



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ
ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΣΗΜΑΤΩΝ, ΕΛΕΓΧΟΥ ΚΑΙ ΡΟΜΠΟΤΙΚΗΣ

Ανάλυση και μοντελοποίηση υφής σε ψηφιακές εικόνες
με χρήση Markov-Gibbs random fields

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

του

Μενέλαου-Κων/νου Δ.
Σταυρόπουλου

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

Αθήνα, Ιούλιος 2005



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ
ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΣΗΜΑΤΩΝ, ΕΛΕΓΧΟΥ ΚΑΙ ΡΟΜΠΟΤΙΚΗΣ

Ανάλυση και μοντελοποίηση υφής σε ψηφιακές εικόνες
με χρήση Markov-Gibbs random fields

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

του

Μενέλαου-Κων/νου Δ.
Σταυρόπουλου

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

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή στις

Ιουλίου 2005.

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

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

.....
Κωνσταντίνος Τζαφέστας
Λέκτορας Ε.Μ.Π.

Αθήνα, Ιούλιος 2005.

.....
Μενέλαος Κων/νος Δ. Σταυρόπουλος

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

Copyright© Μενέλαος Κων/νος Δ. Σταυρόπουλος, 2005.

Με επιφύλαξη παντός δικαιώματος.All rights reserved.

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

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

Περίληψη

Η παρούσα διπλωματική εργασία πραγματεύεται την αποδοτική και αποτελεσματική μοντελοποίηση και ανάλυση της υφής σε ψηφιακές εικόνες με τη βοήθεια των μαρκοβιανών τυχαίων πεδίων (MRF). Αφού γίνει μια μικρή εισαγωγή για την όραση υπολογιστών, επικεντρωνόμαστε στην υφή, στα προβλήματα που την περιβάλλουν, στην αντίληψη της από τον άνθρωπο και κυρίως στα μοντέλα που έχουν χρησιμοποιηθεί μέχρι σήμερα για την ανάλυσή της. Στη συνέχεια γίνεται μια μικρή θεωρητική εισαγωγή στα MRF και σε θέματα που σχετίζονται με αυτά ενώ γίνεται αναφορά και σε μοντέλα υφής που πηγάζουν από αυτά. Με βάση τις παραπάνω ιδέες προτείνεται ένα βελτιωμένο μοντέλο υφής βασισμένο σε MRF με αρκετά καλές επιδόσεις τόσο στη σύνθεση όσο και στην κατηγοριοποίηση υφής. Επίσης προτείνονται αλγόριθμοι για ταχεία σύνθεση, εκτίμηση παραμέτρων και κατηγοριοποίηση ενώ παρατίθενται και πειραματικά δεδομένα και μετρήσεις που αποδεικνύουν την αποτελεσματικότητα της υλοποίησης.

Λέξεις—κλειδιά

όραση υπολογιστών, επεξεργασία εικόνας, μαθηματική μορφολογία, υφή, ανάλυση υφής, μοντελοποίηση υφής, μαρκοβιανά τυχαία πεδία, σύνθεση υφής, κατηγοριοποίηση υφής, εκτίμηση παραμέτρων, αλγόριθμοι

Abstract

This diploma thesis deals with the effective and efficient analysis and modeling of texture appeared in digital images by means of Markov random fields (MRF). After a brief review in general computer vision topics we concentrate in texture, the problems with which it is related, the human perception and mostly in the models used for its analysis until today. The thesis continues with a brief review in MRF theory and all these issues which are useful to manipulate it. Some models based on MRF which are used in texture modeling are discussed. All these ideas lead to an improved model for texture analysis based on MRF with good performance in texture synthesis as well as in texture classification. Finally algorithms for quick synthesis, parameter estimation and texture classification are proposed and we provide experimental data and calculations which prove the potentials of our implementation.

Keywords

computer vision, image processing, Markov random fields, Gibbs random fields, Markov Chains, Monte Carlo Sampling, texture, texture synthesis, texture classification, parameter estimation, mathematical morphology, constrained MRF, algorithms

Πρόλογος

Η όραση υπολογιστών τα τελευταία χρόνια έχει γνωρίσει ραγδαία εξέλιξη καταφέροντας να λύσει πολλά προβλήματα που για χρόνια απασχολούσαν τους επιστήμονες. Η εκμετάλλευση των μέχρι τώρα επιτευγμάτων οδηγεί στην ανάλυση περισσότερων χαρακτηριστικών των φυσικών αντικειμένων που εμφανίζονται σε ψηφιακές εικόνες τα οποία θα μας βοηθήσουν να πετύχουμε υψηλό βαθμό αυτοματοποίησης σε μια πληθώρα σύγχρονων εφαρμογών. Ένα από αυτά τα χαρακτηριστικά είναι η υφή η οποία αποτελεί και το κεντρικό θέμα της παρούσας διπλωματικής εργασίας.

Η απουσία ελληνικής βιβλιογραφίας πάνω στο θέμα αυτό μας οδήγησε να συμπεριλάβουμε εκτενή, κατά τη γνώμη μας, εισαγωγικά κεφάλαια, πλούσια σε πηγές, τα οποία πιστεύουμε θα βοηθήσουν τον αναγνώστη που ενδιαφέρεται να ασχοληθεί επιστημονικά με το πρόβλημα της υφής. Ακόμη παρουσιάζονται διάφοροι αλγόριθμοι σε ψευδοκώδικα ενώ δίνονται οδηγίες για την κατασκευή παρόμοιων με απλές μετατροπές κάτι που διευκολύνει την υλοποίησή τους άρα και τη βαθύτερη κατανόηση του θέματος.

Το κείμενο είναι οργανωμένο σε 6 κεφάλαια. Στόχος μας ήταν κατά το δυνατόν η αυτονομία όσον αφορά τουλάχιστον το θεωρητικό (μαθηματικό) υπόβαθρο που είναι αναγκαίο για την κατανόηση των εννοιών που παρουσιάζονται. Παρόλα αυτά πιστεύουμε ότι ο αναγνώστης πρέπει να είναι εξοικειωμένος με βασικές γνώσεις πάνω στην όραση υπολογιστών, την επεξεργασία εικόνων και σημάτων και την αναγνώριση προτύπων. Τα τρία πρώτα κεφάλαια περιέχουν διάφορα γενικά και χρήσιμα στοιχεία και προετοιμάζουν τον αναγνώστη για το κυρίως μέρος της εργασίας που είναι τα κεφάλαια 4 και 5.

Συγκεκριμένα το **κεφάλαιο 1** περιλαμβάνει μια γενική περιγραφή της όρασης υπολογιστών ενώ γίνεται αναφορά και για τη σημασία της υφής, τα κίνητρα, τα προβλήματα που κληθήκαμε να λύσουμε και τις συνεισφορές της διπλωματικής.

Στο **κεφάλαιο 2** επικεντρωνόμαστε αποκλειστικά στο θέμα της υφής και εξετάζουμε προβλήματα που σχετίζονται με αυτήν, τον τρόπο με τον οποίον την αντιλαμβάνεται ο ανθρώπινος εγκέφαλος, τα μοντέλα που κατα καιρούς έχουν χρησιμοποιηθεί για την ανάλυσή της καθώς και τις εφαρμογές τις οποίες βρίσκει

σε σύγχρονα πρακτικά προβλήματα.

Στη συνέχεια στο **κεφάλαιο 3** πραγματοποιούμε μια μικρή μαθηματική εισαγωγή πάνω στα μαρκοβιανά τυχαία πεδία (MRF) ενώ εστιάζουμε και σε θέματα προσομοίωσης, εκτίμησης παραμέτρων και κατηγοριοποίησης (classification).

Το **κεφάλαιο 4** παρουσιάζει διάφορα μοντέλα βασισμένα σε MRF τα οποία έχουν χρησιμοποιηθεί στη μοντελοποίηση υφής συνοδευόμενα από παραδείγματα σύνθεσης. Αποτελούν μοντέλα που μελετήσαμε οι ίδιοι για την περαιτέρω κατανόηση του προβλήματος αλλά και τη δυνατότητα ανάπτυξης νέων ιδεών που θα οδηγούσαν στη δημιουργία ενός βελτιωμένου μοντέλου.

Όλες οι παραπάνω ιδέες μας οδηγούν στη βελτίωση ενός από τα μοντέλα που παρουσιάζονται στο κεφάλαιο 4 και το οποίο εξηγείται αναλυτικά στο **κεφάλαιο 5**. Επίσης προτείνεται αλγόριθμος σύνθεσης, εκτίμησης παραμέτρων (μέθοδος που αποδεικνύεται και μαθηματικά) και classification ενώ παρατίθενται και αρκετά πειραματικά δεδομένα και μετρήσεις που έγιναν πάνω σε αυτά.

Τέλος στο **κεφάλαιο 6** γίνεται μια ανασκόπηση των συμπερασμάτων μας από τη μελέτη του προβλήματος της υφής ενώ προτείνονται κατευθύνσεις για μελλοντική έρευνα.

Ευχαριστίες

Ένα εκ των βασικότερων κινήτρων για την ενασχόλησή μου με το συγκεκριμένο ερευνητικό αντικείμενο υπήρξε το μάθημα της Όρασης Υπολογιστών το οποίο διδάσκει με ιδιαίτερο ζήλο και ενθουσιασμό ο καθηγητής κ. Πέτρος Μαραγκός. Θα ήθελα σε αυτό το σημείο να τον ευχαριστήσω ιδιαίτερα τόσο για τη βοήθειά του στην επιλογή του θέματος όσο και για την αμέριστη βοήθεια και συμπαράσταση που μου προσέφερε κατά τη διάρκεια εκπόνησης της παρούσας εργασίας.

Ιδιαίτερες ευχαριστίες οφείλω επίσης και στον υποψήφιο διδάκτορα Γιώργο Ευαγγελόπουλο για την προθυμία και την άμεση και ουσιαστική βοήθεια που μου προσέφερε για οποιοδήποτε ζήτημα προέκυψε κατά τη μελέτη των περιοχών που ασχολήθηκα.

Τέλος ήθελα να ευχαριστήσω θερμά την οικογένειά μου για την ουσιαστική συμπαράσταση που μου προσφέρει όλα αυτά τα χρόνια και μου δίνει τη δυνατότητα να συνεχίζω απρόσκοπτα την ακαδημαϊκή μου σταδιοδρομία.

Περιεχόμενα

1	Εισαγωγή	14
1.1	Λίγα λόγια για την όραση υπολογιστών	14
1.2	Σημασία της υφής	15
1.3	Προβλήματα με τα οποία ασχολήθηκε η διπλωματική	16
1.4	Κίνητρα	17
1.5	Συνεισφορές της διπλωματικής	18
2	Υφή	20
2.1	Τί είναι η υφή;	20
2.2	Τα τρία προβλήματα της υφής	22
2.3	Η ανθρώπινη αντίληψη της υφής	23
2.4	Τα πιο διαδεδομένα μοντέλα	26
2.4.1	Μοντέλα που βασίζονται στον πίνακα μεταβάσεων (co-occurrence matrix)	26
2.4.2	Μοντέλα που βασίζονται στην αυτοσυσχέτιση	28
2.4.3	Μοντέλα που βασίζονται στη γραμμική πρόβλεψη	30
2.4.4	Μοντέλα υπολογισμού ενέργειας υφής	31
2.4.5	Μοντέλα υπολογισμού ακμών υφής	32
2.4.6	Μοντέλα μαθηματικής μορφολογίας	32
2.4.7	Μοντέλα που βασίζονται στη γεωμετρική κατάτμηση Voronoi	34
2.4.8	Μοντέλα που βασίζονται στα fractals	35
2.4.9	Μοντέλα που βασίζονται στα Markov-Gibbs random fields	37
2.4.10	Μοντέλα που βασίζονται σε ανάλυση με συστοιχία φίλτρων	40
2.5	Εφαρμογές της υφής σε πρακτικά προβλήματα	41
3	MRF: Μία μαθηματική εισαγωγή	43
3.1	Μαθηματικοί ορισμοί	43
3.1.1	Μαρκοβιανές τυχαίες ακολουθίες	43
3.1.2	Μαρκοβιανά τυχαία πεδία	46

3.1.3	Γκαουσιανές μαρκοβιανές τυχαίες ακολουθίες (Gaussian Markov random sequences)	47
3.1.4	Γκαουσιανά μαρκοβιανά τυχαία πεδία (Gaussian Markov random fields)	48
3.2	Δειγματοληψία των MRF	48
3.2.1	Διακριτά MRF	48
3.2.2	Συνεχή MRF	52
3.2.3	Προσομοιούμενη ανόπτηση (simulated annealing)	53
3.3	Εκτίμηση παραμέτρων	53
3.4	Classification	56
3.5	Constrained MRFs	57
4	Γνωστά μοντέλα υψής που βασίζονται στα MRF	58
4.1	Το διωνυμικό μοντέλο	58
4.2	Το μοντέλο Ising	62
4.3	Το μοντέλο των Efros-Leung	65
4.4	Το μοντέλο FRAME	66
4.5	Το μοντέλο GRF με μορφολογικούς περιορισμούς (Morphologically constrained GRF's)	71
5	Μια νέα πρόταση μοντέλου βασισμένο σε MRF	78
5.1	Θεωρητικό υπόβαθρο	79
5.2	Μια πρόταση αλγορίθμου σύνθεσης	81
5.3	Εκτίμηση παραμέτρων	84
5.4	Classification	86
5.5	Αποτελέσματα πειραμάτων	87
5.5.1	Αποτελέσματα σύνθεσης	87
5.5.2	Αποτελέσματα classification	92
5.6	Μετρήσεις πάνω στις εικόνες που προέκυψαν από σύνθεση	93
5.6.1	Μετρήσεις στατιστικής πρώτης τάξης	93
5.6.2	Μετρήσεις στατιστικής δεύτερης τάξης	96
6	Συμπεράσματα-Κατευθύνσεις για μελλοντική έρευνα	104
6.1	Συμπεράσματα	104
6.2	Κατευθύνσεις για μελλοντική έρευνα	105

Κατάλογος Σχημάτων

2.1	Πειραματικές εικόνες του Julesz	25
2.2	Εικόνα υψής μαζί με το φάσμα ισχύος της	29
2.3	Παράδειγμα τρισδιάστατου fractal τοπίου	38
2.4	Παράδειγμα τρισδιάστατης fractal υψής.	39
4.1	Παράδειγμα ισοτροπικής σύνθεσης με το μοντέλο των Cross και Jain	60
4.2	Παράδειγμα ανισοτροπικής σύνθεσης με το μοντέλο των Cross και Jain	60
4.3	Παράδειγμα σύνθεσης με αρνητικό clustering με το μοντέλο των Cross και Jain	61
4.4	Παράδειγμα σύνθεσης μη διαγώνιας υψής με το μοντέλο των Cross και Jain	61
4.5	Παράδειγμα ισοτροπικής σύνθεσης με το μοντέλο Ising	63
4.6	Παράδειγμα οριζόντιας ανισοτροπικής σύνθεσης με το multi level logistic μοντέλο	63
4.7	Παράδειγμα διαγώνιας ανισοτροπικής σύνθεσης με το multi level logistic μοντέλο	64
4.8	Παράδειγμα σύνθεσης με μηδενική συσσωμάτωση (clustering) με βάση το multi level logistic μοντέλο	64
4.9	Παράδειγμα σύνθεσης τυχαίας υψής με το μοντέλο Efros-Leung	67
4.10	Παράδειγμα σύνθεσης περιοδικής υψής με το μοντέλο Efros-Leung	68
4.11	Παράδειγμα σύνθεσης υψής για διαφορετικά μήκη παραθύρων με το μοντέλο Efros-Leung	69
4.12	Παράδειγμα σύνθεσης περιοδικής υψής με το μοντέλο FRAME	72
4.13	Παράδειγμα σύνθεσης τυχαίας υψής με το μοντέλο FRAME	73
4.14	Παράδειγμα δυαδικής σύνθεσης εικόνας με το μοντέλο των Sivakumar & Goutsias	75
4.15	Παράδειγμα δυαδικής σύνθεσης με το μοντέλο των Sivakumar & Goutsias με ρόμβο ως δομικό στοιχείο	75

4.16	Παράδειγμα δυαδικής σύνθεσης με το μοντέλο των Sivakumar & Goutsias με τετράγωνο ως δομικό στοιχείο	76
4.17	Παράδειγμα σύνθεσης με το μοντέλο των Sivakumar & Goutsias με 16 κλιμακες του γκρί	76
5.1	Παράδειγμα σύνθεσης της εικόνας D12	87
5.2	Παράδειγμα σύνθεσης της εικόνας D18	88
5.3	Παράδειγμα σύνθεσης της εικόνας D37	88
5.4	Παράδειγμα σύνθεσης της εικόνας D42	88
5.5	Παράδειγμα σύνθεσης της εικόνας D48	89
5.6	Παράδειγμα σύνθεσης της εικόνας D51	89
5.7	Παράδειγμα σύνθεσης της εικόνας D56	89
5.8	Παράδειγμα σύνθεσης της εικόνας D71	90
5.9	Παράδειγμα σύνθεσης της εικόνας D75	90
5.10	Παράδειγμα σύνθεσης της εικόνας D88	90
5.11	Παράδειγμα σύνθεσης της εικόνας D91	91
5.12	Παράδειγμα σύνθεσης της εικόνας D101	91
5.13	Σύγκριση σύνθεσης της εικόνας D48	92
5.14	Παράδειγμα σύνθεσης της εικόνας D18	92
5.15	Παράδειγμα σύνθεσης της εικόνας D37	93
5.16	Οι εικόνες D12,D18,D37	94
5.17	Οι εικόνες D47,D10,D96	95
5.18	Οι εικόνες D56,D9,D75	95
5.19	Η εικόνα D32	96
5.20	Σύγκριση ιστογραμμάτων της πραγματική και της συντεθειμένης εικόνας D12	97
5.21	Σύγκριση ιστογραμμάτων της πραγματική και της συντεθειμένης εικόνας D18	98
5.22	Σύγκριση ιστογραμμάτων της πραγματική και της συντεθειμένης εικόνας D37	99
5.23	Σύγκριση ιστογραμμάτων της πραγματική και της συντεθειμένης εικόνας D51	100

Κατάλογος Πινάκων

2.1	Μερικές από τις πιο γνωστες ποσότητες που υπολογίζονται από τον πίνακα μετάβασεων	27
5.1	Αποτελέσματα κατηγοριοποίησης	94
5.2	Αποτελέσματα μετρήσεων second order statistics	101
5.3	Αποτελέσματα μετρήσεων second order statistics	102

Κεφάλαιο 1

Εισαγωγή

1.1 Λίγα λόγια για την όραση υπολογιστών¹

Η όραση υπολογιστών είναι ένας σχετικά νέος επιστημονικός κλάδος. Γεννήθηκε τη δεκαετία του '60 ως υποκλάδος της Τεχνητής Νοημοσύνης από επιστήμονες που είχαν σαν όραμα τη δημιουργία έξυπνων συστημάτων τα οποία θα αντιλαμβάνονταν τα οπτικά ερεθίσματα του φυσικού κόσμου με τον ίδιο τρόπο που τα αντιλαμβάνεται ο άνθρωπος. Κατά τα πρώτα της βήματα επικεντρώθηκε κυρίως στην εξαγωγή και αναγνώριση απλών χαρακτηριστικών σε δυαδικές εικόνες όπως χαρακτήρες κειμένου. Μέχρι σήμερα όμως ο κλάδος αυτός γνώρισε σημαντική εξέλιξη και επέτρεψε στους επιστήμονες να θέτουν διαρκώς νέα προβλήματα και να αναζητούν τη λύση τους. Σήμερα η όραση υπολογιστών επικεντρώνεται στην ανάκτηση χρήσιμων πληροφοριών από ψηφιακές εικόνες οι οποίες θα βοηθούν ένα οποιοδήποτε σύστημα να αντιλαμβάνεται τον κόσμο με ανθρώπινη ευφυΐα. Πιο συγκεκριμένα μερικά από τα τρέχοντα προβλήματα τα οποία καλείται να επιλύσει η όραση υπολογιστών είναι:

- Σχηματισμός εικόνων, προ-επεξεργασία
- Ανάλυση δυαδικών εικόνων (αλγεβρική, γεωμετρική, τοπολογική, στατιστική)
- Ανάλυση γκριζών εικόνων (αλγεβρική, γεωμετρική, στατιστική)
- Ανάλυση σε πολλαπλές κλίμακες
- Βελτίωση ποιότητας και εξαγωγή χαρακτηριστικών
- Τύψη: Μοντελοποίηση και ανάλυση

¹ Η ενότητα αυτή έχει βασιστεί αρκετά στο [53]

- Χρώμα: Μοντελοποίηση και ανάλυση
- Σχήμα: Μοντελοποίηση και ανάλυση
- Οπτική κίνηση 2Δ και 3Δ: ανίχνευση και εκτίμηση
- Κατάτμηση εικόνων σε ομογενείς περιοχές
- Ανακατασκευή τρισδιάστατου σχήματος/ γεωμετρίας, στερέοψη
- Αναγνώριση και κατηγοριοποίηση οπτικών αντικειμένων

Η όραση υπολογιστών είτε μόνη της είτε συνδυαζόμενη με άλλους κλάδους της σύγχρονης επιστήμης έχει πάρα πολλές και χρήσιμες εφαρμογές. Ήδη ευρεία είναι η χρήση της στο βιομηχανικό αυτοματισμό, στη σχεδίαση δηλαδή ενός έξυπνου συστήματος που θα ελέγχει και θα συντονίζει αυτόματα μια γραμμή βιομηχανικής παραγωγής. Πλέον χρησιμοποιείται ιδιαίτερα στη βελτίωση ιατρικών εικόνων καθώς και στην εξαγωγή χαρακτηριστικών και αναγνώριση ιστών του ανθρώπινου σώματος. Έτσι συμβάλλει στην σωστή και έγκαιρη διάγνωση διαφόρων παθήσεων. Εφαρμογες της βλέπουμε επίσης στη ρομποτική, στην ανάπτυξη δηλαδή κίνηση και έλεγχος με οπτική ανάδραση, στην ανάπτυξη ευφύων καμερών (video surveillance), σε θέματα που αφορούν το περιβάλλον (τηλεπισκόπηση, βιοοικολογία) και στην επικοινωνία ανθρώπου και μηχανής. Τέλος ιδιαίτερα ελπιδοφόρες είναι οι εφαρμογές που βρίσκει η όραση υπολογιστών στα μέσα μεταφοράς στην ανάπτυξη δηλαδή συστημάτων αυτόματης πλοήγησης και αποφυγής εμποδίων.

1.2 Σημασία της υψής

Όπως έχουμε ήδη πει ένα από τα αποτελέσματα της ραγδαίας εξέλιξης τα τελευταία χρόνια της Όρασης Υπολογιστών οδήγησε στην αναζήτηση τρόπων για ανάκτηση πληροφοριών και χαρακτηριστικών από ψηφιακές εικόνες τα οποία είναι ιδιαίτερα χρήσιμα στις σύγχρονες πρακτικές εφαρμογές. Ένα από αυτά τα χαρακτηριστικά είναι η υψή. Η μοντελοποίηση και η ανάλυση υψής παραμένει ένα ανοικτό πρόβλημα καθώς είναι ιδιαίτερα δύσκολη η δημιουργία ενός καθολικού μοντέλου στο οποίο να υπακούει. Συνεπώς πολλά μοντέλα έχουν κατακαίρους προταθεί τα οποία στοχεύουν στη μελέτη εκείνων των χαρακτηριστικών της υψής που χρειάζονται στη συγκεκριμένη εφαρμογή.

