ένας αλγόριθμος βελτίωσης της μάσκας που έχει βρεθεί με στοιχεία μαθηματικής μορφολογίας.

2.4.1 Διαδραστική κατάτμηση με χρήση επαναληπτικών γραφοτομών (“GrabCut”)

Η βάση του αλγορίθμου “GrabCut” είναι η μέθοδος κατάτμησης με γραφο-τομές των Boykov και Jolly [35] που περιγράφεται στην ακόλουθη παράγραφο. Η κατάτμησή τους έχει τα καλύτερα αποτελέσματα σε σχέση με κάθε πρότερη προσέγγιση, όσον αφορά τις ιδιότητες του συνόρου και των υποπεριοχών, ζητώντας από το χρήστη να ορίσει απλώς το προσκήνιο και το παρασκήνιο με κάποιες γραμμές. Στα πλεονεκτήματά της μεθόδου τους είναι ότι δεν υπάρχουν περιορισμοί για την τοπολογία της κατάτμησης ενώ τόσο το παρασκήνιο όσο και το προσκήνιο μπορούν να αποτελούνται από πολλές συνεκτικές συνιστώσες.

Κατάτμηση εικόνας με χρήση τομών γράφων

Οι Boykov και Jolly το 2001 χρησιμοποίησαν τομές γράφων για την επίλυση του προβλήματος της κατάτμησης, δεδομένου ότι ο χρήστης παρέχει αρχικά το προσκήνιο και το παρασκήνιο του χάρτη T . Η εικόνα είναι μία διάταξη από N pixels με εντάσεις φωτεινότητας (gray values) $\mathbf{z} = (z_1, \dots, z_n, \dots, z_N)$, ενώ η κατάτμηση της εικόνας εκφράζεται μέσω μίας διάταξης σταθμών (opacity values) $\underline{\alpha} = (\alpha_1, \dots, \alpha_N)$. Γενικά $0 \leq \alpha_n \leq 1$, αλλά για δυαδική κατάτμηση, όπως είναι η εξαγωγή προσκηνίου, απαιτείται $\alpha_n \in \{0, 1\}$, όπου 0 συμβολίζει το παρασκήνιο και 1 το προσκήνιο. Οι παράμετροι $\underline{\theta}$ περιγράφουν τις κατανομές των εντάσεων της εικόνας στο προσκήνιο και στο παρασκήνιο και αναπαριστώνται από ιστογράμματα εντάσεων $\underline{\theta} = \{h(z; \alpha), \alpha = 0, 1\}$, τα οποία κατασκευάζονται κατευθείαν με βάση τις ετικέτες των pixels από το χάρτη T . Η κατάτμηση συνίσταται στο συμπερασμό των αγνώστων μεταβλητών των σταθμών $\underline{\alpha}$, δεδομένων των τιμών έντασης της εικόνας $\mathbf{z}$ και του μοντέλου $\underline{\theta}$.

Στη συνέχεια, ορίζεται μία συνάρτηση ενέργειας έτσι ώστε η ελαχιστοποίησή της να αντιστοιχεί σε μία καλή κατάτμηση της εικόνας. Αυτό σημαίνει ότι η κατάτμηση συμφωνεί με τα παρατηρούμενα ιστογράμματα του προσκηνίου και του παρασκηνίου και η διαδοχή των σταθμών έχει συνοχή, διατηρώντας παράλληλα μία τάση για συνεκτικότητα των αντικειμένων. Έτσι, ορίζεται μία ενέργεια *Gibbs* της μορφής:

$$\mathbf{E}(\underline{\alpha}, \underline{\theta}, \mathbf{z}) = U(\underline{\alpha}, \underline{\theta}, \mathbf{z}) + V(\underline{\alpha}, \mathbf{z}). \quad (2.13)$$

Ο όρος U αξιολογεί το ταίριασμα της κατανομής σταθμών $\underline{\alpha}$ στα δεδομένα $\mathbf{z}$, δεδομένου του μοντέλου ιστογραμμάτων $\underline{\theta}$ και ορίζεται ως:

$$U(\underline{\alpha}, \underline{\theta}, \mathbf{z}) = \sum_n -\log h(z_n; \alpha_n). \quad (2.14)$$

Ο όρος ομαλότητας γράφεται ως

$$V(\underline{\alpha}, \mathbf{z}) = \gamma \sum_{(m,n) \in \mathbf{C}} \text{dis}(m,n)^{-1} [\alpha_n \neq \alpha_m] \exp -\beta(z_m - z_n)^2, \quad (2.15)$$

όπου $\mathbf{C}$ είναι το σύνολο των γειτονικών ζευγών pixels, $\text{dis}(\cdot)$ είναι η ευκλείδεια απόσταση και $[\phi]$ η δείτρια συνάρτηση του κατηγορήματος ϕ :

$$[\phi] = \begin{cases} 1, & \text{αν } \phi \text{ αληθές} \\ 0, & \text{αν } \phi \text{ ψευδές} \end{cases} \quad (2.16)$$

Σημαντικός είναι ο ρόλος της σταθεράς β στην Εξίσωση 2.15. Όταν $\beta = 0$ ο όρος ομαλότητας είναι η γνωστή πρότερη πιθανότητα Ising (*Ising prior*³) και ενθαρρύνει την ομαλότητα παντού στο βαθμό που προσδιορίζεται από τη σταθερά γ . Οι Boykov και Jolly [35] δείξαν ότι είναι πολύ πιο αποτελεσματικό να είναι $\beta > 0$, έτσι ώστε να «χαλαρώνει» η απαίτηση για ομαλότητα στις περιοχές μεγάλης αντίθεσης (*contrast*). Στα πειράματά τους επέλεξαν τη σταθερά β ως εξής:

$$\beta = \frac{1}{2\langle (z_m - z_n)^2 \rangle}, \quad (2.17)$$

όπου το $\langle \cdot \rangle$ προσδιορίζει την αναμενόμενη τιμή.

Έχοντας ορίσει αναλυτικά το μοντέλο ενέργειας και όλες τις παραμέτρους του, η επιθυμητή κατάτμηση προκύπτει μέσω της εύρεσης του ολικού ελαχίστου:

$$\hat{\underline{\alpha}} = \underset{\underline{\alpha}}{\operatorname{argmin}} \mathbf{E}(\underline{\alpha}, \underline{\theta}, \mathbf{z}). \quad (2.18)$$

Η ελαχιστοποίηση γίνεται μέσω ενός κλασικού αλγορίθμου εύρεσης ελάχιστης τομής σε γράφο [35]. Ο αλγόριθμος αυτός αποτελεί θεμέλιο για τη δυαδική κατάτμηση, όπως πολύ βελτιωμένα επιτυγχάνεται μέσω του αλγορίθμου “GrabCut”, τον οποίον πρότειναν οι Rother, Kolmogorov και Blake το 2004 [3] και παρουσιάζεται στην ακόλουθη παράγραφο.

Ο αλγόριθμος “GrabCut”

Οι κύριες καινοτομίες των Rother et al. ως προς τον κλασικό αλγόριθμο κατάτμησης με τομές γράφων των Boykov et al. συνοψίζονται στα ακόλουθα σημεία:

⇒ *Έγχρωμο μοντέλο*

Το μονόχρωμο μοντέλο εικόνας έχει αντικατασταθεί από έγχρωμο με τη χρήση *Gaussian Mixture Model* αντί ιστογραμμάτων.

⇒ *Χρήση επαναληπτικού σχήματος*

Ο αλγόριθμος μίας επανάληψης για την εκτίμηση της ελάχιστης τομής αντικαθίσταται από έναν πιο αποτελεσματικό, επαναληπτικό αλγόριθμο που εκτελεί εναλλάξ εκτίμηση και μάθηση παραμέτρων.

⇒ *Ελαχιστοποίηση εργασίας χρήστη*

Πλέον είναι εφικτό το να προσδιοριστεί από το χρήστη μόνο το φόντο με ένα ορθογώνιο περίβλημα που πρακτικά είναι δύο μόνο σημεία στο επίπεδο της εικόνας (*incomplete labelling*).

Στη συνέχεια σκιαγραφείται σχηματικά ο αλγόριθμος. Για μεγαλύτερη διαδραστικότητα, με κατεύθυνση την περαιτέρω βελτίωση της κατάτμησης, αν ο χρήστης δεν είναι ικανοποιημένος από το τρέχον αποτέλεσμα, μπορεί να ορίσει ρητά την περιοχή που ανήκουν κάποια εσφαλμένα ταξινομημένα pixel και να επανεκτελέσει το τρίτο βήμα του “GrabCut” που αφορά την εκτίμηση κατάτμησης με χρήση ελαχίστων τομών. Ένα παράδειγμα αυτής της αλληλεπίδρασης χρήστη-αλγορίθμου δίνεται στο Σχήμα 2.7.

³Όπως αναφέρθηκε και στο 1^ο κεφάλαιο στην παράγραφο για τα *Τυχαία πεδία*, οι τοπικοί περιορισμοί ενέργειας συχνά εξυπηρετεί τις εφαρμογές να επιβάλλονται με τον ίδιο τρόπο όπως στους κρυστάλλους Ising, μία ιδέα που αντλήθηκε από τα πεδία Gibbs στη Στατιστική Μηχανική.

Αλγόριθμος 2: Κατάτμηση εικόνας με τον αλγόριθμο “GrabCut”.

Αρχικοποίηση

- Ο χρήστης αρχικοποιεί το χάρτη $T = \{T_B, T_U, T_F\}$ προσδιορίζοντας μόνο το T_B . Οι άλλες περιοχές ορίζονται αυτόματα ως $T_F = \emptyset$ και $T_U = \overline{T_B}$.
- Αρχικοποιούνται $a_n = 0$ για $n \in T_B$ και $a_n = 1$ για $n \in T_U$.
- Τα GMMs παρασκηνίου και προσκηνίου αρχικοποιούνται με βάση τα σύνολα $a_n = 0$ και $a_n = 1$ αντίστοιχα.

Επαναληπτικό σχήμα ελαχιστοποίησης

1. Ανάθεση GMM συνιστωσών στα pixel. Για κάθε $n \in T_U$:

$$k_n := \operatorname{argmin}_{k_n} D_n(\alpha_n, k_n, \theta, z_n).$$

2. Εκπαίδευση GMM από τα δεδομένα $\mathbf{z}$:

$$\underline{\theta} := \operatorname{argmin}_{\underline{\theta}} U(\underline{\alpha}, \mathbf{k}, \underline{\theta}, \mathbf{z}).$$

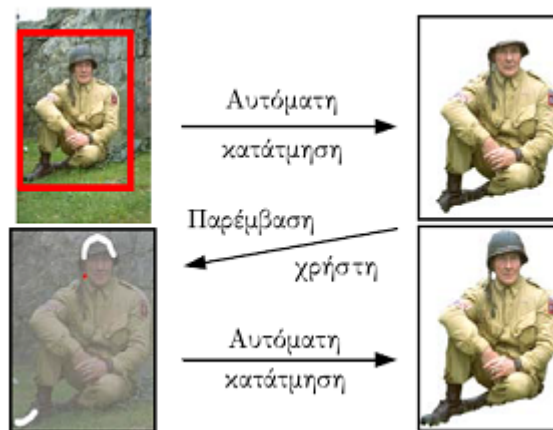
3. Εκτίμηση κατάτμησης με χρήση ελαχίστων τομών:

$$\min_{\{\alpha_n: n \in T_U\}} \min_{\mathbf{k}} \mathbf{E}(\underline{\alpha}, \mathbf{k}, \underline{\theta}, \mathbf{z}).$$

4. Επιστροφή στο Βήμα 1 και επανεκτέλεση μέχρι τη σύγκλιση.

Πρόσθετη επεξεργασία χρήστη

Ο χρήστης ορίζει ρητά την ορθή περιοχή των pixel που ταξινομήθηκαν λάθος μετά την αρχική εκτέλεση του επαναληπτικού σχήματος. Ανανεώνεται ο χάρτης T και εκτελείται το βήμα 3 μία μόνο φορά ή γίνεται επανεκτέλεση του σχήματος για περαιτέρω βελτίωση.



Σχήμα 2.7: Ο χρήστης παρεμβαίνει και ορίζει ρητά (explicitly) τη θέση κάποιων pixel που έχουν ταξινομηθεί σε λάθος περιοχή. Από το [3], στο οποίο έχει βασιστεί η Υποενότητα .

2.4.2 Αξιοποίηση “GrabCut” για την περαιτέρω βελτίωση του μηχανισμού εύρεσης των ελλειπόντων τμημάτων

Ακολούθως, προτείνεται ένας αλγόριθμος που αξιοποιεί τον αλγόριθμο “GrabCut” για τη βελτίωση της εξαχθείσας μάσκας με χρήση μορφολογικών τελεστών και πληροφορίας ακμών από το πρώτο μέρος.

Αλγόριθμος 3: Αξιοποίηση “GrabCut” ώστε η μάσκα που έχει εξαχθεί με μαθηματική μορφολογία και ανίχνευση ακμών να βρίσκει ακριβώς τις ελλείψεις.

- Διαβάζεται η μάσκα M που έχει εξαχθεί με μαθηματική μορφολογία και ανίχνευση ακμών.
- Έστω B κυκλικός δίσκος. Αρχικοποιείται ο χάρτης T ως εξής:

$$T = \{T_B, T_U, T_F\}, \quad T_B = (M \oplus B)^c, \quad T_F = M \ominus B, \quad T_U = (T_B \cup T_F)^c$$

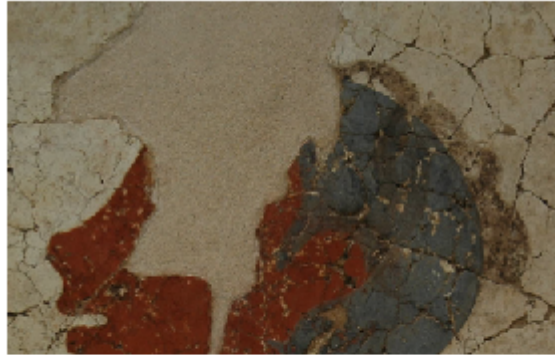
- Εκτελείται ο αλγόριθμος “GrabCut” με είσοδο το χάρτη T και την ελλείπουσα εικόνα.
-

Ο αλγόριθμος “GrabCut” λόγω της επαναληπτικής φύσης του είναι πολύ αποτελεσματικός στην ακριβή αναγνώριση των συνόρων ενός αντικειμένου, αλλά απαιτεί μία μικρή συνεισφορά από το χρήστη, ο οποίος πρέπει να δώσει μία πρόχειρη εκτίμηση του background και ενδεχομένως και του foreground αρχικά. Εκεί βασίζεται η ιδέα του παρόντος αλγορίθμου. Η μάσκα που έχει εξαχθεί με μαθηματική μορφολογία και ανίχνευση ακμών δίνει στην πλειοψηφία των περιπτώσεων μία καλή εκτίμηση της ελλείπουσας περιοχής και το “GrabCut” μπορεί να δώσει ένα σχεδόν τέλειο αποτέλεσμα κατάτμησης.

Το συμπλήρωμα ως προς το πεδίο της εικόνας Ω του dilation της αρχικής μάσκας αποτελεί την περιοχή που με μεγάλη πιθανότητα δεν περιέχει ελλείψεις, ενώ το erosion της αρχικής μάσκας είναι το χωρίο που εκτιμάται ότι ανήκει μέσα στις υπό αναγνώριση ελλείπουσες περιοχές. Η ακτίνα του κυκλικού δίσκου B που θα χρησιμοποιηθεί σχετίζεται με το βαθμό αξιοπιστίας της αρχικής μάσκας. Εκφράζει ουσιαστικά ένα περιθώριο λάθους. Η ακτίνα του δίσκου πρέπει να είναι τουλάχιστον όσο η μέγιστη απόκλιση συνόρου ελλείποντος τμήματος από την πραγματική του θέση και μπορεί να προσδιοριστεί ακριβώς μετά από σχετικό πειραματισμό.

2.4.3 Εφαρμογή του “GrabCut” στις τοιχογραφίες

Ακολουθούν παραδείγματα χρήσης του “GrabCut” στην εύρεση ελλείψεων στις τοιχογραφίες.



Εικόνα με έλλειψη που πρέπει να προσδιοριστεί επ’ ακριβώς



Η μάσκα που προκύπτει από το GrabCut

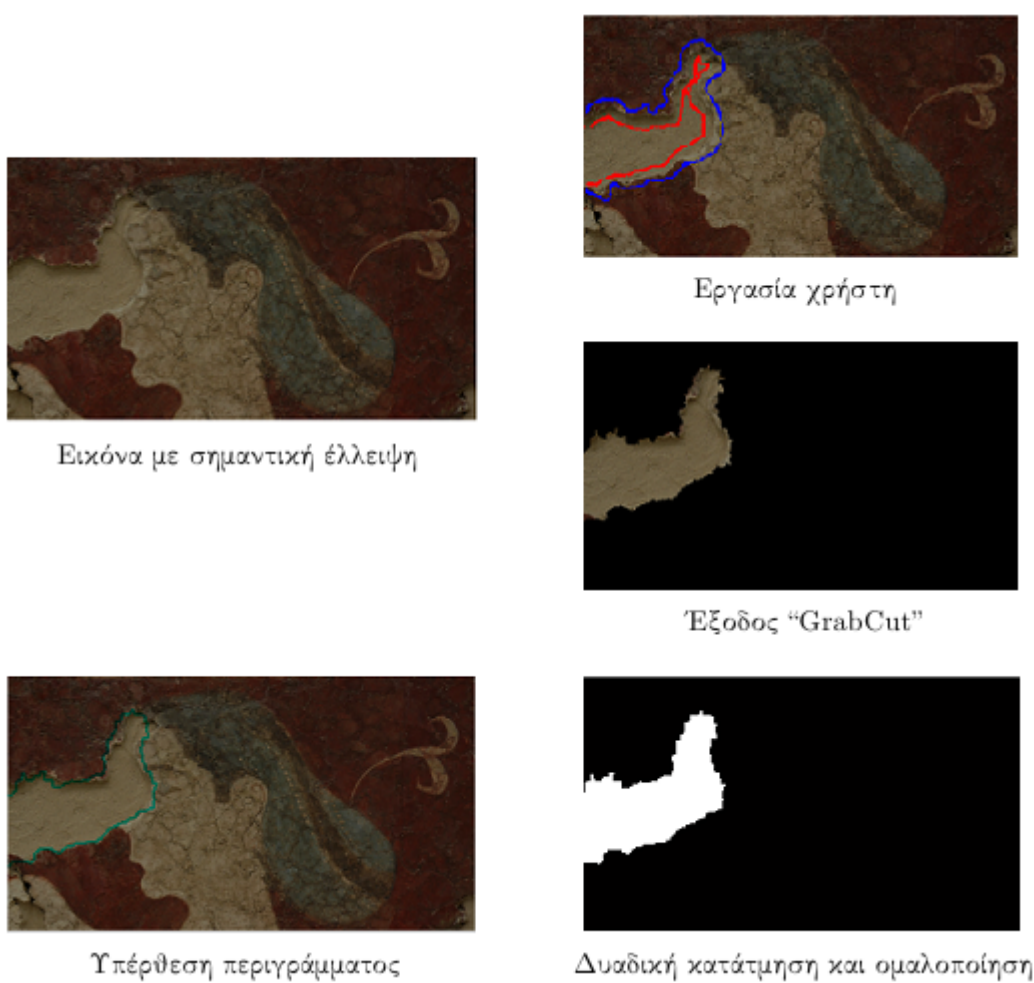


Ομαλοποίησή της με dilation



Υπέρθεσή της επί του τμήματος της τοιχογραφίας

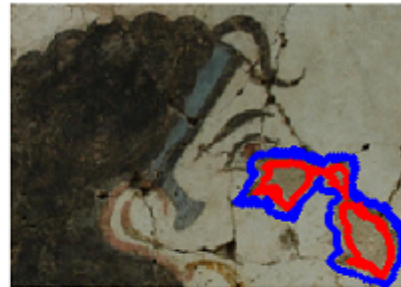
Σχήμα 2.8: Επ’ ακριβώς εύρεση έλλειψης με το “GrabCut” σε περιοχή της τοιχογραφίας *Ακολουθία Ανδρών* που απεικονίζει το κεφάλι του μεγαλύτερου άνδρα. Με μία πρόχειρη περιγραφή του background και του foreground, ως είσοδο του αλγορίθμου, αυτός επιστρέφει μία πολύ ικανοποιητική ταυτοποίηση του ελλείποντος τμήματος. Σε συνδυασμό με την αξιοποίηση του αλγορίθμου μαθηματικής μορφολογίας η αναγνώριση μπορεί να γίνει με ακρίβεια.



Σχήμα 2.9: Ανίχνευση της σημαντικής έλλειψης στην κεφαλή της γυναίκας στην τοιχογραφία *Ακολουθία Γυναικών*.



Εικόνα με σημαντική έλλειψη



Εργασία χρήστη



Έξοδος "GrabCut"



Υπέρθεση περιγράμματος



Διαδική κατάτμηση και ομαλοποίηση

Σχήμα 2.10: Ανίχνευση των ευμεγεθών ελλείψεων στην κροκοσυλλέκτρια από την τοιχογραφία *Κροκοσυλλέκτρια & Πότνια*.

Κεφάλαιο 3

Αποκατάσταση των τοιχογραφιών με ενδοσυμπλήρωση (inpainting)

3.1 Εισαγωγή

Έχουν προταθεί διάφοροι αλγόριθμοι ενδοσυμπλήρωσης κατά καιρούς στο ερευνητικό πεδίο της Όρασης Υπολογιστών. Μία κατηγορία βασίζεται σε γεωμετρική μοντελοποίηση της εικόνας στο χωρικό πεδίο όπως οι μέθοδοι που χρησιμοποιούν μερικές διαφορικές εξισώσεις (Partial Differential Equations, PDEs), όπως η Poisson [38], με συνοριακές συνθήκες Dirichlet ή Neumann για την εξασφάλιση της συνέχειας. Άλλες τεχνικές βασίζονται στη διάδοση των *ισόφωτων* διευθύνσεων της εικόνας, δηλαδή καμπυλών ίσης φωτεινότητας, μέσα στην inpainting περιοχή [23].

Πολλές μέθοδοι σε υπολογιστή που απαιτούν την ανθρώπινη παρέμβαση προσφέρουν τη χρήση ενός εργαλείου δημιουργίας *κλώνων* για την αντιγραφή υπαρχόντων τμημάτων της εικόνας και την ανακατασκευή κατεστραμμένης υψής σε περιοχές της εικόνας. *Σύνθεση υψής* μπορεί επίσης να χρησιμοποιηθεί, όπως ο κλασικός αλγόριθμος Efros-Leung [39] που θα παρουσιαστεί στη συνέχεια του κεφαλαίου. Προσπάθειες ενδοσυμπλήρωσης βάση υποδείγματος (*exemplar-based inpainting*) για την αυτοματοποίηση της διαδικασίας κλωνοποίησης τμημάτων έχουν πρόσφατα διερευνηθεί [40]. Τα κενά της εικόνας «γεμίζονται» αναζητώντας παρόμοια τμήματα (patches) σε μία κοντινή περιοχή της εικόνας και αντιγράφοντας τα pixels του πιο σχετικού patch στο κενό. Το ότι δε γίνεται η συμπλήρωση σε επίπεδο pixel οδηγεί σε μειωμένες παρενέργειες και ατέλειες που πηγάζουν από τη διαδικασία της μεταφοράς τμήματος και έχουν παρουσιαστεί σε πρότερες τεχνικές.

Μια άλλη θεμελιώδης κατηγορία inpainting τεχνικών στηρίζεται στη *στατιστική* μοντελοποίηση στο πεδίο γραμμικών μετασχηματισμών της εικόνας. Το πρόβλημα της ενδοσυμπλήρωσης είναι εξ ορισμού κακώς ορισμένο (*ill-posed* κατά Hadamard) και μία πιθανοτική προσέγγιση βάση μοντέλου συνεισφέρει σημαντικά στην επίλυσή του. Αναπαραστάσεις της εικόνας τόσο στο πεδίο του μετασχηματισμού Fourier όσο και στο πεδίο χωρικά εντοπισμένων κυματιδιακών μετασχ-

ματισμών πλεονεκτούν για την περιγραφή μη στάσιμων φυσικών εικόνων, ενώ έχουν αποδειχθεί ιδιαίτερα αποδοτικοί σε εφαρμογές συμπίεσης. Γενικά, τα κυματίδια χώρου κλίμακας είναι κατάλληλα για αναπαράσταση τμηματικά ομαλών εικόνων, ενώ τα κυματίδια χώρου συχνότητας ενδείκνυνται για την αναπαράσταση τοπικά περιοδικών παραστάσεων υφής.

Οι Papandreou, Maragos, Kokaram [41] στα πλαίσια της ψηφιακής αποκατάστασης των τοιχογραφιών της Θήρας εστίασαν σε υπερπλήρεις μιγαδικές κυματιδιακές αναπαραστάσεις πλαισίου και συγκεκριμένα στο δίδεντρο κυματιδιακό μετασχηματισμό που εμφανίζει σημαντικά βελτιωμένες ιδιότητες μεταφορικής και περιστροφικής συμμετρίας σε σχέση με τους κρίσιμα δειγματοποιημένους ορθοκανονικούς μετασχηματισμούς και απέδειξαν την υπεροχή του για την εφαρμογή αυτή. Είναι γεγονός όμως ότι η χρήση wavelet μετασχηματισμού έφερε σημαντική υπολογιστική πολυπλοκότητα στο πρόβλημα, χωρίς ανάλογη βελτίωση στα αποτελέσματα, γι' αυτό και δεν επιλέχθηκε στην παρούσα εργασία. Αντίθετα, προτιμήθηκε η χρήση PDEs που εμφανίζει σημαντικά πλεονεκτήματα, τα οποία παρουσιάζονται στην επόμενη ενότητα.

3.2 Οφέλη χρήσης PDEs στην Όραση Υπολογιστών

Η χρήση PDEs στην επεξεργασία εικόνας είναι μία κλασική πλέον τάση στην Όραση Υπολογιστών. Βασικό χαρακτηριστικό τους είναι το ότι αποτελούν συνεχή και όχι διακριτά μοντέλα. Έτσι τα προβλήματα μοντελοποιούνται σε συνεχείς χώρους και η διακριτοποίηση εισέρχεται μόνο κατά την αριθμητική αναπαράσταση και επίλυση στον υπολογιστή.

Πλεονεκτήματα των PDEs είναι η συσχέτιση με άλλα κλασικά προβλήματα μερικών διαφορικών εξισώσεων από τη φυσική, τα οποία έχουν μελετηθεί εκτενώς και επιλυθεί, η εκμετάλλευση φυσικών νόμων στην υπηρεσία της Όρασης Υπολογιστών και η αξιοποίηση των αποτελεσμάτων της αριθμητικής ανάλυσης στη διακριτοποίηση των συνεχών μοντέλων.

Η χρήση των συνεχών μοντέλων των μερικών διαφορικών εξισώσεων αποδείχθηκε πολύ χρήσιμη για τη μη-γραμμική επεξεργασία εικόνας. Η μη-γραμμική εξίσωση ανισοτροπικής διάχυσης ή η ανάλυση εικόνας σε χώρους κλίμακας (scale spaces) έδωσαν σπουδαία αποτελέσματα, ενώ όπως θα εξεταστεί ακολούθως η μέθοδος που χρησιμοποιήθηκε στην παρούσα εργασία για το structure inpainting είναι ένας συνδυασμός μη-γραμμικών μοντέλων από PDEs και του λογισμού μεταβολών (variational calculus).

3.3 Ενδοσυμπλήρωση περιεχομένου (structure inpainting) με το TV μοντέλο

3.3.1 Αποθορυβοποίηση εικόνας και μοντέλο ολικής μεταβολής

Στην επεξεργασία εικόνας η αποθορυβοποίηση ολικής μεταβολής (*total variation denoising*) είναι μία διεργασία προορισμένη για την αφαίρεση του θορύβου. Η μέθοδος στηρίζεται στην ιδέα ότι οι περιοχές της εικόνας με μεγάλη λεπτομέρεια, η οποία με μεγάλη πιθανότητα είναι θόρυβος, παρουσιάζουν μεγάλη ολική μεταβολή και συνεπώς το ολοκλήρωμα του μέτρου των παραγώγων της εικόνας λαμβάνει μεγάλες τιμές στις περιοχές αυτές. Έτσι, ελαχιστοποιώντας το ολοκλήρωμα αυτό, κρατώντας παράλληλα την εξαχθείσα εικόνα «κοντά» στην αυθεντική, αποβάλλεται η ανεπιθύμητη λεπτομέρεια, αλλά διατηρείται η σημαντική πληροφορία όπως οι ακμές και έτσι πετυχαίνεται η αποθορυβοποίηση. Οι περιορισμοί επιβάλλονται με τη χρήση πολλαπλασιαστών

Lagrange. Τελικά η λύση του προβλήματος ανάγεται στην επίλυση μίας μερικής διαφορικής εξίσωσης στο πολύπτυχο (manifold) που ορίζεται από τους περιορισμούς. Για $t \rightarrow \infty$ η λύση συγκλίνει στη μόνιμη της κατάσταση και αποτελεί την αποθορυβοποιημένη εικόνα. Οι πρωτεργάτες της ιδέας αυτής ήταν οι Rudin, Osher και Fatemi το 1992 [29]. Ακολούθως παρουσιάζονται τα κύρια σημεία της μεθόδου με μαθηματικό φορμαλισμό βάσει μίας πιο πρόσφατης εργασίας των Chan, Osher και Shen το 2001 [42].

Έστω $u^0(x)$ η θορυβώδης εκδοχή μίας «καθαρής» εικόνας $u(x)$

$$u^0(x) = u(x) + n(x), \quad (3.1)$$

όπου ο θόρυβος $n(x)$ είναι τυχαίος με μηδενική μέση τιμή και διασπορά σ^2 :

$$E[n(x)] = 0, \quad E[n^2(x)] = \sigma^2. \quad (3.2)$$

Οι γραμμικές μέθοδοι εν προκειμένω δεν είναι αποτελεσματικές, διότι μαζί με το θόρυβο «καταστρέφουν» τα μεμονωμένα χαρακτηριστικά υψηλής συχνότητας, όπως ακμές και ασυνέχειες. Έτσι, οι ερευνητές έχουν στραφεί σε μη γραμμικούς τελεστές, όπως συμβαίνει στους αλγόριθμους ανισοτροπικής διάχυσης που είναι πολύ δημοφιλείς. Σε αυτούς ανήκει και το επιτυχημένο μοντέλο των Rudin et al. [29], το οποίο στηρίζεται στην ελαχιστοποίηση της L^1 (total variation, TV) νόρμας:

$$TV[u] = \int_{\Omega} |\nabla u| dx, \quad (3.3)$$

όπου Ω είναι το πεδίο ορισμού της εικόνας και dx το στοιχειώδες εμβαδόν επί του Ω . Σε σχέση με ένα τυπικό L_2 μοντέλο έχει λιγότερο απλή υλοποίηση, ενώ εμφανίζεται πολυπλοκότητα λόγω της μη γραμμικότητας, αλλά το κέρδος είναι μεγάλο ειδικά όσον αφορά την αποτελεσματικότητα στη διατήρηση των ακμών.

Οι υποθέσεις για το θόρυβο 3.2 οδηγούν στους εξής περιορισμούς για την ελαχιστοποίηση της TV νόρμας:

$$\int_{\Omega} u dx = \int_{\Omega} u^0 dx, \quad \frac{1}{|\Omega|} \int_{\Omega} (u - u^0)^2 dx = \sigma^2, \quad (3.4)$$

όπου $|\Omega|$ είναι το εμβαδόν του πεδίου ορισμού της εικόνας Ω . Έτσι, οι Εξισώσεις 3.3 και 3.4 μαζί ορίζουν ένα πρόβλημα βελτιστοποίησης υπό συνθήκες.

Επειδή η TV νόρμα είναι αναλλοίωτη ως προς τη μετατόπιση του ορίσματος, δηλαδή

$$TV[u + c] = TV[u] \quad (3.5)$$

για κάθε σταθερά c , ο πρώτος περιορισμός ενσωματώνεται αυτόματα. Για την ικανοποίηση του δεύτερου περιορισμού, εισάγεται ένας πολλαπλασιαστής Lagrange λ και απαιτείται η ελαχιστοποίηση του εξής συναρτησιακού ενέργειας:

$$J[u] = \int_{\Omega} |\nabla u| dx + \frac{\lambda}{2} \int_{\Omega} (u - u^0)^2 dx. \quad (3.6)$$

Η Euler-Lagrange εξίσωση του J γράφεται

$$-\nabla \left(\frac{\nabla u}{|\nabla u|} \right) + \lambda(u - u^0) = 0 \quad (3.7)$$

και επιλέγοντας την κατεύθυνση της πιο απότομης μεταβολής (steepest descent marching) προκύπτει

$$\frac{\partial u}{\partial t} = \nabla \left(\frac{\nabla u}{|\nabla u|} \right) + \lambda(u^0 - u). \quad (3.8)$$

Για την αποφυγή της διαίρεσης με το $|\nabla u| = 0$ σε επίπεδες περιοχές ή σε ακρότατα της εικόνας, γίνεται η κανονικοποίηση

$$|\nabla u|_\alpha = \sqrt{|\nabla u|^2 + \alpha^2} \quad (3.9)$$

με κάποια μικρή σταθερά α . Έτσι, η Euler-Lagrange Εξίσωση 3.7, τροποποιημένη πλέον, ελαχιστοποιεί το κανονικοποιημένο συναρτησιακό ενέργειας

$$J_\alpha[u] = \int_\Omega |\nabla u|_\alpha dx + \frac{\lambda}{2} \int_\Omega (u - u^0)^2 dx. \quad (3.10)$$

Στις ακόλουθες δύο υποενότητες παρουσιάζονται οι βασικές αρχές που πρέπει να πληροί ένα μοντέλο ενδοσυμπλήρωσης και περιγράφεται μία παραλλαγή του μοντέλου αποθορυβοποίησης που είναι κατάλληλη για την επίλυση του προβλήματος της ενδοσυμπλήρωσης, βάση και αυτών των αρχών. Οι υποενότητες αυτές έχουν βασιστεί στην εργασία των Chan και Shen [11].

3.3.2 Βασικές αρχές που πρέπει να ικανοποιεί ένα inpainting μοντέλο

Αν και τα ομαλά inpainting μοντέλα έχουν επιτρέψει την αυστηρή μαθηματική ανάλυση σχετικά με την ακρίβεια και τη φύση τους, σε εφαρμογές του πραγματικού κόσμου δεν είναι πολύ πρακτικά, γιατί αφενός οι εικόνες, ακόμη και αυτές χωρίς υφή, ντετερμινιστικά δεν είναι ομαλές, καθώς περιέχουν ακμές και ασυνέχειες, και αφετέρου οι εικόνες είναι στατιστικά αλλοιωμένες από θόρυβο. Ο άνθρωπος-τεχνίτης που καλείται να κάνει αποκατάσταση σε ένα έργο τέχνης δεν συναντά τέτοιες δυσκολίες.

Έτσι, ένα πρακτικό inpainting μοντέλο αξιώνεται να υπακούει τις ακόλουθες αρχές, ώστε να είναι αποτελεσματικό. Πρέπει:

⇒ *Να αντλεί την τοπική πληροφορία.*

Αφού δεν απαιτείται από τα μοντέλα καθολική μάθηση (global learning), η εξαχθείσα πληροφορία u_D πρέπει να καθορίζεται από την διαθέσιμη πληροφορία του u_0 στη γειτονιά του inpainting πεδίου D .

⇒ *Να παρουσιάζει ευρωστία στο θόρυβο.*

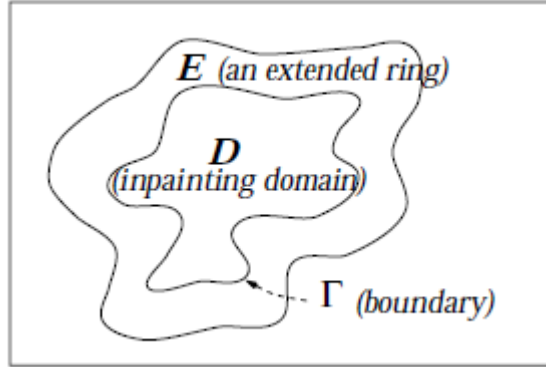
Για το ανθρώπινο μάτι είναι εύκολο να διαχωρίσει τα αντικείμενα όταν ο θόρυβος βρίσκεται κάτω από ένα ανεκτό επίπεδο. Συνεπώς, είναι μία ιδιότητα που αναμένεται να έχει το inpainting μοντέλο, ώστε να συμπληρώνει σωστά την ελλείπουσα πληροφορία.

⇒ *Να αποκαθιστά τις λεπτές ομαλές ακμές που έχουν «σπάσει».*

Οι ακμές είναι κρίσιμος παράγοντας στην αναγνώριση των αντικειμένων και στην κατάτμηση και συχνά αποchrύπτεται μέρος τους. Έως ένα βαθμό αλλοίωσης αναμένουμε το μοντέλο να μπορεί να τις ανασκευάζει.

3.3.3 Το μοντέλο ενδοσυμπλήρωσης ολικής μεταβολής

Για τις ανάγκες της ακόλουθης ανάλυσης ορίζεται ξανά η τοπολογία του προβλήματος ενδοσυμπλήρωσης (Σχήμα 3.1). Έστω, λοιπόν, D η περιοχή από όπου λείπει πληροφορία, με ομαλό κατά τμήματα σύνορο Γ και E ένα κλειστό πεδίο στη συμπληρωματική περιοχή D^c , έτσι ώστε $\Gamma \in E \cup D$. Η τοπολογία αυτή υπακούει στις αρχές του inpainting σχετικά με την τοπικότητα και την ευρωστία σε θόρυβο.



Σχήμα 3.1: Το μοντέλο ενδοσυμπλήρωσης ολικής μεταβολής υπολογίζει την ελλείπουσα περιοχή $u|_D$ βάση της TV νόρμας στο πεδίο $E \cup D$ και του περιορισμού θορύβου στο E [11].

Γίνεται η υπόθεση ότι η εικόνα $u^0|_E$ περιέχει λευκό θόρυβο. Το TV μοντέλο ενδοσυμπλήρωσης συνίσταται στην εύρεση μίας συνάρτησης u στο πεδίο $E \cup D$ που ελαχιστοποιεί ένα κατάλληλο ρυθμιστικό συναρτησιακό

$$R[u] := \int_{E \cup D} r(|\nabla u|) dx, \quad (3.11)$$

ύπο τον περιορισμό ταιριάσματος (ή αποθορυβοποίησης) στο E

$$\frac{1}{\text{Area}(E)} \int_E |u - u^0|^2 dx = \sigma^2. \quad (3.12)$$

Η συνάρτηση r είναι μη αρνητική για μη αρνητικές εισόδους, ενώ το σ είναι η τυπική απόκλιση του λευκού θορύβου.