Η επιτυχημένη μοντελοποίηση της υψής θα σηματοδοτούσε μια νέα επανάσταση στο χώρο της Όρασης Υπολογιστών. Με κάτι τέτοιο θα μπορούσαμε να απεικονίσουμε πολύ πιο ρεαλιστικά γραφικά σε υπολογιστικές εφαρμογές χωρίς

ιδιαίτερο κόστος σε μνήμη καθώς η υφή θα γεννιόταν με τη βοήθεια υπολογισμών. Ακόμα θα μας επέτρεπε υψηλό βαθμό αυτοματοποίησης σε πάρα πολλές εφαρμογές όπως παρακολούθηση χώρων, αναγνώριση αντικειμένων, αναγνώριση κατάστασης αντικειμένων, αναγνώριση σχήματος περιοχών. Επίσης σημαντική βοήθεια θα μας παρείχε ένα τέτοιο μοντέλο στην αποκατάσταση εικόνων που έχουν φθορές ή εμφανίζονται θορυβώδεις. Κάτι τέτοιο έχει πολλές εφαρμογές στην αρχαιολογία καθώς πολλά αρχαιολογικά ευρήματα που ανασύρονται έχουν υποστεί σημαντική φθορά από το χρόνο. Πρωτεύοντα ρόλο παίζει επίσης η ανάλυση υφής όταν θέλουμε να μελετήσουμε απομακρυσμένες περιοχές, περιοχές πολύ μικρές στις οποίες δεν μπορεί να εισέλθει ο άνθρωπος, περιοχές πολύ μεγάλες που χρειάζονται πολύ χρόνο να εξερευνηθούν αλλά και περιοχές τις οποίες δεν μπορεί να προσεγγίσει ο άνθρωπος για λόγους ασφαλείας. Σε τέτοιες περιπτώσεις εκμεταλλευόμαστε την υφή των αντικειμένων ώστε να εξαχθούν χρήσιμα συμπεράσματα για τις περιοχές αυτές.

Όλα τα παραπάνω συμπληρώνονται από το γεγονός ότι πλέον σήμερα το κόστος για τη λήψη μιας ψηφιακής φωτογραφίας έχει μειωθεί δραστικά. Άρα μπορούμε να έχουμε πληθώρα δεδομένων τα οποία χρειάζονται γρήγορη και αξιόπιστη επεξεργασία.

1.3 Προβλήματα με τα οποία ασχολήθηκε η διπλωματική

Σε πρώτη φάση κατά την ανάληψη αυτής της διπλωματικής εργασίας, στόχος μας ήταν η αναγνώριση διαφορετικών συστατικών εδάφους από εικόνες εδαφοτομών υψηλής ανάλυσης οι οποίες είχαν τραβηχτεί από ηλεκτρονικό μικροσκόπιο. Η πρώτη μας ιδέα ήταν να χρησιμοποιήσουμε ένα ήδη υπάρχον μοντέλο βασισμένο σε MRF. Τα ιδιαίτερα χαρακτηριστικά όμως αυτών των συστατικών μας οδήγησαν στην ανάπτυξη ενός νέου μοντέλου το οποίο βασίζεται στις ιδέες ενός παλαιότερου.

Με βάση το μοντέλο αυτό ασχοληθήκαμε γενικότερα για την επίλυση δύο από τα τρία βασικότερα προβλήματα που σχετίζονται με την υφή: τη σύνθεση υφής και την κατηγοριοποίηση υφής τα οποία αποτέλεσαν και τα κριτήριά μας για τις βελτιώσεις που πραγματοποιήσαμε στο μοντέλο μας. Επίσης συμπληρώσαμε τα κριτήρια αυτά με μετρήσεις στατιστικής πρώτης και δεύτερης τάξης με στόχο να διασφαλίσουμε την εγκυρότητα των αποτελεσμάτων.

Επίσης κατά την κατασκευή του μοντέλου προέκυψε πρόβλημα σωστής προσομοίωσης. Έτσι αναγκαστήκαμε να αναπτύξουμε έναν αλγόριθμο σύνθεσης ο οποίος συγκλίνει γρηγορότερα από άλλους που θεωρούν αλλαγή ενός pixel σε κάθε κύκλο. Στη συνέχεια λόγω της φύσης του μοντέλου έπρεπε να βρούμε

έναν αξιόπιστο τρόπο εκτίμησης παραμέτρων. Ο αλγόριθμος που αναπτύξαμε για το λόγο αυτό αποδεικνύεται μαθηματικά, ενώ η αποτελεσματικότητά του καταδεικνύεται και πρακτικά με επιτυχημένα παραδείγματα συνθέσεων διαφόρων φυσικών υφών.

Τέλος κρίναμε σκόπιμο να συμπεριλάβουμε κάποια βασικά στοιχεία για τη θεωρία των MRF, θέματα που αφορούν την υφή γενικότερα καθώς και μια ευρεία συλλογή μοντέλων που έχουν χρησιμοποιηθεί για την ανάλυσή της με ιδιαίτερη έμφαση σε αυτά που βασίζονται στα MRF.

1.4 Κίνητρα

Τα κίνητρα για την επιλογή του συγκεκριμένου ερευνητικού θέματος είναι αρκετά και σημαντικά. Συνοψίζονται κυρίως στα παρακάτω:

- **Η σημασία επίλυσης του προβλήματος της υφής.** Όπως αναφέραμε και προηγουμένως η ανάλυση και η μοντελοποίηση της υφής με αξιόπιστο και αποδοτικό τρόπο θα βοηθούσε στην επίλυση πολλών πρακτικών προβλημάτων που απασχολούν τη σύγχρονη τεχνολογία
- **Η δυσκολία του προβλήματος.** Όπως θα δούμε αναλυτικότερα στη συνέχεια το πρόβλημα της ανάλυσης και μοντελοποίησης της υφής είναι κάτι ιδιαίτερα δύσκολο καθώς περιέχει χαρακτηριστικά αντιφατικά μεταξύ τους (άλλες φορές είναι ντετερμινιστική και άλλες στατιστική, άλλες φορές αποτελείται από ίδια μοτίβα σε διαφορετικές κλίμακες και άλλες απο ακανόνιστα σχήματα, άλλες φορές περιλαμβάνει απλά γεωμετρικά σχήματα σε συγκεκριμένες θέσεις που καθορίζονται με κάποιον περιοδικό τρόπο και άλλες τα σχήματα τοποθετούνται με εντελώς τυχαίο τρόπο κ.α). Στόχος λοιπόν για τη σχεδίαση ενός τέτοιου μοντέλου είναι ο συγκερασμός όλων αυτών των παραμέτρων που χαρακτηρίζουν την υφή με αναγκαίους βέβαια συμβιβασμούς.
- **Το άλυτο του προβλήματος.** Το πρόβλημα που εξετάζουμε αποτελεί ένα ακόμα ανοιχτό ερευνητικό πεδίο. Δεν έχει μέχρι στιγμής βρεθεί ένας καθολικός τρόπος χειρισμού της υφής άρα η ενασχόληση με αυτό και κάθε νέα πρόταση παρουσιάζουν μεγάλο επιστημονικό ενδιαφέρον.
- **Η ανίχνευση της ανθρώπινης αντίληψης της υφής.** Στη συνέχεια θα δούμε αναλυτικότερα τις προσπάθειες και τα πειράματα που αφορούσαν στην κατανόηση του τρόπου αντίληψης της υφής από τον ανθρώπινο εγκέφαλο. Κάτι τέτοιο αποτελεί αντικείμενο ενός άλλου κλάδου της επιστήμης που ονομάζεται ψυχοφυσική. Η διαρκείς προσπάθειες που έγιναν

και που γίνονται ακόμα πάνω στον τομέα αυτό μαρτυρούν τη σημασία που έχει η υφή ακόμα και για την ανθρώπινη αντίληψη.

- **Η σημασία των MRF μοντέλων.** Είναι πλέον ευρέως αποδεκτό ότι τα μοντέλα που πηγάζουν από τα MRF χρησιμοποιούνται ευρύτατα όχι μόνο στη μοντελοποίηση υφής αλλά και σε μία πληθώρα άλλων φυσικών φαινομένων με στατιστικό ή τυχαίο χαρακτήρα. Αξίζει να σημειωθεί η άποψη ενός επιστήμονα ότι η έλευση των MRF τον 20^ο αιώνα είναι τόσο επαναστατική και έχει τέτοια χρήση στην ανάλυση στατιστικών φαινομένων η οποία μπορεί να συγκριθεί με την έλευση των διαφορικών εξισώσεων τον 18^ο αιώνα!
- **Οι ενδιαφέρουσες μαθηματικές ιδιότητες που χαρακτηρίζουν τα MRF.** Όπως θα δούμε και στη συνέχεια τα MRF έχουν διάφορες ιδιαίτερα ενδιαφέρουσες μαθηματικές ιδιότητες τις οποίες μπορούμε να εκμεταλλευτούμε για τη δημιουργία απλών και αξιόπιστων μοντέλων για την υφή.
- **Η εισαγωγή νέων περιορισμών στη συνάρτηση ενέργειας ενός MRF.** Η σύνδεση ενός MRF με άλλες παραδοσιακές μεθόδους της όρασης υπολογιστών και επεξεργασίας εικόνων είναι ένα πεδίο που αποτελεί πρόκληση για μελέτη, καθώς οδηγεί στην ανάπτυξη μοντέλων που λαμβάνουν υπόψη τους πολλά διαφορετικά μεταξύ τους χαρακτηριστικά.

1.5 Συνεισφορές της διπλωματικής

Οι πρωτότυπες συνεισφορές της παρούσας διπλωματικής εργασίας εντοπίζονται κυρίως στα παρακάτω σημεία:

- Χρήση σταθερού δομικού στοιχείου για σύνθεση, εκτίμηση παραμέτρων και κατηγοριοποίηση υφής σε μοντέλο GRF με μορφολογικούς περιορισμούς.
- Εισαγωγή για πρώτη φορά της έννοιας της κατεύθυνσης των πρωταρχικών στοιχείων που συνθέτουν μια φυσική υφή σε μοντέλο GRF με μορφολογικούς περιορισμούς.
- Εισαγωγή περιορισμού τιμής φωτεινότητας των δομικών στοιχείων στο GRF με μορφολογικούς περιορισμούς με τη βοήθεια ειδικού φίλτρου στα openings και closings που χρησιμοποιούνται κατά τον υπολογισμό της συνάρτησης ενέργειας.

- Πρόταση αλγορίθμων σύνθεσης, εκτίμησης παραμέτρων και κατηγοριοποίησης υφής με βάση το προτεινόμενο μοντέλο.
- Ορισμός του διαφοροποιημένου oriented pattern spectrum και χρήση αυτού για την εκτίμηση παραμέτρων του μοντέλου.

Τα παραπάνω σημεία μας ενδιαφέρουν ιδιαίτερα και θα συνεχίσουμε να τα μελετάμε με στόχο να τα δημοσιεύσουμε.

Εκτός από τα παραπάνω η διπλωματική εργασία περιλαμβάνει:

- Γενική αναφορά στην υφή και τα προβλήματα που σχετίζονται με αυτήν.
- Μια μικρή συλλογή μοντέλων που χρησιμοποιήθηκαν και χρησιμοποιούνται στην ανάλυση υφής.
- Ένα βασικό μαθηματικό υπόβαθρο για τα MRF καθώς και ανάλυση και επεξήγηση συναφών θεμάτων όπως δειγματοληψία των MRF και αλγόριθμοι σύνθεσης, εκτίμηση παραμέτρων και προσομοιούμενη ανόπτηση.
- Αναλυτική επεξήγηση μοντέλων που βασίζονται στα MRF που χρησιμοποιούνται για μοντελοποίηση υφής με παραδείγματα σύνθεσης και παραδείγματα αλγορίθμων σύνθεσης που χρησιμοποιούν.
- Παρουσίαση πειραματικών αποτελεσμάτων σύνθεσης με το νέο μοντέλο που προτείνεται και πραγματοποίηση υπολογισμών που αποδεικνύουν τη αποτελεσματικότητα του μοντέλου.

Κεφάλαιο 2

Υφή

2.1 Τί είναι η υφή;

Η έννοια της υφής είναι λίγο πολύ διαισθητικά γνωστή σε όλους μας. Η χρήση της αποσκοπεί στην περιγραφή της αίσθησης που αποκομίζουμε από την επαφή μας με αντικείμενα του φυσικού κόσμου. Η αίσθηση αυτή ανακτάται κυρίως μέσω της όρασης (αλλαγές στο χρώμα της επιφάνειας) και της αφής (σύσταση της επιφάνειας, μικροανωμαλίες που υπάρχουν σ' αυτή κ.τ.λ). Παρόλο όμως που είναι εύκολο να αντιληφθούμε την υφή σε διάφορα αντικείμενα που παρατηρούμε καθημερινά είναι εξαιρετικά δύσκολο να την ορίσουμε. Κατα καιρούς έχουν προταθεί πολλοί διαφορετικοί ορισμοί από επιστήμονες οι οποίοι ασχολήθηκαν με το θέμα αυτο ,που δείχνουν τον ξεχωριστό τρόπο με τον οποίον ο καθένας μας την αντιλαμβάνεται. Μερικοί από τους ορισμούς αυτούς είναι¹:

- “Θα μπορούσαμε να ορίσουμε την υφή ως αυτό που συνιστά μια μακροσκοπική περιοχή.Η δομή της απλά χαρακτηρίζεται από επαναλαμβανόμενα πρότυπα των οποίων τα στοιχεία ή τα πρωταρχικά στοιχεία (primitives) είναι διατεταγμένα σύμφωνα με ένα κανόνα τοποθέτησης.”[45]
- “Μία περιοχή εντός μιας εικόνας έχει σταθερή υφή εάν ένα σύνολο τοπικών στατιστικών ή άλλων τοπικών ιδιοτήτων στη συνάρτηση της εικόνας είτε είναι σταθερές,είτε αλλάζουν με αργό ρυθμό,είτε είναι κατά προσέγγιση περιοδικές .”[35]
- “Η υφή μιας εικόνας όπως εμείς την αντιλαμβανόμαστε είναι μη σχηματική και κυψελοειδής... Περιγράφεται από τον αριθμό και τον τύπο των πρωταρχικών στοιχείων που περιέχει καθώς και την τοπική οργάνωση

¹Οι ορισμοί που παρατίθενται βρίσκονται στο [72]

(διάταξη) αυτών... Ένα θεμελιώδες χαρακτηριστικό της υφής: δεν μπορεί να αναλυθεί χωρίς να έχουμε ή να υποθέτουμε ως δεδομένο ένα στιγμιότυπο αναφοράς των πρωταρχικών στοιχείων της. Για οποιαδήποτε ομαλή γκριζα επιφάνεια υπάρχει κλίμακα στην οποία όταν εξετάζεται η επιφάνεια αυτή, δεν έχει υφή. Κατόπιν αυξάνοντας την ανάλυση η επιφάνεια παίρνει μία απαλή υφή και στη συνέχεια μια τραχιά υφή.”[27]

- “Η υφή στην περιπτώσή μας ορίζεται ως μια ιδιότητα ενός πεδίου το οποίο δεν έχει συστατικά αριθμήσιμα. Οι σχέσεις διάταξης ανάμεσα στα συστατικά είναι συνεπώς μη προφανής ενώ στο πεδίο δεν εμφανίζεται ξεκάθαρη κατεύθυνση (orientation). Η λογική αυτού του ορισμού είναι ότι η προσοχή του παρατηρητή οδηγείται κυρίως στις γενικές ιδιότητες της εικόνας -π.χ. της συνολικής τραχύτητας, ανωμαλίας ή ομαλότητας. Στη φύση μη αριθμήσιμα (απεριοδικά) πρότυπα γεννώνται με στοχαστικές αντί για ντετερμινιστικές διαδικασίες. Κατά περίπτωση, εντούτοις, το σύνολο όλων των προτύπων χωρίς προφανή μη αριθμήσιμα συστατικά περιλαμβάνουν πολλές ντετερμινιστικές (ακόμα και περιοδικές) υφές.”[58]
- “Η υφή είναι προφανώς μια παράδοξη έννοια. Από τη μια μεριά χρησιμοποιείται συχνά στις μέρες μας στην επεξεργασία της οπτικής πληροφορίας, ειδικότερα για πρακτικές περιπτώσεις ταξινόμησης. Από την άλλη όμως κανείς δεν έχει επιτύχει στο να δημιουργήσει μια κοινώς αποδεκτή ερμηνεία. Η ανάλυση αυτού του παράδοξου, πιστεύουμε, θα εξαρτάται από ένα πλουσιότερο, πιο εξελιγμένο μοντέλο για την επεξεργασία της οπτικής πληροφορίας, κεντρική παραδοχή του οποίου θα είναι αντιπροσωπευτικά συστήματα σε πολλά και διαφορετικά αφαιρετικά επίπεδα. Τα επίπεδα αυτά θα περιλαμβάνουν κυρίως πραγματικές στάθμες φωτεινότητας στον πυθμένα και θα εξελισσονται με τη βοήθεια συντελεστών ακμών, κατεύθυνσης, ίσως και μεγέθους. Συνεπώς φάνεται προφανές ότι αυτές οι πολυεπίπεδες δομές πρέπει να συμπεριληφθούν τόσο στην ερμηνεία της υφής, όσο και στον υπολογισμό αυτών των συντελεστών.”[76]
- “Η έννοια της υφής φαίνεται να εξαρτάται από τρεις παράγοντες:
 - α) ένα τοπικό μοτίβο επαναλαμβάνεται εντός μιας περιχής και είναι μεγάλο συγκριτικά με το μέγεθος της περιοχής
 - β) το μοτίβο αυτό αποτελείται από πρωταρχικά στοιχεία που διατάσσονται με μη τυχαίο τρόπο και
 - γ) τα στοιχεία αυτά είναι βασικά ομοιόμορφες μονάδες που έχουν σχεδόν τις ίδιες διαστάσεις εντός της περιοχής υφής.”[28]

Παρατηρώντας τους παραπάνω ορισμούς θα δούμε ότι δεν είναι γενικοί δηλαδή εστιάζουν σε συγκεκριμένα χαρακτηριστικά της υφής τα οποία χρησιμοποιούν οι ερευνητές για τη συγκεκριμένη εφαρμογή πάνω στην οποία εργάζονται. Μια κοινή συνισταμένη που θα μπορούσαμε να εξάγουμε από αυτούς είναι ότι η υφή βρίσκεται συνήθως στην επιφάνεια των αντικειμένων και χαρακτηρίζεται από τοπικές μεταβολές της φωτεινότητας των pixels. Αποτελείται από μικρά συστατικά τα οποία είναι διατεταγμένα είτε με ορισμένο είτε με τυχαίο τρόπο. Τα μικρά αυτά συστατικά ή πρωταρχικά στοιχεία (primitives) είναι δυνατόν είτε να συνθέτουν όλη την υφή του αντικειμένου είτε μικρότερες περιοχές οι οποίες επαναλαμβάνονται ώστε να μας δώσουν την τελική δομή της.

2.2 Τα τρία προβλήματα της υφής

Η υφή θεωρούμενη ως μια συνάρτηση που εκφράζει τοπικές μεταβολές στις στάθμες του γκριζου στην επιφάνεια ενός αντικειμένου αποτελεί αντικείμενο εκτενούς μελέτης των επιστημόνων σήμερα. Τα βασικά προβλήματα που ανακύπτουν κατά την ανάλυση και τη μελέτη της υφής είναι τρία:

- Η κατηγοριοποίηση της υφής (texture classification)
- Η κατάτμηση ομογενών² περιοχών (texture segmentation)
- Η σύνθεση υφής (texture synthesis)

Πολλές φορές η υφή που έχουν κάποια αντικείμενα μπορεί να μας φανερώσει την κατάσταση στην οποία βρίσκονται (π.χ. ένα καρκινογόνο κύτταρο έχει διαφορετική υφή από ένα υγιές) ή άλλες φορές είναι μοναδική και μας βοηθά να το ξεχωρίζουμε από άλλα (π.χ η υφή του ξύλου ή η υφή του τούβλινου τοίχου). Σε τέτοιες περιπτώσεις μπορούμε να εκμεταλλευτούμε την υφή που απεικονίζεται σε ψηφιακές φωτογραφίες ώστε να αναγνωρίσουμε τα αντικείμενα που υπάρχουν σ' αυτήν, είτε να ανιχνεύσουμε την κατάσταση στην οποία βρίσκονται (π.χ. σε ιατρικές εικόνες). Για να γίνει αυτό πρέπει να έχουμε μια βάση δεδομένων με υφές οι οποίες αντιστοιχούν σε αντικείμενα. Εμείς πρέπει να συγκρίνουμε τη νέα υφή που δεχόμαστε ως είσοδο με όλες τις υφές της βάσης και να την κατατάξουμε σε μία από τις υπάρχουσες κατηγορίες. Άρα τελικά μπορούμε να αποφανθούμε για το είδος ή τη φύση του αντικειμένου στο οποίο ανήκει η υφή εισόδου. Το πρόβλημα αυτό λέγεται κατάταξη ή κατηγοριοποίηση υφής (texture segmentation).

Υπάρχουν όμως περιπτώσεις στις οποίες η υφή που δίνεται ως είσοδος είτε να μην υπάρχει στη βάση δεδομένων είτε για κάποιους λόγους να μην μπορεί να

²περιοχών με την ίδια υφή

κατηγοριοποιηθεί σε μια από τις υπάρχουσες κατηγορίες. Είναι επίσης δυνατό να μη μας ενδιαφέρει η αναγνώριση των υφών που υπάρχουν σε μια ψηφιακή εικόνα αλλά η εύρεση ομογενών περιοχών δηλαδή περιοχών με την ίδια υφή. Σε αυτές τις περιπτώσεις μας ενδιαφέρει η αφηρημένη κατηγοριοποίηση των περιοχών (περιοχή 1 με συγκεκριμένα σύνορα, περιοχή 2 πάλι με συγκεκριμένα σύνορα κ.ο.κ). Το πρόβλημα τότε λέγεται κατάτμηση περιοχών υφής (texture segmentation).

Σε πολλές πάλι εφαρμογές μας ζητείται η δημιουργία τεχνητής υφής. Καλούμαστε δηλαδή να μιμηθούμε τη φύση δημιουργώντας αντικείμενα που μοιάζουν όσο το δυνατόν περισσότερο με τα πραγματικά. Το επίπεδο ρεαλισμού κάθε αντικειμένου όμως το προσδίδει σε μεγάλο βαθμό η πιστή αναπαράσταση της υφής του. Για να γίνει αυτό έχουμε συλλέξει απαιτούμενες πληροφορίες από φωτογραφία που περιέχει τη φυσική υφή που θέλουμε να αναπαράγουμε και κατόπιν με βάση κάποιο μοντέλο συνθέτουμε την τεχνητή υφή χρησιμοποιώντας τις πληροφορίες που έχουμε συλλέξει. Το πρόβλημα λέγεται σύνθεση υφής (texture synthesis).

2.3 Η ανθρώπινη αντίληψη της υφής

Πριν εξετάσουμε τα πιο διαδεδομένα μοντέλα με τα οποία προσπαθούμε σήμερα να αναλύσουμε την υφή είναι χρήσιμο σε αυτό το σημείο να δούμε πως ένας άνθρωπος την αντιλαμβάνεται. Αυτό είναι εξαιρετικά χρήσιμο καθώς με αυτόν τον τρόπο μπορούμε να βρούμε τα χαρακτηριστικά εκείνα της υφής που χρησιμοποιεί η ίδια η φύση για να λύσει κάποιο από τα τρία παραπάνω προβλήματα. Βέβαια πρέπει να πούμε ότι και το ανθρώπινο οπτικό σύστημα αδυνατεί πολλές φορές να ξεχωρίσει αντικείμενα. Χαρακτηριστικό παράδειγμα είναι διάφορα ψάρια που παίρνουν το χρώμα του βυθού για να καμουφλάρονται από τους εχθρούς τους. Τέτοιες περιπτώσεις είναι αντικείμενο εντατικής μελέτης καθώς αναλύονται οι οπτικές διεργασίες που επιτρέπουν το διαχωρισμό υφών άλλες φορές με επιτυχία και άλλες όχι. Δεδομένου ότι το ανθρώπινο οπτικό σύστημα είναι το τελειότερο που γνωρίζουμε μέχρι σήμερα πολλοί νέοι αλγόριθμοι υφής συγκρίνονται από πλευράς απόδοσης με τμηματά του που πραγματοποιούν την ίδια εργασία.

Ο τομέας που εξετάζει την αντίληψη των οπτικών ερεθισμάτων από τον ανθρώπινο εγκέφαλο ονομάζεται *ψυχοφυσιολογία*. Η Σχολή *Gestalt* θεωρείται ως μια από τις σημαντικότερες στον τομέα αυτό καθώς επικεντρώνεται στη διερεύνηση των γενικών αρχών που διέπουν την αντίληψη του ανθρώπινου συστήματος όρασης.