Προφανώς, αν το σύνολο D είναι κενό, ο παραπάνω φορμαλισμός αφορά το πρόβλημα της αποθορυβοποίησης που αναλύθηκε στην Υποενότητα 3.3.1. Συγκεκριμένα για $r(s) = s$ προκύπτει ακριβώς το κλασικό μοντέλο των Rudin et al. Οι Εξισώσεις 3.11 και 3.12, όπως έχουν καταστρωθεί, ικανοποιούν τις αρχές της τοπικότητας και της ευρωστίας σε θόρυβο. Για να ικανοποιηθεί και η τρίτη αρχή του inpainting που αφορά την ικανότητα αποκατάστασης «σπασμένων» ακμών, απαιτείται ένα κατάλληλο συναρτησιακό $r(s)$. Κατά μήκος των ακμών ο όρος ∇u είναι μία μονοδιάστατη συνάρτηση δ . Έτσι, απαιτείται το ολοκλήρωμα

$$\int_{E \cup D} r(\delta) dx \quad (3.13)$$

να είναι πεπερασμένο. Κάτι που συνεπάγεται ότι αν

$$r(s) = s^\alpha \quad (3.14)$$

για κάποια δύναμη α καθώς $s \rightarrow \infty$, τότε $\alpha \leq 1$. Επιλέγεται $\alpha = 1$ για να επιτευχθεί *σχεδόν κυρτότητα*. Έτσι, προκύπτει η επιλογή $r(s) = s$ που συναντάται στο μοντέλο των Rudin et al. και θα υιοθετηθεί και στο TV inpainting μοντέλο που θα περιγραφεί στην παρούσα ενότητα.

Έτσι, αντίστοιχα με το πρόβλημα της αποθρομβοποίησης, πρέπει τώρα να ελαχιστοποιηθεί το ακόλουθο συναρτησιακό:

$$J_\lambda[u] = \int_{E \cup D} |\nabla u| dx + \frac{\lambda}{2} \int_E (u - u^0)^2 dx, \quad (3.15)$$

όπου και πάλι το λ έχει το ρόλο του πολλαπλασιαστή Lagrange για το αντίστοιχο πρόβλημα υπό περιορισμούς (Εξισώσεις 3.11 και 3.12).

Για την επίλυση του προβλήματος ελαχιστοποίησης πρέπει να βρεθεί που μηδενίζεται η Euler-Lagrange παράγωγος του συναρτησιακού $J_\lambda[u]$:

$$\frac{\partial J_\lambda[u]}{\partial u} = -\nabla \cdot \left(\frac{\nabla u}{|\nabla u|} \right) + \lambda_e(u - u^0) = 0 \quad (3.16)$$

για όλα τα σημεία $z = (x, y) \in E \cup D$ συν μια συνοριακή συνθήκη τύπου Neumann. Ο πολλαπλασιαστής Lagrange δίνεται από τη σχέση

$$\lambda_e = \begin{cases} \lambda, & z \in E \\ 0, & z \in D \end{cases} \quad (3.17)$$

Έτσι προκύπτει η διαφορική εξίσωση της πιο απότομης καθόδου για το $J_\lambda[u]$:

$$\frac{\partial u}{\partial t} = \nabla \cdot \left(\frac{\nabla u}{|\nabla u|} \right) + \lambda_e(u - u^0). \quad (3.18)$$

Κατά την επίλυση της 3.18, για την αποφυγή της διαίρεσης με το 0 σε περιοχές με μηδενική παράγωγο, αντικαθιστάται ο όρος

$$\nabla \cdot \left(\frac{\nabla u}{|\nabla u|} \right) \text{ με τον «ανυψωμένο» όρο } \nabla \cdot \left(\frac{\nabla u}{|\nabla u|_\alpha} \right), \text{ όπου } |s|_\alpha := \sqrt{s^2 + \alpha^2} \quad (3.19)$$

και α μία μικρή (συνήθως) θετική σταθερά ανύψωσης. Ουσιαστικά ελαχιστοποιείται το συναρτησιακό ενέργειας:

$$J_\lambda^\alpha[u] = \int_{E \cup D} \sqrt{|\nabla u|^2 + \alpha^2} dx + \frac{\lambda}{2} \int_E (u - u^0)^2 dx. \quad (3.20)$$

Η σταθερά α έχει και το ρόλο του *κατωφλίου*. Στις ομαλές περιοχές, όπου $|\nabla u| \ll \alpha$ το μοντέλο προσομοιώνει το αρμονικό inpainting, ενώ κατά μήκος των ακμών όπου $|\nabla u| \gg \alpha$ το μοντέλο είναι το κλασικό TV inpainting. Στο σημείο αυτό πρέπει να υπογραμμιστεί ότι το TV inpainting μοντέλο, όπως προκύπτει και από την προηγούμενη ανάλυση, είναι πάνω από όλα και ένα φίλτρο αφαίρεσης θορύβου. Συνεπώς, ταυτόχρονα με το inpainting λαμβάνει χώρα και αποθρομβοποίηση στην περιοχή E , δηλαδή έξω από το inpainting χωρίο.

Η όλη διαδικασία ελέγχεται μέσω της παραμέτρου λ , η οποία εκφράζει το πόσο αλλάζουν οι τιμές των pixels που ανήκουν στην περιοχή αυτή. Σε περίπτωση απουσίας θορύβου τίθεται $\lambda = \infty$ και η περιοχή E μένει αναλλοίωτη, ενώ αν υπάρχει θόρυβος στην παράμετρο λ δίνεται μία πεπερασμένη τιμή και γίνεται παράλληλα με την ενδοσυμπλήρωση και η αποθρομβοποίηση.

3.3.4 Μαθηματική θεμελίωση της μεθόδου

Για την απόδειξη ύπαρξης βέλτιστης λύσης του TV inpainting μοντέλου, στο [12] διατυπώνονται και αποδεικνύονται τα δυο ακόλουθα θεωρήματα:

Θεώρημα 3.1. Υπαρξη βέλτιστης λύσης σε εικόνα χωρίς θόρυβο

Έστω u_0 μια εικόνα στο χώρο $BV(\Omega)^1$ που παίρνει τιμές στο διάστημα $[0, 1]$ (μαύρο) με 1 (λευκό). Τότε το δίχως θόρυβο TV Inpainting μοντέλο $TV(u) = \int_{\Omega} |\nabla u|$, $u|_{\Omega \setminus D} = u_0|_{\Omega \setminus D}$, μαζί με τον περιορισμό $u \in [0, 1]$, έχει τουλάχιστον μία βέλτιστη λύση.

Απόδειξη.

Αφού η αυθεντική εικόνα u_0 ικανοποιεί τους περιορισμούς, πάντα μπορεί να βρεθεί μία ακολουθία από inpaintings $(u_n)_n$ που ικανοποιούν το μοντέλο. Τότε οι όροι

$$\int_{\Omega} |\nabla u_n| \quad \text{και} \quad \int_{\Omega} |u_n(x)| dx$$

είναι φραγμένοι για όλα τα n , αφού Ω φραγμένο και $u_n \in [0, 1]$. Με βάση την ασθενή ιδιότητα συμπίεσης των BV συναρτήσεων, υπάρχει μία υπακολουθία, έστω $(u_n)_n$, που συγκλίνει σε κάποιο $u_{tv} \in L_1(\Omega)$ στην L_1 νόρμα. Φυσικά το u_{tv} πληροί τους περιορισμούς

$$u_{tv}|_{\Omega \setminus D} = u_0|_{\Omega \setminus D} \quad \text{και} \quad u_{tv} \in L_1(\Omega).$$

Και με χρήση της ιδιότητας της L_1 συνέχειας από κάτω, προκύπτει ότι

$$\int_{\Omega} |\nabla u_{tv}| \leq \liminf \int_{\Omega} |\nabla u_n| = \min_u \int_{\Omega} |\nabla u|.$$

Έτσι, η εικόνα u_{tv} ελαχιστοποιεί το συναρτησιακό.

□

Θεώρημα 3.2. Υπαρξη βέλτιστης λύσης σε εικόνα με θόρυβο

Έστω μια εικόνα u_0 στο $\Omega \setminus D$. Υποθέτουμε ότι υπάρχει μία εικόνα u_c στο Ω που ανήκει στο χώρο $BV(\Omega)$ και η οποία ικανοποιεί τον περιορισμό αποθορυβοποίησης

$$\frac{1}{\text{Area}(\Omega \setminus D)} \int_{\Omega \setminus D} (u_c - u_0)^2 dx = \sigma^2$$

και τον περιορισμό $u_c \in [0, 1]$. Τότε υπάρχει τουλάχιστον ένα βέλτιστο TV inpainting u_{tv} στο Ω που κάνει ενδοσυμπλήρωση στο D και αποθορυβοποίηση στο $\Omega \setminus D$.

Απόδειξη.

Δεδομένης της υπόθεσης για το u_c , υπάρχει μία ακολουθία από inpaintings $(u_n)_n$ που ικανοποιούν τους περιορισμούς και ελαχιστοποιούν το συναρτησιακό $J_{\lambda}[\cdot]$. Λόγω του περιορισμού στο διάστημα $[0, 1]$, η $(u_n)_n \in BV(\Omega)$. Έτσι, υπάρχει μία υπακολουθία, έστω $(u_n)_n$, η οποία συγκλίνει στην L_1 νόρμα σε κάποια εικόνα $u_{tv} \in L_1(\Omega)$. Και η ιδιότητα της L_1 συνέχειας από κάτω εγγυάται ότι:

$$\int_{\Omega} |\nabla u_{tv}| \leq \liminf \int_{\Omega} |\nabla u_n| = \min_u \int_{\Omega} |\nabla u|.$$

¹Ο χώρος των Bounded Variation συναρτήσεων περιγράφεται στην Ενότητα 1.2.4.

Επίσης απαιτείται $u_n \rightarrow u_{tv}$ σχεδόν παντού στο Ω . Έτσι, πρέπει $u_{tv} \in [0, 1]$, αλλά και ακόμη να ισχύει με βάση το *θεώρημα κυρίαρχης σύγκλισης του Lebesgue*,

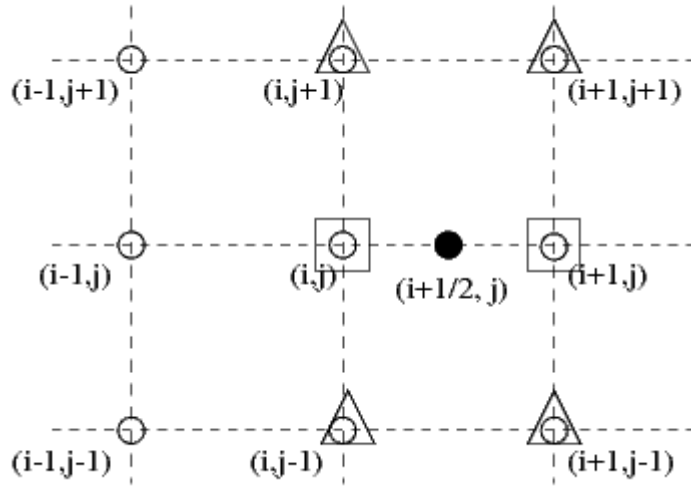
$$\int_{\Omega \setminus D} (u_{tv} - u_0)^2 dx = \lim_n \int_{\Omega \setminus D} (u_n - u_0)^2 dx.$$

Συνεπώς, το u_{tv} είναι πράγματι ένα βέλτιστο TV inpainting.

□

3.3.5 Αριθμητικό σχήμα υλοποίησης της μεθόδου

Στην παρούσα ενότητα που έχει βασιστεί στο [11] περιγράφεται ένα αριθμητικό σχήμα για την υλοποίηση του TV inpainting model. Έστω οι συμβολισμοί για τα γειτονικά σημεία ενός pixel στο διακριτό πλέγμα (grid) της εικόνας όπως αυτοί φαίνονται στο Σχήμα 3.2. Ορίζονται και τα ενδιάμεσα σημεία για τη διευκόλυνση στους υπολογισμούς, χωρίς όμως να αντιστοιχούν σε πραγματικά pixel της εικόνας.



Σχήμα 3.2: Τα γειτονικά σημεία ενός pixel επί του διακριτού πλέγματος της εικόνας [12].

Στη συνέχεια εκφράζονται τα μεγέθη της Εξίσωσης 3.16 σε διακριτή μορφή. Έστω $\mathbf{v} = (v^1, v^2) = \frac{\nabla u}{\|\nabla u\|}$. Η απόκλισή του είναι:

$$\nabla \cdot \mathbf{v} = \frac{\partial v^1}{\partial x} + \frac{\partial v^2}{\partial y} \simeq \frac{v^1_{(i+1/2, j)} - v^1_{(i-1/2, j)}}{h} + \frac{v^2_{(i, j+1/2)} - v^2_{(i, j-1/2)}}{h} \quad (3.21)$$

όπου h η διακριτική ικανότητα που λαμβάνεται πάντα μονάδα στην επεξεργασία εικόνας.

Έτσι, ορίζονται οι προσεγγίσεις στα ενδιάμεσα βοηθητικά σημεία, όπου η πληροφορία δεν

είναι απευθείας διαθέσιμη. Για το σημείο $(i+1/2, j)$ για παράδειγμα ισχύουν:

$$\begin{aligned}
 v_{(i+1/2,j)}^1 &= \frac{1}{|\nabla u_{(i+1/2,j)}|} \left[\frac{\partial u}{\partial x} \right]_{(i+1/2,j)} \simeq \frac{1}{|\nabla u_{(i+1/2,j)}|} \frac{u_{(i+1,j)} - u_{(i,j)}}{h} \\
 |\nabla u_{(i+1/2,j)}| &\simeq \frac{\sqrt{(u_{(i+1/2,j)} - u_{(i,j)})^2 + [(u_{(i+1,j+1)} + u_{(i,j+1)} - u_{(i,j-1)} - u_{(i-1,j+1)})/4]^2}}{h} \\
 \left[\frac{\partial n}{\partial x} \right] &\simeq \frac{u_{(i+1,j)} - u_{(i,j)}}{h} \\
 \left[\frac{\partial n}{\partial y} \right] &\simeq \frac{u_{(i+1,j+1)} + u_{(i+1,j)} - u_{(i-1,j)} - u_{(i-1,j+1)}}{4h}
 \end{aligned} \tag{3.22}$$

Αντίστοιχες εκφράσεις με τις 3.22, ισχύουν και για τα άλλα ενδιάμεσα σημεία $(i, j+1)$, $(i-1, j)$ και $(i, j-1)$. Έτσι για ένα σημείο (i, j) η Εξίσωση 3.16 γίνεται

$$\sum_{(i',j') \in N_{(i,j)}} \frac{1}{|\nabla u_{(i'-1/2,j'-1/2)}|} (u_{(i,j)} - u_{(i',j')}) + \lambda_e(i, j) (u_{(i,j)} - u_{(i,j)}^0) = 0, \tag{3.23}$$

όπου $N_{(i,j)}$ η γειτονιά του σημείου (i, j) .

Ορίζονται:

$$w_{(i',j')} = \frac{1}{|\nabla u_{(i',j')}|} \tag{3.24}$$

$$h_{(i'-i,j'-j)} = \frac{w_{(i',j')}}{\sum_{Q \in N_{(i,j)}} w_Q + \lambda_e(i, j)} \tag{3.25}$$

$$h_{(i,j)} = \frac{\lambda_e(i, j)}{\sum_{Q \in N_{(i,j)}} w_Q + \lambda_e(i, j)} \tag{3.26}$$

Έτσι, η 3.23 γράφεται:

$$u_{(i,j)} = \sum_{(i',j') \in N_{(i,j)}} h_{(i'-i,j'-j)} u_{(i',j')} + h_{(i,j)} u_{(i,j)}^0, \tag{3.27}$$

με

$$\sum_{(i',j') \in N_{(i,j)}} h_{(i'-i,j'-j)} + h_{(i,j)} = 1. \tag{3.28}$$

Για την αριθμητική επίλυση της μη γραμμικής εξίσωσης 3.27 χρησιμοποιείται η επαναληπτική μέθοδος Gauss-Jacobi:

$$u_{(i,j)}^{(n)} = \sum_{(i',j') \in N_{(i,j)}} h_{(i'-i,j'-j)}^{(n-1)} u_{(i',j')}^{(n-1)} + h_{(i,j)}^{(n-1)} u_{(i,j)}^{(n-1)}, \tag{3.29}$$

όπου $h^{(n-1)} = h(u^{(n-1)})$.

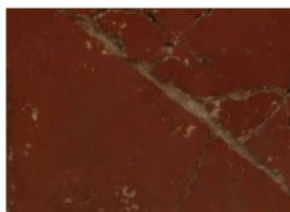
Τέλος, αυτοματοποιήθηκε η διαδικασία σύγκλισης του αλγορίθμου με τον έλεγχο αν η απόσταση Hamming μεταξύ δύο διαδοχικών στιγμιότυπων της υπό ενδοσυμπλήρωση περιοχής είναι κάτω από το 0.5% της πληροφορίας του τελευταίου στιγμιότυπου. Όταν ικανοποιηθεί αυτή η συνθήκη θεωρείται ότι ο αλγόριθμος έχει συγχλίνει. Στιγμιότυπα λαμβάνονται κάθε 400 επαναλήψεις του αλγορίθμου.

3.3.6 Εφαρμογή του TV inpainting μοντέλου στις τοιχογραφίες

Ακολουθούν κάποια παραδείγματα εφαρμογής του TV inpainting μοντέλου στις τοιχογραφίες.



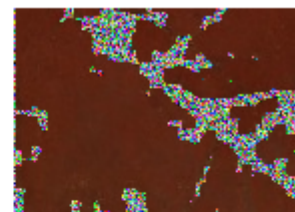
Εντός πλαισίου η περιοχή ενδιαφέροντος



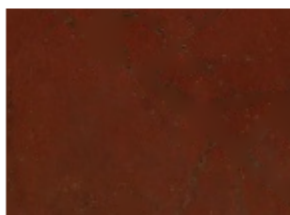
Αρχική εικόνα



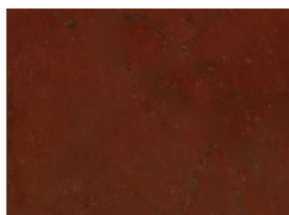
Μάσκα



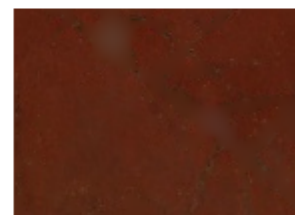
Αρχικοποίηση με τυχαίες τιμές
εντάσεων στα κενά



Τελικό αποτέλεσμα (στις
1200 επαναλήψεις)



Μετά από 800 επαναλήψεις



Μετά από 400 επαναλήψεις

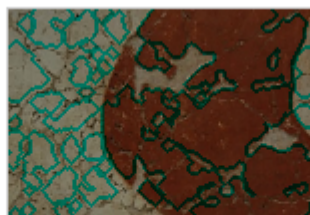
Σχήμα 3.3: Εφαρμογή inpainting σε τμήμα της τοιχογραφίας *Ακολουθία Ανδρών* από την πλάτη του μεγαλύτερου άνδρα. Το τμήμα αυτό παρουσιάζει μικρές ελλείψεις, δηλαδή ρωγμές μικρού ή μέσου πλάτους, οι οποίες αντιμετωπίζονται αποτελεσματικά από το TV inpainting μοντέλο.