Όσον αφορά ειδικά την υφή μεγάλη είναι η συνεισφορά του Bela Julesz[38]. Ο Julesz πραγματοποιούσε πειράματα με συνθετική υφή προσπαθώντας να ανακα-

λύφει πως θα την επεξεργαστεί το ανθρώπινο μάτι.

Σύμφωνα με τη μελέτη του η υφή αποτελείται από θεμελιώδη στοιχεία τα οποία ονόμασε τεξτόνια (textons) . Αυτά χωρίζονται σε τρία βασικά είδη:

- Προσανατολισμένα Blobs με επιμήκυνση. Σ' αυτά εντάσσονται τα παραλληλόγραμμα, οι ελλείψεις, ευθύγραμμα τμήματα με συγκεκριμένα χρώματα, γωνιακές κατευθύνσεις, παχη και μήκη .
- Terminators δηλαδή άκρα ευθυγράμμων τμημάτων και
- crossings (διασταυρώσεις) ευθυγράμμων τμημάτων.

Κατόπιν όρισε τις παρακάτω στατιστικές ποσότητες:

- Στατιστική πρώτης τάξης (first order statistics) ονομάζεται η πιθανότητα να παρατηρήσει κανείς την γκριζα στάθμη ενός pixel τυχαία επιλεγμένου. Η πιθανότητα αυτή υπολογίζεται από το ιστόγραμμα της εικόνας ενώ εξαρτάται από μεμονωμένες τιμές των pixels και όχι από τις τιμές της γειτονιάς τους.
- Στατιστική δεύτερης τάξης (Second order statistics) ονομάζεται η πιθανότητα παρατήρησης ενός ζεύγους pixels που βρίσκεται στα άκρα ενός διπόλου το οποίο έχει τυχαίο μήκος, τυχαίο προσανατολισμό και βρίσκεται σε τυχαίο σημείο μέσα στην εικόνα.

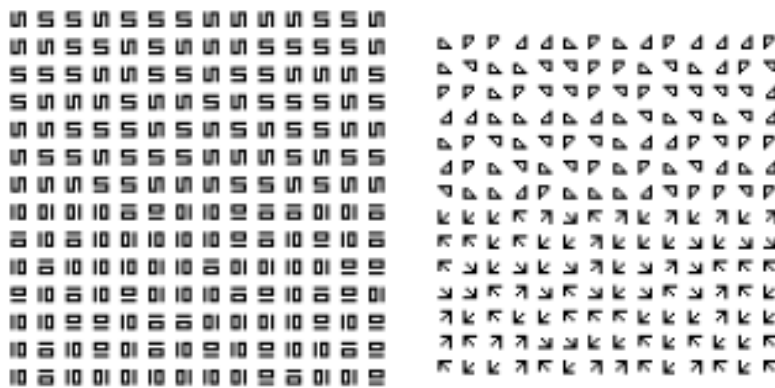
Επίσης διέκρινε την ανθρώπινη όραση σε δύο είδη που αλληλοσυμπληρώνονται:

- Η Preattentive όραση είναι η “πρώτη ματιά” θα λέγαμε. Είναι παράλληλη ,στιγμιαία ,χωρίς προσεκτική εξέταση ενώ καλύπτει μεγάλο οπτικό πεδίο. Διαρκεί περίπου 150msec.
- Η Attentive όραση χαρακτηρίζεται από σειριακή έρευνα ,με εστιασμένη προσοχή ,σε βήματα 50msec,περιορισμένη σε μικρό διάφραγμα.

Τα δύο παραπάνω είδη είναι αλληλένδετα καθώς η preattentive όραση που προηγείται κατευθύνει την attentive όραση σε θέσεις όπου υπάρχουν διαφορές στο είδος ή στην πυκνότητα (αριθμό) των τεξτονίων αλλά αγνοεί τις σχέσεις μεταξύ των χωρικών θέσεων αυτών.

Ο Julesz προχώρησε στη διατύπωση μιας εικασίας την οποία στη συνέχεια προσπάθησε να υποστηρίξει με πρωτοποριακά πειράματα. Η εικασία συνοψίζεται στο ότι δύο υφές δεν είναι διαχωρίσιμες με τη βοήθεια της preattentive όρασης εφ' όσον η στατιστική δεύτερης τάξης των τεξτονίων που περιέχουν είναι παρόμοια.

Για παράδειγμα ας δούμε την πρώτη εικόνα του σχήματος 2.1 που κατασκευάστηκε από τον Julesz και η στατιστική δεύτερης τάξης των υφών που περιέχει είναι η ίδια. Παρατηρούμε ότι με μια γρήγορη ματιά δεν μπορούμε να αναγνωρίσουμε τις δύο διαφορετικές περιοχές υφής. Με αφορμή αυτήν την εικασία ο Julesz προχώρησε και σε άλλα πειράματα με προσεκτικά κατασκευασμένα ζεύγη υφής των οποίων η στατιστική δεύτερης τάξης είναι ίδια.



Σχήμα 2.1: Δυο εικόνες συνθετικής υφής κατασκευασμένες από τον Julesz. Κάθε μία περιέχει περιέχει δύο υφές: μία στο πάνω μισό και μία στο κάτω. Στην πρώτη εικόνα το ανθρώπινο σύστημα όρασης δεν μπορεί να ξεχωρίσει αμέσως (με preattentive όραση) τις δύο περιοχές κάτι που συμβαίνει στη δεύτερη. Από το [72]

Το αποτέλεσμα των πειραμάτων διαφοροποιούνται λίγο από την αρχική εικασία. Συγκεκριμένα παρατηρώντας τις δύο εικόνες του σχήματος 2.1, οι οποίες όπως είπαμε περιέχουν ζεύγη υφών με ίδια στατιστική δεύτερης τάξης, ανακαλύπτουμε ότι στη δεύτερη εικόνα μπορούμε κατευθείαν να αναγνωρίσουμε τις δύο περιοχές. Αυτό συμβαίνει λόγω της διαφορετικής φύσης των τεξτονίων σε κάθε μία. Στην πρώτη παρατηρούμε ότι ο αριθμός των terminators (τελικά σημεία κάθε γραμμής) είναι ίδιος (τέσσερα) για τα δύο είδη τεξτονίων ενώ στη δεύτερη εικόνα τα τεξτόνια της άνω περιοχής έχουν αριθμο terminators τρία ενώ στην κάτω τέσσερα. Αυτό το χαρακτηριστικό μας δίνει τη δυνατότητα να οριοθετήσουμε τις δύο περιοχές αμέσως.

Νεότερες μελέτες στον τομέα της ψυχοφυσιολογίας έδειξαν ότι ο εγκέφαλος πραγματοποιεί αποσύνθεση του οπτικού ερεθίσματος με φίλτρα συχνότητας και

κατεύθυνσης. Αυτή η ιδέα οδήγησε στην υιοθέτηση μοντέλων υφής που βασίζονται στο πολυκάναλο φιλτράρισμα όπως για παράδειγμα το FRAME που θα περιγράψουμε παρακάτω. Περισσότερα για τις μελέτες αυτές καθώς και γενικότερα για την ανθρώπινη αντίληψη της υφής μπορούν να αναζητηθούν πληροφορίες στα [37],[7],[8],[9],[70],[4].

2.4 Τα πιο διαδεδομένα μοντέλα

Θέτοντας σαν στόχο την επίλυση των τριών βασικών προβλημάτων υφής που προαναφέραμε είμαστε υποχρεωμένοι να μοντελοποιήσουμε την υφή, να θεωρήσουμε δηλαδή ότι είναι προϊόν ενός μαθηματικού μοντέλου οπότε γνωρίζοντας το μοντέλο γνωρίζουμε και την ίδια την υφή. Η ανεύρεση του κατάλληλου μοντέλου είναι κάτι ιδιαίτερα σημαντικό στη διαδικασία, καθώς με αυτόν τον τρόπο γίνεται η επικοινωνία του υπολογιστικού συστήματος με τα δεδομένα.

Η δυσκολία εξεύρεσης ενός κοινά αποδεκτού ορισμού για την υφή (ενός θα λέγαμε ντετερμινιστικού ορισμού) φανερώνει μια σημαντική δυσκολία που προκύπτει: εφ' όσον μιλάμε για κάτι τόσο ασαφές, που δεν μπορεί να περιγραφεί απόλυτα, με ποιόν τρόπο θα μπορέσουμε να το μοντελοποιήσουμε; Πώς δηλαδή θα κάνουμε τον υπολογιστή να αντιληφθεί με ακρίβεια κάτι που εμείς αντιλαμβανόμαστε τελείως διαισθητικά και άρα πολλές φορές υποκειμενικά;

Κάτι τέτοιο ουσιαστικά σημαίνει ότι δύσκολα μπορούμε να βρούμε ένα βέλτιστο μοντέλο που θα περιγράφει όλα τα χαρακτηριστικά της υφής που χρειαζόμαστε. Για το λόγο αυτό έχουν προταθεί πολλά μοντέλα τα οποία εξάγουν χαρακτηριστικά χρήσιμα για την εφαρμογή για την οποία προορίζονται. Στη συνέχεια θα προσπαθήσουμε να παρουσιάσουμε συνοπτικά τα πιο σημαντικά, κατηγοριοποιημένα ανάλογα με τα μαθηματικά εργαλεία που χρησιμοποιούν. Για περισσότερες πληροφορίες ο αναγνώστης μπορεί να απευθυνθεί στο [57]

2.4.1 Μοντέλα που βασίζονται στον πίνακα μεταβάσεων (co-occurrence matrix)

Τα μοντέλα αυτά ονομάζονται έτσι διότι προσπαθούν να αντλήσουν πληροφορίες για την υφή μιας εικόνας μέσω του υπολογισμού του πίνακα μεταβάσεων αυτής και κατόπιν υπολογισμού ποσοτήτων που σχετίζονται με τον πίνακα αυτό.

Πριν προχωρήσουμε παραθέτουμε έναν γενικό ορισμό για τον πίνακα μεταβάσεων:

Ορισμός 2.1. Έστω εικόνα I διαστάσεων $m \times n$ έτσι ώστε

$I(x, y) \in [0 \cdots G - 1], \forall (x, y) \in [1 \cdots m, 1 \cdots n]$ (έχουμε G στάθμες του γκριζου) καθώς και διάνυσμα $\vec{d} = (dx, dy)$. Τότε ο πίνακας μεταβάσεων (co-occurrence matrix) $P(i, j)$ διαστάσεων $G \times G$ ορίζεται ως εξής:

Χαρακτηριστικά υφής	Τύπος
Ενέργεια	$\sum_{i,j} P_{ij}^2$
Εντροπία	$-\sum_{i,j} P_{ij} \log P_{ij}$
Μέγιστη Πιθανότητα	$\max_{i,j} P_{ij}$
Αντίστροφη στιγμιαία διαφορά	$\sum_{i,j,i \neq j} \frac{(P_{ij})^l}{ i-j ^k}$
Κοντράστ	$\sum_{i,j} i-j ^k (P_{ij})^l$
Συσχέτιση	$\sum_{i,j} \frac{(i-\mu_x)(j-\mu_y)P_{ij}}{\sigma_x \sigma_y}$
Ομοιογένεια	$\sum_{i,j} \frac{P(i,j)}{1+ i-j }$
Τάση για συσσωμάτωση (clustering)	$\sum_{i,j} (i+j-\mu_x-\mu_y)^k P_{ij}$

Πίνακας 2.1: Μερικές από τις πιο γνωστες ποσότητες που υπολογίζονται από τον πίνακα μεταβάσεων

$P_d(i, j) = |\{(r, s), (t, v) : I(r, s) = i, I(t, v) = j\}|$ όπου $(r, s), (t, v) \in N \times N$ $(t, v) = (r \pm dx, s \pm dy)$ και $|\cdot|$ συμβολίζει τον πληθάρημο συνόλου.

Πρακτικά ο πίνακας μεταβάσεων απεικονίζει τον αριθμό μεταβάσεων από το pixel στάθμης i σε pixel στάθμης j κατά το μέτρο και τη διεύθυνση που ορίζει το διάνυσμα \vec{d} . Με βάση τον πίνακα μεταβάσεων μπορούν να υπολογιστούν και άλλα μεγέθη τα οποία χρησιμοποιούνται στην ανάλυση υφής. Κάποια από αυτά βρίσκονται στον πίνακα 2.1. Στον πίνακα οι ποσότητες μ_x, μ_y και σ_x, σ_y είναι οι μέσες τιμές και οι διασπορές αντίστοιχα των $P_d(x)$ και $P_d(y)$ όπου $P_d(x) = \sum_j P_d(x, j)$ και $P_d(y) = \sum_i P_d(i, y)$.

Πολλοί είναι οι επιστήμονες που χρησιμοποίησαν είτε τον ίδιο τον πίνακα μεταβάσεων είτε κάποια από τις παράγωγες ποσότητές του για τη μοντελοποίηση υφής. Ο πρώτος που χρησιμοποίησε τα στατιστικά του πίνακα μεταβάσεων για πειράματα διαχωρισμού υφής του ανθρώπινου οπτικού συστήματος ήταν ο Julesz το 1961. Από τότε ακολούθησαν αρκετοί. Ενδεικτικά θα αναφέρουμε τα ονόματα των Darling & Joseph (1968) [19], Deutsch & Belknap (1972) [18], Zobrist & Thompson (1985) [1], Bartles & Wied (1975) [52], Rosenfeld & Troy (1970) [3], Haralick (1971) [59], Galloway (1975) [44], Chen & Pavlidis (1979)

[50] καθώς και πολλοί άλλοι.

Οι περισσότεροι από τους παραπάνω επιστήμονες ασχολήθηκαν με το πρόβλημα μοντελοποίησης υφής εκτός από τον Haralick που το 1975 [60] έκανε προσπάθειες να παράγει μια εικόνα στην οποία η τιμή κάθε κυττάρου ανάλυσης resolution cell είναι συνάρτηση της υφής των γειτονικών κυττάρων ανάλυσης και τους Chen & Pavlidis οι οποίοι χρησιμοποίησαν τον πίνακα μετάβασης σε συνδυασμό με έναν “διαίρει και συγχώνευε” (split-and-merge) αλγόριθμο ώστε να καταστήσουν μια εικόνα σε περιοχές παρόμοιας υφής.

Τα χαρακτηριστικά που υπολογίζονται από τον πίνακα μετάβασης έχουν το πλεονέκτημα ότι μπορούν να χαρακτηρίσουν τις τοπικές εξαρτήσεις των κλιμάκων του γκρι σε ένα πρότυπο υφής με τέτοιο τρόπο ώστε να παραμένουν αμετάβλητα σε μονοτονικούς μετασχηματισμούς. Έχουν όμως και σημαντικά μειονεκτήματα όπως το ότι δεν λαμβάνουν υπόψιν το σχήμα των πρωταρχικών στοιχείων της υφής. Επίσης δεν υπάρχει ενδεδειγμένη μέθοδος επιλογής του διανύσματος \vec{d} οπότε είμαστε αναγκασμένοι να έχουμε πληθώρα δεδομένων αδυνατώντας να ξεχωρίσουμε ποιά από αυτά είναι σημαντικά .

2.4.2 Μοντέλα που βασίζονται στην αυτοσυσχέτιση

Ένα ιδιαίτερα χρήσιμο μέγεθος για τη μελέτη της υφής μιας εικόνας είναι η συναρτησιακή αυτοσυσχέτιση της εικόνας αυτής. Η αυτοσυσχέτιση μπορεί να μας δώσει ένα μέτρο της κανονικότητας της θέσης των πρωταρχικών στοιχείων που συνθέτουν την υφή όπως και το πόσο λεία ή τραχιά είναι η υφή που εξετάζουμε. Με αυτόν τον τρόπο εκμεταλλευόμαστε και την ιδιότητα πολλών υφών στις οποίες ο τρόπος τοποθέτησης των πρωταρχικών στοιχείων είναι επαναλαμβανόμενος. Παρακάτω δίνουμε έναν ορισμό για την αυτοσυσχέτιση μιας εικόνας.

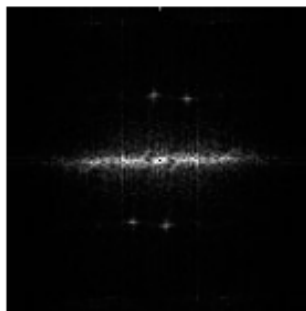
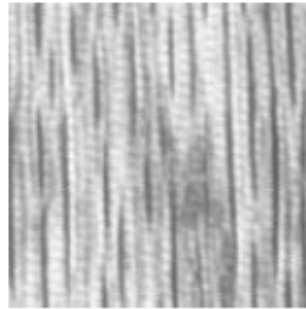
Ορισμός 2.2. Έστω εικόνα I διαστάσεων $m \times n$ η οποία αποτελείται από G στάθμες του γκριζου. Τότε η αυτοσυσχέτιση (autocorrelation) $\rho(x, y)$ ορίζεται ως εξής:

$$\rho(x, y) = \frac{\sum_{u=0}^{m-x} \sum_{v=0}^{n-y} I(u, v) I(u+x, v+y)}{\sum_{u=0}^{m-1} \sum_{v=0}^{n-1} I^2(u, v)}$$

Στον παραπάνω ορισμό δεν έχουμε περιλάβει τις οριακές συνθήκες (στα άκρα της εικόνας ο παραπάνω τύπος δεν δίνει αποτέλεσμα). Η συνάρτηση αυτή όπως προείπαμε μπορεί να απεικονίσει το μέγεθος των στοιχείων υφής (άρα πόσο λεία είναι η υφή). Εάν η υφή είναι τραχιά τότε η αυτοσυσχέτιση ελαττώνεται

αργά συναρτήσει της απόστασης. Αντίθετα αν η υφή μας είναι λεία τότε η αυτοσυσχέτιση ελαττώνεται απότομα. Αν πάλι έχουμε περιοδικότητα στην υφή η ίδια περιοδικότητα μεταφέρεται και στην συνάρτηση αυτοσυσχέτισης.

Επίσης είναι γνωστό ότι η αυτοσυσχέτιση σχετίζεται με το φάσμα ισχύος (είναι ζεύγος μετασχηματισμού Fourier). Με βάση αυτό διερευνήθηκαν οι ιδιότητες του φάσματος ισχύος μιας εικόνας με υφή. Βρέθηκε ότι υφές που περιέχουν κατεύθυνση επηρεάζουν την κατανομή του φάσματος ισχύος κάτι που φαίνεται και στην εικόνα 2.2. Αργότερα έγιναν διάφορες προσπάθειες να χρησιμοποιηθούν αυτά τα χαρακτηριστικά του φάσματος χωρίζοντας το χώρο συχνοτήτων σε δακτυλίους (ώστε να λάβουμε το συχνοτικό περιεχόμενο) και σε σφήνες (ώστε να λάβουμε το κατευθυντικό περιεχόμενο).



Σχήμα 2.2: Μια εικόνα υφής και το φάσμα ισχύος αυτής. Βλέπουμε πως η κατεύθυνση της υφής επηρεάζει την κατανομή του φάσματος ισχύος. Από το [72]

2.4.3 Μοντέλα που βασίζονται στη γραμμική πρόβλεψη

Το διδιάστατο μοντέλο γραμμικής πρόβλεψης μας δίνει τη δυνατότητα να χρησιμοποιήσουμε γραμμικές σχέσεις ώστε να εκτιμήσουμε την τιμή φωτεινότητας ενός pixel δεδομένων των τιμών της γειτονιάς του με σκοπο να χαρακτηρίσουμε την υφή της γειτονιάς αυτής. Για τραχιές υφές οι συντελεστές του γραμμικού μοντέλου είναι σχεδόν σταθεροί ενώ για λείες υφές οι συντελεστές διαφέρουν αρκετά.

Η γραμμική εξάρτηση την οποία έχει ένα pixel μιας εικόνας σε ένα άλλο γνωρίζουμε ότι μπορεί να αναπαρασταθεί κατάλληλα από τη συνάρτηση αυτοσυσχέτισης. Αυτήν την γραμμική εξάρτηση εκμεταλλεύεται το μοντέλο γραμμικής πρόβλεψης για την υφή και πρωτοχρησιμοποιήθηκε από τους Mc Cormick και Jayaramamurthy(1974) [6]. Οι ερευνητές αυτοι εφάρμοσαν τη μέθοδο χρονοσειρών των Box & Jenkins(1970)[31] για εκτίμηση παραμέτρων μιας δοθείσης υφής. Κατόπιν χρησιμοποιούσαν τις εκτημηθείσες παραμέτρους και ένα δοσμένο σύνολο αρχικών τιμών για να δείξουν ότι η συντεθείσα υφή πλησιάζει σε εμφάνιση την πραγματική. Αρκετοί επιστήμονες που ασχολήθηκαν μετέπειτα με το μοντέλο αυτό χρησιμοποίησαν παρόμοια τεχνική.

Στη συνέχεια θα κάνουμε μια σύντομη περιγραφή του μοντέλου. Έστω εικόνα που έχει παραχθεί από τυχαίο θόρυβο και τυχαία ακολουθία K pixels τα οποία έχουν ήδη συντεθεί, η τιμή του επόμενου pixel μπορεί να προκύψει από το γραμμικό συνδυασμό των τιμών των προηγούμενων, προστιθέμενο στο γραμμικό συνδυασμό των προηγούμενων L τιμών τυχαίου θορύβου. Οι συντελεστές των δύο γραμμικών συνδυασμών αποτελούν τις παραμέτρους του μοντέλου. Άρα εάν θέσουμε a_1, a_2, \dots, a_K τις ήδη γνωστές τιμές των K pixels και b_1, b_2, \dots, b_L οι τιμές τυχαίου θορύβου τότε θα έχουμε:

$$a_{K+1} = \sum_{k=0}^{K-1} \alpha_k a_{K-k} + \sum_{\ell=0}^{L-1} \beta_\ell b_{K-\ell}$$

Παρόλο που το μονοδιάστατο μοντέλο που χρησιμοποιήθηκε από τους Reed και Jayaramamurthy (1972) [36] δούλευε αρκετά καλά για τις δύο κάθετες δομές υφής που εμφάνιζαν στην εργασία τους, η απόδοση θα έπεφτε κατά πολύ εάν δοκίμαζαν υφές με διαγώνια κατεύθυνση. Καλύτερη απόδοση σε γενικότερες υφές έχει το διδιάστατο μοντέλο γραμμικής πρόβλεψης. Σ' αυτό κάθε pixel εξαρτάται από μια διδιάστατη γειτονιά $N(i, j)$ η οποία ορίζεται ως εξής:

$$N(i, j) = \{(k, l) | ((i - D \leq k < i) \wedge (j - D \leq l < j + D)) \vee ((k = i) \wedge (j - D \leq l < j))\} \quad (2.1)$$

Συνεπώς το pixel $a(i, j)$ υπολογίζεται ως εξής:

$$a(i, j) = \sum_{k, \ell \in N(i, j)} \alpha(i - k, \ell - j) a(k, \ell) + \sum_{k, \ell \in N(i, j)} \beta(i - k, j - \ell) b(k, \ell)$$

Το μοντέλο αυτό μπορεί να χρησιμοποιηθεί τόσο σε εφαρμογές σύνθεσης όσο και κατάτμησης. Ας υποθέσουμε ότι τα $\{\alpha_c(m, n), \beta_c(m, n)\}$ είναι οι συντελεστές για την υφή c και έστω θ μια τιμή κατωφλίου. Η εκτιμώμενη τιμή του pixel $a(i, j)$ θα είναι:

$$\hat{a}_c(i, j) = \sum_{k, \ell \in N(i, j)} \alpha_c(i - k, \ell - j) a_c(k, \ell) + \sum_{k, \ell \in N(i, j)} \beta_c(i - k, j - \ell) [a_c(k, \ell) - \hat{a}_c(k, \ell)]$$

Υποθέτοντας επομένως μια ομοιόμορφη πρότερη κατανομή, αποφασίζουμε αν το pixel (i, j) ανήκει στην υφή k αν:

$$|a(i, j) - \hat{a}_k(i, j)| \leq |a(i, j) - \hat{a}_l(i, j)| \wedge (|a(i, j) - \hat{a}_k(i, j)| \leq \theta)$$

Εάν $|a(i, j) - \hat{a}_k(i, j)| > \theta$ τότε μπορούμε να αποφασίσουμε ότι το pixel (i, j) ανήκει στο σύνορο.

Το πλεονέκτημα του μοντέλου γραμμικής πρόβλεψης είναι ότι είναι εύκολη η χρήση του εκτιμητή (του διανύσματος των συντελεστών) για σύνθεση υφής. Συνεπώς αυτή η μέθοδος είναι επαρκής να περιγράψει τα πάντα σε μία υφή. Το μειονέκτημα του είναι ότι μπορεί να περιγράψει υφές που αποτελούνται κυρίως από μικρά πρωταρχικά στοιχεία.

2.4.4 Μοντέλα υπολογισμού ενέργειας υφής

Το μοντέλο αυτό προτάθηκε από τον Laws (1980) [39] και έχει σαν ιδέα το φιλτράρισμα της αρχικής εικόνας με διάφορα φίλτρα και κατόπιν τον υπολογισμό της ενέργειας της υφής σε κάθε έξοδο του κάθε φίλτρου με τη βοήθεια μη γραμμικού τελεστή. Υποθέτοντας γειτονιά 7×7 pixels η ενέργεια που αντιστοιχεί στο n -οστό φίλτρο θα είναι:

$$S_n(r, c) = \frac{1}{49} \sum_{i=-3}^3 \sum_{j=-3}^3 |J_n(r + 1, c + j)|$$

Τα φίλτρα που χρησιμοποίησε ο Laws δημιουργήθηκαν από τον ίδιο ώστε να απομονώνουν συγκεκριμένα χαρακτηριστικά. Συγκεκριμένα οι πυρήνες του απομονώνουν το επίπεδο της υφής (level), την τραχύτητά της (edgeness), το σχήμα της (shape), κυματικές ιδιότητες (waves), απότομες μεταβολές (ripples) και

ταλαντώσεις (oscillations). Κατόπιν πραγματοποίησε πείραμα κατηγοριοποίησης υφής και έδειξε ότι το μοντέλο αυτό μπορεί να ξεχωρίσει οκτώ συγκεκριμένες υφές με ακρίβεια 94% αντίθετα από το μοντέλο πινάκων μετάβασης που έχει ακρίβεια 72%.

Το μεγάλο μειονέκτημα όμως του μοντέλου είναι ότι λόγω των μεγάλων γειτονιών που χρησιμοποιεί αποτυγχάνει να χειριστεί με επιτυχία σύνορα μεταξύ υφών καθώς πάντα θα περιλαμβάνει ένα μείγμα τους. Είναι λοιπόν πολύ πιθανό στα σημεία αυτά το διάνυσμα των πινάκων ενέργειας (περιλαμβάνει τους πίνακες ενέργειας που παράγονται από όλα τα φίλτρα) να είναι κοντά όχι σε κάποια από τις ήδη υπάρχουσες υφές του συνόρου αλλά από μια τρίτη, άσχετη με τις δύο πραγματικές.

2.4.5 Μοντέλα υπολογισμού ακμών υφής

Στα μοντέλα αυτά μετρώνται οι ακμές της υφής ανα μονάδα εμβαδού. Μια ακμή που περνά από ένα pixel μπορεί να ανιχνευθεί συγκρίνοντας τις τιμές των pixels της γειτονιάς του αρχικού για τοπικές ιδιότητες. Για την ανίχνευση μικροακμών χρησιμοποιούνται μικρές γειτονιές ενώ για την ανίχνευση μεγάλων ακμών χρησιμοποιούνται μεγαλύτερες.

Η ιδέα αυτή πρωτοδιατυπώθηκε από τους Rosenfeld & Troy (1970)[3] και η τοπική ιδιότητα που πρότειναν για την ανίχνευση ακμών ήταν η συνάρτηση κλίσης (gradient) υπολογισμένη με τη μέθοδο του Roberts. Κατόπιν οι Sutton & Hall (1972)[61] επέκτειναν την ιδέα των Rosenfeld & Troy ορίζοντας μια διαφοροποιημένη συναρτησιμότητα κλίσης η οποία εξαρτάται από την απόσταση ανάμεσα στα pixels. Άρα για κάθε απόσταση d και υποσύνολο I της αρχικής εικόνας που ορίζεται από τη γειτονιά N η διαφοροποιημένη συνάρτηση κλίσης $g(d)$ θα είναι:

$$g(d) = \sum_{(i,j) \in N} \{ |I(i,j) - I(i-d,j)| + |I(i,j) - I(i+d,j)| + |I(i,j) - I(i,j-d)| + |I(i,j) - I(i,j+d)| \}$$

Η καμπύλη της συνάρτησης κλίσης μοιάζει με την καμπύλη της αρνητικής συνάρτησης αυτοσυσχέτισης. Οι Sutton & Hall αυτή τη μέθοδο σε ένα πείραμα αναγνώρισης ασθενειών και πέτυχαν ακρίβεια αναγνώρισης 80% στο διαχωρισμό υγιών από μη υγιείς πνεύμονες.

2.4.6 Μοντέλα μαθηματικής μορφολογίας

Η μαθηματική μορφολογία είναι μια κλασική μαθηματική περιοχή που βασίζεται στη θεωρία συνόλων (set theory) και στην άλγεβρα πλεγμάτων (lattice

algebra) με πολλές εφαρμογές στην όραση υπολογιστών. Ουσιαστικά είναι ένα αλγεβρικό σύστημα μη γραμμικών τελεστών οι οποίοι εφαρμόζονται σε εικόνες και απομονώνουν χρήσιμα χαρακτηριστικά που αφορούν τη γεωμετρική τους δομή. Εκτενέστερα για τη μαθηματική μορφολογία θα μιλήσουμε στο κεφάλαιο οπότε εδώ αναφέρουμε περιληπτικά τις γενικές αρχές που διέπουν αυτήν την ιδέα.

Πριν προχωρήσουμε καλό είναι να δώσουμε συνοπτικούς ορισμούς για τους βασικούς μορφολογικούς τελεστες. Οι δύο βασικότεροι είναι η διαστολή (dilation) και η συστολή (erosion). Στην περίπτωση δυαδικής εικόνας X ο τελεστής διαστολής ισοδυναμεί με την Minkowski πρόσθεση του συνόλου X με ένα σύνολο (παράθυρο) B :

$$X \oplus B = \{x + y : x \in X, y \in B\} = \bigcup_{y \in B} X_{+y}$$

όπου $X_{+y} = \{x + y : x \in X\}$ είναι η μετάθεση του συνόλου X κατά διάνυσμα y .

Αντίστοιχα η συστολή ισοδυναμεί με τη Minkowski αφαίρεση του συνόλου X με το συμμετρικό του συνόλου B , $B^r \equiv \{x : -x \in B\}$

$$X \ominus B = \{x : B_{+x} \subseteq X\} = \bigcap_{y \in B} X_{-y}$$

σε περίπτωση grayscale εικόνων οι παραπάνω τελεστές γενικεύονται ως εξής:

$$(f \oplus B)(x) = \bigvee_{y \in B} f(x - y)$$

$$(f \ominus B)(x) = \bigwedge_{y \in B} f(x + y)$$

Από τους δύο παραπάνω τελεστές παράγονται δύο επίσης πολύ χρήσιμοι: το opening

$$X \circ B = (X \ominus B) \oplus B$$

και το closing

$$X \bullet B = (X \oplus B) \ominus B$$

Η μορφολογική προσέγγιση για την ανάλυση υφής προτάθηκε από τον Matheron (1975) [25] καθώς και τους Serra & Verchery (1973) [34]. Η ιδέα βασίζεται στην έννοια του δομικού στοιχείου (π.χ. ευθ. τμήμα, ρόμβος, κυκλικός δίσκος, τετράγωνο κ.α) το οποίο αναπαρίσταται με ένα δυαδικό σύνολο (όπως το σύνολο B στους παραπάνω ορισμούς των μορφολογικών τελεστών. Αφού έχουμε επιλέξει το δομικό στοιχείο δημιουργούμε ένα σύνολο από δομικά στοιχεία

ίδιου σχήματος με το αρχικό αλλά διαφορετικού μεγέθους. Κατόπιν εφαρμόζουμε τους μορφολογικούς τελεστές opening και closing στην εικόνα υφής με δομικά στοιχεία όλα τα στοιχεία του συνόλου που έχουμε αρχικά δημιουργήσει.

Στη συνέχεια ορίζουμε τη συνάρτηση $G(s)$ ως εξής:

$$G(s) = vol(X \circ sB)$$

όπου sB είναι το δομικό στοιχείο B διάστασης s και με vol συμβολίζουμε τον αριθμό των στοιχείων του συνόλου. Από τη συνάρτηση G μπορούν να υπολογιστούν πολλές χρήσιμες ποσότητες όπως το pattern spectrum ή η χωρική κατανομή size distribution. Τα μεγέθη αυτά απεικονίζονται σε ιστογράμματα με βάση τη διάσταση του δομικού στοιχείου και εξάγονται χρήσιμα συμπεράσματα για την υφή των εικόνων.

Πολλοί ερευνητές έχουν ασχοληθεί και με διαφοροποιημένους μορφολογικούς τελεστές όπως τα reconstruction opening-closing και τα area opening-closing εκμεταλλευόμενοι τις ειδικές ιδιότητες που αυτοί έχουν.

2.4.7 Μοντέλα που βασίζονται στη γεωμετρική κατάτμηση Voronoi

Τα μοντέλα αυτά ανήκουν στα γεωμετρικά μοντέλα τα οποία αρχικά ανιχνεύουν τα πρωταρχικά στοιχεία της υφής και κατόπιν προσπαθούν να βρουν τους νόμους που διέπουν την τοποθέτησή τους στο επίπεδο. Εδώ από κάθε πρωταρχικό στοιχείο υφής εξάγεται ένα σημείο (token) με έναν αλγόριθμο που παρουσιάζεται παρακάτω και κατόπιν προχωρούμε σε κατάτμηση του επιπέδου με τη μέθοδο γεωμετρικής κατάτμησης Voronoi σε πολύγωνα. Η προσεκτική επιλογή των tokens μπορεί μας οδηγήσει από τη μέθοδο Voronoi σε κατάτμηση των περιοχών υφής.

Ο αλγόριθμος για την επιλογή των tokens διατυπώθηκε από τους Tuceryan & Jain και δίνεται συνοπτικά παρακάτω:

1. Εφαρμόζουμε στο αρχικό σχήμα φίλτρισμα με φίλτρο τη λαπλασιανή ενός γκαουσιανού φίλτρου (Laplacian of Gaussian LoG).

2. Επιλογή εκείνων των pixel τα οποία εμφανίζουν τοπικά μέγιστα στη φιλτραρισμένη εικόνα. Η τιμή ενός pixel λέμε ότι εμφανίζει τοπικό μέγιστο όταν η τιμή του είναι μεγαλύτερη από αυτή των έξι από τους οκτώ γείτονές του. Το αποτέλεσμα αυτής της επιλογής είναι μια δυαδική εικόνα.

3. Εφαρμογή ανάλυσης συνδεδεμένων στοιχείων (connected component analysis) στη δυαδική αυτή εικόνα με βάση τους οκτώ κοντινότερους γείτονες. Κάθε συνδεδεμένο στοιχείο αποτελεί και ένα token.

Αφού έχουμε εξάγει τα tokens προχωράμε στην κατάτμηση Voronoi. Στη συνέχεια από τα πολύγωνα που προκύπτουν εξάγουμε διάφορα χρήσιμα χαρα-

κτηριστικά όπως π.χ. το εμβαδό τους. Με βάση κάποιο κριτήριο ομαδοποιούμε τα "όμοια" πολύγωνα. Αυτό έχει σαν συνέπεια να εμφανίζονται στην εικόνα μας μεγάλες περιοχές που συνορεύουν μεταξύ τους. Αυτά τα σύνορα είναι τα σύνορα των υφών της αρχικής εικόνας. Για περισσότερες πληροφορίες ο αναγνώστης μπορεί να ανατρέξει στα [46],[47],[73]

2.4.8 Μοντέλα που βασίζονται στα fractals

Σε πολλές φυσικές αγριες υφές έχουμε σίγουρα διακρίνει πρωταρχικά στοιχεία με έντονες ακμές που επαναλαμβάνονται σε πολλαπλές κλίμακες. Ένας πολύ καλός τρόπος να μοντελοποιηθούν αυτές οι ιδιότητες στην επεξεργασία εικόνων γενικότερα είναι τα *fractals*. Στις εικόνες 2.3 και 2.4 μπορούμε να δούμε δύο fractal συνθέσεις. Ο Mandelbrot[40] ήταν αυτός που αχολήθηκε πρώτος με τη γεωμετρία των fractals βλέποντας την έντονη παρουσία της στη φύση.

Σε πρώτη φάση θα ορίσουμε ένα ντετερμινιστικό fractal και κατόπιν θα αναφερθούμε σε μια βασική του ιδιότητα. Προτού όμως γίνει αυτό θα πρέπει να ορίσουμε κάποια άλλα μεγέθη.

Ορισμός 2.3. Έστω φραγμένο σύνολο A υποσύνολο ενός μετρικού χώρου X . Η διάμετρος του A ορίζεται ως:

$$\text{diam}(A) = \sup_{a,b \in A} d(a,b)$$

όπου d είναι η ευκλείδεια ή κάποια ισοδύναμη απόσταση.

Ορισμός 2.4. Έστω σύνολο A , υποσύνολο ενός μετρικού χώρου X και έστω αριθμός $d \geq 0$. Το d -διάστατο μέτρο Hausdorff του A ($H^d(A)$) ορίζεται ως το *infimum* θετικών αριθμών y τέτοιων ώστε για κάθε $r > 0$, το A καλύπτεται από αριθμήσιμη οικογένεια κλειστών συνόλων με διάμετρο μικρότερη του r έτσι ώστε το άθροισμα των των διαμέτρων τους υψωμένων στην d δύναμη είναι μικρότερο του y .

Ορισμός 2.5. Έστω φραγμένο σύνολο A υποσύνολο ενός μετρικού χώρου X . Η διάσταση Hausdorff του A ($D_H(A)$) είναι το *infimum* αριθμών $d \geq 0$ έτσι ώστε το d -διάστατο μέτρο Hausdorff του A να είναι 0.

Ορισμός 2.6. Έστω φραγμένο σύνολο A υποσύνολο ενός μετρικού χώρου X . Η τοπολογική διάσταση του A ($D_T(A)$) έστω d είναι ο μικρότερος αριθμός για τον οποίο ισχύει: για κάθε κάλυψη C του συνόλου A υπάρχει κάλυψη $C' \subseteq C$ τέτοια ώστε κάθε κάθε σημείο του A εμφανίζεται το πολύ σε $d + 1$ σύνολα της κάλυψης C' .

Ο Mandelbrot όρισε το fractal ως ένα φραγμένο σύνολο στον ευκλείδιο n -διάστατο χώρο για το οποίο ισχύει ότι η διάσταση Hausdorff D_H είναι μεγαλύτερη από την τοπολογική διάσταση D_T δηλαδή ισχύει $D_H > D_T$.

Ας δούμε τώρα την πιο σημαντική ιδιότητα που έχουν τα fractals. Λέγεται αυτοομοιότητα (self-similarity) και εκφράζει την ομοιότητα των δομικών μοτίβων σε διάφορες κλίμακες. Ο επίσημος μαθηματικός της ορισμός δίνεται παρακάτω:

Ορισμός 2.7. Έστω φραγμένο σύνολο A στον n -διάστατο ευκλείδιο χώρο. Το σύνολο αυτό θα λέμε ότι είναι self-similar εάν ισούται με την ένωση N διαφορετικών ξένων μεταξύ τους συνόλων, αντιγράφων του εαυτού του, καθένα από τα οποία έχει υποστεί σμίκρυνση κατά λόγο r .

Η διάσταση D ενός fractal σχετίζεται με το αριθμό N και το λόγο r ως εξής:

$$D = \frac{\log N}{\log (1/r)}$$

Το μέγεθος αυτό (η διάσταση του fractal) μας δίνει ένα μέτρο του πόσο άγρια είναι η επιφάνεια του fractal. Όσο μεγαλύτερη διάσταση τόσο πιο άγρια η επιφάνεια.

Για να μπορέσουμε όμως να μοντελοποιήσουμε πραγματικές υφές με τη βοήθεια των fractals πρέπει με κάποιο τρόπο να εκτιμήσουμε τη διαστασή τους σε μια εικόνα που θεωρούμε ότι έχει παραχθεί με αυτά. Πολλές μέθοδοι έχουν προταθεί για το σκοπό αυτό. Μία από αυτές είναι η εκτίμηση της διάστασης κουτιού (box dimension). Έστω φραγμένο σύνολο A στο n -διάστατο ευκλείδιο επίπεδο και ας φανταστούμε κουτιά διαστάσεως L_{max} τα οποία καλύπτουν το σύνολο A . Μία εκδοχή του συνόλου A σε σμίκρυνση κατά λόγο r θα χρειαστεί $1/r^D$ τέτοια σύνολα. Το νέο λοιπόν σύνολο καλύπτεται από κουτιά διαστάσεως $L = rL_{max}$. Ο αριθμός των κουτιών σχετίζεται με τη fractal διάσταση ως εξής:

$$N(L) = 1/r^D = \left[\frac{L_{max}}{L} \right]^D$$

Από την παραπάνω σχέση μπορούμε να εκτιμήσουμε τη fractal διάσταση λαμβάνοντας διάφορες τιμές για το L και υπολογίζοντας την κλίση της ευθείας

$$\ln (N(L)) = -D \ln (L) + D \ln (L_{max})$$

Ένας καλύτερος τρόπος προτάθηκε από τον Voss. υποθέτουμε ότι θέλουμε να εκτιμήσουμε τη fractal διάσταση της επιφάνειας μιας εικόνας. Έστω $P(m, L)$ η πιθανότητα εύρεσης m σημείων μέσα σε κουτί διάστασης L κεντραρισμένο σε αυθαίρετο σημείο της επιφάνειας A . Έστω M ο συνολικός αριθμός σημείων της εικόνας. Όταν κάποιος καλύπτει την εικόνα με κουτιά πλευράς L

τότε $(M/m)P(m, L)$ είναι το προσδοκώμενο πλήθος κουτιών με m σημεία εντός. Ο συνολικός αριθμός κουτιών που αναμένεται να καλύπτει όλη την εικόνα θα είναι:

$$E[N(L)] = M \sum_{m=1}^N (1/m)P(m, L)$$

Η προσδοκώμενη τιμή $E[N(L)]$ είναι ανάλογη του L^{-D} άρα απο εδώ μπορούμε να εκτιμήσουμε την fractal διάσταση.

Για την περαιτέρω διαφοροποίηση των fractal υφών προτάθηκε και ένα άλλο μέγεθος για το fractal, η *lacunarity*. Η *lacunarity* Λ ορίζεται ως εξής:

$$\Lambda = E \left[\left(\frac{M}{E(M)} - 1 \right)^2 \right]$$

όπου $M = M(L) = KL^D$ είναι η μάζα του fractal. Η *lacunarity* είναι μικρή όταν η υφή είναι λεία και μεγάλη όταν η υφή είναι τραχιά.

Η μελέτη των fractals από τους επιστήμονες δεν σταματά εδώ καθώς πολλά από τα χαρακτηριστικά τους συνεχίζονται να μελετώνται. Επίσης πολλοί έχουν στραφεί στη χρήση τους σε συνδυασμό με άλλα μοντέλα. Για περισσότερες πληροφορίες ο αναγνώστης ας απευθυνθεί στα [40], [51], [10], [64].

2.4.9 Μοντέλα που βασίζονται στα Markov-Gibbs random fields

Τα μαρκοβιανά τυχαία πεδία χρησιμοποιούνται σήμερα ευρύτατα στη μοντελοποίηση φαινομένων τα οποία παρουσιάζουν αφ' ενός μεν μια τυχειότητα αφέτερου δε χαρακτηρίζονται από την εξάρτηση της τιμής τους σε ένα σημείο από γειτονικές τιμές του σημείου αυτού. Στην όραση υπολογιστών βρίσκει πολλές εφαρμογές τόσο στη μοντελοποίηση (σύνθεση, κατηγοριοποίηση, κατάτμηση) της υφής όσο και σε γενικότερα θέματα κατάτμησης, αποκατάστασης και συμπίεσης εικόνων.

Στην περίπτωση μας για τη μοντελοποίηση υφής θεωρούμε την εικόνα μας I (έστω $m \times n$) σαν ένα $m \times n$ αλγεβρικό πλέγμα

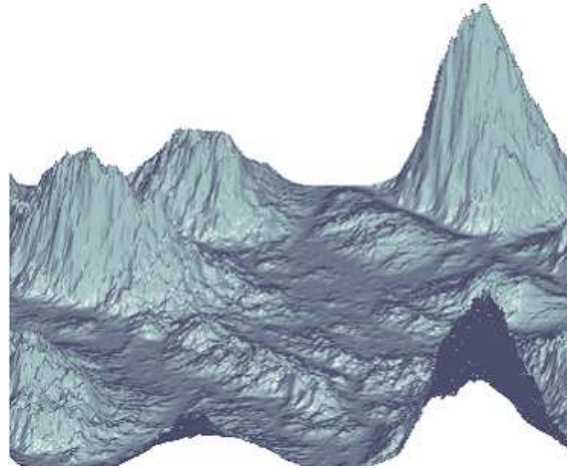
$$L = \{(i, j) | 1 \leq i \leq m, 1 \leq j \leq n\}.$$

Με $I(i, j)$ θα συμβολίζουμε μια τυχαία μεταβλητή η οποία αντιπροσωπεύει την τιμή του pixel (i, j) του πλεγματος. Έστω G το πεδίο τιμών αυτών των τυχαίων μεταβλητών όπου εδώ ουσιαστικά αντιπροσωπεύουν το πλήθος των σταθμών του γκρίζου που έχει η εικόνα μας. Επίσης συμβολίζουμε με

$$\Omega = \{(x_{11}, x_{12}, \dots, x_{1n}, \dots, x_{m1}, \dots, x_{mn}) | x_{ij} \in G, \forall (i, j) \in [1 \dots m, 1 \dots n]\}$$



Σχήμα 2.3: Παράδειγμα τρισδιάστατου fractal τοπίου. Από το [64]



Σχήμα 2.4: Παράδειγμα τρισδιάστατης fractal υψής.

το σύνολο όλων των χρωματισμών της εικόνας I .

Ένα τυχαίο μαρκοβιανό πεδίο είναι ένα πεδίο το οποίο έχει επιπλέον τις ιδιότητες της θετικότητας, της μαρκοβιανής δομής (markovianity) και ομοιογένειας.

Αποδεικνύεται (περισσότερα στο Κεφάλαιο 3) ότι δεδομένου ενός συστήματος γειτονιάς και ενός MRF υπάρχει μοναδικό τυχαίο πεδίο Gibbs και αντίστροφως. Άρα ουσιαστικά το μαρκοβιανό τυχαίο πεδίο μοντελοποιείται με μια συνάρτηση πυκνότητας πιθανότητας Gibbs. Η κατανομή Gibbs δίνεται από τον τύπο

$$P(X = x) = \frac{1}{Z} \exp(-U(x)), \quad x \in \Omega$$

όπου $U(x)$ είναι μια συνάρτηση ενέργειας και Z είναι η σταθερά κανονικοποίησης που ονομάζεται partition function.

Εκείνο που ουσιαστικά χαρακτηρίζει τη μορφή που θα έχουν οι εικόνες που παράγονται από ένα MRF είναι η συνάρτηση ενέργειας και ο τρόπος που ορίζεται. Συνήθως ορίζεται με βάση τις κλίκες που προκύπτουν από το σύστημα γειτονιάς. Οπότε η ενέργεια ισούται με ένα άθροισμα συναρτήσεων που ορίζονται σε κάθε μία από τις κλίκες και λέγονται potential functions. Είναι δηλαδή

$$U(x) = \sum_{c \in Q} V_c(x)$$

όπου $V_c(x)$ είναι οι potential functions και Q το σύνολο που περιέχει τις κλίκες που προκύπτουν από τη γειτονία.

Στο κεφάλαιο 3 θα εξετάσουμε αναλυτικότερα το θεωρητικό υπόβαθρο των MRF ενώ στο κεφάλαιο 4 θα εξετάσουμε τα κυριότερα μοντέλα που έχουν προταθεί για μοντελοποίηση υφής με βάση τα MRF.