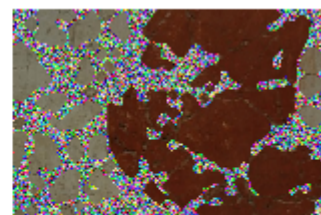
Εντός πλαισίου η περιοχή ενδιαφέροντος



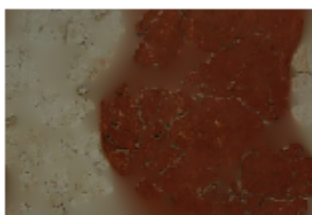
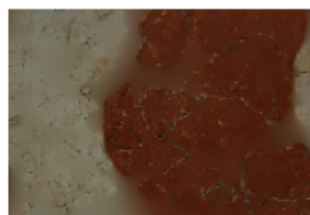
Αρχική μορφή



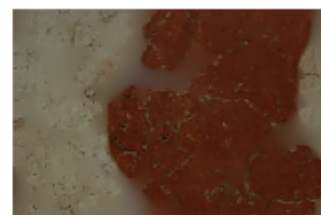
Μάσκα



Αρχικοποίηση με τυχαίες τιμές

Τελική μορφή (σύγκλιση
στις 1200 επαναλήψεις)

Μετά από 800 επαναλήψεις

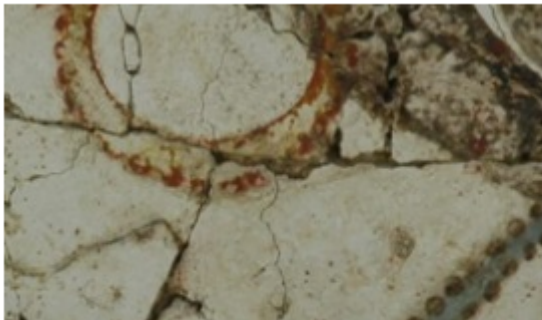


Μετά από 400 επαναλήψεις

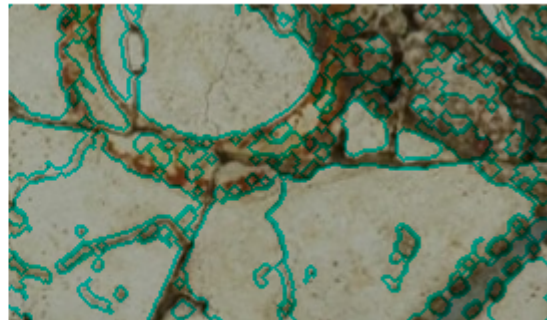
Σχήμα 3.4: Εφαρμογή inpainting σε τμήμα της ίδιας τοιχογραφίας που αφορά μέρος του σώματος του μεγαλύτερου άνδρα. Η περιοχή εμφανίζει πολλές ρωγμές, αλλά και κάποιες ελλείψεις μέσης έκτασης ως προς το μέγεθος της εικόνας, οι οποίες δεν αντιμετωπίζονται εύκολα. Παρόλ' αυτά, σε σχετικά μικρό πλήθος επαναλήψεων η παράσταση επιδιορθώνεται αρκούντως ικανοποιητικά. Ένας αλγόριθμος που αξιοποιεί το συναρτησιακό *elastica* του Euler θα μπορούσε να αποφέρει σαφώς βελτιωμένη αποκατάσταση.



Εντός πλαισίου η περιοχή ενδιαφέροντος



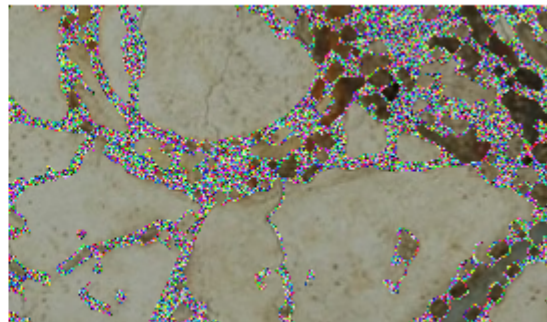
Αρχικό τμήμα



Μάσκα



Σύγκλιση στις 800 επαναλήψεις



Αρχικοποίηση

Σχήμα 3.5: Εφαρμογή inpainting σε τμήμα της τοιχογραφίας *Κροκοσυλλέκτρια & Πότνια* που απεικονίζει το σκουλαρίκι της Πότνιας. Αν και το περίγραμμα της μάσκας δείχνει ικανοποιητικό, παρόλ' αυτά αρχικοποιείται μεγαλύτερη επιφάνεια με τυχαίες τιμές από την ελάχιστη δυνατή και έτσι χάνεται σημαντική για την αποκατάσταση της τοιχογραφίας πληροφορία. Βασική αιτία είναι ότι αρκετές ακμές της παράστασης αναγνωρίζονται ως ελλείψεις. Η εκπαίδευση ενός μοντέλου, ώστε να αναγνωρίζει ποιες ακμές δεν αφορούν ελλείψεις, θα μπορούσε να προσφέρει σημαντική βελτίωση στο αποτέλεσμα.

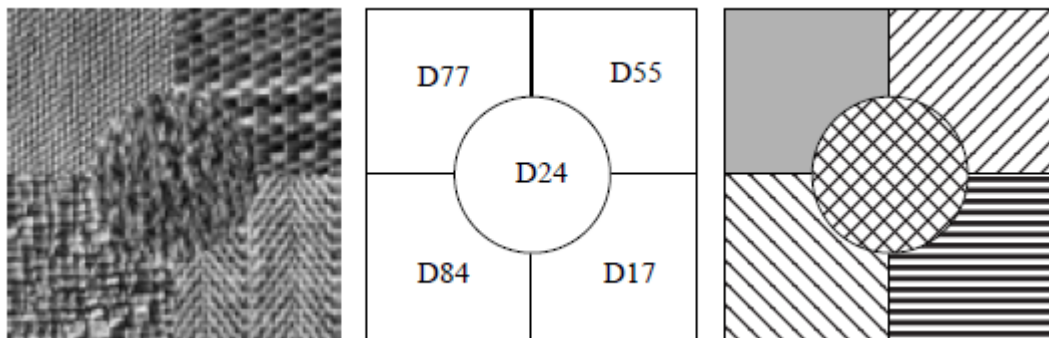
3.4 Ενδοσυμπλήρωση με σύνθεση υφής (texture inpainting)

Οι αλγόριθμοι για structure inpainting δε δίνουν άριστα αποτελέσματα όταν η εικόνα έχει υφή, γιατί συμπληρώνουν την inpainting περιοχή με λείο και ομογενή τρόπο. Αυτό συνεπάγεται αλλοίωση της αρχικής εικόνας, της οποίας δίνεται ουσιαστικά μόνο μία αδρή περιγραφή, ένα σύνολο λείων επιφανειών διαχωριζομένων από ακμές. Για την ολοκληρωμένη αντιμετώπιση του inpainting προβλήματος απαιτείται η ενιαία αποκατάσταση τόσο του περιεχομένου όσο και της υφής της εικόνας (*structure & texture inpainting*). Στην παρούσα ενότητα γίνεται μία εισαγωγή στην έννοια της υφής, περιγράφεται το πρόβλημα της σύνθεσης υφής και παρουσιάζεται ο αλγόριθμος των Efros και Leung, μαζί με κάποια αποτελέσματα εφαρμογής του στις τοιχογραφίες.

3.4.1 Υφή

Σε πολλούς αλγορίθμους επεξεργασίας εικόνας λαμβάνονται απλουστευτικές υποθέσεις σχετικά με την ομοιομορφία των εντάσεων σε τοπικές περιοχές της εικόνας. Όμως στον αληθινό κόσμο λίγες επιφάνειες είναι ομογενείς. Για παράδειγμα, μία ξύλινη επιφάνεια δεν είναι ομοιομορφη αλλά περιέχει διακυμάνσεις στην ένταση των pixel της, οι οποίες σχηματίζουν συγκεκριμένα επαναλαμβανόμενα μοτίβα (patterns) που είναι αυτό που χαρακτηρίζεται ως υφή (texture). Τα μοτίβα αυτά μπορεί να είναι το αποτέλεσμα της φύσης της επιφάνειας του αντικειμένου, αλλά μπορεί να σχετίζονται με διαφοροποιήσεις στην ανάκλαση όπως το χρώμα σε μία επιφάνεια.

Η υφή είναι μία θεμελιώδης έννοια στην Όραση Υπολογιστών, καθώς είναι μία ιδιότητα που παρουσιάζεται στην πλειοψηφία των εικόνων. Έχει αναπτυχθεί ένα ολόκληρο παρακλάδι στην ΟΥ, αυτό της *Ανάλυσης Υφής*. Ένα βασικό πρόβλημα της περιοχής είναι η *κατάτμηση υφής* που αφορά το διαχωρισμό των διαφορετικών ειδών υφής σε μια εικόνα. Μπορεί να αποτελέσει το πρώτο στάδιο του προβλήματος της *ταξινόμησης υφής*, η οποία έχει να κάνει με την αντιστοίχιση εικόνων σε ήδη καθορισμένες κατηγορίες βάση της υφής τους. Τέλος, η *σύνθεση υφής* έχει να κάνει με την κατασκευή μίας εικόνας υφής με υπόδειγμα μία άλλη εικόνα παρόμοιας υφής. Το συγκεκριμένο πρόβλημα αποτελεί και τη βάση για το inpainting υφής που θα εξεταστεί αναλυτικά στην Υποενότητα 3.4.4. Στο Σχήμα 3.6 παρουσιάζεται παράδειγμα εικόνας υφής.



Σχήμα 3.6: Μία εικόνα με πολλούς τύπους υφής. Η ταξινόμηση υφής αφορά την ταυτοποίηση της υφής καθεμιάς περιοχής, ενώ στην κατάτμηση υφής διαχωρίζονται οι περιοχές υφής σε μία εικόνα, μέσω του προσδιορισμού των συνόρων τους, χωρίς την ακριβή αναγνώριση του τύπου υφής ανά περιοχή. Έχει αντληθεί από το [13], στο οποίο βασίζεται η Υποενότητα 3.4.1.

Η υφή έχει χαρακτηριστεί ως παράδοξη έννοια, διότι αν και χρησιμοποιείται συχνά στην προεπεξεργασία οπτικής πληροφορίας, όπως στην κατάτμηση, και παρόλο που είναι απλό να αντιληφθεί την υφή το ανθρώπινο μάτι δεν έχει επιτευχθεί ως τώρα να δοθεί ένας γενικά αποδεκτός ορισμός της από τους ερευνητές της Όρασης Υπολογιστών. Κάποιοι την έχουν περιγράψει ως το περιεχόμενο μίας μακροσκοπικής εικόνας. Η υφή εν αντιθέσει με άλλες ιδιότητες της εικόνας, όπως η φωτεινότητα και το χρώμα, είναι μία ιδιότητα που αφορά μία ομάδα από pixel και όχι διακριτά μεταξύ τους pixel. Ο Haralick στο [43] θεωρεί την υφή ως κυτταρική δομή που περιγράφεται από το πλήθος και τη μορφή των στοιχειωδών δομών που την αποτελούν, καθώς και τη χωρική οργάνωση αυτών στο χώρο. Επίσης, τονίζει ένα θεμελιώδες χαρακτηριστικό της υφής, το ότι δε μπορεί να αναλυθεί εάν δεν υπάρχει ένα χωρικό πλαίσιο αναφοράς. Για κάθε επιφάνεια, ακόμη και αυτές που εμφανίζονται ως, ομοιόμορφες υπάρχει μία κλίμακα όπου παρουσιάζεται υφή.

Τυπικά, οι εικόνες υφής αποτελούνται από μοτίβα που επαναλαμβάνονται μέσα στην εικόνα σε μία ή περισσότερες κλίμακες και που εμφανίζουν χωρική περιοδικότητα είτε σταθερή είτε όχι. Τα μοτίβα αυτά είναι γνωστά και ως *texels* (texture elements). Είναι οι δομικές μονάδες των εικόνων υφής και εμφανίζουν κάποιο στοχαστικό χαρακτήρα. Ανάλογα με αυτόν μπορεί να χαρακτηριστεί από τελείως στοχαστική (π.χ. λευκός θόρυβος) έως τελείως κανονική (όπως είναι η σκακιέρα).

3.4.2 Σύνθεση υφής (texture synthesis)

Δεδομένου ενός δείγματος εικόνας υφής, το πρόβλημα της *σύνθεσης υφής* έγκειται ουσιαστικά στη σύνθεση μιας άλλης εικόνας με παρόμοια² υφή. Μόνη προϋπόθεση για την αρχική εικόνα είναι να είναι τέτοιου μεγέθους ώστε να διακρίνεται η χωρική της περιοδικότητα.

Οι μέθοδοι σύνθεσης υφής μπορούν να ταξινομηθούν [1] στις ακόλουθες κατηγορίες:

⇒ *Στατιστικές μέθοδοι*

Σε κάποιους αλγόριθμους η υφή μοντελοποιείται είτε ως ένα Τυχαίο Πεδίο Gibbs είτε ως ένα Τυχαίο Πεδίο Markov, όπως συμβαίνει στον αλγόριθμο Efros-Leung που περιγράφεται στην Ενότητα 3.4.3. Σε άλλους χρησιμοποιούνται 2Δ μοντέλα γραμμικής πρόβλεψης οδηγούμενα από λευκό Gaussian θόρυβο. Επίσης, σε αυτή την κατηγορία εμπίπτουν τα στατιστικά Fractal.

⇒ *Γεωμετρικές και δομικές μέθοδοι*

Αξιοποιούν το γεγονός ότι λίγες ομάδες τοπικών κυρίαρχων χαρακτηριστικών, τα λεγόμενα *textons*, αποτελούν διακριτικοί παράγοντες μεταξύ των περιοχών υφής για το ανθρώπινο μάτι.

⇒ *Μέθοδοι ενέργειας*

Η κατάτμηση υφής με συστοιχίες φίλτρων Gabor και η σύνθεση υφής με πολυδιάστατους τελεστές ενέργειας είναι κλασικά παραδείγματα μεθόδων της κατηγορίας.

Ο αλγόριθμος σύνθεσης υφής των Efros-Leung είναι αρκετά απλός και πολύ αποτελεσματικός συνάμα, γι' αυτό και επιλέχθηκε για την παρούσα εφαρμογή.

² Αν η νέα εικόνα αποτελείται από ένα ή περισσότερα αντίγραφα της εικόνας δείγματος (*tiling*), δηλαδή έχει περιοχές με την ίδια ακριβώς υφή, το αποτέλεσμα σπάνια είναι ικανοποιητικό οπτικά.

3.4.3 Ο αλγόριθμος σύνθεσης υφής Efros-Leung

Η υφή μοντελοποιείται ως ένα *Τυχαίο Μαρκοβιανό Πεδίο*, δηλαδή η συνάρτηση κατανομής της φωτεινότητας ενός pixel, δεδομένων των τιμών των φωτεινότητων των γειτονικών pixel, είναι ανεξάρτητη από την υπόλοιπη εικόνα. Ως γειτονιά ενός pixel θεωρείται ένα τετράγωνο περιττής διάστασης με κέντρο το συγκεκριμένο pixel. Η διάσταση του τετραγώνου επιλέγεται μεγαλύτερη ή ίση από τη μεγαλύτερη διάσταση του texel, έτσι ώστε να περιέχει όλα τα χαρακτηριστικά του δείγματος υφής. Η σύνθεση υφής γίνεται pixel προς pixel.

Έστω I_s η εικόνα-δείγμα (sample) και I η εικόνα που προκύπτει από τη σύνθεση. Υποθέτουμε ότι έχουμε συνθέσει όλη την εικόνα I εκτός από ένα pixel p . Έστω $\omega(p) \subset I$ η τετραγωνική γειτονιά του p . Ακόμη ορίζεται ένα μέτρο διαφοράς $d(\omega_1, \omega_2)$ μεταξύ δύο τετραγωνικών εικόνων. Τέλος θεωρούμε μία τετραγωνική εικόνα $\omega \subset I_s$. Το καλύτερο ταίριασμα ω της εικόνας ω_p δίνεται από τη σχέση $\omega_{best} = \underset{\omega}{\operatorname{argmin}} d(\omega(p), \omega)$. Στη συνέχεια, από το σύνολο εικόνων ω έτσι ώστε $d(\omega(p), \omega) < (1 + \epsilon)d(\omega(p), \omega_{best})$, όπου ϵ μικρή θετική σταθερά, λαμβάνεται τυχαία μία εικόνα ω_{can} . Έτσι, στο pixel p ανατίθεται η τιμή του pixel που βρίσκεται στο κέντρο της επιλεγμένης εικόνας ω_{can} . Η διαφορά d μεταξύ δύο εικόνων ορίζεται ως

$$d(\omega(p), \omega) = \sum_i \sum_j [M(i, j) \cdot (\omega(p)(i, j) - \omega(i, j))^2] * G(i, j) \quad (3.30)$$

μόνο για τα pixel που οι δύο εικόνες έχουν πάρει τιμές ($M(i, j) = 1$, αλλιώς $M(i, j) = 0$). Ο Gaussian πυρήνας δίνει μεγαλύτερη έμφαση στα pixel που είναι κοντά στο τρέχον υπό σύνθεση pixel.

Για την παράλληλη σύνθεση περισσότερων του ενός pixel θα πρέπει να υπολογιστούν τα από κοινού μεγέθη σύμφωνα με τη μοντελοποίηση της υφής ως MRF. Επειδή κάτι τέτοιο είναι ιδιαίτερα «βαρύ» υπολογιστικά στην πράξη, οι Efros και Leung ακολουθούν μία διαφορετική προσέγγιση. Αρχικά τοποθετείται το δείγμα υφής μέσα στην εικόνα I και η σύνθεση γίνεται pixel προς pixel αρχίζοντας από το όριο όπου βρίσκεται τον δείγμα I_s και πηγαίνοντας προς το όριο της εικόνας I μέσω μιας σπειροειδούς διαδρομής.

Ο αλγόριθμος Efros-Leung, έχοντας δώσει πολύ καλά αποτελέσματα στη σύνθεση υφής, αξιοποιήθηκε κατάλληλα τροποποιημένος στην ενδοσυμπλήρωση υφής και περιγράφεται στην επόμενη υποενότητα.

3.4.4 Ενδοσυμπλήρωση υφής βάση του αλγορίθμου Efros-Leung

Έστω η εικόνα I και $D \subset I$ η ελλείπουσα περιοχή της. Η εικόνα-δείγμα I_s είναι το συμπλήρωμα του D : $I_s = \{(x, y) : (x, y) \in I \setminus D\}$. Κάθε φορά γίνεται σύνθεση στο σύνολο

$$P = \{(x, y) : (x, y) \in (W \oplus B) \setminus W\}, \quad (3.31)$$

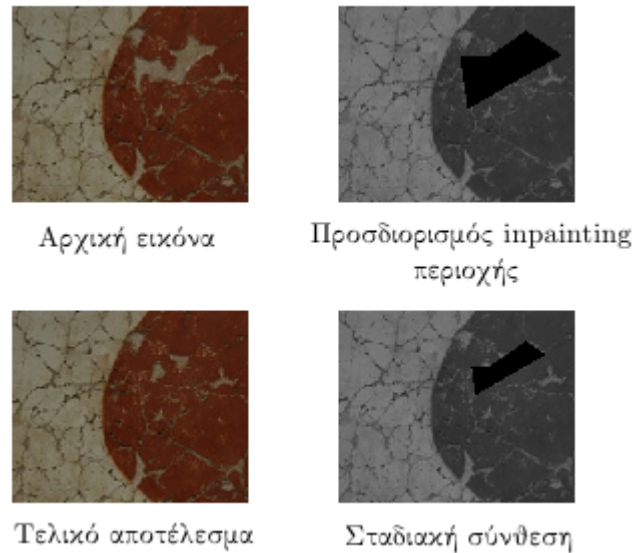
όπου W είναι η δυαδική συνάρτηση που είναι 1 για τα σημεία που έχουν ήδη συνθεθεί και 0 διαφορετικά, ενώ B είναι ο μοναδιαίος δίσκος. Τα pixel του συνόλου P ανακατεύονται με τυχαίο τρόπο και μετά ταξινομούνται με φθίνουσα σειρά ως προς τον αριθμό των γειτονικών τους pixel που έχουν συνθεθεί. Η ταξινόμηση γίνεται γιατί όσο πιο πολλά γειτονικά συμπληρωμένα pixel

έχει ένα σημείο τόσο περισσότερη πληροφορία είναι διαθέσιμη για το ποιο μέρος του texel αντιπροσωπεύει, οπότε επιλέγεται και ορθότερη τιμή.