2.4.10 Μοντέλα που βασίζονται σε ανάλυση με συστοιχία φίλτρων

Η ανάλυση με συστοιχία φίλτρων (filterbanks) χρησιμοποιείται ευρύτατα σε πολλές επιστήμες αλλά και στην όραση υπολογιστών όταν έχουμε να επεξεργαστούμε δεδομένα τα οποία δεν είναι ομοιόμορφα στο πεδίο της συχνότητας αλλά παρουσιάζουν δραστηριότητα μόνο σε συγκεκριμένες συχνότητες. Κάτι αντίστοιχο χρησιμοποιείται και για την ανάλυση υφής κυρίως για περιπτώσεις όπου αυτή παρουσιάζει περιοδικότητα. Πρέπει να πούμε ότι η έμπνευση για τη χρήση αυτής της τεχνικής οφείλεται στη φύση. Είχαμε ήδη πει σε προηγούμενο μέρος αυτού του κεφαλαίου ότι ο ανθρώπινος εγκέφαλος αποσυνθέτει την οπτική πληροφορία μέσω συστοιχίας φίλτρων διαφόρων συχνοτήτων και κατευθύνσεων.

Σαν φίλτρο συνήθως χρησιμοποιείται το πραγματικό φίλτρο Gabor το οποίο έχει κρουστική απόκριση της μορφής:

$$f(x, y) = \exp\left(-\frac{(x - x_c)^2}{2\sigma_x^2} - \frac{(y - y_c)^2}{2\sigma_y^2}\right) \cos(u(x - x_c) + v(y - y_c))$$

Η επιλογή του φίλτρου Gabor γίνεται αφένός μεν διότι μπορεί να συλλάβει δραστηριότητα σε διάφορες κατευθύνσεις και συχνότητες αφ' ετέρου δε διότι είναι βέλτιστο τόσο στο χρόνο όσο και στη συχνότητα (δηλαδή για ένα Gabor φίλτρο ισχύει πάντα $(\Delta t)_f (\Delta \omega)_f = \frac{1}{2}$ όπου Δt είναι η μέση τετραγωνική διάκριση του φίλτρου και $\Delta \omega$ είναι το μέσο τετραγωνικό εύρος ζώνης του φίλτρου).

Από τη στιγμή που πραγματοποιηθεί η συνέλιξη της εικόνας με τη συστοιχία, στην έξοδο των φίλτρων εφαρμόζεται ένας μη γραμμικός τελεστής που συνήθως αποδίδει μια εκτίμηση της ενέργειας. Το τελικό διάνυσμα των ενεργειών που λαμβάνουμε είναι θα λέγαμε το αποτύπωμα της εικόνας που μελετάμε και με βάση αυτό μπορούμε να προχωρήσουμε σε αποφάσεις κατηγοριοποίησης και κατάτμησης εικόνων.

Ιδιαίτερο ενδιαφέρον παρουσιάζει η προσέγγιση των Κοκκινου, Ευαγγελόπουλου και Μαραγκου (2004) [20] όπου οι εικόνες υφής θεωρούνται ως σήματα πλάτους που έχουν υποστεί διαμόρφωση AM-FM. Πραγματοποιώντας αποδιαμόρφωση στις εικόνες εξόδου του filterbank λαμβάνουν το πλάτος και τη συχνότητα και με βάση αυτά λαμβάνονται οι αποφάσεις του μοντέλου για κατάτμηση περιοχών. Για θέματα αποδιαμόρφωσης που συνδιάζονται με την υφή ο αναγνώστης μπορεί επίσης να αναζητήσει τα [54],[32],[15].

2.5 Εφαρμογές της υφής σε πρακτικά προβλήματα

Οι εφαρμογές της ανάλυσης της υφής σε πρακτικά προβλήματα είναι πολλές και σημαντικές. Από τη στιγμή που σήμερα μπορούμε με άνεση να λαμβάνουμε ψηφιακές φωτογραφίες, ένας τεράστιος όγκος δεδομένων μπορεί να αξιοποιηθεί. Σε πολλές περιπτώσεις η ανάλυση υφής ήδη έχει παίξει σημαντικό ρόλο (όπως π.χ. στην ανάλυση εικόνων από απομακρυσμένους ανιχνευτές) ενώ σε άλλες (όπως στην επιθεώρηση επιφανειών) βρίσκονται συνεχώς νέες εφαρμογές.

Θα μπορούσαμε να χωρίσουμε τις εφαρμογές της υφής σε τέσσερις μεγάλες κατηγορίες: την επιθεώρηση, την ανάλυση ιατρικών εικόνων, την ανάλυση και επεξεργασία εγγράφων και την επεξεργασία εικόνων από απομακρυσμένο ανιχνευτή.

Στην πρώτη κατηγορία μπορούμε να περιλάβουμε εφαρμογές όπως αυτόματη ανίχνευση δυσμορφιών ή αστοχιών διαφόρων υλικών, εκτίμηση παλαιότητας και φθοράς σε χαλιά ή μπογιά αυτοκινήτου, επίβλεψη βιομηχανικής παραγωγής, έλεγχος ποιότητας κ.α. Σε όλες αυτές τις περιπτώσεις μας ενδιαφέρει κυρίως η σωστή κατάτμηση ομογενών περιοχών ώστε να μπορεί το σύστημά μας να αποφασίσει αυτόματα αν και που υπάρχει πρόβλημα και να αντιδράσει. Εδώ η κατάτμηση δεν είναι τόσο δύσκολη καθώς οι εφαρμογές είναι συγκεκριμένες και η είσοδος στο σύστημα περιορισμένη σε εύρος.

Οι ιατρικές εικόνες που λαμβάνονται από αξονικούς ή μαγνητικούς τομογράφους, υπερηχογραφήματα και άλλα μέσα ιατρικής απεικόνισης περιέχουν έναν εξαιρετικό αριθμό δεδομένων για την υγεία του ασθενή. Η ανάλυση υφής βοήθησε και βοηθά στην αξιοποίηση αυτών των δεδομένων κάτι που πολλές φορές είναι δύσκολο με το γυμνό μάτι. Μπορούμε να χρησιμοποιήσουμε την ανάλυση υφής στην αναγνώριση δυσμορφιών (π.χ. ένα κάταγμα), στην αναγνώριση μιας πάσχουσας περιοχής και απομόνωσή της από την υγιή (π.χ. αναγνώριση και οριοθέτηση μιας καρκινογόνου περιοχής) και στην αυτόματη ανάλυση και μέτρηση διαφόρων δεικτών υγείας. Χαρακτηριστικό παράδειγμα της σχέσης της ανάλυσης υφής με τη διάγνωση ασθενειών αποτελεί μια πάθηση των πνευμόνων η οποία αλλάζει την υφή τους απεικονιζόμενη σε μια ακτινογραφία. Άρα η ανίχνευση της ασθένειας μεταφράζεται άμεσα σε κατηγοριοποίηση υφής. Στις ιατρικές εικόνες τα δεδομένα έχουν σαφώς μεγαλύτερο εύρος σε σχέση με αυτές της βιομηχανικής επιθεώρησης. Επίσης θέλουμε άριστα ποσοστά ακρίβειας καθώς καλούμαστε να αποφασίσουμε για θέματα ζωτικής σημασίας.

Η επεξεργασία εγγράφων έγινε στις μέρες μας επιτακτική ανάγκη λόγω της πληθώρας των πληροφοριών που έχουμε στη διάθεσή μας. Η επεξεργασία αυτή μπορεί να αφορά περιπτώσεις αναγνώρισης διεύθυνσης από ταχυδρομικό φάκελο

μέχρι επεξήγηση χαρτών. Σε πολλές περιπτώσεις αναγνώρισης διεύθυνσης το πρώτο βήμα είναι η ικανότητα να ξεχωρίσουμε την περιοχή που περιέχει την πολύτιμη πληροφορία από το υπόλοιπο κομμάτι. Βέβαια σε αυτού του είδους τις εφαρμογές μπορούμε να εκμεταλλευτούμε στο έπακρο τον τρόπο με τον οποίο έχει τυπωθεί ή γραφτεί το προς εξέταση έγγραφο(π.χ. σε ένα γράμμα γνωρίζουμε τις θέσεις που γράφει κανείς τα στοιχεία του παραλήπτη και του αποστολέα). Τέλος μια ακόμη εφαρμογή έχει να κάνει με την επεξεργασία εντύπων και χωρισμό τους σε περιοχές ενδιαφέροντος.

Η ανάλυση εικόνων απομακρυσμένων ανιχνευτών αποτελεί έναν ιδιαίτερα ενδιαφέρον χώρο εφαρμογών ανάλυσης υφής. Στις περιπτώσεις αυτές η υφή παίζει πρωτεύοντα ρόλο καθώς είναι εκείνη η ιδιότητα που σε μεγάλο βαθμό μας βοηθά να ξεχωρίσουμε διαφορετικά υλικά. Χαρακτηριστικό παράδειγμα αποτελεί η ανάλυση διαστημικών εικόνων που απεικονίζουν την επιφάνει ενός απομακρυσμένου πλανήτη. Επίσης η υφή χρησιμοποιείται και στην ανάλυση εικόνων της γης που ελήφθησαν από δορυφόρο. Εδώ μας ενδιαφέρει η αναγνώριση και κατάτμηση κατοικημένων περιοχών, ποταμών, λιμνών και βουνών. Χαρακτηριστικό αυτών των εικόνων είναι η ομοιογένεια που εμφανίζουν οι περιοχές.

Κεφάλαιο 3

MRF: Μία μαθηματική εισαγωγή

Όπως αναφέραμε και στο κεφάλαιο 2 ένας από τους βασικότερους τρόπους μοντελοποίησης της υψής είναι η χρήση μαρκοβιανών τυχαίων πεδίων. Επειδή το θέμα αυτό (η μοντελοποίηση) παρουσιάζεται εκτενώς στα επόμενα κεφάλαια κρίναμε σκόπιμο να παραθέσουμε στο σημείο αυτό μια μαθηματική εισαγωγή η οποία θα βοηθήσει στην αρτιότερη κατανόηση των επόμενων κεφαλαίων. Αρχικά δίνουμε τους βασικούς ορισμούς και τα βασικά θεωρήματα για μαρκοβιανές ακολουθίες και για μαρκοβιανά πεδία και στη συνέχεια παρουσιάζονται διάφορες ιδέες για δειγματοληψία και εκτίμηση παραμέτρων. Επίσης γίνεται μια συνοπτική περιγραφή των συνεχών μοντέλων όσο και μοντέλων με περιορισμούς. Για περισσότερες πληροφορίες πάνω στη θεωρία των μαρκοβιανών τυχαίων πεδίων ο αναγνώστης μπορεί να ανατρέξει στα [71],[33],[29] και [22].

3.1 Μαθηματικοί ορισμοί

3.1.1 Μαρκοβιανές τυχαίες ακολουθίες

Στην παρούσα ενότητα θα δώσουμε κάποιους ορισμούς οι οποίοι θα μας χρησιμεύσουν στη συνέχεια στα μαρκοβιανά τυχαία πεδία καθώς και τον ορισμό της τυχαίας μαρκοβιανής μονοδιάστατης ακολουθίας για πληρέστερη κατανόηση του θέματος.

Ορισμός 3.1. Έστω $A = \{n \in \mathbb{Z} : 1 \leq n \leq N\}$ $N < +\infty$ μια μονοδιάστατη και πεπερασμένη συλλογή από σημεία. Για κάθε σημείο r της συλλογής αυτής θεωρούμε σύνολο $N_r \subset A$ για το οποίο ισχύει:

$$r \notin N_r \quad \text{και} \quad r \in N_k \Leftrightarrow k \in N_r. \quad (3.1)$$

Το σύνολο N_r θα αναφέρεται από εδώ και στο εξής ως γειτονιά του σημείου r και επίσης κάθε ζεύγος σημείων (r, k) που ικανοποιεί τη σχέση (3.1) θα αποκαλούνται στο εξής γείτονες.

Για την περίπτωση μας θα χρησιμοποιήσουμε συμμετρικές ομογενείς γειτονιές οι οποίες ορίζονται ως

$$N_k^{(r)} = \{n \in \mathbb{Z} : n \neq k, k - r \leq n \leq k + r\}, \text{ με } r = 1, 2, \dots$$

Ορισμός 3.2. Μια συλλογή η οποία αποτελείται είτε από ένα σημείο είτε από σημεία τα οποία είναι όλα γείτονες μεταξύ τους ονομάζεται κλίκα.

Θα συμβολίζουμε με $C_p^{(r)}$, όπου $1 \leq p \leq r + 1$, το σύνολο που περιλαμβάνει όλες τις κλίκες που δημιουργούνται από το σύστημα γειτονιάς $N_n^{(r)}$ και περιέχουν ακριβώς p σημεία.

Ας υποθέσουμε τώρα ότι X_n είναι μια τυχαία μεταβλητή η οποία σχετίζεται με το σημείο $n \in A$. Έστω επίσης ότι η τυχαία αυτή μεταβλητή λαμβάνει τιμές από ένα χώρο καταστάσεων $G \in \mathbb{R}$ τέτοιον ώστε $0 \in G$. Στην περίπτωση που ο χώρος καταστάσεων είναι πεπερασμένος και διακριτός έχουμε χωρίς βλάβη της γενικότητας $G = \{0 = a_1 < a_2 < \dots < a_l\}$ και $2 < l < +\infty$. Το σύνολο $X = \{X_n, n = 1, 2, \dots, N\}$ ορίζει μια μονοδιάστατη τυχαία ακολουθία η οποία λαμβάνει τιμές $x = \{x_n, n = 1, 2, \dots, N\}$ στο καρτεσιανό γινόμενο G^N . Παραδείγματος χάρη έστω ότι είναι $X = \{X_1, X_2\}$ και $G = \{0, 1\}$. Τότε αυτή η τυχαία ακολουθία λαμβάνει τιμές $x = \{x_1, x_2\}$ από το σύνολο $G^2 = \{(0, 0), (0, 1), (1, 0), (1, 1)\}$. Κάθε τυχαία ακολουθία χαρακτηρίζεται από μια συνάρτηση πιθανότητας $P(x)$ η οποία εκφράζει την πιθανότητα να λάβει η τυχαία ακολουθία την τιμή x δηλαδή να ισχύει $X = x$. Η μαρκοβιανή τυχαία ακολουθία ουσιαστικά χαρακτηρίζεται από τις ιδιαίτερες ιδιότητες που έχει η συνάρτηση πιθανότητάς της.

Ορισμός 3.3. Μία τυχαία ακολουθία X ορισμένη στο σύνολο A θα λέμε ότι είναι μαρκοβιανή (Markov random sequence) τάξης r αν ισχύουν οι παρακάτω περιορισμοί:

(α) Η πιθανότητα $P(x)$ είναι θετική για κάθε τιμή του x δηλαδή

$$P(h) > 0, \text{ για κάθε } x \in G^N \quad (3.2)$$

(β) Για κάθε $n = 1, 2, \dots, N$, η δεσμευμένη πιθανότητα του X_n , δεδομένων των τιμών του X για όλα τα σημεία του συνόλου $A \setminus \{n\}$ εξαρτάται μόνο από τις τιμές του X στα σημεία που ανήκουν στη γειτονιά του σημείου n δηλαδή

$$P(x_n | x_k, k \in A \setminus \{n\}) = P(x_n | x_k, k \in N_n^{(r)}) \quad (3.3)$$

για κάθε $n = 1, 2, \dots, N$ και κάθε $x \in G^N$

Ο ορισμός αυτός της μαρκοβιανής ακολουθίας είναι πολύ σημαντικός ώστε να καταλάβουμε τις ιδιότητές της αλλάδεν μας δίνει μια σαφή εικόνα της φύσης της συνάρτησης πιθανότητας. Στη συνέχεια παραθέτουμε τον ορισμό της Gibbs τυχαίας ακολουθίας και ένα θεώρημα το οποίο συνδέει την μαρκοβιανή με την Gibbs τυχαία ακολουθία.

Ορισμός 3.4. Κάθε τυχαία ακολουθία X της οποίας η συνάρτηση πιθανότητας $P(x)$ έχει τη μορφή

$$P(x) = \frac{1}{Z_0} \exp \{-\Psi(x)\}, \quad \text{για κάθε } x \in G^N \quad (3.4)$$

όπου

$$Z_0 = \sum_x \exp \{-\Psi(x)\} \quad (3.5)$$

ονομάζεται Gibbs τυχαία ακολουθία (*Gibbs random sequence*).

Θεώρημα 3.1 (Hammersley-Clifford). Μια τυχαία ακολουθία X ορισμένη στο A είναι μαρκοβιανή τυχαία ακολουθία τάξης r αν και μόνο αν υπάρχουν μοναδικές συναρτήσεις ψ (οι οποίες είναι μηδενικές αν τουλάχιστον ένα από τα ορίσματά της είναι μηδενικό), ορισμένες στις κλίκες του A έτσι ώστε

$$P(x) = \frac{1}{Z_0} \exp \left\{ - \sum_{\{n\} \in C_1^{(r)}} \psi_n(x_n) - \sum_{\{n,k\} \in C_2^{(r)}} \psi_{n,k}(x_n, x_k) - \dots - \sum \sum \dots \sum_{\{n,k,\dots,l\} \in C_{r+1}^{(r)}} \psi_{n,k,\dots,l}(x_n, x_k, \dots, x_l) \right\},$$

για κάθε $x \in G^N$ (3.6)

όπου

$$Z_0 = \sum_x \exp \left\{ - \sum_{\{n\} \in C_1^{(r)}} \psi_n(x_n) - \sum_{\{n,k\} \in C_2^{(r)}} \psi_{n,k}(x_n, x_k) - \dots - \sum \sum \dots \sum_{\{n,k,\dots,l\} \in C_{r+1}^{(r)}} \psi_{n,k,\dots,l}(x_n, x_k, \dots, x_l) \right\}. \quad (3.7)$$

Από τις σχέσεις 3.6, 3.7 και τον ορισμό 3.4 συνάγουμε ότι οποιαδήποτε τυχαία ακολουθία X στο A είναι τυχαία μαρκοβιανή ακολουθία εάν και μόνο εάν είναι τυχαία ακολουθία Gibbs με πιθανότητα $P(x)$ η οποία δίνεται από τις εξισώσεις (3.4) και (3.5), με

$$\Psi(h) = \sum_{\{n\} \in C_1^{(r)}} \psi_n(x_n) + \sum_{\{n,k\} \in C_2^{(r)}} \psi_{n,k}(x_n, x_k) + \dots + \sum \sum \dots \sum_{\{n,k,\dots,l\} \in C_{r+1}^{(r)}} \psi_{n,k,\dots,l}(x_n, x_k, \dots, x_l) \quad (3.8)$$

Τέλος πρέπει να προσθέσουμε ότι οι συναρτήσεις ψ αναφέρονται στη βιβλιογραφία ως *local potential functions*.

3.1.2 Μαρκοβιανά τυχαία πεδία

Οι μαρκοβιανές τυχαίες ακολουθίες είναι ιδιαίτερα χρήσιμες στη μοντελοποίηση μονοδιάστατων ποσοτήτων. Στην περίπτωση όμως που το φαινόμενο που θέλουμε να περιγράψουμε είναι διδιάστατο τότε θα χρειαστούμε τα μαρκοβιανά τυχαία πεδία. Αυτά δεν είναι τίποτα άλλο παρά μια διδιάστατη επέκταση των μαρκοβιανών ιδιοτήτων σε δύο διαστάσεις.

Ορισμός 3.5. Έστω $A_2 = \{a \in \mathbb{Z}^2 : a = (m, n), 1 \leq m \leq M, 1 \leq n \leq N\}$, $M, N < +\infty$ μια διδιάστατη και πεπερασμένη συλλογή από σημεία. Για κάθε σημείο r της συλλογής αυτής θεωρούμε σύνολο $N_r \subset A_2$ για το οποίο ισχύει:

$$r \notin N_r \quad \text{και} \quad r \in N_k \Leftrightarrow k \in N_r. \quad (3.9)$$

Το σύνολο N_r θα αναφέρεται από εδώ και στο εξής ως γειτονιά του σημείου r και επίσης κάθε ζεύγος σημείων (r, k) που ικανοποιεί τη σχέση (3.9) θα αποκαλούνται στο εξής γείτονες.

Όπως και πρίν θεωρούμε τυχαία μεταβλητή X_n η οποία αντιστοιχίζεται σε σημείο $s \in A_2$ και παίρνει τιμές από τον χώρο καταστάσεων G . Η συλλογή $X = \{X_s, s \in A_2\}$ είναι ένα διδιάστατο τυχαίο πεδίο και λαμβάνει τιμές $x = \{x_s, s \in A_2\}$ στο καρτεσιανό γινόμενο G^{MN} .

Ορισμός 3.6. Ένα τυχαίο πεδίο X ορισμένο στο σύνολο A_2 θα λέμε ότι είναι μαρκοβιανό (*Markov random field*) τάξης r αν ισχύουν οι παρακάτω περιορισμοί:

(α) Η πιθανότητα $P(x)$ είναι θετική για κάθε τιμή του x δηλαδή

$$P(h) > 0, \text{ για κάθε } x \in G^{MN} \quad (3.10)$$

(β) Για κάθε $s \in A_2$, η δεσμευμένη πιθανότητα του X_s , δεδομένων των τιμών του X για όλα τα σημεία του συνόλου $A_2 \setminus \{s\}$ εξαρτάται μόνο από τις τιμές του X στα σημεία που ανήκουν στη γειτονιά του σημείου s δηλαδή

$$P(x_s | x_t, t \in A \setminus \{s\}) = P(x_s | x_t, t \in N_s^{(r)}) \quad (3.11)$$

για κάθε $s \in A_2$ και κάθε $x \in G^{MN}$

Προφανώς το θεώρημα 3.1 ισχύει και στις δύο διαστάσεις συνεπώς κάθε τυχαίο πεδίο H στο A_2 θα είναι μαρκοβιανό τάξης r αν και μόνο εάν είναι ένα

τυχαίο πεδίο Gibbs του οποίου η συνάρτηση πιθανότητας δίνεται από (3.4) και (3.5) με

$$\begin{aligned} \Psi(h) = & \sum_{\{s_1\} \in C_1^{(r)}} \psi_{s_1}(x_{s_1}) + \sum_{\{s_1, s_2\} \in C_2^{(r)}} \psi_{s_1, s_2}(x_{s_1}, x_{s_2}) + \cdots + \\ & + \sum_{\{s_1, s_2, \dots, s_{r+1}\} \in C_{r+1}^{(r)}} \psi_{s_1, s_2, \dots, s_{r+1}}(x_{s_1}, x_{s_2}, \dots, x_{s_{r+1}}) \end{aligned} \quad (3.12)$$

3.1.3 Γκαουσιανές μαρκοβιανές τυχαίες ακολουθίες (Gaussian Markov random sequences)

Οι γκαουσιανές μαρκοβιανές τυχαίες ακολουθίες (Gaussian MRS) (όπως και τα γκαουσιανά μαρκοβιανά τυχαία πεδία που θα δούμε παρακάτω) ανήκουν σε μια ευρύτερη κατηγορία μοντέλων που ονομάζονται συνεχή μοντέλα. Τα μοντέλα αυτά έχουν σαν χώρο καταστάσεων όχι ένα πεπερασμένο σύνολο αλλά ολόκληρο το \mathbb{R} .

Ένα Gaussian MRS θα έχει συνάρτηση πιθανότητας:

$$P(x) = \frac{[\det \mathbb{B}]^{1/2}}{(\pi k_B T)^{N/2}} \exp\left\{-\frac{x^T \mathbb{B} x}{k_B T}\right\} \quad (3.13)$$

με $\det \mathbb{B} > 0$ όπου

$$\mathbb{B} = \begin{pmatrix} 1 & \beta & 0 & 0 & \dots & 0 & \beta \\ \beta & 1 & \beta & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \beta & 0 & 0 & 0 & \dots & \beta & 1 \end{pmatrix} \quad (3.14)$$

είναι ένας $N \times N$ αραιός, κυκλικά συμμετρικός πίνακας.

Αποδεικνύεται ότι αν $W = \{W_n, n = 1, 2, \dots, N\}$ είναι μια τυχαία γκαουσιανή ακολουθία ορισμένη ως

$$W_n = X_n - E[X_n | X_{n-1}, X_{n+1}] \quad (3.15)$$

τότε θα έχουμε ότι

$$X_n = -\beta(X_{n-1} + X_{n+1}) + W_n \quad (3.16)$$

που αποτελεί την αναδρομική σχέση ορισμού της Gaussian MRS.

3.1.4 Γκαουσιανά μαρκοβιανά τυχαία πεδία (Gaussian Markov random fields)

Αντίστοιχα με τα GMRS επεκτείνοντας το μοντέλο στις δύο διαστάσεις θα έχουμε:

$$P(x) = \frac{[\det \mathbb{B}]^{1/2}}{(\pi k_B T)^{MN/2}} \exp\left\{-\frac{x^T \mathbb{B} x}{k_B T}\right\} \quad (3.17)$$

όπου x είναι ένας πίνακας $MN \times 1$ που καλύπτεται από τα x_{mn} τοποθετημένα λεξιλογραφικά, ο \mathbb{B} είναι ένας $MN \times MN$ μπλόκ κυκλικά συμμετρικός πίνακας με M^2 μπλόκ από $N \times N$ κυκλικά συμμετρικούς πίνακες. Θεωρούμε ότι $\det \mathbb{B} > 0$.

Αποδεικνύεται ότι η αναδρομική αναπαράσταση της GMRS επεκτείνεται και στα GMRF οπότε έχουμε:

$$X_s = -\beta_1(X_t + X_{t'}) - \beta_1(X_u + X_{u'}) - \beta_1(X_v + X_{v'}) - \beta_1(X_w + X_{w'}) + W_s, \quad (3.18)$$

με $s \in A_2$
και

$$s = (m, n), t = (m-1, n), u = (m-1, n-1), v = (m, n-1),$$

$$w = (m-1, n+1), t' = (m+1, n), u' = (m+1, n+1), v' = (m, n+1), w' = (m+1, n-1)$$

όπου $W = \{W_s, s \in A_2\}$ είναι γκαουσιανό τυχαίο πεδίο τέτοιο ώστε:

$$E[W_s W_{s'}] = \begin{cases} k_B T/2 & \text{αν } s = s' \\ 0 & \text{αλλιώς} \end{cases}$$

3.2 Δειγματοληψία των MRF

3.2.1 Διακριτά MRF

Ας υποθέσουμε ότι μας δίνεται ένα MRF και η κατανομή Gibbs στην οποία αυτό αντιστοιχεί. Θέλουμε να βρούμε την τιμή $x_o = \{x_s, s \in A_2\}$ για την οποία ισχύει:

$$P(X = x_o) = \arg \max\{P(X = x)\}$$

ή

$$U(X = x_o) = \arg \min\{U(X = x)\}.$$

Για να το κάνουμε αυτό ένας τρόπος είναι να εξετάσουμε μία μία τις τιμές που μπορεί να πάρει το x και από αυτές να διαλέξουμε εκείνη ή εκείνες

οι οποίες ελαχιστοποιούν τη συνάρτηση ενέργειας. Με έναν απλό συλλογισμό συνδυαστικής ανακαλύπτουμε ότι σε μια τέτοια διαδικασία θα χρειαζόταν να εξετάσουμε $|G|^{|A_2|} = (|G|)^{MN}$ πιθανές τιμές. Για μια τυπική εφαρμογή σε μια εικόνα έχουμε $M = 64, N = 64, G = 16$ άρα μιλάμε για $(16)^{4096} \approx 1,1 \times 10^{4932}$ τιμές!!! Προφανώς κάτι τέτοιο είναι απαγορευτικό υπολογιστικά οπότε προσπαθούμε να λύσουμε το πρόβλημα πραγματοποιώντας δειγματοληψία της κατανομής. Αυτό σημαίνει να κατασκευάσουμε μια ακολουθία από τιμές με τέτοιο τρόπο έτσι ώστε η ακολουθία αυτή να συγκλίνει στην επιθυμητή τιμή x_o .

Μια διαδεδομένη τεχνική για να κάνουμε δειγματοληψία ενός διακριτού MRF (MRF με G πεπερασμένο σύνολο) είναι η δημιουργία μαρκοβιανών αλυσίδων. Αυτές είναι ακολουθίες απεικονίσεων των τυχαίων μεταβλητών X και αν επιλέξουμε σωστά τις πιθανότητες μεταβάσεων αυτών θα συγκλίνουν στην επιθυμητή τιμή x_o . Η μαρκοβιανή ακολουθία $\{X_k, k = 1, 2, \dots\}$ δημιουργείται με χώρο καταστάσεων το G^{MN} και πιθανότητες μεταβάσεων που δίνονται από:

$$p_k(i, j) = P(X_{k+1} = x_j | X_k = x_i), \text{ για } i, j = 1, 2, \dots, r, k = 1, 2, \dots$$

όπου $r = G^{MN}$. Εάν

$$\lim_{k \rightarrow +\infty} P(X_k = x_i | X_1 = x_s) = \pi(x_i), \text{ για κάθε } i = 1, 2, \dots, r, x_s \in G^{MN} \quad (3.19)$$

τότε και για αρκετά μεγάλο k κάθε κατάσταση της μαρκοβιανής αλυσίδας θα είναι ένα δείγμα που θα λαμβάνεται από το π . Συνεπώς η δειγματοληψία ενός MRF μπορεί να γίνει με τη βοήθεια μιας εργοδικής μαρκοβιανής αλυσίδας της οποίας η κατανομή ηρεμίας (equilibrium) είναι η κατανομή Gibbs του αρχικού μας MRF. Το παρακάτω θεώρημα μας εξασφαλίζει την απαιτούμενη εργοδικότητα:

Θεώρημα 3.2. Έστω $\{p_k(i, j), i, j = 1, 2, \dots, r\}$ μια συλλογή πιθανοτήτων μεταβάσεων μιας μαρκοβιανής αλυσίδας $\{X_k, k = 1, 2, \dots\}$, τέτοια ώστε

$$\lim_{k \rightarrow +\infty} p_k(i, j) = p(i, j) \text{ για κάθε } i, j = 1, 2, \dots, r \quad (3.20)$$

όπου $\{p(i, j), i, j = 1, 2, \dots, r\}$ είναι οι πιθανότητες μετάβασης μιας μη ανάγωγης και απεριοδικής μαρκοβιανής αλυσίδας και, εάν

$$\pi_j^{(k)} = \sum_{i=1}^r \pi_i^{(k)} p_k(i, j), \text{ για κάθε } j = 1, 2, \dots, r, k = 1, 2, \dots \quad (3.21)$$

όπου $\{\pi_i^{(k)}, i = 1, 2, \dots, r\}$ είναι μια συλλογή από πιθανότητες τότε

$$\lim_{k \rightarrow +\infty} P(X_k = x_i | X_1 = x_s) = \lim_{k \rightarrow +\infty} \pi_i^{(k)} = \pi_i, \text{ για κάθε } i = 1, 2, \dots, r, k = 1, 2, \dots \quad (3.22)$$

όπου π_i είναι η μοναδική πιθανότητα ηρεμίας της μαρκοβιανής αλυσίδας με πιθανότητες μετάβασης $\{p(i, j), i, j = 1, 2, \dots, r\}$.

Μια χρήσιμη περίπτωση λαμβάνεται εάν μια συλλογή $\{p_k(i, j), i, j = 1, 2, \dots, r\}$ ικανοποιεί εκτός από το θεώρημα (3.2) και την γενική εξίσωση ισορροπίας

$$\pi(x_j) = \sum_{i=1}^r \pi(x_i) p_k(i, j), \text{ για κάθε } j = 1, 2, \dots, r, k = 1, 2, \dots \quad (3.23)$$

Συνεπώς σε αυτήν την περίπτωση $\pi_i = \pi(x_i)$, για κάθε $i = 1, 2, \dots, r$ και η (3.19) προφανώς ικανοποιείται.

Οι βασικές στρατηγικές αλγορίθμων που θα δούμε στη συνέχεια βασίζονται στην αμφίδρομη μαρκοβιανή αλυσίδα για την οποία ικανοποιείται η ειδική εξίσωση ισορροπίας:

$$\pi(x_j) p_k(j, i) = \pi(x_i) p_k(i, j) \text{ για κάθε } i, j = 1, 2, \dots, r, k = 1, 2, \dots \quad (3.24)$$

Προσοχή χρειάζεται στο ότι οι παραπάνω αλυσίδες δεν μας εξασφαλίζουν τις απαιτήσεις του θεωρήματος (3.2) για μη ανάγωγη και απεριοδική μαρκοβιανή αλυσίδα συνεπώς πρέπει να είμαστε ιδιαίτερα προσεκτικοί στις επιλογές μας.

Οι πιο γνωστοί αλγόριθμοι δειγματοληψίας ενός MRF βασίζονται σε πιθανότητες μετάβασης $p_k(i, j)$ της μορφής:

$$p_k(i, j) = s_k(i, j) a_k(i, j), \text{ για } i \neq j \quad (3.25)$$

και

$$p_k(i, j) = 1 - \sum_{i=1, j \neq i}^r p_k(i, j) \quad (3.26)$$

όπου

$$s_k(i, j) = P(X_{k+1/2} = x_j | X_k = x_i), \text{ για } i \neq j \quad (3.27)$$

και

$$a_k(i, j) = P(X_{k+1} = x_j | X_{k+1/2} = x_j, X_k = x_i) \text{ για } i \neq j \quad (3.28)$$

Η πιθανότητα s_k είναι πιθανότητα επιλογής ενώ η πιθανότητα a_k είναι πιθανότητα αποδοχής. Η διαδικασία ουσιαστικά έχει ως εξής: Απο μια απεικόνιση $X_k = x_i$ διαλέγουμε μία απεικόνιση $X_{k+1/2} = x_j$ με πιθανότητα $s_k(i, j)$. Κατόπιν αποδεχόμαστε αυτήν την απεικόνιση ώστε να αποτελέσει μέλος της αλυσίδας μας με πιθανότητα $a_k(i, j)$. Το ενδιάμεσο βήμα $k + 1/2$ ουσιαστικά δηλώνει την αβεβαιότητα για το αν η απεικόνιση που έχει επιλεγεί θα αποτελέσει μέλος της αλυσίδας ή όχι.

Οι πιθανότητες a_k και s_k μπορούν να εκλεγούν με τους παρακάτω τρόπους:

(α) Για την a_k :

1.

$$a_k(i, j) = \begin{cases} 1 & \text{για } t_k(i, j) \geq 1 \\ t_k(i, j) & \text{για } t_k(i, j) < 1 \end{cases}$$

2.

$$a_k(i, j) = \frac{1}{1 + t_k(i, j)}$$

3.

$$a_k(i, j) = \frac{g(\min\{t_k(i, j), t_k(j, i)\})}{1 + t_k(i, j)}$$

όπου $g(x) = 1 + 2(\frac{x}{2})^\gamma$ για $\gamma \geq 1$

4.

$$a_k(i, j) = a_k(j, i) = 1$$

όπου $t_k(i, j) = \frac{\pi(x_j)s_k(j, i)}{\pi(x_i)s_k(i, j)}$.

(β) Για την s_k :

1. Η απεικόνιση x_j διαφέρει από την x_i μόνο στο σημείο $d_k \in A_2$ το οποίο επιλέγεται λεξικογραφικά από όλα τα σημεία του A_2 . Η τιμή του x_j επιλέγεται τυχαία με ίση πιθανότητα από όλες τις πιθανές τιμές εξαιρουμένης αυτής που είχε αρχικά το σημείο d_k .
2. Η μόνη διαφορά με προηγούμενως είναι ότι το σημείο d_k επιλέγεται με ίση πιθανότητα από όλα τα σημεία του A_2 .
3. Η μόνη διαφορά με την πρώτη περίπτωση είναι ότι η τιμή του x_j επιλέγεται με πιθανότητα

$$P(X_{d_k} = x_{j, d_k} | X_d = x_{i, d}, d \in N_{d_k})$$

από όλες τις πιθανές τιμές του G .

4. Η μόνη διαφορά με την τρίτη περίπτωση είναι ότι το σημείο d_k επιλέγεται τυχαία με ίση πιθανότητα από όλα τα σημεία του A_2 .

Ο συνδυασμός των περιπτώσεων 1 για την a_k και 1 για την s_k μας δίνει τον λεξικογραφικό αλγόριθμο Metropolis ενώ ο συνδυασμός 1 και 2 μας δίνει τον αλγόριθμο Metropolis με τυχαία επιλογή. Αντίστοιχα η επιλογή 2 για την a_k και 1 για την s_k μας δίνει τον λεξικογραφικό αλγόριθμο Barker ενώ η επιλογή 2 και 2 αντίστοιχα μας δίνει τον αλγόριθμο Barker με τυχαία επιλογή. Πρόσθετα η επιλογή 3 για την a_k και 1 για την s_k μας δίνει τον λεξικογραφικό αλγόριθμο Hasting ενώ η επιλογή 3 και 2 αντίστοιχα μας δίνει τον αλγόριθμο Hasting με

τυχαία επιλογή. Τέλος επιλογή των περιπτώσεων 4 και 3 ή 4 και 4 μας δίνουν τον αλγόριθμο Gibbs sampler με λεξικογραφική ή τυχαία επιλογή.

Σαν παράδειγμα παραθέτουμε τον ψευδοκώδικα του αλγορίθμου Metropolis με τυχαία επιλογή σημείων. Με αντίστοιχο τρόπο υλοποιούνται και οι υπόλοιποι.

Αλγόριθμος Metropolis (με τυχαία επιλογή σημείου)

1. Διαλέγω ένα σημείο s του A_2 με τυχαίο τρόπο και ίση πιθανότητα από όλα τα σημεία που ανήκουν σ' αυτό και έστω $X_k = x_i$ η αρχική μας απεικόνιση.
2. Θεωρώ νέα απεικόνιση x_j η οποία είναι ίδια με την x_i εκτός από την τιμή του σημείου s . Η νέα τιμή του σημείου s επιλέχθηκε τυχαία με ίση πιθανότητα ανάμεσα στις πιθανές τιμές εκτός από την τιμή που είχε το s στην απεικόνιση x_i .
3. Επιλέγω τη νέα απεικόνιση να ανήκει στη μαρκοβιανή αλυσίδα (να ισχύει δηλαδή $X_{k+1} = x_j$) με πιθανότητα

$$a_k(i, j) = \begin{cases} 1 & \text{για } t_k(i, j) \geq 1 \\ t_k(i, j) & \text{για } t_k(i, j) < 1 \end{cases}$$

4. Ελέγχω το κριτήριο σύγκλισης και αν δεν ικανοποιείται ξαναγυρίζω στο βήμα 1, αλλιώς σταματώ.

3.2.2 Συνεχή MRF

Οι αλγόριθμοι που συζητήθηκαν προηγουμένως να επεκταθούν και στα συνεχή μοντέλα, αυτά δηλαδή που έχουν χώρο καταστάσεων το \mathbb{R} ή κάποιο άλλο μη αριθμησιμο σύνολο. Συγκεκριμένα αν έχουμε ότι $G = \mathbb{R}$ πρέπει να προσέξουμε μια μικρή λεπτομέρεια. Υποθέτουμε ότι βρισκόμαστε στο βήμα k της μαρκοβιανής αλυσίδας με απεικόνιση $X_k = x_i$ και επιλέγουμε ένα σημείο s με ίση πιθανότητα από όλα τα σημεία του A_2 . Θεωρούμε επίσης ότι η νέα απεικόνιση x_j θα διαφέρει από την αρχική μόνο στην τιμή του σημείου s . Πώς όμως θα επιλέξουμε τιμή για το $x_{j,s}$ από τη στιγμή που ο χώρος καταστάσεων δεν είναι πεπερασμένος; Στην περίπτωση αυτή η τιμή του $x_{j,s}$ λαμβάνεται δειγματοληπτώντας την γκαουσιανή δεσμευμένη συνάρτηση πυκνότητας πιθανότητας της μορφής:

$$r_\tau(x_{i,s}, x_{j,s}) = \frac{1}{\sqrt{2\pi\tau}} \exp\left\{-\frac{(x_{i,s} - x_{j,s})^2}{2\tau}\right\} \quad (3.29)$$

όπου $\tau > 0$.

Αποδεικνύεται ότι η παραγόμενη μαρκοβιανή ακολουθία συγκλίνει στην επιθυμητή κατάσταση ελάχιστης ενέργειας.

3.2.3 Προσομοιούμενη ανόπτηση (simulated annealing)

Πριν προχωρήσουμε στην εκτίμηση παραμέτρων ενός μοντέλου MRF θα αναφέρουμε απλώς μια τεχνική η οποία αναφέρεται στη βιβλιογραφία και χρησιμοποιείται για τη γρηγορότερη σύγκλιση των αλγορίθμων που περιγράφηκαν παραπάνω. Στη διεθνή βιβλιογραφία η τεχνική αυτή αναφέρεται ως προσομοιούμενη ανόπτηση (simulated annealing).

Η τεχνική της ανόπτησης χρησιμοποιείται σήμερα ευρέως στη μεταλλουργία ως τρόπος σκλήρυνσης των μετάλλων. Βασίζεται στην απότομη θέρμανση του μετάλλου σε υψηλή θερμοκρασία και κατόπιν την αργή ψύξη του μέχρι μια χαμηλή θερμοκρασία.

Στην περίπτωση μας η τεχνική που χρησιμοποιούμε ονομάστηκε έτσι καθώς σε κάθε βήμα του αλγορίθμου η μεταβλητή T η οποία λογίζεται ως η “θερμοκρασία” του συστήματος τείνει στο μηδέν αφού είχε ξεκινήσει από μια μεγάλη τιμή. Στην απλή περίπτωση υλοποίησης των αλγορίθμων δειγματοληψίας πολλές φορές τείνουν να μένουν για πολλή ώρα σε τοπικά ελάχιστα παρόλο που χρησιμοποιούμε την πιθανότητα αποδοχής a_k η οποία βοηθά σημαντικά στην αποφυγή τέτοιων καταστάσεων. Με την προσομοιωμένη ανόπτηση όμως αρχικά όταν η θερμοκρασία είναι ακόμα υψηλή επιλέγονται απεικονίσεις, οι οποίες δεν είναι οι τοπικά βέλτιστες, η επιλογή τους όμως εξασφαλίζει την τελική σύγκλιση στην κατάσταση ελάχιστης ενέργειας. Μια ακολουθία που μπορούμε να επιλέξουμε για το T είναι η

$$T_k = \frac{A}{\ln(k+1)}, \text{ για } k = 1, 2, \dots \quad (3.30)$$

επιλέγοντας κατάλληλα το A έτσι ώστε να εξασφαλίζεται η σύγκλιση ανεξάρτητα από την αρχική κατάσταση της μαρκοβιανής αλυσίδας. Τέτοιες επιλογές υπάρχουν στη βιβλιογραφία στα [26],[23],[24],[21] και [16]

3.3 Εκτίμηση παραμέτρων

Ένα βασικό πρόβλημα που αντιμετωπίζουμε κατά τη μοντελοποίηση υφής με MRF είναι ο υπολογισμός των παραμέτρων του μοντέλου που βέλτιστα περιγράφουν την πραγματική υφή της εικόνας που δίνουμε ως είσοδο. Θεωρώντας δηλαδή ότι κάθε υφή που θέλουμε να μελετήσουμε αποτελεί προϊόν ενός μοντέλου MRF ψάχνουμε για εκείνο το διάνυσμα συντελεστών του μοντέλου αυτού

το οποίο μας εξασφαλίζει ότι η υφή εισόδου είναι το ground state του μοντέλου μας δεδομένου του συγκεκριμένου διανύσματος συντελεστών.

Αυτό είναι ιδιαίτερα χρήσιμο καθώς έτσι μπορούμε με έναν αλγόριθμο σύνθεσης πόσο κοντά είναι η πραγματική υφή με αυτήν που παράγεται από το μοντέλο άρα και πόσο σωστή είναι η αρχική μας υπόθεση για το ότι η πραγματική μας υφή αποτελεί προϊόν του μοντέλου μας. Επίσης ένας τέτοιος υπολογισμός θα μας βοηθούσε ιδιαίτερα στο στάδιο του classification καθώς είναι λογικό παρόμοιες υφές να έχουν παρόμοια διανύσματα συντελεστών ενώ υφές που διαφέρουν μεταξύ τους θα διαφέρουν και στο διάνυσμα συντελεστών.

Εφ' όσον το μοναδικό μας δεδομένο είναι η εικόνα πραγματικής υφής εκείνο που μπορούμε να κάνουμε είναι να εκτιμήσουμε στατιστικά το διάνυσμα αυτό. Η κλασική τεχνική που χρησιμοποιείται σε τέτοιες περιπτώσεις είναι η εκτιμήτρια μεγίστης πιθανοφάνειας (Maximum Likelihood Estimator) η οποία στην περίπτωση μας υλοποιείται ως εξής:

$$\hat{\theta}_{ML}(X) = \arg \max_{\theta} \left\{ \frac{1}{MN} \ln \pi_{\theta}(X) \right\} = \arg \max_{\theta} \left\{ -\frac{1}{MN} \ln Z_{\theta} - \frac{1}{MN} U_{\theta}(X) \right\} \quad (3.31)$$

Όπως βλέπουμε η εκτίμηση με τη μέθοδο μεγίστης πιθανοφάνειας απαιτεί τον υπολογισμό ή την εκτίμηση της partition function κάτι που είναι ιδιαίτερα δύσκολο στα περισσότερα μοντέλα. Για το λόγο αυτό έχουν προταθεί πιο απλοί τρόποι εκτίμησης των παραμέτρων που μας ενδιαφέρουν. Στο σημείο αυτό θα παρουσιάσουμε συνοπτικά τους πιο βασικούς, αυτούς δηλαδή που είναι πιο διαδεδομένοι:

1. **Coding estimator:** Εξετάζοντας ξανά τον ορισμό (3.6) και ειδικότερα τον περιορισμό (β) ανακαλύπτουμε ότι λόγω αυτής της ιδιότητας έχουμε τη δυνατότητα να διαμερίσουμε το σύνολο A_2 σε δύο σύνολα $A_2^{(1)}$ και $A_2^{(2)}$ έτσι ώστε αφ' ενός $A_2^{(1)} \cap A_2^{(2)} = \emptyset$ και $A_2^{(1)} \cup A_2^{(2)} = A_2$ αφ' ετέρου δε υποθέτοντας ότι διαμερίζουμε και το σύνολο X των τυχαίων μεταβλητών σε X_1 και X_2 αντίστοιχα με τη διαμέριση του συνόλου και δεδομένου του X_2 , το X_1 περιλαμβάνει τυχαίες μεταβλητές ανεξάρτητες μεταξύ τους. Στην περίπτωση αυτή η coding εκτιμήτρια $\hat{\theta}_{CD}$ θα δίνεται από τον τύπο:

$$\begin{aligned} \hat{\theta}_{CD} &= \arg \max_{\theta} \left\{ \frac{1}{MN} \ln \pi_{\theta}(X_1|X_2) \right\} \\ &= \arg \max_{\theta} \left\{ \frac{1}{MN} \sum_{s \in A_2^{(1)}} \ln P_{\theta}(X_s|X_w, w \in N^{(r)}) \right\} \quad (3.32) \end{aligned}$$

Για να υπολογιστεί η παραπάνω παράσταση χρειαζόμαστε γνώση μόνο των τοπικών χαρακτηριστικών συναρτήσεων τις οποίες ούτως ή άλλως γνωρίζουμε ενώ δεν χρειάζεται να υπολογίσουμε την partition function. Άρα με τον τρόπο αυτό μπορούμε εύκολα να εκτιμήσουμε το διάνυσμα παραμέτρων του μοντέλου μας. Παρόλα αυτά έχει αποδειχτεί στην πράξη ότι σε πολλά μοντέλα η coding εκτιμήτρια αποτυγχάνει να δώσει ικανοποιητικά αποτελέσματα.

2. **Maximum Pseudo-Likelihood estimator:** Η ιδέα στην οποία στηρίζεται ο maximum pseudo-likelihood estimator είναι η ίδια με αυτή στην οποία βασίζεται ο coding estimator. Η μόνη διαφορά τους έγκειται στο ότι ο maximum pseudo-likelihood estimator πραγματοποιεί αθροίσεις όχι μόνο στο σύνολο $A_2^{(2)}$ αλλά σε όλο το πλέγμα A_2 . Συνεπώς ο maximum pseudo-likelihood estimator $\hat{\theta}_{MPL}$ θα είναι:

$$\begin{aligned}\hat{\theta}_{MPL} &= \arg \max_{\theta} \left\{ \frac{1}{MN} \ln \prod_{s \in A_2} P_{\theta}(X_s | X_w, w \in N^{(r)}) \right\} \\ &= \arg \max_{\theta} \left\{ \frac{1}{MN} \sum_{s \in A_2} \ln P_{\theta}(X_s | X_w, w \in N^{(r)}) \right\} \quad (3.33)\end{aligned}$$

Στην πράξη έχει βρεθεί ότι στατιστικά αλλά και υπολογιστικά ο maximum pseudo-likelihood estimator είναι πιο ελκυστικός από τον coding estimator.