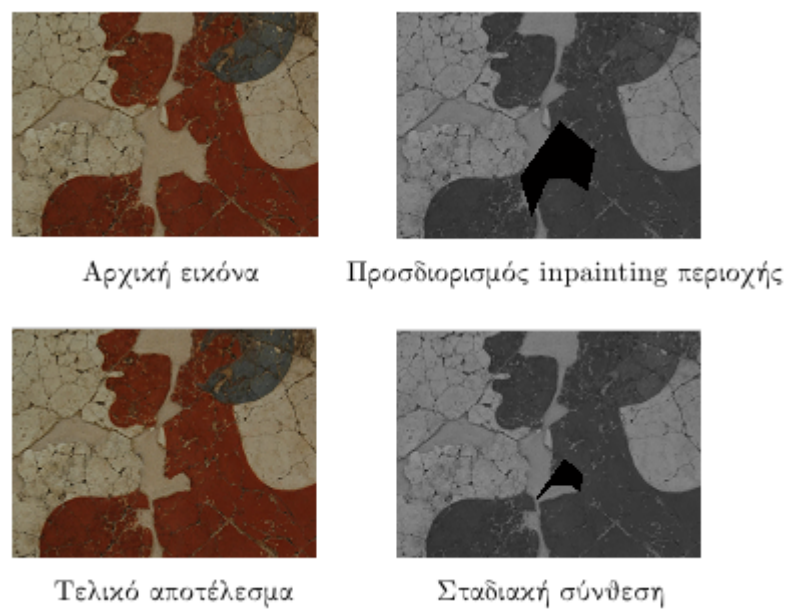
Για κάθε pixel της ταξινομημένης λίστας εφαρμόζεται η ίδια διαδικασία με την περίπτωση της σύνθεσης. Βρίσκεται το σύνολο των εικόνων $\omega \in I_s$ που ικανοποιούν τη σχέση $d(\omega(p), \omega) < (1 + \epsilon)d(\omega(p), \omega_{best})$ και επιλέγεται μία από αυτές ω_{can} με τυχαίο τρόπο. Ο λόγος που δεν επιλέγεται η καλύτερα εικόνα ω_{best} είναι για να αποφευχθεί η αντιγραφή της ίδιας εικόνας πολλές φορές, όπως στην τεχνική tiling. Στην επόμενη ενότητα παρουσιάζονται κάποια παραδείγματα από texture inpainting.

3.4.5 Εφαρμογή του αλγορίθμου Efros-Leung στις τοιχογραφίες

Στα Σχήματα 3.7 και 3.8 απεικονίζονται δύο παραδείγματα εφαρμογής του αλγορίθμου Efros-Leung στις τοιχογραφίες.



Σχήμα 3.7: Σύνθεση υψής σε περιοχή του σώματος του μεγαλύτερου άνδρα στην τοιχογραφία *Ακολουθία Ανδρών*. Είναι το ίδιο τμήμα με το Σχήμα 3.4, όπου είχε εφαρμοστεί το TV inpainting model. Όπως βλέπουμε πλέον δεν εμφανίζεται θόλωμα και το αποτέλεσμα είναι αισθητικά καλύτερο. Με διαδοχικές εφαρμογές σύνθεσης υψής σε όλες τις περιοχές της εικόνας μπορεί να γίνει πλήρης αποκατάσταση.



Σχήμα 3.8: Σύνθεση υφής στο λαιμό του μεγαλύτερου άνδρα στην τοιχογραφία *Ακολουθία Ανδρών*. Η γραμμή του λαιμού επεκτείνεται και η κόκκινη περιοχή συμπληρώνεται σταδιακά. Ο αλγόριθμος των Efros-Leung μπορεί να αποφέρει αισθητικά άρτια αποκατάσταση.

Κεφάλαιο 4

Αξιοποίηση πληροφορίας όλης της επιφάνειας των τοιχογραφιών

4.1 Εισαγωγή

Στην ενδοσυμπλήρωση περιεχομένου με τις κλασικές PDE μεθόδους, λόγω της φύσης της μεθόδου, χρησιμοποιείται ως επί το πλείστον πληροφορία από τη γειτονιά της ελλείπουσας περιοχής. Στην περίπτωση της ενδοσυμπλήρωσης υψής συμπληρώνεται η επιθυμητή περιοχή με επαναλαμβανόμενη χρήση κομματιών συγκεκριμένης υψής από την υπόλοιπη εικόνα. Συνεπώς, σε κάθε περίπτωση δε μπορεί να αξιοποιηθεί ικανοποιητικά μέρος από την ευρύτερη εικόνα και από οποιαδήποτε θέση της, ανεξαρτήτως αν περιλαμβάνει κάποιο texel ή όχι.

Για την καλύτερη δυνατή συμπλήρωση των κενών στις τοιχογραφίες απαιτείται να ανακτάται πληροφορία από οποιαδήποτε περιοχή της επιφάνειάς τους. Στην παρούσα εργασία δεν έχει αυτοματοποιηθεί το κομμάτι του εντοπισμού της πληροφορίας αυτής και αφήνεται ως υποκείμενο μελλοντικής έρευνας. Από τη στιγμή που βρεθεί η κατάλληλη εικόνα, κάτι που εν προκειμένω γίνεται από το χρήστη, για την προσάρτησή της στην ελλείπουσα περιοχή απαιτείται ένας αποτελεσματικός αλγόριθμος συνένωσης εικόνων χωρίς «ραφές» και άλλες παρενέργειες (*seamless image stitching*).

Έχουν χρησιμοποιηθεί διάφορες προσεγγίσεις σε αυτή την κατεύθυνση με πιο πρόσφατη και επιτυχημένη την εργασία των Tao, Johnson, Paris [6] που ακολουθεί την πρωτοποριακή εργασία των Levin, Zomet, Peleg, Weiss [5] και έρχεται να προτείνει νέες τεχνικές για περαιτέρω βελτίωση της εφαρμογής. Στην επόμενη ενότητα περιγράφονται οι δύο εργασίες, με παραπομπές στην ευρύτερη βιβλιογραφία της περιοχής και στο τέλος του κεφαλαίου παρατίθενται αποτελέσματα της εφαρμογής της μεθόδου των Tao et al. στην αποκατάσταση των τοιχογραφιών.

Αρκετά πρόσφατα έχουν δημοσιευτεί εργασίες όπου τόσο η ανάκτηση της πληροφορίας όσο και η επακόλουθη συνένωση στην ελλείπουσα περιοχή γίνονται αυτοματοποιημένα. Η δημοσίευση των Hays και Efros το 2008 [44] έχει να επιδείξει εντυπωσιακά αποτελέσματα. Έχοντας στη διά-

θεση τους μια βάση δεδομένων με πάνω από 2.000.000 εικόνες σχεδίασαν και υλοποίησαν μια υπολογιστικά αποδοτική εφαρμογή με σημασιολογικά ορθές επιλογές κομματιών σε πολύ μεγάλο ποσοστό, χωρίς προεπεξεργασία για την απόδοση σημασιολογίας στην κάθε εικόνα (*labelling*), αλλά αντίθετα υλοποιώντας μία εφαρμογή οδηγούμενη καθαρά από τα δεδομένα.

Τόσο η σημασιολογική ορθότητα όσο και το χαμηλό υπολογιστικό κόστος στη μέθοδο των Hayes και Efros επιτυγχάνονται με μία αναζήτηση δύο σταδίων. Πρώτα ανιχνεύονται σκηνές που σχετίζονται σημασιολογικά (*scene matching*) και στη συνέχεια μέσα σε αυτές αναζητούνται «μπαλώματα» που ταιριάζουν καλύτερα με τα εγγύτερα στο σύνορο της ελλείπουσας περιοχής δεδομένα (*template matching*). Στη συνέχεια για τη συνένωση χρησιμοποιείται ανάμειξη Poisson, με παρόμοιο τρόπο με τις εργασίες που περιγράφονται στην Ενότητα 4.2, αφού όμως αρχικά βρίσκεται το σύνορο με την ελάχιστη δυνατή μεταβολή της παραγωγού των υπό ένωση εικόνων μέσω κατάτμησης με τομές γράφων.

Συμπλήρωση σκηνής αξιοποιώντας πολλαπλές εικόνες πραγματοποίησαν και οι Wilczkowiak et al. το 2005 [45]. Όμως σε αυτή την εργασία η βάση δεδομένων είναι σαφώς πιο περιορισμένη και αφορά επιλεγμένες από το χρήστη εικόνες. Άλλες σχετικές μέθοδοι για τη σύνθεση σημασιολογικά έγκυρων εικόνων, είτε ως προς το αναγραφόμενο κείμενο είτε ως προς της ιδιότητες της ανακτώμενης περιοχής, χρησιμοποιούν βάσεις δεδομένων με επισημειωμένες ως προς το περιεχόμενο περιοχές. Στην εργασία των Diakopoulos et al. [46] υπάρχει ένα στάδιο απόδοσης ετικετών από το χρήστη στις εικόνες της βάσης δεδομένων, δηλαδή αποτελεί μία επιβλεπόμενη διεργασία (*supervised labelling process*). Τέλος ημι-επιβλεπόμενη επισημείωση συναντάται στη μέθοδο των Johnson et al. [47].

Όσον αφορά την επίλυση του προβλήματος της συνένωσης εικόνων δύο είναι οι κύριες κατηγορίες μεθόδων. Προϋπόθεση σε κάθε περίπτωση είναι η *ευθυγράμμιση* των εικόνων. Οι αλγόριθμοι βέλτιστης συνένωσης [48] αναζητούν μία καμπύλη στην περιοχή επικάλυψης των εικόνων όπου η διαφορά τους είναι *minimal*, δηλαδή στη γειτονιά της καμπύλης δεν υπάρχει κανένα pixel με μικρότερη διαφορά στην ένταση της φωτεινότητας των αντίστοιχων σημείων στις αρχικές εικόνες σε σχέση με την αντίστοιχη διαφορά στο εκάστοτε pixel που ανήκει στην καμπύλη. Εν συνεχεία, κάθε εικόνα αντιγράφεται στην αντίστοιχη πλευρά της καμπύλης αυτής. Όταν πάνω στην καμπύλη η διαφορά είναι μηδενική, το μωσαϊκό που παράγεται είναι οπτικά οπτικά άριστο. Συνήθως δε συμβαίνει όμως αυτό και εμφανίζονται ορατές «ραφές» στη σύνθεση.

Στη δεύτερη κατηγορία μεθόδων ελαχιστοποιούνται οι ατέλειες μέσω της ομαλοποίησης της μετάβασης μεταξύ των εικόνων. Στην τεχνική *feathering* [49] η εικόνα-μωσαϊκό I είναι ένας συνδυασμός των εικόνων εισόδου I_1 και I_2 με βάρη που υπολογίζονται συναρτήσει της απόστασής τους από τη ραφή. Στην τεχνική *pyramid blending* [50] συναντάται ξεχωριστός χειρισμός για τις διαφορετικές ζώνες συχνότητας. Η επεξεργασία στις χαμηλές συχνότητες γίνεται με δεδομένα μίας μεγαλύτερης επιφάνειας, ενώ οι περιοχές έντονων μεταβολών αναμειγνύονται σε μία στενή χωρική ζώνη. Έτσι, προκύπτει σταδιακή μετάβαση στις χαμηλές συχνότητες και αποφεύγεται η αναπαραγωγή των ακμών στις περιοχές υψής και έντονων λεπτομερειών. Στην κατηγορία αυτή ανήκουν και οι αλγόριθμοι που περιγράφονται στις Ενότητες 4.2.2 και 4.2.3.

4.2 Οπτικά βέλτιστη συνένωση εικόνων (seamless image stitching)

4.2.1 Εργασία στο πεδίο των παραγώγων

Τόσο στην εργασία των Tao et al. όσο και στην εργασία των Levin et al. το «ράψιμο» των εικόνων γίνεται στο πεδίο των παραγώγων. Αν και απλή ως εφαρμογή, αυτή η διαδικασία συχνά προκαλεί ατέλειες στα τμήματα της «ραφής» που οι δύο υπό συνένωση εικόνες διαφέρουν σημαντικά. Έτσι, αναζητούνται τεχνικές για την εξάλειψη στο μέγιστο δυνατό βαθμό αυτών των ατελειών για το καλύτερο οπτικό αποτέλεσμα.

Γενικότερα, η εργασία στο πεδίο των παραγώγων εμφανίζεται στον πυρήνα πολλών εφαρμογών της Όρασης Υπολογιστών, όπως η σύνθεση πανοράματος [5], η συμπλήρωση σκηνής [38] η δημιουργία φωτογραφιών πολύ υψηλής ανάλυσης [51], η σύνθεση υψής [52] και φυσικά η ενδο-συμπλήρωση [53]. Όλες αυτές οι μέθοδοι πρώτα ευθυγραμμίζουν τις προς συνένωση περιοχές, στη συνέχεια υπολογίζουν ένα πεδίο παραγώγων στόχο και τις συνοριακές συνθήκες αυτών των περιοχών και τελικά επιλύουν την εξίσωση Poisson για την εύρεση της τελικής εικόνας.

Συγκεκριμένα, στην εφαρμογή της συνένωσης δύο εικόνων στόχος είναι η δημιουργία ενός οπτικά «πειστικού» μωσαϊκού με τις εξής ιδιότητες: το μωσαϊκό να είναι όσο ομοιότερο γίνεται με τις αρχικές εικόνες τόσο στη γεωμετρία όσο και στη φωτομετρία του, αλλά και η «ραφή» μεταξύ των δυο εικόνων να είναι κατά το δυνατόν αόρατη, δηλαδή η μετάβαση μεταξύ των δύο εικόνων να είναι ομαλή. Αν και οι απαιτήσεις αυτές είναι ευρέως αποδεκτές, τα κριτήρια ποιότητας διαφοροποιούνται μεταξύ των δημοσιεύσεων. Το πλεονέκτημα εργασίας στο πεδίο των παραγώγων είναι η ανεξαρτησία από τη μέση τιμή της έντασης των εικόνων και η χαμηλότερη ευαισθησία σε μικρές διαφορές μεταξύ των δύο εικόνων που οφείλονται σε μη ομοιόμορφη φωτομετρική απόκριση της κάμερας και σε αποκλίσεις στη σκίαση των απεικονιζόμενων σκηνών. Ως παράδειγμα συνένωσης δύο εικόνων παρατίθεται το Σχήμα 4.1, το οποίο έχει ληφθεί από την εργασία των Levin et al., όπου δύο εικόνες I_1 και I_2 απεικονίζουν γειτονικές ως προς τον ορίζοντα περιοχές του ίδιου τοπίου και συνενώνονται για τη δημιουργία του μωσαϊκού I .

Στον αλγόριθμό τους οι Levin et al. υπολογίζουν την εικόνα-μωσαϊκό I μέσω μίας διαδικασίας βελτιστοποίησης που χρησιμοποιεί τις παραγώγους των εικόνων. Η βελτιστοποίηση γίνεται ως προς κάποιες συναρτήσεις κόστους που αξιολογούν την προσαρμογή του μωσαϊκού στις δύο υπό σύνθεση εικόνες και την ομαλότητα της μετάβασης μεταξύ των δύο εικόνων. Στην ακόλουθη υποενότητα περιγράφεται εκτενώς η μέθοδός τους.

4.2.2 Η μέθοδος των Levin, Zomet, Peleg, Weiss

Το 2004 οι Levin et al. εισήγαγαν δύο παρεμφερείς προσεγγίσεις συνένωσης εικόνων. Σε κάθε περίπτωση η βελτιστοποίηση έγινε στο πεδίο των παραγώγων, κάτι που είχε πολλαπλά οφέλη και απέδωσε οπτικά ανώτερα αποτελέσματα σε σχέση με τα προϋπάρχοντα στην περιοχή.

Η πρώτη προσέγγιση υπολογίζει τη σύνθεση ελαχιστοποιώντας μία συνάρτηση κόστους E_p , η οποία είναι ένα μέτρο απόστασης μεταξύ των παραγώγων της εικόνας σύνθεσης και των παραγώγων των αρχικών εικόνων. Πιο συγκεκριμένα, θεωρούμε I_1 και I_2 δύο ευθυγραμμισμένες εικόνες εισόδου. Έστω τ_1 , τ_2 οι περιοχές που ανήκουν αποκλειστικά στις δύο εικόνες και ω η



Σχήμα 4.1: Δίνονται δύο φωτογραφικές λήψεις I_1 και I_2 του ίδιου τοπίου που απεικονίζουν γειτονικές ως προς τον ορίζοντα περιοχές του $\tau_1 \cup \omega$ και $\omega \cup \tau_2$. Η επικόλλησή τους δίπλα-δίπλα φυσικά δεν έχει το επιθυμητό αποτέλεσμα και δημιουργείται μία «ραφή» (π.χ. λόγω μη ομοιόμορφης φωτομετρικής απόκρισης της κάμερας). Ο αλγόριθμος των Levin et al. αποδίδει ένα φυσικό οπτικά αποτέλεσμα [5].

περιοχή επικάλυψής τους, όπου $\tau_1 \cap \tau_2 = \tau_1 \cap \omega = \omega \cap \tau_2 = \emptyset$. Και έστω W μία μάσκα με βάρη. Το αποτέλεσμα της συνένωσης I ορίζεται μέσω ελαχιστοποίησης του E_p ως προς το $\hat{I}$:

$$E_p(\hat{I}; I_1, I_2, W) = d_p(\nabla \hat{I}, \nabla I_1, \tau_1 \cup \omega, W) + d_p(\nabla \hat{I}, \nabla I_2, \omega \cup \tau_2, U - W) \quad (4.1)$$

όπου U ομοιόμορφη εικόνα, $d_p(J_1, J_2, \phi, W)$ η απόσταση μεταξύ των J_1, J_2 στο χωρίο ϕ :

$$d_p(J_1, J_2, \phi, W) = \sum_{\mathbf{q} \in \phi} W(\mathbf{q}) \|J_1(\mathbf{q}) - J_2(\mathbf{q})\|_p^p \quad (4.2)$$

και $\|\cdot\|_p$ η l_p νόρμα. Έτσι, στην κοινή περιοχή των δύο εικόνων ω η συνάρτηση κόστους «τιμωρεί» τις τιμές των παραγώγων που είναι ασυνεπείς με οποιαδήποτε από τις δύο εικόνες. Επιπλέον «τιμωρούνται» οι μεγάλες τιμές παραγώγων για την εικόνα-μωσαϊκό εκεί που και οι δύο εικόνες έχουν χαμηλές τιμές παραγώγου, χρήσιμη ιδιότητα για την αποφυγή δημιουργίας λάθους ακμών.

Η επιλογή της παραμέτρου p έχει συνέπειες αρχικά στον αλγόριθμο βελτιστοποίησης και εν συνεχεία στην παραγόμενη εικόνα. Η ελαχιστοποίηση του E_p για $p \geq 1$ είναι κυρτή κι έτσι μπορούν να χρησιμοποιηθούν αποδοτικοί αλγόριθμοι βελτιστοποίησης. Οι Levin et al. χρησιμοποιούν ένα σχήμα ελαχιστοποίησης του E_2 με βάση γνωστούς αλγορίθμους και ένα καινοτόμο αλγόριθμο γρήγορης ελαχιστοποίησης του E_1 . Το W είναι είτε ομοιόμορφη μάσκα (για το E_1)

είτε μάσκα *feathering* (για το E_2).

Στη συνέχεια παρουσιάζεται σχηματικά ο πρώτος αλγόριθμος των Levin et al. Αμέσως από κάτω ακολουθεί ο δεύτερος και απλούστερος αλγόριθμός τους για τη σύνθεση μωσαϊκού, μία διαδικασία δύο σταδίων:

Αλγόριθμος 4: Ο 1ος αλγόριθμος για σύνθεση μωσαϊκού των Levin et al.