3. **Monte Carlo Maximum Likelihood Estimator:** Ξαναγυρίζοντας στην αρχική μας ιδέα για χρήση του εκτιμητή μεγίστης πιθανοφάνειας είπαμε ότι κάτι τέτοιο απαιτεί τον υπολογισμό της partition function. Όμως στη βιβλιογραφία έχουν προταθεί διάφοροι εκτιμητές για αυτήν. Ένας από τους πιο βασικούς είναι ο εκτιμητής Jerrum-Sinclair. Συνδυαζόμενος λοιπόν με την κλασική εξίσωση μεγίστης πιθανοφάνειας μας δίνει τη δυνατότητα να εκτιμήσουμε τις παραμέτρους $\hat{\theta}_n$ με τη βοήθεια των εξισώσεων:

$$\hat{\theta}_n(x) = \arg \max_{\theta \in D_n} \hat{l}_{\theta}^{(n)}(x) \quad (3.34)$$

όπου

$$\hat{l}_{\theta}^{(n)}(X, K_n) = \frac{1}{MN} [U_{\theta_{n-1}}(X) - U_{\theta}(X)] - \frac{1}{MN} \ln \frac{1}{K_n} \sum_{k=1}^{K_n} \frac{\exp\{-U_{\theta}(X_k)\}}{\exp\{-U_{\theta_{n-1}}(X_k)\}} \quad (3.35)$$

και $D_n = \{\theta \in \mathbb{R}^d : \|\theta - \theta_{n-1}\| \leq \delta\}$, για μικρό $\delta > 0$ ενώ $\{H_k, k = 1, 2, \dots, K_n\}$ είναι μια συλλογή από MRF που λαμβάνονται με τη βοήθεια των αλγορίθμων σύνθεσης που περιγράψαμε παραπάνω τα οποία συγκλίνουν στο $\pi_{\theta_{n-1}}$. Αξίζει να σημειώσουμε ότι ο εκτιμητής αυτός είναι ιδιαίτερα χρονοβόρος υπολογιστικά και σε μερικές περιπτώσεις ο maximum pseudo-likelihood estimator αποτελεί μια καλή προσέγγισή του. Συνεπώς πρέπει να προσέχουμε σε ποιές εφαρμογές χρειάζεται να τον χρησιμοποιούμε.

3.4 Classification

Η έννοια της κατηγοριοποίησης (classification) αποτελεί κομμάτι ενός ευρύτερου επιστημονικού πεδίου που ονομάζεται αναγνώριση προτύπων. Στο πεδίο αυτό το πρόβλημα της αντιστοίχισης δεδομένων που εμφανίζονται σαν είσοδοι με μία κάθε φορά κατηγορία από ένα σύνολο πολλών κατηγοριών εξετάζεται αφηρημένα και δεν εξαρτάται από το είδος των δεδομένων. Εκείνο που έχει σημασία είναι αφενός μεν η φύση των αντικειμενικών συναρτήσεων (είναι συναρτήσεις που μετρούν κάποιες ποσότητες στα δεδομένα εισόδου και χρησιμοποιούνται διαδικασία λήψης αποφάσεων) και το κριτήριο το οποίο χρησιμοποιείται στη λήψη αποφάσεων.

Όταν εμείς μιλάμε για κατηγοριοποίηση υφής ουσιαστικά εννοούμε εφαρμογή μιας από αυτές τις γενικές μεθόδους της αναγνώρισης προτύπων στο συγκεκριμένο πρόβλημα. Η δημιουργία classifier, μιας αφηρημένης δηλαδή οντότητας η οποία δέχεται σαν είσοδο υφές και με βάση κάποια μαθηματική σχέση αποφασίζει και αντιστοιχίζει την είσοδο σε μία από τις δυνατές κατηγορίες, χωρίζεται σε δύο στάδια: το στάδιο εκμάθησης (learning) και το στάδιο εξέτασης (testing).

Στο πρώτο στάδιο καλούμαστε να “διδάξουμε” τον classifier, δίνοντάς του δηλαδή παραδείγματα εισόδων λέγοντάς του σε ποια κατηγορία αυτές ανήκουν. Ουσιαστικά αυτό είναι ένα στάδιο εκτίμησης παραμέτρων των υφών εισόδου και αντιστοιχισή κάθε διανύσματος συντελεστών στην αντίστοιχη κλάση. Έτσι δημιουργείται ένα σύνολο $C = \{\theta_n^k, k = 1, 2, \dots, P\}$ όπου P ο αριθμός των κλάσεων. Το σύνολο αυτό θα αποτελέσει το βασικό εργαλείο με βάση το οποίο θα ληφθούν οι αποφάσεις.

Στο δεύτερο στάδιο ο classifier είναι πλέον έτοιμος να δεχθεί τυχαίες εισόδους από υφές και να τις κατηγοριοποιεί σε μία από τις γνωστές κλάσεις. Τα κριτήρια που προτείνει η θεωρία στην αναγνώριση προτύπων για λήψη απόφασης είναι αρκετά. Συνήθως όμως χρησιμοποιείται το παρακάτω:

$$k = \arg \max_{1 \leq m \leq P} \{P(X_{test} | \theta_n^m)\} \quad (3.36)$$

όπου X_{test} είναι η εικόνα εισόδου. Ουσιαστικά ο παραπάνω τύπος σημαίνει ότι κατατάσσουμε την εικόνα εισόδου X_{test} στην κλάση k αν για το διάνυσμα παραμέτρων θ_n^k η πιθανότητα $P(X_{test}|\theta_n^k)$ μεγιστοποιείται. Στη βιβλιογραφία προτείνονται πολλά και διαφορετικά κριτήρια απόφασης όπως και τρόποι δημιουργίας αποτελεσματικών αντικειμενικών συναρτήσεων. Μερικά παραδείγματα μπορεί ο αναγνώστης να αναζητήσει στο [62].

3.5 Constrained MRFs

Παρατηρώντας την εξίσωση (3.4) και λαμβάνοντας βέβαια υπόψη το θεώρημα (3.1) που εκφράζει την ισοδυναμία ενός Markov Random Field με ένα και μοναδικό Gibbs Random Field βλέπουμε την ελευθερία που έχει κανείς να ορίσει τη συνάρτηση ενέργειας $\Psi(X)$. Ουσιαστικά εφόσον η κατανομή Gibbs συγκλίνει πάντοτε σε μία εικόνα που ελαχιστοποιεί τη συνάρτηση ενέργειας είναι δυνατόν να εκμεταλλευτούμε αυτό το γεγονός ώστε να προσδώσουμε στην παραγόμενη εικόνα επιθυμητές ιδιότητες. Τότε λέμε ότι έχουμε κατασκευάσει ένα περιορισμένο μαρκοβιανό τυχαίο πεδίο (Constrained MRF) όπου οι περιορισμοί αφορούν στις ιδιότητες τις οποίες έχουμε ενσωματώσει στην συνάρτηση της ενέργειας.

Κάτι τέτοιο πολλές φορές περιορίζει το εύρος των δυνατών παραγόμενων υφών (π.χ όταν έχουμε θέσει έναν στατικό περιορισμό όπως περιορισμό σταθερής μορφολογικής δομής) ενώ άλλες φορές ο περιορισμός είναι δυναμικός οπότε εξαρτάται και επηρεάζεται από τους εκάστοτε συντελεστές του μοντέλου.

Ένα τέτοιο παράδειγμα constrained MRF είναι το μοντέλο των Sivakumar & Goutsias . Το μοντέλο αυτό ονομάζεται Morphologically constrained GRF και η συνάρτηση ενέργειας του ελαχιστοποιείται για εικόνες μορφολογικά λείες (morphologically smooth) τόσο όσο επιτρέπει ένα διάνυσμα συντελεστών \vec{v} . Διαφορετικές υφές επιτυγχάνονται αφένος με την αλλαγή του διανύσματος \vec{v} αλλά και από τη φύση της μορφολογικής συνάρτησης που βρίσκεται στη συνάρτηση ενέργειας. Το μοντέλο αυτό παρουσιάζεται λεπτομερώς στο επόμενο κεφάλαιο.

Ένα άλλο παράδειγμα είναι τα Strong MRFs τα οποία διατυπώθηκαν από τον R. Paget. Αυτά ορίζουν ένα άλλο σύστημα γειτονιάς και έτσι αλλάζει και ο τρόπος ορισμού των τοπικών χαρακτηριστικών του MRF. Με αυτόν τον τρόπο γίνεται πιο εύκολη τόσο η σύνθεση υφών αλλά κυρίως προτείνεται ένας αποτελεσματικός τρόπος εκτίμησης παραμέτρων του μοντέλου. Για περισσότερες λεπτομέρειες ο αναγνώστης ας ανατρέξει στο [48].

Κεφάλαιο 4

Γνωστά μοντέλα υφής που βασίζονται στα MRF

Αφου κάναμε μια μικρή μαθηματική εισαγωγή πάνω στα MRF και συζητήσαμε βασικά θέματα που ανακύπτουν από τη χρήση τους για μοντελοποίηση υφής προχωράμε να δούμε αναλυτικά κάποιες υλοποιήσεις μοντέλων με βάση τα MRF και έχουν σαν στόχο την ανάλυση υφής. Από τα πολλά που έχουν κατα καιρούς δημοσιευτεί διαλέξαμε πέντε απο αυτά αφένός μεν διότι θεωρούμε ότι αντιπροσωπεύουν διαφορετικές κατηγορίες μοντέλων και αφέτέρου διότι το καθένα στην εποχή του αποδείχτηκε αποτελεσματικό στα προβλήματα που κλήθηκε να λύσει. Λόγω έλλειψης χώρου δεν μπορέσαμε να συμπεριλάβουμε περισσότερα μοντέλα. Ο αναγνώστης μπορεί να ανατρέξει στα [11],[55],[2],[74],[66],[75],[56],[68],[30],[41], [65],[43],[63],[13]

4.1 Το διωνυμικό μοντέλο

Το διωνυμικό μοντέλο είναι ένα από τα πρώτα μοντέλα βασισμένα σε MRF που χρησιμοποιήθηκε στη μοντελοποίηση υφής. Οι πρώτοι που έκαναν κάτι τέτοιο είναι οι επιστήμονες Cross και Jain. Θεωρεί ότι η πιθανότητα $P(X = k | neighbors)$ δηλαδή η πιθανότητα ένα pixel στο σημείο (x, y) να έχει την τιμή k δεδομένης της τιμής των γειτόνων του (δηλαδή η local characteristics συνάρτηση) ακολουθεί διωνυμική κατανομή. Συγκεκριμένα έστω x η τιμή του κεντρικού μας pixel στο σημείο (x, y) και $t, t', u, u', v, v', z, z'$ οι τιμές των pixels στα σημεία $(x, y - 1), (x, y + 1), (x - 1, y), (x + 1, y), (x - 1, y - 1), (x + 1, y + 1), (x - 1, y + 1), (x + 1, y - 1)$ αντίστοιχα (θεωρούμε γειτονιά δεύτερης τάξης αν και το μοντέλο μπορεί να επεκταθεί και σε μεγαλύτερες τάξεις) τότε για δυαδικές εικόνες ($G = 2$) θα έχουμε:

$$P(X = x | neighbors(2^{\text{nd}} \text{ order})) = \frac{\exp(xT)}{1 + \exp(T)} \quad (4.1)$$

όπου

$$T = a + b(1, 1)(t + t') + b(1, 2)(u + u') + b(2, 1)(v + v') + b(2, 2)(z + z'). \quad (4.2)$$

Γενικά όταν έχουμε G στάθμες του γκριζου τότε η συνάρτηση ενέργειας θα είναι:

$$U(X) = - \sum_{x_s \in X} (\ln \binom{G-1}{x_s} + x_s T) \quad (4.3)$$

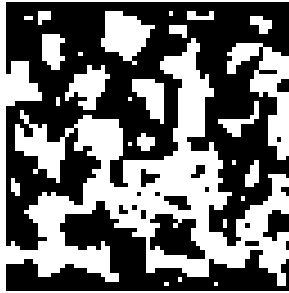
Το μοντέλο αυτό πρωτοδιατυπώθηκε από τον Besag[5] ο οποίος το ονόμασε auto-binomial model.

Κατόπιν οι συγγραφείς Cross και Jain μας δίνουν έναν ενδιαφέρον αλγόριθμο με τον οποίο πραγματοποίησαν σύνθεση υφής με βάση το μοντέλο αυτό. Τον παραθέτουμε και εμείς σε ψευδοκώδικα:

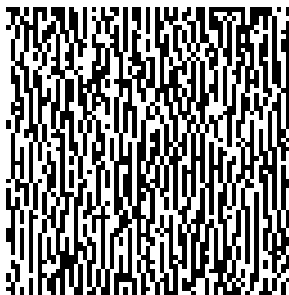
1. Έλεγξε αν ικανοποιείται η συνθήκη τερματισμού.
2. Αν ναι τελείωσε, αλλιώς
3. Διάλεξε δύο σημεία $X(1)$ και $X(2)$ με $X(1) \neq X(2)$ (δηλαδή τα δύο σημεία έχουν διαφορετικό χρώμα)
4. Διάλεξε έναν αριθμό u τυχαία ανάμεσα στο διάστημα $[0 \dots 1]$ θεωρώντας ότι ακολουθεί ομοιόμορφη κατανομή.
5. Θέσε $r = \frac{P(Y)}{P(X)}$ (όπου $P(Y)$ είναι ο χρωματισμός που επιτυγχάνεται ανταλλάσσοντας τις τιμές των $X(1)$ και $X(2)$).
6. Εάν $r > u$ τότε $X = Y$
7. Αλλιώς διατήρησε τον χρωματισμό X

Παρακάτω φαίνονται κάποιες εικόνες υφής οι οποίες έχουν συντεθεί με το μοντέλο των Cross και Jain. Πρέπει να πούμε ότι η σύνθεση δεν πραγματοποιήθηκε με τον παραπάνω αλγόριθμο καθώς παρατηρήσαμε ότι ήταν ιδιαίτερα αργός όσον αφορά τη σύγκλιση.

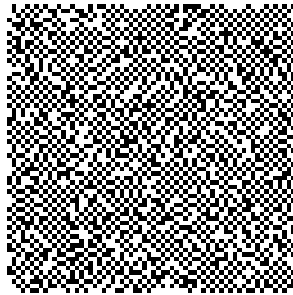
Σαν μέθοδος εκτίμησης παραμέτρων υιοθετήθηκε η μέθοδος του coding estimator. Στην προσπάθεια εκτίμησης παραμέτρων πραγματικών υφών από το άλμπουμ του Brodatz αποκαλύφθηκαν δύο προβλήματα. Αφενός υπήρχε πρόβλημα στην επιλογή της τάξης της γειτονιάς (άλλες υφές είχαν καλύτερο fitting σε γειτονιά πρώτης τάξης ενώ άλλες σε γειτονιά δεύτερης τάξης) και



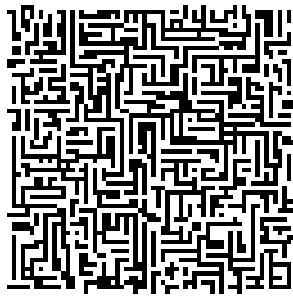
Σχήμα 4.1: Η παραπάνω εικόνα έχει συντεθεί με συντελεστές $a = -6$, $b(1,1) = b(1,2) = 1$, $b(2,1) = b(2,2) = 0$. Παρατηρούμε το έντονο ισοτροπικό clustering που δημιουργείται λόγω των υψηλών τιμών των συντελεστών $b(1,1)$ και $b(1,2)$ που επηρεάζουν το οριζόντιο και το κάθετο clustering αντίστοιχα.



Σχήμα 4.2: Η παραπάνω εικόνα έχει συντεθεί με συντελεστές $a = -0.26$, $b(1,1) = -2$, $b(1,2) = 2.1$, $b(2,1) = 0.13$, $b(2,2) = 0.05$. Το κάθετο clustering είναι έντονο όπως περιμέναμε.



Σχήμα 4.3: Η παραπάνω εικόνα έχει συντεθεί με συντελεστές $a = 5.09$, $b(1,1) = -2.25$, $b(1,2) = -2.16$, $b(2,1) = 0$, $b(2,2) = 0$. Εδώ καταφέρνουμε να περιορίσουμε το clustering θέτοντας αρνητικές τιμές στους αντίστοιχους συντελεστές.



Σχήμα 4.4: Η παραπάνω εικόνα έχει συντεθεί με συντελεστές $a = 0.17$, $b(1,1) = 2.06$, $b(1,2) = 2.05$, $b(2,1) = -2.13$, $b(2,2) = -2.1$. Παρατηρούμε ότι το πάχος των οριζόντιων και κάθετων γραμμών που συνιστούν την εικόνα είναι μικρό λόγω των αρνητικών διαγώνιων συντελεστών.

αφέτερου ότι σε συγκεκριμένο πείραμα που έγινε υπήρχε απόκλιση περίπου 13% κατά μέσο όρο στην εκτίμηση συντελεστών σε διαφορετικές περιοχές της ίδιας εικόνας υφής.

Αυτά τα προβλήματα φάνηκαν κυρίως στο classification (αναφέρεται ως texture matching) πραγματικών υφών. Από τα αποτελέσματα οι επιστήμονες συνέλαξαν ότι το μοντέλο χειρίζεται ικανοποιητικά υφές με μικρά πρωταρχικά στοιχεία ενώ αντίθετα αποτυγχάνει να αποδώσει σωστά υφές με μεγάλο clustering και όχι ιδιαίτερα σαφή κατεύθυνση. Επίσης τα πρωταρχικά στοιχεία ενός τέτοιου τύπου MRF είναι μη γεωμετρικά ενώ όσο αυξάνεται το μέγεθος της εικόνας υφής που θέλουμε να συνθέσουμε τόσο πιο ακριβείς πρέπει να είμαστε στην εκτίμηση των παραμέτρων.

Παρόλα αυτά αναγνωρίζονται διάφορα πλεονεκτήματα του μοντέλου όπως ότι δεν έχει a priori κατεύθυνση αλλά μπορούμε να ορίσουμε γειτονίες σε οποιαδήποτε κατεύθυνση θέλουμε, ότι οι παράμετροι από μόνοι τους είναι επαρκείς ώστε να συνθέσουν μια εικόνα και διάφορα άλλα. Για περισσότερες λεπτομέρειες όσον αφορά το μοντέλο αυτό ο αναγνώστης μπορεί να ανατρέξει στο [12].

4.2 Το μοντέλο Ising

Το μοντέλο Ising πρωτοδημοσιεύτηκε από τον E. Ising το 1925 με σκοπό την εξήγηση του φαινομένου του φερρομαγνητισμού. Στη συνέχεια όμως πολλοί ερευνητές χρησιμοποίησαν το μοντέλο αυτό για τη μοντελοποίηση υφής. Πρέπει να πούμε ότι το αυτούσιο μοντέλο Ising το οποίο είναι δυαδικό ($G = [0, 1]$) συμπίπτει με τη δυαδική περίπτωση του διωνυμικού μοντέλου που εξετάσαμε παραπάνω. Συγκεκριμένα το μοντέλο Ising με χώρο καταστάσεων $G = [-1, 1]$ θα έχει συνάρτηση ενέργειας:

$$U(X) = \sum_{x_s \in X} x_s (a + b_1 t + b_2 t' + b_3 u + b_4 u') \quad (4.4)$$

Επειδή όμως στις δυαδικές εικόνες ο χώρος καταστάσεων είναι $G = [0, 1]$ η παραπάνω εξίσωση θα γίνει:

$$U(X) = \sum_{x_s \in X} (2x_s - 1)(a + b_1(2t - 1) + b_2(2t' - 1) + b_3(2u - 1) + b_4(2u' - 1)) \quad (4.5)$$

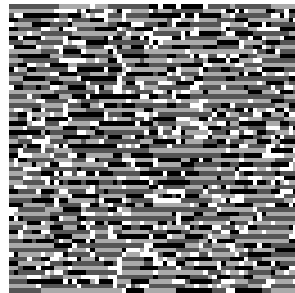
Κατόπιν έγινε επέκταση του μοντέλου αυτού για περισσότερες των δύο στάθμες του γκρίζου και προέκυψε το λεγόμενο multi-level-logistic model με εξίσωση (θεωρώντας γειτονιά δεύτερης τάξης):

$$\begin{aligned}
 U(X) = & \sum_{x_s \in X} ax_s + b_1\delta(x_s - t) + b_2\delta(x_s - t') + b_3\delta(x_s - u) + b_4\delta(x_s - u') + \\
 & + b_5\delta(x_s - v) + b_6\delta(x_s - v') + b_7\delta(x_s - z) + b_8\delta(x_s - z') \quad (4.6)
 \end{aligned}$$

Στις εικόνες 4.5, 4.6, 4.7 και 4.8 φαίνονται διάφορα παραδείγματα σύνθεσης με βάση το multi-level-logistic model.

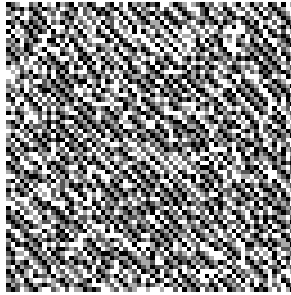


Σχήμα 4.5: Η παραπάνω εικόνα έχει συντεθεί με το μοντέλο Ising. Η ομοιότητα με ιστροπικές δυαδικές εικόνες του διωνυμικού μοντέλου είναι εμφανής.

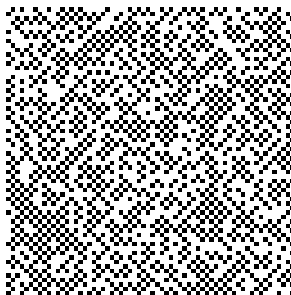


Σχήμα 4.6: Η παραπάνω εικόνα έχει συντεθεί με βάση το multi level logistic μοντέλο. Έχει τέσσερις στάθμες του γκριζου ενώ η δομή της είναι ανιστροπική με έμφαση στον οριζόντιο άξονα.

Γενικά οι εικόνες που προκύπτουν από το multi level logistic μοντέλο μοιάζουν αρκετά με αυτές του διωνυμικού. Παρατηρούμε ότι οι συντελεστές b_i και εδώ ελέγχουν το clustering στις βασικές διευθύνσεις που ορίζουν οι κλίκες της γειτονιάς. Από την άλλη η ιστροπική σύνθεση είναι ευαίσθητη σε μικρές αλλαγές των τιμών των συντελεστών.



Σχήμα 4.7: Ίδια περίπτωση με την προηγούμενη μόνο που εδώ έχει δοθεί έμφαση στον κύριο διαγώνιο άξονα.



Σχήμα 4.8: Η σύνθεση εδώ έγινε με γνώμονα την αποφυγή του clustering. Η εικόνα αποτελείται από δύο στάθμες του γκριζου(δυναδική).

Το multi level logistic μοντέλο χρησιμοποιήθηκε αρκετά από πολλούς επιστήμονες για απλές εφαρμογές μοντελοποίησης υψής. Ενδεικτικά αναφέρουμε τους S.Geman και D.Geman που χρησιμοποιούν το μοντέλο για αποκατάσταση εικόνων με θόρυβο [21], καθώς και τους Derin και Elliot [14] που χρησιμοποιούν το μοντέλο για κατάτμηση ομογενών περιοχών.

4.3 Το μοντέλο των Efros-Leung

Ένα ιδιαίτερα ενδιαφέρον μοντέλο που βασίζεται στα MRF είναι αυτό των ερευνητών Efros και Leung. Η διαφορά του με τα προηγούμενα μοντέλα που εξετάσαμε είναι ότι επιτρέπει μόνο τη σύνθεση υψής δεδομένου ενός δείγματος της πραγματικής υψής που θέλουμε να συνθέσουμε με τη βοήθεια ενός γρήγορου αλγόριθμου.