- Έστω I_1 και I_2 δύο ευθυγραμμισμένες εικόνες εισόδου, τ_1 και τ_2 οι περιοχές που ανήκουν αποκλειστικά σε μία εκ των δύο και ω η περιοχή επικάλυψής τους, όπου $\tau_1 \cap \tau_2 = \tau_1 \cap \omega = \omega \cap \tau_2 = \emptyset$. Και έστω W μία μάσκα με βάρη.
- Το αποτέλεσμα της συνένωσης I ορίζεται μέσω ελαχιστοποίησης του E_p ως προς το $\hat{I}$:

$$E_p(\hat{I}; I_1, I_2, W) = d_p(\nabla \hat{I}, \nabla I_1, \tau_1 \cup \omega, W) + d_p(\nabla \hat{I}, \nabla I_2, \omega \cup \tau_2, U - W)$$

όπου U ομοιόμορφη εικόνα και $d_p(J_1, J_2, \phi, W)$ η απόσταση μεταξύ των J_1, J_2 στο χωρίο ϕ :

$$d_p(J_1, J_2, \phi, W) = \sum_{\mathbf{q} \in \phi} W(\mathbf{q}) \|J_1(\mathbf{q}) - J_2(\mathbf{q})\|_p^p$$

και $\|\cdot\|_p$ η l_p νόρμα. Επιλέγεται $p \geq 1$.

Αλγόριθμος 5: Ο 2ος αλγόριθμος για σύνθεση μωσαϊκού των Levin et al.

$\Rightarrow$ *Σύνθεση παραγώγων*

Υπολογίζονται οι παράγωγοι των αρχικών εικόνων και συνενώνονται ώστε να δημιουργηθεί ένα πεδίο F . Χρησιμοποιείται ένας κλασικός αλγόριθμος, όπως pyramid blending, feathering, εύρεσης βέλτιστης καμπύλης.

$\Rightarrow$ *Ελαχιστοποίηση*

Υπολογίζεται η εγγύτερη στο πεδίο F εικόνα-μωσαϊκό, ελαχιστοποιώντας την απόσταση $d_p(\nabla I, F, \pi, U)$, όπου π ολόκληρη η εικόνα και U η ομοιόμορφη εικόνα.

Το ποια προσέγγιση εκ των δύο επιλέγεται εξαρτάται από την εκάστοτε περίπτωση και είναι κάτι που πέρα από τα αντικειμενικά κριτήρια αξιολογείται αισθητικά από το χρήστη. Στη συνέχεια, σε μία προσπάθεια μαθηματικής θεμελίωσης παρατίθενται δύο θεωρήματα μαζί με τις αποδείξεις τους που έχουν αντληθεί από το [54]. Έτσι, εξάγονται κάποια τυπικά συμπεράσματα για την επιλογή κατάλληλου αλγορίθμου κάθε φορά.

Θεώρημα 4.1. Έστω I_1 και I_2 δύο εικόνες εισόδου για έναν αλγόριθμο συνένωσης και έστω ότι υπάρχει μία καμπύλη $x = f(y)$, τέτοια ώστε για κάθε $\mathbf{q} \in \{(f(y), y)\}$, $I_1(\mathbf{q}) = I_2(\mathbf{q})$. Και έστω U μία ομοιόμορφη εικόνα. Τότε η λύση βέλτιστης ραφής I που ορίζεται παρακάτω είναι ίδια με την εικόνα-μωσαϊκό που παράγεται μέσω ελαχιστοποίησης του $E_p(I; I_1, I_2, U)$, το οποίο ορίζεται μέσω της Εξίσωσης 4.1 για κάθε $0 < p \leq 1$.

$$I = \begin{cases} I_1(x, y) & x < f(y) \\ I_2(x, y) & x \geq f(y) \end{cases}$$

Απόδειξη.

Έστω ω η περιοχή επικάλυψης των δύο εικόνων, ψ_1 η περιοχή όπου $x < f(y)$ και ψ_2 η περιοχή όπου $x \geq f(y)$. Η τιμή του $E_p(I; I_1, I_2, U)$ επί της εικόνας I είναι:

$$\begin{aligned} E_p(I; I_1, I_2, U) &= 0 + d_p(\nabla I, \nabla I_1, \omega, U) + d_p(\nabla I, \nabla I_2, \omega, U) \\ &= \sum_{\mathbf{q} \in \omega \cap \psi_2} \|\nabla I(\mathbf{q}) - \nabla I_1(\mathbf{q})\|_p^p + \sum_{\mathbf{q} \in \omega \cap \psi_1} \|\nabla I(\mathbf{q}) - \nabla I_2(\mathbf{q})\|_p^p \\ &= \sum_{\mathbf{q} \in \omega} \|\nabla I_2(\mathbf{q}) - \nabla I_1(\mathbf{q})\|_p^p \end{aligned}$$

Και έστω $\hat{I}$ μία άλλη λύση για την εικόνα-μωσαϊκό. Χρησιμοποιώντας το ότι l_p κοίλο για $0 < p \leq 1$, δείχνουμε ότι $E_p(\hat{I}; I_1, I_2, U) \geq E_p(I; I_1, I_2, U)$:

$$\begin{aligned} E_p(\hat{I}; I_1, I_2, U) &= d_p(\nabla \hat{I}, \nabla I_1, \omega, U) + d_p(\nabla \hat{I}, \nabla I_2, \omega, U) + d_p(\nabla \hat{I}, \nabla I_1, \tau_1, U) + d_p(\nabla \hat{I}, \nabla I_2, \tau_2, U) \\ &\geq \sum_{\mathbf{q} \in \omega} (\|\nabla \hat{I}(\mathbf{q}) - \nabla I_1(\mathbf{q})\|_p^p + \|\nabla \hat{I}(\mathbf{q}) - \nabla I_2(\mathbf{q})\|_p^p) \\ &\geq \sum_{\mathbf{q} \in \omega} \|\nabla I_2(\mathbf{q}) - \nabla I_1(\mathbf{q})\|_p^p \\ &= E_p(I; I_1, I_2, U) \end{aligned}$$

□

Το Θεώρημα 4.1 δηλώνει ότι η πρώτη προσέγγιση των Levin et al. με χρήση l_1 νόρμας είναι το ίδιο αποτελεσματική ως προς την εξάλειψη γεωμετρικών ατελειών όσο και οι μέθοδοι εύρεσης της βέλτιστης «ραφής» όταν αυτή υπάρχει. Βέβαια, η προσέγγισή τους πλεονεκτεί των μεθόδων αυτών όταν η «τέλεια ραφή» δεν υπάρχει, όπως για παράδειγμα λόγω φωτομετρικών ασυνεπειών.

Το επόμενο θεώρημα δείχνει την ισοδυναμία της 1ης προσέγγισης των Levin et al. υπό την l_2 νόρμα και της τεχνικής *feathering* υπό την l_2 νόρμα, δηλαδή της 2ης προσέγγισής τους με χρήση του αλγορίθμου feathering στο πρώτο στάδιό της.

Θεώρημα 4.2. Έστω I_1 και I_2 δύο εικόνες εισόδου για έναν αλγόριθμο σύνθεσης και έστω W μία μάσκα feathering. Υποθέτουμε ότι ω , η περιοχή επικάλυψης των I_1 και I_2 , είναι ολόκληρη η εικόνα (χωρίς βλάβη της γενικότητας, καθώς $W(\mathbf{q}) = 1$ για $\mathbf{q} \in \tau_1$ και $W = 0$ για $\mathbf{q} \in \tau_2$). Έστω I η λύση που προκύπτει από ελαχιστοποίηση του $E_p(I; I_1, I_2, W)$ μέσω της Εξίσωσης 4.1 και F το ακόλουθο πεδίο:

$$F = W(\mathbf{q})\nabla I_1(\mathbf{q}) + (1 - W(\mathbf{q}))\nabla I_2(\mathbf{q})$$

Τότε η I είναι η εικόνα με το εγγύτερο πεδίο παραγώγων στο F υπό την l_2 νόρμα.

Απόδειξη.

Έστω f_x και f_y οι τελεστές παραγώγου συνέλιξης και έστω $F_{xj} = I_j * f_x$, $F_{yj} = I_j * f_y$. Τότε:

$$E_2(I; I_1, I_2, W) = \sum_{\mathbf{q} \in \omega} W(\mathbf{q}) [((I * f_x)(\mathbf{q}) - F_{x1}(\mathbf{q}))^2 + ((I * f_y)(\mathbf{q}) - F_{y1}(\mathbf{q}))^2] + \\ \sum_{\mathbf{q} \in \omega} (1 - W(\mathbf{q})) [((I * f_x)(\mathbf{q}) - F_{x2}(\mathbf{q}))^2 + ((I * f_y)(\mathbf{q}) - F_{y2}(\mathbf{q}))^2]$$

Ορίζουμε P_W το γραμμικό τελεστή που πολλαπλασιάζει κάθε σημείο της εικόνας $\mathbf{q}$ με ένα βάρος $W(\mathbf{q})$ και έστω f_x^r , f_y^r οι συμπληρωματικοί τελεστές των f_x , f_y . Λαμβάνοντας τις παραγώγους του $E_2(I; I_1, I_2, W)$ ως προς I και αφού τις εξισώσουμε με το 0, παίρνουμε για κάθε σημείο της εικόνας $\mathbf{q}$:

$$f_x^r * P_W(I * f_x) + f_x^r * P_{1-W}(I * f_x) + f_y^r * P_W(I * f_y) + f_y^r * P_{1-W}(I * f_y) = \\ f_x^r * P_W(F_x^1) + f_x^r * P_{1-W}(F_x^2) + f_y^r * P_W(F_y^1) + f_y^r * P_{1-W}(F_y^2)$$

Παρατηρώντας ότι $P_{1-W} + P_W$ είναι ο ταυτοτικός τελεστής, παίρνουμε την εξίσωση Poisson:

$$\Delta I(\mathbf{q}) = \text{div}(W(\mathbf{q}))F_1(\mathbf{q}) + (1 - W(\mathbf{q}))F_2(\mathbf{q})$$

όπου Δ είναι ο τελεστής Laplace, $F_1 = (F_{x1}, F_{y1})$, $F_2 = (F_{x2}, F_{y2})$. Έτσι, το ελάχιστο του $E_2(I; I_1, I_2)$ λαμβάνεται με την εικόνα που έχει τις παραγώγους της εγγύτερα στο πεδίο F υπό την l_2 νόρμα.

□

Το Θεώρημα 4.2 υπαγορεύει ότι η 1η προσέγγιση των Levin et al. υπό την l_2 νόρμα είναι ισοδύναμη με την τεχνική *feathering* υπό την l_2 νόρμα, ακολουθούμενη από την επίλυση της εξίσωσης Poisson.

4.2.3 Η μέθοδος των Tao, Johnson, Paris

Ένα πρόβλημα στη σύνθεση παραγώγων είναι ότι το πεδίο που προκύπτει από το συνδυασμό της πληροφορίας των δύο εικόνων (combined gradient field) ενδέχεται να μην είναι ολοκληρώσιμο. Μία εικόνα με παραγώγους που ταιριάζουν στο πεδίο στόχο και στις συνοριακές συνθήκες μπορεί να μην υπάρχει. Κάτι τέτοιο οδηγεί στη διαρροή χρωμάτων, στην εμφάνιση «φωτοστέφανων», σε ασυνέχειες και άλλες ατέλειες. Σε κάποιες εργασίες γίνεται απλώς μετακίνηση του συνόρου σε κατάλληλη περιοχή, όπως οι αλγόριθμοι βέλτιστης καμπύλης, κάτι που δεν αποτελεί ασφαλή λύση όταν οι εικόνες διαφέρουν σημαντικά ή όταν το σύνορο είναι προκαθορισμένο.

Στην εργασία τους το 2010 οι Tao et al. παρουσιάζουν έναν αλγόριθμο δύο βημάτων για την ελαχιστοποίηση των ατελειών αυτών, χαρακτηρίζοντας την προέλευση των τυπικών ατελειών και εντοπίζοντας τις περιοχές που θα ήταν περισσότερο ή λιγότερο εμφανείς. Στο πρώτο βήμα της μεθόδου επεξεργάζονται τις τιμές των παραγώγων στο σύνορο για να ελαχιστοποιήσουν τις ατέλειες στις περιοχές που θα ήταν περισσότερο εμφανή. Στο δεύτερο βήμα τους χρησιμοποιούν ένα σταθμισμένο σχήμα ολοκλήρωσης για την ανακατασκευή της εικόνας από το πεδίο παραγώγων έτσι ώστε τα υπόλοιπα να τοποθετούνται στις περιοχές υψής, όπου είναι λιγότερο ορατά. Ο συνδυασμός των δύο βημάτων έδωσε σπουδαία αποτελέσματα. Επιπλέον, η μέθοδός τους συγκροτείται ως βελτιστοποίηση ελαχίστων τετραγώνων που επιλύεται με χρήση ενός αραιού γραμμικού

συστήματος, κάτι που καθιστά τη μέθοδο υπολογιστικά αποδοτική.

Η γενική μορφή του προβλήματος ανακατασκευής αφορά την εύρεση μιας εικόνας I που προσεγγίζει το πεδίο-στόχο $\mathbf{v}$ με την έννοια των ελαχίστων τετραγώνων

$$\operatorname{argmin}_I \int \|\nabla I - \mathbf{v}\|^2 \quad (4.3)$$

που ελαχιστοποιείται λύνοντας την εξίσωση Poisson

$$\Delta I - \operatorname{div}(\mathbf{v}) = 0 \quad (4.4)$$

όπου Δ είναι ο τελεστής Laplace $\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$ και div είναι ο τελεστής απόκλισης $\frac{\partial}{\partial x} + \frac{\partial}{\partial y}$.

Για την επίλυση της εξίσωσης 4.4 απαιτούνται συνοριακές συνθήκες που εξαρτώνται από την εφαρμογή. Στην περίπτωση της συνένωσης εικόνων το πεδίο στόχος $\mathbf{v}$ υπολογίζεται με βάση την εικόνα παρασκήνιου (background) B , την εικόνα προσκήνιου (foreground) F και μια επιλογή για το προσκήνιο S με ένα σύνορο β ως εξής:

$$\mathbf{v}(x, y) = \begin{cases} \nabla F & \text{αν } (x, y) \in S, (x, y) \notin \beta \\ \nabla B & \text{αν } (x, y) \notin S \\ \frac{1}{2}(\nabla F + \nabla B) & \text{αν } (x, y) \in \beta \end{cases} \quad (4.5)$$

Οι παράγωγοι στο προσκήνιο και στο παρασκήνιο είναι προφανώς ολοκληρώσιμες, καθώς υπολογίζονται κατευθείαν από τις εικόνες. Αυτό όμως δεν είναι εξασφαλισμένο στο σύνορό τους, δημιουργώντας σφάλματα που πρέπει να αντιμετωπιστούν κατά την ολοκλήρωση. Στην εργασία των Farbman et al. [55] το συγκεκριμένο ζήτημα αντιμετωπίζεται από το χρήστη, ο οποίος αναγνωρίζει τις περιοχές και τις μαρκάρει. Οι παράγωγοι των μαρκαρισμένων περιοχών αγνοούνται, κάτι που οδηγεί στην απομάχρυνση των διαρροών. Αντιθέτως, στην εργασία των Tao et al. αναλύεται η εικόνα ώστε αυτόματα να προσαρμόζεται η διαδικασία ολοκλήρωσης. Η μέθοδος τους φέρει ομοιότητες με την εργασία των Lalonde et al. [56] που προτείνουν η διαδικασία ανακατασκευής να γίνεται με βάση το μέτρο των παραγώγων της εικόνας. Και πάλι όμως χρωματικές διαρροές είναι πιθανό να εμφανιστούν στην προσέγγιση αυτή όταν τα σύνορα δεν είναι ορισμένα ακριβώς. Ακολούθως, παρουσιάζονται οι καινοτομίες του αλγορίθμου των Tao et al.

Ελαχιστοποίηση στροβιλισμού (*curl*) στο σύνορο των δύο εικόνων

Αναγκαία συνθήκη ώστε ένα πεδίο παραγώγων $\mathbf{u}$ να είναι ολοκληρώσιμο είναι να έχει μηδενικό στροβιλισμό, κάτι που δεν ισχύει απαραίτητα στο κοινό σύνορο των δύο εικόνων, όπου η παράγωγος προκύπτει από την ανάμειξη των (ολοκληρώσιμων) παραγώγων δύο εικόνων. Όταν ο στροβιλισμός είναι μη μηδενικός, η εξίσωση Poisson δεν ελαχιστοποιείται ακριβώς και παραμένουν υπόλοιπα που δημιουργούν ορατές ατέλειες.

Στόχος λοιπόν είναι η ελαχιστοποίηση του στροβιλισμού στο σύνορο των δύο υπό ταίριασμα εικόνων. Έστω $\mathbf{v}$ το πεδίο παραγώγων στόχος. Μία αφελής λύση θα ήταν να απαιτηθεί $\operatorname{curl}(\mathbf{v}) = 0$ οπουδήποτε. Όμως το ακόλουθο αντιπαράδειγμα καταδεικνύει ότι αυτή η προσέγγιση δε θα πετύχαινε το στόχο. Έστω ότι η εικόνα προσκήνιου τοποθετείται όπως είναι στην εικόνα παρασκήνιου, δημιουργώντας την εικόνα I_{seam} που προφανώς έχει ευδιάκριτες «ραφές»

Αλγόριθμος 6: Οι κύριες συνεισφορές του αλγορίθμου των Tao et al. για εύρωστη σύνθεση μωσαϊκού

⇒ *Σύνορα χαμηλού στροβιλισμού*

Ελαχιστοποιώντας το στροβιλισμό των παραγώγων-στόχων στο σύνορο προσκηνίου-παρασκηνίου, περιορίζονται οι ατέλειες.

⇒ *Σταθμισμένη εξίσωση Poisson*

Με κατάλληλη στάθμιση της εξίσωσης Poisson τα υπόλοιπα της ολοκλήρωσης συγκεντρώνονται στις περιοχές υψής, όπου είναι λιγότερο εμφανή.

⇒ *Αποτελεσματική συνένωση εικόνων*

Οι δύο τεχνικές συνθέτουν έναν αποτελεσματικό αλγόριθμο συνένωσης εικόνων, ο οποίος παραμένει γραμμικός ως προς την αυθεντική εξίσωση Poisson και εξαλείφει σε μεγάλο βαθμό τις ατέλειες που δημιουργούνται στην περίπτωση αυτή.

στη γενική περίπτωση. Ο στροβιλισμός είναι μηδενικός οπουδήποτε στη νέα εικόνα αφού υπολογίζεται με τα δεδομένα των αρχικών εικόνων, αλλά η εικόνα που παράγεται δεν είναι σε καμία περίπτωση η επιθυμητή.

Για την αντιμετώπιση του παραπάνω προβλήματος απαιτείται η εύρεση τιμών παραγώγων που ελαχιστοποιούν το στροβιλισμό και είναι κοντά στις τιμές παραγώγων των αυθεντικών εικόνων εισόδου. Χρησιμοποιείται ενέργεια ελαχίστων τετραγώνων, ενώ οι παράγωγοι-στόχοι $\mathbf{v}$ στο σύνορο των εικόνων είναι οι άγνωστοι. Ο πρώτος όρος στην παράσταση 4.6 ελαχιστοποιεί το στροβιλισμό και ο δεύτερος κρατάει τις ζητούμενες τιμές κοντά στις αντίστοιχες των αρχικών εικόνων:

$$\arg \min_{\mathbf{v}} \int_{\beta} ([\text{curl}(\mathbf{v})]^2 + W_{\beta} [\mathbf{v} - \frac{1}{2}(\nabla B + \nabla F)]^2) \quad (4.6)$$

όπου η σταθερά W_{β} ελέγχει τη σημασία του δεύτερου όρου. Για μεγάλες τιμές του W_{β} ελαχιστοποιείται κυρίως η απόσταση του $\mathbf{v}$ από τις παραγώγους εισόδου, κάτι που οδηγεί σε απόλυτη έλλειψη ραφών αλλά ατέλειες διαρροών χρώματος, θόλωμα κ.λπ., ενώ για μικρές τιμές του W_{β} ελαχιστοποιείται κυρίως ο στροβιλισμός του $\mathbf{v}$, κάτι που οδηγεί σε εικόνα που μοιάζει με εκείνη που παράγει η προαναφερθείσα αφελής λύση, όπου παρουσιάζονται ευδιάκριτες ραφές. Στη μέθοδο των Tao et al. γίνεται συνδυασμός των δυο μηχανισμών μεταβάλλοντας κατάλληλα τις τιμές του W_{β} , ανάλογα με την τοπική δομή της εικόνας (adaptive approach).

Διαισθητικά, ένα σύνορο δίχως ραφές (μεγάλη W_{β}) απαιτείται όταν εκατέρωθεν της ραφής είτε και οι δύο περιοχές είναι ομαλές είτε έχουν υψή, διότι η διαρροή είναι ένα φαινόμενο χαμηλής συχνότητας που αποχρύπτεται σε αυτές τις περιπτώσεις. Όταν όμως μία περιοχή υψής συνενώνεται με μία ομαλή περιοχή υπάρχει σοβαρός κίνδυνος δημιουργίας διαρροών και άλλου τύπου ατελειών στην ομαλή περιοχή. Σε αυτή την περίπτωση απαιτείται κυρίως ελαχιστοποίηση του στροβιλισμού (μικρή W_{β}). Έτσι, και οι διαρροές ελαχιστοποιούνται, ενώ σε μία τέτοια σύνθεση η διακριτική εμφάνιση «ραφής» δε δημιουργεί προβληματικό οπτικά αποτέλεσμα. Ακολούθως αναλύεται πώς γίνεται ο υπολογισμός της σταθεράς W_{β} .

Η στρατηγική των Tao et al. στηρίζεται στην ανίχνευση υψής. Για το σκοπό αυτό χρησιμο-

ποιούν το ακόλουθο σχήμα:

$$T_{\sigma_1, \sigma_2}(\mathbf{g}) = \frac{G_{\sigma_1} \otimes |\mathbf{g}|}{G_{\sigma_2} \otimes |\mathbf{g}|} n(|\mathbf{g}|) \quad (4.7)$$

όπου $\mathbf{g}$ είναι ένα πεδίο παραγώγων, G_σ είναι μία Gaussian πλάτους σ , σ_1 και σ_2 είναι δύο παράμετροι έτσι ώστε $\sigma_1 < \sigma_2$, $\otimes$ είναι ο τελεστής συνέλιξης και $n(\cdot)$ μία συνάρτηση που ελέγχει το θόρυβο. Το ανωτέρω σχήμα ελέγχει τις τιμές των παραγώγων στις δύο γειτονιές που καθορίζονται από τα σ_1 και σ_2 . Όταν $T > 1$ η εικόνα έχει τοπικά περισσότερη υφή, ενώ όταν $T < 1$ η εικόνα έχει τοπικά λιγότερη υφή σε σχέση με την ευρύτερη περιοχή. Το σχήμα αυτό θα ήταν ευαίσθητο σε θόρυβο στις ομαλές περιοχές. Το συγκεκριμένο ζήτημα αντιμετωπίζεται με τη συνάρτηση ελέγχου θορύβου n που ορίζεται ως εξής:

$$n = \begin{cases} 0, & \text{για περιοχές με πολύ χαμηλές παραγώγους} \\ 1, & \text{διαφορετικά} \end{cases} \quad (4.8)$$

Στόχος είναι η διάκριση της περιοχής χωρίς υφή από την περιοχή με υφή. Μεταξύ των περιοχών με υφή είναι αδιάφορο το ποια έχει περισσότερη ή λιγότερη υφή. Έτσι, χρησιμοποιείται ο εκτιμητής $\bar{T} = \min(1, T)$ που ταξινομεί όλες τις περιοχές με υφή στην ίδια κλάση. Με βάση την ανάλυση του αλγορίθμου παραπάνω, απαιτούνται μεγάλες τιμές του W_β όταν προσκίνηιο και παρασκήνιο έχουν υφή είτε είναι και οι δύο ομαλές και αντίθετα χαμηλές τιμές του W_β όταν η μια περιοχή έχει υφή και η άλλη όχι. Έστω η διαφορά $D = |\bar{T}(\nabla F) - \bar{T}(\nabla B)|$ και w μία αρκούντως μικρή τιμή. Το W_β ορίζεται ως εξής:

$$W_\beta = \begin{cases} w & \text{αν } D > \tau \\ \min(1, w + \lambda(1 - D/\tau)) & \text{διαφορετικά} \end{cases} \quad (4.9)$$

όπου λ και τ ελέγχουν τη γραμμική μετάβαση. Οι Tao et al. μέσω των πειραμάτων τους βρήκαν ότι για $\lambda = 4$, $w = 0.05$, $\tau = 0.005$, $\sigma_1 = 0.5$ και $\sigma_2 = 2$ ο αλγόριθμος αποδίδει καλά. Συνεπώς, όπως ήταν επιθυμητό, στο W_β αποδίδεται μία μικρή τιμή w για μεγάλο D , η μονάδα για μικρό D και γραμμικές μεταβάσεις ενδιάμεσα.

Ελέγχοντας την τοποθεσία των υπολοίπων της ολοκλήρωσης

Αν και ο αλγόριθμος ελαχιστοποιεί το στροβιλισμό του πεδίου παραγώγων $\mathbf{v}$, γενικά το πεδίο παραμένει μη ολοκληρώσιμο. Η στρατηγική των Tao et al. είναι να τροποποιήσουν των εξίσωση Poisson έτσι ώστε τα υπόλοιπα της ολοκλήρωσης να τοποθετηθούν στις περιοχές που είναι λιγότερο εμφανή. Διαισθητικά, πρέπει να αποφευχθούν τα σφάλματα στις ομαλές περιοχές όπου είναι εμφανή και τοποθετούνται στις περιοχές υφής όπου η οπτική μάσκα (*visual masking*) θα τα αποκρύψει.

Έστω μία σταθερά W_P με υψηλές τιμές στις περιοχές που τα σφάλματα θα ήταν ορατά και χαμηλές τιμές αλλού. Δεδομένου του W_P , χρησιμοποιείται η ακόλουθη έκφραση, έτσι ώστε να «τιμωρούνται» λιγότερο οι περιοχές όπου τα υπόλοιπα είναι προτιμότερο να εμφανίζονται:

$$\operatorname{argmin}_I \int W_P \|\nabla I - \mathbf{v}\|^2 \quad (4.10)$$

Φυσικά, αφού πρέπει να ελαχιστοποιηθεί η διαφορά των ∇I και $\mathbf{v}$, το W_P λαμβάνεται θετικό. Για την αποδοτικότητα της μεθόδου το W_P σχεδιάζεται ανεξάρτητο της εικόνας I . Έτσι προκύπτει ένα κλασικό συναρτησιακό ελαχίστων τετραγώνων, το οποίο ελαχιστοποιείται λύνοντας

ένα γραμμικό σύστημα. Εφαρμόζοντας τη φόρμουλα Euler-Lagrange στην 4.10, προκύπτει το ακόλουθο γραμμικό σύστημα:

$$\operatorname{div}(W_P(\nabla I - \mathbf{v})) = 0 \quad (4.11)$$

Για να διατηρηθεί γραμμικό καμία ποσότητα δε σχετίζεται με την εικόνα I . Για την εκτίμηση της υψής χρησιμοποιείται το πεδίο παραγώγων $\mathbf{v}$, το οποίο δεν περιλαμβάνει τις παραγώγους της τελικής εικόνας αλλά είναι μία ικανοποιητική προσέγγισή της. Καθώς είναι επιθυμητές υψηλές τιμές για το W_P στις ομαλές περιοχές και χαμηλές στις περιοχές υψής, χρησιμοποιείται το ακόλουθο σχήμα:

$$W_P = 1 - p\bar{T}(\mathbf{v}) \quad (4.12)$$

όπου p είναι μία παράμετρος που έχει να κάνει με το πόσος έλεγχος ασκείται στην τοποθεσία των υπολοίπων. Για παράδειγμα, για $p = 0$ δεν ασκείται καθόλου έλεγχος που είναι η κλασική εξίσωση Poisson, ενώ για μεγαλύτερες τιμές ελέγχεται η διαδικασία. Βέβαια πρέπει $p < 1$ ώστε $W_P > 0$. Ο Tao et al. διαπίστωσαν ότι για $p = 0.999$ προκύπτουν τα καλύτερα αποτελέσματα.

Ανεξάρτητα του ορισμού του W_P μπορεί ναδειχθεί ότι τα παραγόμενα υπόλοιπα της μεθόδου έχουν δομή που είναι ευθυγραμμισμένη με το περιεχόμενο της εικόνας. Οι Wang et al. [57] έχουν δείξει ότι αυτή η ευθυγράμμιση παράγει καλύτερα αποτελέσματα. Για τη βαθύτερη κατανόηση του ρόλου του W_P , για $W_P \neq 0$, η Εξίσωση 4.4 ξαναγράφεται ως εξής:

$$\underbrace{\Delta I - \operatorname{div}(\mathbf{v})}_{\text{όρος Poisson}} + \underbrace{\nabla \log W_P \cdot (\nabla I - \mathbf{v})}_{\text{πρόσθετος όρος}} = 0 \quad (4.13)$$

Ο πρώτος όρος είναι ο γνωστός, ενώ ο δεύτερος είναι μηδενικός για σταθερό W_P , όταν το σχήμα συμπεριφέρεται σαν την κλασική εξίσωση Poisson και τα υπόλοιπα επεκτείνονται τοπικά ομοιόμορφα στο χώρο. Για μεταβλητό W_P , όμως, το σχήμα ξεφεύγει από την κλασική Poisson και επιτρέπει ασυνέχειες στα υπόλοιπα. Καθώς το W_P μετρά την ποσότητα της υψής, οι μεταβολές των υπολοίπων ευθυγραμμίζονται με τις ακμές υψής, κάτι που εξασφαλίζει την ομοιότητα δομής που έχει υποδειχθεί ως επιθυμητή από τους Wang et al. [57].



Σχήμα 4.2: Για την αποφυγή των οπτικών ατελειών που ανακύπτουν στην απλή σύνθεση Poisson, ο αλγόριθμος των Tao, Johnson & Paris υλοποιείται σε δύο φάσεις. Στην πρώτη φάση προσαρμόζονται οι παράγωγοι της έντασης στο σύνορο μέσω της ελαχιστοποίησης του στροβιλισμού και στη δεύτερη φάση σταθμίζεται η διαδικασία ανακατασκευής έτσι ώστε τα υπόλοιπα να συγκεντρώνονται κυρίως στις περιοχές υψής. Ο αλγόριθμος των Tao et al. αποδίδει ένα άψογο οπτικά αποτέλεσμα [6].

4.3 Εφαρμογή οπτικά βέλτιστης συνένωσης εικόνων στις τοιχογραφίες

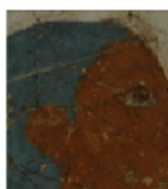
Η εφαρμογή του αλγορίθμου των Tao et al. για την οπτικά βέλτιστη συνένωση εικόνων στις τοιχογραφίες έδωσε κάποια πολύ ενδιαφέροντα αποτελέσματα, καθώς υπήρξε η δυνατότητα συμπλήρωσης μεγάλων κενών με πολύ πειστικό στο ανθρώπινο μάτι τρόπο, τα οποία χωρίς την ανάκτηση πληροφορίας από άλλες περιοχές των τοιχογραφιών θα ήταν αδύνατο να καλυφθούν. Τα Σχήματα 4.3 και 4.4 αναπαριστούν το ανώτερο τμήμα του σώματος του μεγαλύτερου άνδρα της τοιχογραφίας *Ακολουθία Ανδρών*, ο οποίος βρίσκεται στο μέσο της, ενώ το Σχήμα 4.5 έχει ληφθεί από την ίδια τοιχογραφία και αφορά την ψηφιακή επιδιόρθωση τμήματος του υφάσματος που κρατά ο συγκεκριμένος άνδρας. Το Σχήμα 4.6 παρουσιάζει την επιδιόρθωση τμήματος της κορδέλας στα μαλλιά της Πότνιας από την τοιχογραφία *Κροκοσυλλέκτρια και Πότνια*.

Όπως φαίνεται στη διαδοχή των εικόνων στο Σχήμα 4.3, ανασύρεται το πρόσθιο τμήμα της κεφαλής από το νεότερο άνδρα, αναστρέφεται ως προς τον κάθετο άξονα της εικόνας και σε κατάλληλη κλίμακα προσαρτάται στην ελλείπουσα περιοχή της κεφαλής του μεγαλύτερου άνδρα. Για την καλύτερη προσαρμογή το ανασυρόμενο ανεστραμμένο τμήμα περιστράφηκε ωρολογιακά κατά 10 μοίρες. Το Σχήμα 4.4 αφορά την ίδια ανδρική φιγούρα και συγκεκριμένα το λαιμό της. Μέσω δύο σταδίων πραγματοποιείται η ψηφιακή αποκατάσταση αντλώντας πληροφορία κατάλληλης μορφολογίας από τη μέση και την πλάτη αντίστοιχα. Η επιδιόρθωση είναι ενδεικτική, αλλά με ανάλογες διαδικασίες μπορεί να αποκατασταθεί όλος ο κορμός του νεαρού άνδρα.

Τέλος, τα Σχήματα 4.5 και 4.6 είναι δύο ακόμη παραδείγματα εφαρμογής της μεθόδου για την ψηφιακή επιδιόρθωση τμήματος του υφάσματος που κρατάει ο μεγαλύτερος άνδρας στην τοιχογραφία *Ακολουθία Ανδρών* και ενός ενδεικτικού τμήματος της κορδέλας στα μαλλιά της Πότνιας στην τοιχογραφία *Κροκοσυλλέκτρια και Πότνια*.



Εντός πλαισίου βρίσκεται η σημασιολογικά και μορφολογικά κατάλληλη πληροφορία για την αποκατάσταση της ελλειπούσας περιοχής



Η εντός πλαισίου περιοχή



Οριζόντια αναστροφή



Ωρολογιακή περιστροφή 10°



Η μάσκα

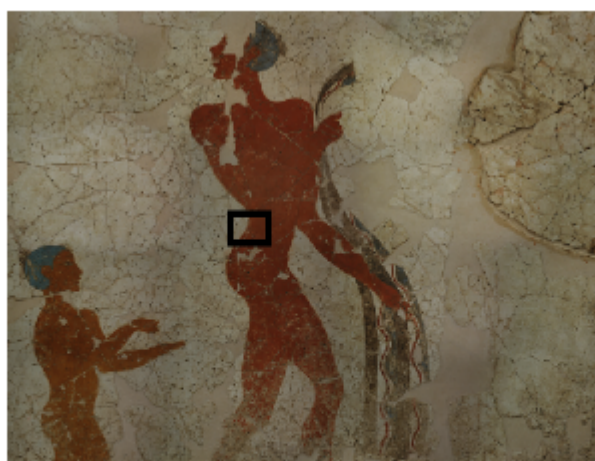


Τρέχουσα μορφή τοιχογραφίας



Μετά την ψηφιακή αποκατάσταση

Σχήμα 4.3: Προσάρτηση τμήματος της κεφαλής του νεότερου άνδρα, μετά από αναστροφή, κατάλληλη περιστροφή και μετατόπιση, στην ελλείπουσα περιοχή της κεφαλής του μεγαλύτερου άνδρα (τοιχογραφία *Ακολουθία Ανδρών*).



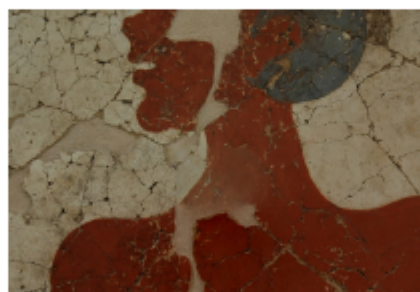
Η περιοχή στη μέση του γηραιότερου άνδρα από όπου αντλείται πληροφορία



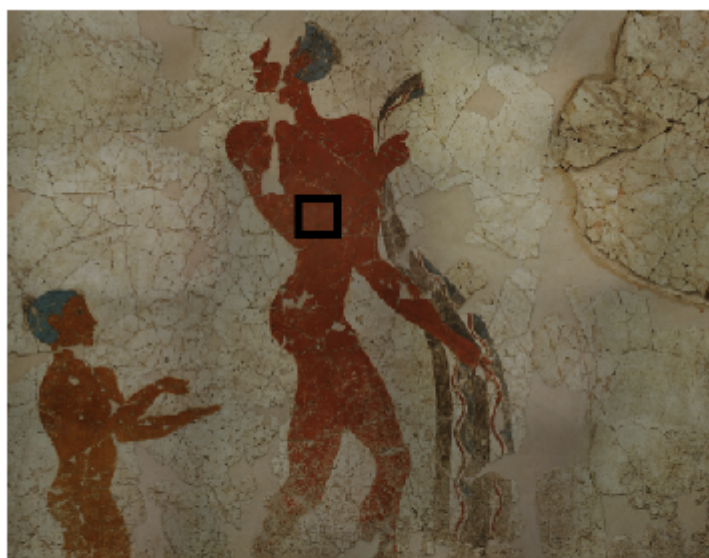
Η περικλειόμενη από το πλαίσιο περιοχή Η μάσκα στο πρώτο στάδιο αποκατάστασης



Αρχική εικόνα



Μετά το πρώτο στάδιο



Η περιοχή στην πλάτη του γηραιότερου άνδρα από όπου αντλείται νέα πληροφορία



Η περικλειόμενη από το νέο πλαίσιο περιοχή



Η μάσκα στο δεύτερο στάδιο της αποκατάστασης

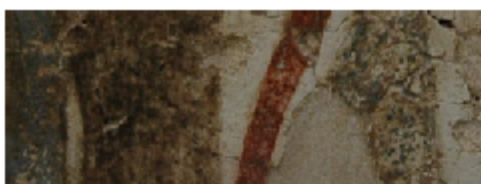


Το αποτέλεσμα μετά την αποκατάσταση δύο σταδίων

Σχήμα 4.4: Ψηφιακή αποκατάσταση του λαιμού της ανδρικής φιγούρας με πληροφορία που αντλείται από τη μέση και την πλάτη (τοιχογραφία *Ακολουθία Ανδρών*).



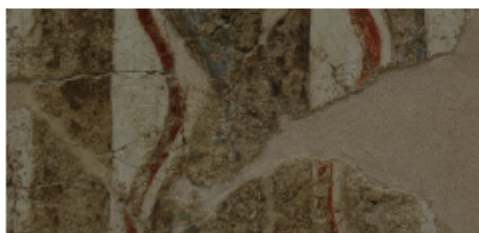
Μέρος της τοιχογραφίας με την υποπεριοχή που αξιοποιείται σε πλαίσιο



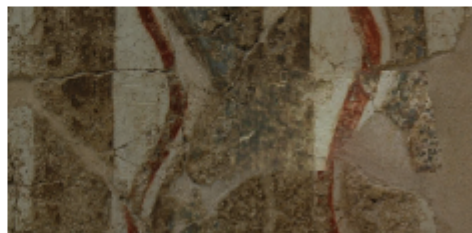
Η πληροφορία εντός πλαισίου



Η μάσκα πλαισίου



Ελλείπον τμήμα αρχικής εικόνας



Αποτέλεσμα μετά τη ψηφιακή του αποκατάσταση

Σχήμα 4.5: Ψηφιακή επιδιόρθωση τμήματος του υφάσματος που κρατάει ο μεγαλύτερος άνδρας με τη λήψη υποπεριοχής της εικόνας αντίστοιχης πληροφορίας (τοιχογραφία *Ακολουθία Ανδρών*).



Σε πλαίσιο τμήμα της τοιχογραφίας από όπου αντλείται πληροφορία για τη διόρθωση τμήματος παρόμοιας φύσης



Κοντινότερο πλάνο της περιοχής



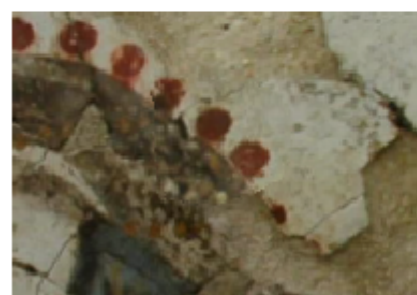
Η πληροφορία εντός του πλαισίου



Η μάσκα πλαισίου



Περιοχή ελλειπούσας πληροφορίας



Αποτέλεσμα μετά τη μερική ψηφιακή της αποκατάσταση

Σχήμα 4.6: Αποκατάσταση ενδεικτικού τμήματος της κορδέλας στα μαλλιά της Πότνιας (τοιχογραφία *Κροκοσυλλέκτρια & Πότνια*).

Κεφάλαιο 5

Συμπεράσματα

5.1 Συμβολή της διπλωματικής εργασίας

Όσον αφορά την πρωτοτυπία της, η εργασία αυτή συνεισφέρει κυρίως στα εξής σημεία:

- Συνδυασμός τεχνικών αυτομάτου εντοπισμού των φθορών και inpainting για την ενιαία αντιμετώπιση του προβλήματος της αποκατάστασης των τοιχογραφιών.

- Πρόταση ενός αλγορίθμου που χρησιμοποιεί μαθηματική μορφολογία και ανίχνευση ακμών για την αυτόματη ανίχνευση των φθορών και των ελλειπόντων τμημάτων στις τοιχογραφίες. Οδήγησε στην υλοποίηση μιας αποτελεσματικής μεθόδου για την εξαγωγή κατάλληλης μάσκας που βρίσκει τις περιοχές που έχουν υποστεί φθορές και αποτελεί το πρώτο βήμα της επίλυσης του προβλήματος της ψηφιακής επιδιόρθωσης των τοιχογραφιών.

- Πρόταση ενός αλγορίθμου που βελτιώνει την εξαχθείσα μάσκα με τις προαναφερθείσες μεθόδους, ώστε να «κόβει» ακριβώς το σύνορο των ελλειπουσών περιοχών, ειδικά σε περιπτώσεις που η μαθηματική μορφολογία δίνει μία αδρή εικόνα τους. Αξιοποιεί επαναληπτικές τομές γράφων και συγκεκριμένα τον αλγόριθμο “GrabCut”.

Γενικότερα, η διπλωματική εργασία είχε τις εξής ακόμη κατευθύνσεις:

- Επισκόπηση της βιβλιογραφίας σχετικά με τα προβλήματα της ενδοσυμπλήρωσης, της κατάρτισης εικόνας και της σύνθεσης μωσαϊκού.

- Τροποποίηση διαθέσιμης υλοποίησης του total variation inpainting μοντέλου για ενδοσυμπλήρωση περιεχομένου. Έγινε παραμετροποίηση μέσω πειραμάτων για την προσαρμογή της υλοποίησης στη συγκεκριμένη εφαρμογή και αυτοματοποιήθηκε η σύγκλιση του αλγορίθμου.

- Παρουσίαση μίας εισαγωγής στην υφή και το πρόβλημα της σύνθεσης υφής και αξιοποίηση διαθέσιμης υλοποίησης του αλγορίθμου Efros-Leung για ενδοσυμπλήρωση υφής στα πειράματα.

- Περιγραφή σχετικών με την κατάρτιση εργασιών και εφαρμογή του “GrabCut” στην κατάρτιση εικόνων από τις τοιχογραφίες, ειδικά σε περιπτώσεις που η μαθηματική μορφολογία

αποτυγχάνει ή που απαιτείται μεγάλη ακρίβεια στον προσδιορισμό του συνόρου.

- Εκτενής παρουσίαση σύγχρονων αλγορίθμων σύνθεσης μωσαϊκού και αξιοποίηση της υλοποίησης του αλγορίθμου των Tao et al. στα πειράματα, έπειτα από κατάλληλη τροποποίηση του πηγαίου κώδικα ώστε να μπορεί να λειτουργήσει για την παρούσα εφαρμογή.

5.2 Κατευθύνσεις για μελλοντική έρευνα

Παρόλο που η παρούσα εργασία έδωσε ικανοποιητικά αποτελέσματα, υπάρχουν πολλά περιθώρια για περαιτέρω διερεύνηση του θέματος και πληρέστερη ψηφιακή αποκατάσταση των τοιχογραφιών. Ορισμένες κατευθύνσεις όπου μπορεί να στραφεί η μελλοντική έρευνα είναι οι εξής:

- Αυτόματη εύρεση περιοχών σχετικής φύσης και σημασιολογίας στις τοιχογραφίες για την ανάκτηση και συνένωση εκεί όπου υπάρχει ευμεγέθης έλλειψη. Μία στρατηγική είναι να γίνει εκ των προτέρων (offline) απόδοση ετικετών από το χρήστη, σχετικές με τη σημασιολογία και το χρώμα για παράδειγμα, σε διάφορες υποπεριοχές των τοιχογραφιών και στη συνέχεια με ταίριασμα προτύπου (template matching) να αναζητείται σε περιοχές με την ίδια ετικέτα, σε όλες τις κλίμακες και γωνίες, ένα χωρίο που η πληροφορία του είναι αρκούντως κοντά σε αυτή της γειτονιάς της έλλειψης. Το ότι η αναζήτηση γίνεται μόνο στις υποπεριοχές με την ίδια επισημείωση καθιστά τον αλγόριθμο πιο γρήγορο και πιο αποτελεσματικό. Οι αλγόριθμοι κατάτμησης είναι απαραίτητοι τόσο για τον προσδιορισμό της έλλειψης στην υπό διόρθωση περιοχή, αλλά και στο ανακτώμενο τμήμα σε συνδυασμό με inpainting για αποκατάσταση των μικρών ατελειών του (αν οι ελλείψεις δεν είναι ασήμαντες το τμήμα θεωρείται μη κατάλληλο για αξιοποίηση). Έτσι, γίνεται μέσω μίας αυτόματης διαδικασίας η εκμετάλλευση της πληροφορίας όλης της τοιχογραφίας και όχι μόνο της γειτονιάς της inpainting περιοχής. Είναι ένα αρκετά φιλόδοξο εγχείρημα που σε κάποιες περιπτώσεις θα μπορούσε να φέρει θετικά αποτελέσματα, αν και η μικρή βάση δεδομένων (οι τοιχογραφίες είναι μικρού πλήθους) το καθιστά σε κάποιες περιπτώσεις ανεδαφικό.

- Αξιοποίηση πιο αποτελεσματικών αλγορίθμων για την ενδοσυμπλήρωση. Το Total Variation inpainting μοντέλο για structure inpainting και ο αλγόριθμος Efros-Leung για texture inpainting έδωσαν πολύ ικανοποιητικά αποτελέσματα, αλλά σίγουρα υπάρχουν περιθώρια βελτιωμένης ενδοσυμπλήρωσης. Η χρήση τεχνικών κυματιδίων είναι μία λύση, αν και εισάγει μεγάλη πολυπλοκότητα στο πρόβλημα.

- Χρήση τεχνικών 3D Υπολογιστικής Όρασης για βελτιωμένη ανίχνευση των ελλειπουσών περιοχών, αξιοποιώντας το γεγονός ότι με βάση τη μεθοδολογία συντήρησης που έχει ακολουθηθεί έως τώρα στις τοιχογραφίες τα ευρεθέντα τμήματα προεξέχουν ελαφρώς του κονιάματος που τοποθετείται στα κενά. Μία τέτοια προσπάθεια, όμως, προϋποθέτει τη λήψη νέου φωτογραφικού υλικού υπό καλά ελεγχόμενες συνθήκες φωτισμού.

Βιβλιογραφία

- [1] P. Maragos, “Image analysis & computer vision,” 2005. In Greek.
- [2] G. Papandreou, *Image Analysis and Computer Vision: Theory and Applications in the Restoration of Ancient Wall Paintings*. PhD thesis, National Technical University of Athens, 2009. In Greek.
- [3] C. Rother, V. Kolmogorov, and A. Blake, “Grabcut”: Interactive foreground extraction using iterated graph cuts,” *ACM Transactions on Graphics*, vol. 23, pp. 309–314, 2004.
- [4] E. A. Pnevmatikakis, “Inpainting models for automatic structure - texture restoration and text removal from images,” 2004. Diploma thesis. In Greek.
- [5] A. Levin, A. Zomet, S. Peleg, and Y. Weiss, “Seamless image stitching in the gradient domain,” in *ECCV (4)*, pp. 377–389, 2004.
- [6] M. W. Tao, M. K. Johnson, and S. Paris, “Error-tolerant image compositing,” in *ECCV (1)*, pp. 31–44, 2010.
- [7] M. Bertalmío, A. L. Bertozzi, and G. Sapiro, “Navier-stokes, fluid dynamics, and image and video inpainting,” in *CVPR (1)*, pp. 355–362, 2001.
- [8] S. Kondo and T. Miura, “Reaction-diffusion model as a framework for understanding biological pattern formation,” *Science*, vol. 329, September 2010.
- [9] S. Draves and E. Reckase, “The Fractal Flame Algorithm,” tech. rep., Nov. 2008.
- [10] J. J. Shen, “Inpainting and the fundamental problem of image processing,” *SIAM News*, vol. 36, no. 5, 2003.
- [11] J. Shen and T. F. Chan, “Mathematical models for local nontexture inpaintings,” *SIAM Journal of Applied Mathematics*, vol. 62, no. 3, pp. 1019–1043, 2002.
- [12] T. F. Chan, S. H. Kang, and J. Shen, “Euler’s elastica and curvature based inpaintings,” *SIAM J. Appl. Math.*, vol. 63, pp. 564–592, 2002.
- [13] M. Tuceryan and A. K. Jain, “Texture analysis,” 1998.
- [14] X. Ntoumas, “Akrotiri theras: Twenty years of research (1967-1987),” *BAE*, no. 116, 1992.
- [15] S. Marinatos, “Excavations at theras i-vii, 1967-1973 seasons,” *BAE*, no. 178-180, the Archaeological Society at Athens, 1999.

- [16] P. Rousopoulos, C. Papaodyseus, D. Arabadjis, M. Exarhos, and M. Panagopoulos, "Image and pattern analysis for the determination of the method of drawing celebrated thera wall-paintings circa 1650 b.c.," *JOCCH*, vol. 3, no. 2, p. 3, 2010.
- [17] S. Sotiropoulou, V. Perdikatsis, C. Apostolaki, A. G. Karydas, A. Devetzi, and K. Birtacha, "Lead pigments and related tools at akrotiri, thera, greece. provenance and application techniques," *Journal of Archaeological Science*, Feb. 2010.
- [18] S. Sherratt, "The wall paintings of thera: Proc. of the first int. symposium.," vol. 2, 2000.
- [19] B. J. Brown, C. Toler-Franklin, D. Nehab, M. Burns, D. P. Dobkin, A. Vlachopoulos, C. Doumas, S. Rusinkiewicz, and T. Weyrich, "A system for high-volume acquisition and matching of fresco fragments: reassembling theran wall paintings," *ACM Trans. Graph.*, vol. 27, no. 3, 2008.
- [20] H. Shin, C. Doumas, T. A. Funkhouser, S. Rusinkiewicz, K. Steiglitz, A. Vlachopoulos, and T. Weyrich, "Analyzing fracture patterns in theranwall paintings," in *VAST*, pp. 71–78, 2010.
- [21] I. Giakoumis, N. Nikolaidis, and I. Pitas, "Digital image processing techniques for the detection and removal of cracks in digitized paintings," *IEEE Transactions on Image Processing*, vol. 15, no. 1, pp. 178–188, 2006.
- [22] C. Mpoulotis, "Aspects of religious expression in akrotiri," *ALS*, vol. 3, 2005. In Greek.
- [23] M. Bertalmío, G. Sapiro, V. Caselles, and C. Ballester, "Image inpainting," in *SIGGRAPH*, pp. 417–424, 2000.
- [24] B. Dacorogna, *Introduction to the Calculus of Variations*. London: Imperial College Press, 2004.
- [25] M. F. Barnsley, *Fractals everywhere*. Academic Press, 1988.
- [26] S. Geman and D. Geman, "Stochastic relaxation, gibbs distributions, and the bayesian restoration of images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 6, no. 6, pp. 721–741, 1984.
- [27] S. C. Zhu, Y. N. Wu, and D. Mumford, "Minimax entropy principle and its application to texture modeling," *Neural Computation*, vol. 9, no. 8, pp. 1627–1660, 1997.
- [28] D. Marr and E. Hildreth, "Theory of edge detection," *Proceedings of the Royal Society of London. Series B, Biological Sciences*, vol. 207, no. 1167, pp. 187–217, 1980.
- [29] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," in *Proceedings of the eleventh annual international conference of the Center for Nonlinear Studies on Experimental mathematics : computational issues in nonlinear science: computational issues in nonlinear science*, (Amsterdam, The Netherlands, The Netherlands), pp. 259–268, Elsevier North-Holland, Inc., 1992.
- [30] D. Mumford and J. Shah, "Optimal approximations by piecewise smooth functions and associated variational problems," *Communications on Pure and Applied Mathematics*, vol. 42, no. 5, pp. 577–685, 1989.

- [31] V. Perdikatsis, V. Kilikoglou, S. Sotiropoulou, and E. Chryssikopoulou, “Physicochemical characterization of pigments from theran wall paintings,” *[She00]*, pp. 103–118.
- [32] V. Caselles, R. Kimmel, and G. Sapiro, “Geodesic active contours,” 1997.
- [33] M. Kass, A. P. Witkin, and D. Terzopoulos, “Snakes: Active contour models,” *International Journal of Computer Vision*, pp. 321–331, 1988.
- [34] Y. Boykov and V. Kolmogorov, “Computing geodesics and minimal surfaces via graph cuts,” in *International Conference on Computer Vision*, pp. 26–33, 2003.
- [35] Y. Boykov and M.-P. Jolly, “Interactive graph cuts for optimal boundary and region segmentation of objects in n-d images,” in *ICCV*, pp. 105–112, 2001.
- [36] P. Maragos, “Morphological filtering for image enhancement and feature detection,” *The Image and Video Processing Handbook, Second Edition*, pp. 135–156, 2005. Edited by A.C.Bovik, Elsevier Academic Press.
- [37] J. Canny, “A Computational Approach to Edge Detection,” *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. PAMI-8, pp. 679–698, Nov. 1986.
- [38] P. Pérez, M. Gangnet, and A. Blake, “Poisson image editing,” *ACM Trans. Graph.*, vol. 22, no. 3, pp. 313–318, 2003.
- [39] A. Efros and T. Leung, “Texture synthesis by non-parametric sampling,” in *International Conference on Computer Vision*, pp. 1033–1038, 1999.
- [40] A. Criminisi, P. Pérez, and K. Toyama, “Region filling and object removal by exemplar-based image inpainting,” *IEEE Transactions on Image Processing*, vol. 13, no. 9, pp. 1200–1212, 2004.
- [41] G. Papandreou, P. Maragos, and A. Kokaram, “Image inpainting with a wavelet domain hidden markov tree model,” in *ICASSP*, pp. 773–776, 2008.
- [42] T. F. Chan, S. Osher, and J. Shen, “The digital tv filter and nonlinear denoising,” *IEEE Transactions on Image Processing*, pp. 231–241, 2001.
- [43] R. M. Haralick, “Statistical and structural approaches to texture,” *Proceedings of the IEEE*, vol. 67, pp. 786–804, May 1979.
- [44] J. Hays and A. A. Efros, “Scene completion using millions of photographs,” *Commun. ACM*, vol. 51, no. 10, pp. 87–94, 2008.
- [45] M. Wilczkowiak, G. J. Brostow, B. Tordoff, and R. Cipolla, “Hole filling through photomontage,” in *BMVC*, 2005.
- [46] N. Diakopoulos, I. Essa, and R. Jain, “Content based image synthesis,” in *In CIVR*, pp. 299–307, 2004.
- [47] M. Johnson, G. J. Brostow, J. Shotton, O. Arandjelovic, V. Kwatra, and R. Cipolla, “Semantic photo synthesis,” 2006.
- [48] J. Davis, “Mosaics of scenes with moving objects,” pp. 354–360, 1998.

- [49] M. Uyttendaele, A. Eden, and R. Szeliski, "Eliminating ghosting and exposure artifacts in image mosaics," in *CVPR (2)*, pp. 509–516, 2001.
- [50] P. J. Burt, Edward, and E. H. Adelson, "The laplacian as a compact image code," *IEEE Transactions on Communications*, vol. 31, pp. 532–540, 1983.
- [51] W. T. Freeman, E. C. Pasztor, and O. T. Carmichael, "Learning low-level vision," *Intl. J. Computer Vision*, 2000.
- [52] A. A. Efros and W. T. Freeman, "Image quilting for texture synthesis and transfer," in *Proceedings of SIGGRAPH 2001*, pp. 341–346, 2001.
- [53] C. Ballester, M. Bertalmío, V. Caselles, G. Sapiro, and J. Verdera, "Filling-in by joint interpolation of vector fields and gray levels," *IEEE Transactions on Image Processing*, vol. 10, no. 8, pp. 1200–1211, 2001.
- [54] A. Levin, A. Zomet, S. Peleg, and Y. Weiss, "Seamless image stitching in the gradient domain 1," Tech. Rep. 2003-82, the Hebrew University of Jerusalem, 2003.
- [55] Z. Farbman, G. Hoffer, Y. Lipman, D. Cohen-Or, and D. Lischinski, "Coordinates for instant image cloning," *ACM Trans. Graph.*, vol. 28, no. 3, 2009.
- [56] J.-F. Lalonde, D. Hoiem, A. A. Efros, C. Rother, J. M. Winn, and A. Criminisi, "Photo clip art," *ACM Trans. Graph.*, vol. 26, no. 3, p. 3, 2007.
- [57] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE TRANSACTIONS ON IMAGE PROCESSING*, vol. 13, no. 4, pp. 600–612, 2004.
- [58] E. A. Pnevmatikakis and P. Maragos, "An inpainting system for automatic image structure - texture restoration with text removal," in *ICIP*, pp. 2616–2619, 2008.
- [59] P. Maragos, "A representation theory for morphological image and signal processing.," *IEEE Trans. Pattern Anal. Mach. Intell.*, pp. 586–599, 1989.
- [60] V. Caselles, J.-M. Morel, and C. Sbert, "An axiomatic approach to image interpolation," *IEEE Transactions on Image Processing*, vol. 7, no. 3, pp. 376–386, 1998.
- [61] S. Masnou and J.-M. Morel, "Level lines based disocclusion," in *ICIP (3)*, pp. 259–263, 1998.
- [62] S. Rane, M. Bertalmío, and G. Sapiro, "Structure and texture filling-in of missing image blocks in wireless transmission and compression," in *ICIP (1)*, pp. 317–320, 2002.
- [63] S. Esedoglou and J. Shen, "Digital inpainting based on the mumford-shah-euler image model," *European Journal of Applied Mathematics*, no. 13, pp. 353–370, 2002.
- [64] J. Wang and M. F. Cohen, "Image and video matting: a survey," *Found. Trends. Comput. Graph. Vis.*, vol. 3, pp. 97–175, January 2007.