Έστω I_{real} μια πραγματική εικόνα υψής και έστω $I_{sample} \subset I_{real}$ ένα δείγμα αυτής. Έστω επίσης I η εικόνα που θέλουμε να συνθέσουμε. Θεωρούμε ένα pixel $p \in I$ για το οποίο ορίζουμε την τετραγωνική εικόνα $\omega(p)$ με κέντρο το p και πλάτος w . Η εικόνα αυτή αποτελεί τη γειτονιά πάνω στην οποία ορίζουμε το MRF μοντέλο. Ορίζουμε επίσης την ποσότητα $d(\omega_1, \omega_2)$ ως την "απόσταση" μεταξύ δύο τετραγωνικών γειτονιών ω_1 και ω_2 . Υποθέτουμε ότι έχουμε συνθέσει όλα τα pixels της εικόνας I και μένει να συντεθεί μόνο το pixel p . Για να βρούμε την τιμή του p αρχικά κάνουμε μια εκτίμηση της δεσμευμένης πιθανότητας $P(p|\omega(p))$ (προφανώς λόγω της ιδιότητας των μαρκοβιανών πεδίων η τιμή του p είναι ανεξάρτητη των τιμών των pixels που βρίσκονται εκτός της γειτονιάς του) και κατόπιν πραγματοποιούμε δειγματοληψία αυτής της πιθανότητας.

Θεωρητικά θα μπορούσαμε να ορίσουμε το σύνολο

$$\Omega(p) = \{\omega' \subset I_{real} : d(\omega(p), \omega') = 0\} \quad (4.7)$$

και με βάση αυτό θα μπορούσαμε να εκτιμήσουμε την ζητούμενη πιθανότητα από το ιστόγραμμα των κεντρικών τιμών των γειτονιών που περιλαμβάνονται στο σύνολο Ω_p . Κάτι τέτοιο όμως στην περιπτωσή μας δεν είναι εφικτό καθώς έχουμε στη διάθεσή μας μόνο ένα μικρό τμήμα της πραγματικής υψής συνεπώς είναι πιθανό να μην υπάρχουν σ' αυτό γειτονιές που να απέχουν μηδενική "απόσταση" από το ω_p . Άρα είμαστε υποχρεωμένοι να φτιάξουμε ένα σύνολο $\Omega'(p)$ ως εξής:

$$\Omega'(p) = \{\omega' \in I_{sample} : d(\omega(p), \omega') < (1 + \epsilon)d(\omega_p, \omega_{best})\} \quad (4.8)$$

όπου

$$\omega_{best} = \arg \min_{\omega} d(\omega, \omega(p)) \subset I_{sample} \quad (4.9)$$

και $\epsilon = 0.1$. Άρα πλέον μπορούμε να φτιάξουμε το ιστόγραμμα των τιμών των κεντρικών pixels των γειτονιών που περιλαμβάνει το σύνολο $\Omega'(p)$ και να προβλέψουμε την τιμή που πρέπει να έχει το pixel p είτε δειγματοληπτώντας ομοιόμορφα το ιστόγραμμα είτε χρησιμοποιώντας τις αποστάσεις d ως βάση.

Το μόνο που μας μένει είναι να ορίσουμε την "απόσταση" d . Σε πρώτη φάση θα μπορούσαμε να σκεφτούμε την ευκλείδεια κανονικοποιημένη απόσταση δηλαδή ένα κανονικοποιημένο άθροισμα των τετραγώνων των διαφορών των τιμών των αντίστοιχων pixels. Όμως κάτι τέτοιο θα μετρούσε με το ίδιο βάρος ένα μη ταίριασμα στις τιμές των κεντρικών pixels με ένα μη ταίριασμα στις τιμές pixels του συνόρου. Άρα προσανατολιζόμαστε στη χρήση της ευκλείδειας απόστασης συνελισσόμενη με έναν γκαουσιανό διδιάστατο πυρήνα. Βέβαια από τον υπολογισμό της απόστασης πρέπει να παραλείψουμε pixels τα οποία δεν έχουν συντεθεί ακόμα. Άρα θα πρέπει να εφαρμόσουμε μια "μάσκα" έναν πίνακα δηλαδή ο οποίος θα έχει την τιμή 1 εάν το pixel στη θέση αυτή έχει συντεθεί και 0 αλλιώς. Συνεπώς ο ορισμός της απόστασης d θα είναι:

$$d(\omega(p), \omega) = \sum_i \sum_j \{M(i, j)(\omega(p)(i, j) - \omega(i, j))^2\} * G(i, j) \quad (4.10)$$

όπου G είναι ο γκαουσιανός πυρήνας και M είναι η μάσκα.

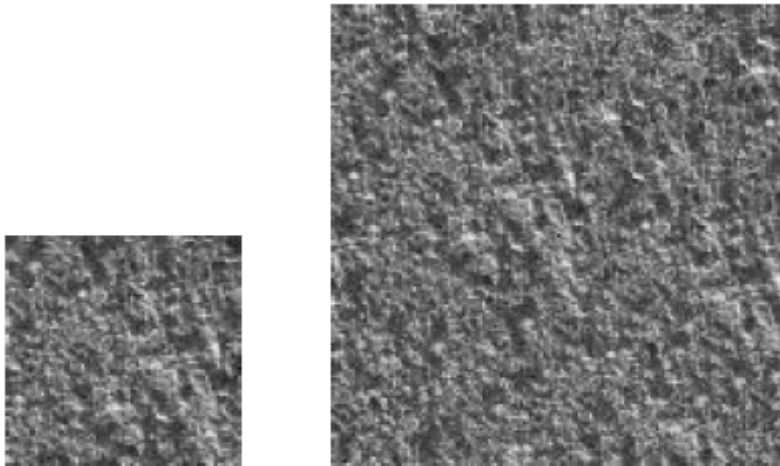
Συνεπώς για να συνθέσουμε την εικόνα I τοποθετούμε μέσα της το δείγμα που μας έχει δοθεί και με μια σπειροειδή κίνηση από τα άκρα του δείγματος προς τα άκρα της προς σύνθεση εικόνας εφαρμόζουμε τον αλγόριθμο που περιγράφηκε παραπάνω.

Στις εικόνες 4.9, 4.10 και 4.11 βλέπουμε διάφορα αποτελέσματα του παραπάνω αλγορίθμου.

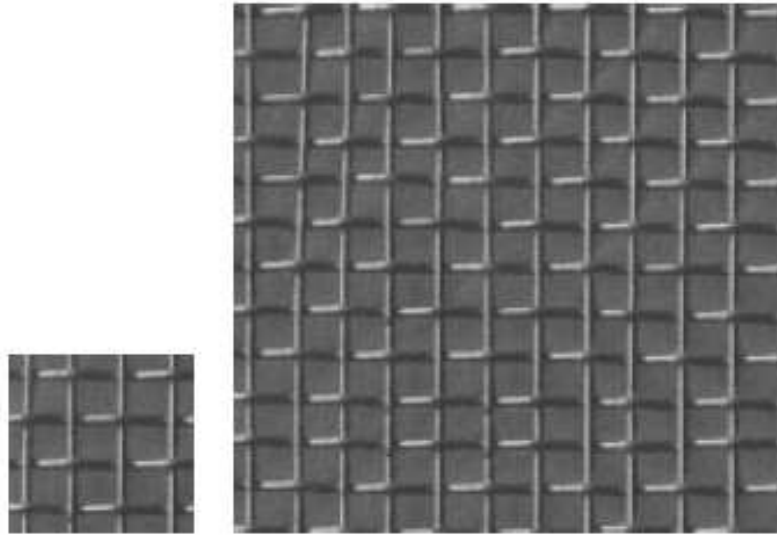
Θα πρέπει να πούμε ότι τα αποτελέσματα σύνθεσης είναι ιδιαίτερα ρεαλιστικά. Το μόνο πρόβλημα που έχει αυτό το μοντέλο είναι ότι δεν μπορεί να χρησιμοποιηθεί για κατάτμηση ή κατηγοριοποίηση υφών. Οι βασικές του χρήσεις αφορούν αποκατάσταση υφής ή απλώς σύνθεση μεγάλων περιοχών υφής.

4.4 Το μοντέλο FRAME

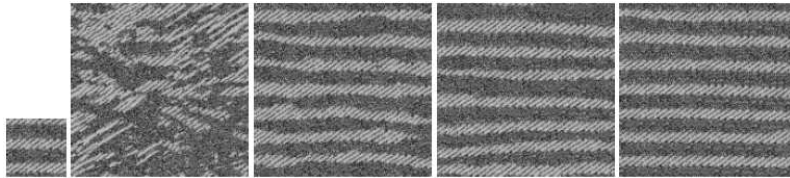
Όταν σε προηγούμενο κεφάλαιο είχαμε μιλήσει για την ανθρώπινη αντίληψη της υφής είχαμε πει ότι νεότερες μελέτες δείχνουν ότι ο εγκέφαλος πραγματοποιεί διαδικασίες που μπορούν να προσομοιωθούν με ανάλυση με συστοιχία φίλτρων. Αυτή η ιδέα ενέπνευσε την ομάδα των ερευνητών Zhu, Wu και Mumford ώστε να αναπτύξουν το μοντέλο FRAME (Filters, Random fields And Maximum Entropy). Γνωρίζουμε ότι τα MRF μοντέλα ορίζονται με βάση μια γειτονιά. Η γειτονιά αυτή είναι συνήθως μικρή και για λόγους υπολογιστικής



Σχήμα 4.9: Πάνω αριστερά μπορούμε να δούμε το δείγμα της υφής που μας δίνεται (γρανίτης) και δεξιά το αποτέλεσμα της σύνθεσης. Από το [17].



Σχήμα 4.10: Το συγκεκριμένο παράδειγμα αφορά σύνθεση περιοδικής υφής. Παρατηρούμε ότι το μοντέλο απεικονίζει πολύ καλά τόσο την περιοδική όσο και την τυχαία υφή. Από το [17].



Σχήμα 4.11: Εδώ παρατηρούμε πόσο επηρεάζει τη σύνθεση το μήκος του παραθύρου. Στην πρώτη συντεθειμένη εικόνα έχουμε μήκος 5 ενώ οι επόμενες έχουν μήκος 11,15,23 αντίστοιχα. Από το [17].

αποτελεσματικότητας χρησιμοποιούνται μόνο κλίκες των δύο pixels. Κάτι τέτοιο έχει σαν αποτέλεσμα την αδυναμία τέτοιων μοντέλων να αναλύσουν υφές με μεγάλα δομικά στοιχεία. Από την άλλη η ανάλυση με συστοιχία φίλτρων είναι εγγενώς τοπική διαδικασία άρα και πάλι δεν λύνεται το παραπάνω πρόβλημα.

Για να αντιμετωπιστεί αυτή η ανεπάρκεια προτάθηκε από τους παραπάνω ερευνητές το χτίσιμο της κατανομής πιθανότητας του μοντέλου με τις αποκρίσεις της προς συνθέσης εικόνας συνελισσόμενης με ένα filterbank χρησιμοποιώντας τον περιορισμό της μέγιστης εντροπίας (Maximum Entropy).

Συγκεκριμένα τα βήματα που απαιτούνται για τη δημιουργία της συνάρτησης πιθανότητας αναφέρονται συνοπτικά παρακάτω:

- Θεωρούμε εικόνα I και έστω $H = \{h_i, 1 \leq i \leq n\}$ μια συστοιχία φίλτρων. Έστω $\phi_i(x) = I(x) * h_i(x)$ είναι η απόκριση της εικόνας I ως προς το φίλτρο ϕ_i . Έστω επίσης $H(\phi_i)$ το ιστόγραμμα της εικόνας ϕ_i . Θεωρώντας λοιπόν συνάρτηση πιθανότητας p απαιτούμε η προσδοκώμενη τιμή της ως προς ϕ_i να ακολουθεί την κατανομή του ιστογράμματος της ϕ_i δηλαδή

$$E_p[\phi_i] = \int p(x)\phi_i(x)dx \sim H(\phi_i) \quad (4.11)$$

- Όσες συναρτήσεις p ικανοποιούν τη σχέση (4.11) τοποθετούνται σε ένα σύνολο Ω . Από αυτές λοιπόν διαλέγουμε πιο απλή, εκείνη δηλαδή που έχει τη μεγαλύτερη εντροπία. Συνεπώς το πρόβλημα μπορεί να διατυπωθεί ως η εκτίμηση της συνάρτησης \hat{p} έτσι ώστε

$$\hat{p} = \arg \max \left\{ - \int p(x) \log p(x) dx \right\} \quad (4.12)$$

όπου $p \in \Omega$ και $\int p(x) dx = 1$.

- Χρησιμοποιώντας τους πολλαπλασιαστές Lagrange η συνάρτηση που επιλύει το παραπάνω πρόβλημα ελαχιστοποίησης είναι της μορφής

$$p(x; \Lambda_k) = \frac{1}{Z} \exp \left\{ - \sum_x \sum_{n=1}^N \lambda_n(\phi_n(x)) \right\} \quad (4.13)$$

όπου $\Lambda = \{\lambda_1, \lambda_2, \dots, \lambda_N\}$ είναι ένα σύνολο συναρτήσεων που ορίζονται από τους πολλ/στες Lagrange οι οποίοι μπαίνουν στο πρόβλημα ελαχιστοποίησης.

- Από τη μορφή της συνάρτησης πιθανότητας συμπεραίνουμε ότι μιλάμε για κατανομή Gibbs δηλαδή για ένα μαρκοβιανό πεδίο

- Οι τοπικές αλληλεπιδράσεις δεν υλοποιούνται μέσω των συναρτήσεων με βάση τις κλίκες της γειτονιάς όπως γνωρίζαμε αλλά ως τη “μάθηση” των περιορισμών στις εξόδους των φίλτρων.

Για την προσέγγιση των λαανγκρανζιανών συναρτήσεων οι Zhu, Wu, Mumford επινόησαν έναν αλγόριθμο μάθησης. Για περισσότερες πληροφορίες για αυτόν καθώς και γενικότερα για το FRAME μοντέλο ο αναγνώστης μπορεί να ανατρέξει στο [69].

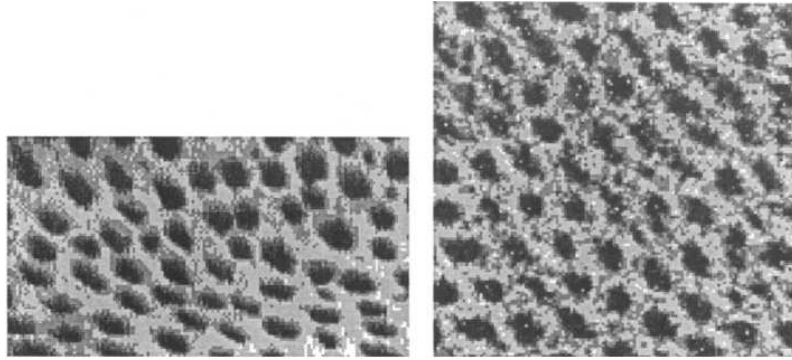
Στη συνέχεια στις εικόνες (4.12) και (4.13) δίνουμε κάποια παραδείγματα σύνθεσης με το FRAME.

4.5 Το μοντέλο GRF με μορφολογικούς περιορισμούς (Morphologically constrained GRF's)

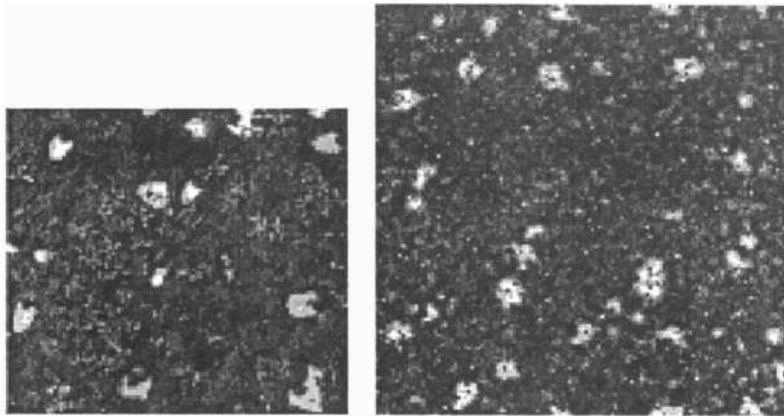
Μέχρι τώρα τα μοντέλα που έχουμε εξετάσει τα οποία βασίζονται στα MRF περιέχουν συναρτήσεις ενέργειας οι οποίες επιτρέπουν αλληλεπιδράσεις μόνο μεταξύ pixels τα οποία χαρακτηρίζονται ως κοντινοί γείτονες θεωρώντας ένα σύστημα γειτονιάς. Ουιδιότητες που ικανοποιούν οι καταστάσεις ηρεμίας(καταστάσεις ελάχιστης ενέργειας (ground states)αυτών των συναρτήσεων δεν είναι προφανείς.Εκείνο που σίγουρα γνωρίζουμε είναι ότι θα έχουν μια σταθερή δομή λόγω των ισχυρών συσχετίσεων μεταξύ των pixels αυτών. Παρόλα αυτά για να θεωρήσουμε πετυχημένη μια μοντελοποίηση με GRF πρέπει αυτή να λαμβάνει υπόψη την κρίσιμη δομική πληροφορία των εικόνων που μας ενδιαφέρουν.

Κάτι τέτοιο θα ήταν δυνατό αν εισάγαμε στην συνάρτηση ενέργειας κάποια σχέση η οποία όταν ελαχιστοποιείται θα προσδίδει στην προς σύνθεση εικόνα κάποια συγκεκριμένα δομικά χαρακτηριστικά. Η προσέγγιση των K. Sivakumar & Goutsias [67] αφορούσε στην εισαγωγή στην συνάρτηση ενέργειας ποσοτήτων που λαμβάνουν υπόψη γεωμετρικές ιδιότητες των εικόνων υλοποιούμενες με τη βοήθεια μορφολογικών περιορισμών.

Συγκεκριμένα στη συνάρτηση ενέργειας εισάγονται βασικά εργαλεία της μαθηματικής μορφολογίας που λέγονται granulometries.Μια granulometry (συμβολίζεται με $\{\gamma_s\}_{s=0,1,\dots}$) είναι μια παραμετρική οικογένεια απεικονίσεων από το \mathcal{X} στον εαυτό του (όπου με \mathcal{X} συμβολίζουμε το σύνολο που περιέχει όλες τις εικόνες με πεπερασμένο χώρο καταστάσεων και πεπερασμένες τιμές στα pixels τους), τέτοια ώστε $\gamma_0(X) = X$ και $\{\gamma_s\}_{s=1,2,\dots}$ είναι μια οικογένεια από openings για την οποία ισχύει : για κάθε $r > s \Rightarrow \gamma_r \leq \gamma_s$. Αντίστοιχα ορίζεται η δυαδική παραμετρική οικογένεια απεικονίσεων που ονομάζεται antigranulometry , συμβολίζεται με $\{\phi_s\}_{s=0,1,\dots}$ και ισχύει: $\phi_s(X) = (\gamma_s(X^*))^*$ (όπου A^*



Σχήμα 4.12: Το συγκεκριμένο παράδειγμα αφορά σύνθεση περιοδικής υφής (δέρμα λεοπάρδαλης). Παρατηρούμε μια μικρή αδυναμία στην απεικόνιση της διαβάθμισης των χρωμάτων (σε μεγάλες αλλαγές). Από το [69].



Σχήμα 4.13: Εδώ παρατηρούμε τη σύνθεση τυχαίας υφής (άμμος). Οι διαφορές ανάμεσα στη συνθετική και την πραγματική υφή είναι αμελητέες. Από το [69].

συμβολίζει το συμπλήρωμα ενός συνόλου A). Αποδεικνύεται ότι η $\{\phi_s\}_{s=0,1,2,\dots}$ είναι antigranulometry αν και μόνο αν $\phi_0(X) = X$ και $\{\gamma_s\}_{s=1,2,\dots}$ είναι μια οικογένεια από closings για την οποία ισχύει : για κάθε $r > s \Rightarrow \gamma_r \geq \gamma_s$.

Δύο πολύ χρήσιμες περιπτώσεις είναι η granulometry που ορίζεται ως:

$$\begin{aligned}\gamma_0(X) &= X \\ \gamma_s(X) &= X \circ sB, s \geq 1\end{aligned}\quad (4.14)$$

όπου $0B = \{(0, 0)\}$ και $sB = (s-1)B \oplus B$, ενώ B είναι κάποιο δομικό στοιχείο (structuring element), και η αντίστοιχη antigranulometry που ορίζεται ως:

$$\begin{aligned}\phi_0(X) &= X \\ \phi_s(X) &= X \bullet s\check{B}, s \geq 1.\end{aligned}\quad (4.15)$$

Με βάση τα παραπάνω οι Sivakumar & Goutsias χρησιμοποίησαν συνάρτηση ενέργειας για το μοντέλο τους την

$$\begin{aligned}U(X) &= \sum_{i=0}^I a_i \left| \left| \gamma_i(X) - \gamma_{i+1}(X) \right|_W - v_i G|W| \right|^p + \\ &+ \sum_{j=0}^J a_{-j} \left| \left| \phi_j(X) - \phi_{j-1}(X) \right|_W - v_{-j} G|W| \right|^p\end{aligned}\quad (4.16)$$

όπου $I \geq 0, J \geq 1$ δύο ακέραιοι, $p \geq 1$ μια πραγματική παράμετρος, $\{a_i, -J \leq i \leq I\}$ είναι δυαδικοί παράμετροι και $\{v_i, -J \leq i \leq I\}$ είναι πραγματικές μη αρνητικές παράμετροι τέτοιες ώστε

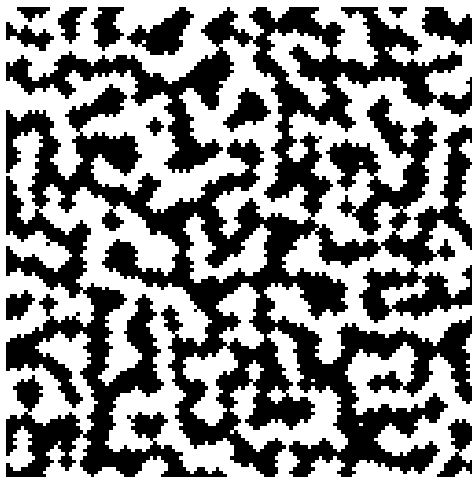
$$\sum_{i=-J}^I v_i \leq 1 \quad .$$

Επίσης $\|X\|_W = \sum_{w \in W} |X(w)|$ και $W \subset \mathbb{Z}^2$ είναι ένα μη κενό και φραγμένο παράθυρο παρατήρησης. Συνεπώς σε χαμηλές θερμοκρασίες οι εικόνες χαμηλής ενέργειας του μοντέλου θα έχουν χωρική πυκνότητα

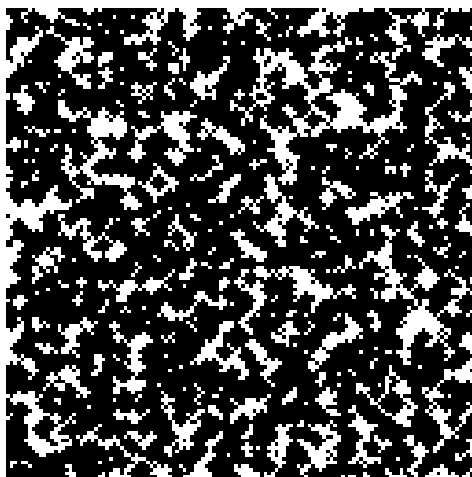
$$s_X(s) = \frac{1}{G|W|} \begin{cases} E \left[\left| \gamma_s(X) - \gamma_{s+1}(X) \right|_W \right] & \text{για } s \geq 0 \\ E \left[\left| \phi_{|s|}(X) - \phi_{|s|-1}(X) \right|_W \right] & \text{για } s \leq -1 \end{cases}$$

η οποία θα καθορίζεται από τις τιμές των παραμέτρων a_i και v_i .

Οι εικόνες (4.14), (4.15), (4.16), (4.17) περιέχουν κάποια παραδείγματα σύνθεσης με το μοντέλο αυτό.



Σχήμα 4.14: Η παραπάνω εικόνα έχει συντεθεί με παραμέτρους $I = 1, J = 2, v = [0, 0, 0, 0], a = [1, 1, 1, 1]$ και $G = 1$ (δυαδική). Δομικό στοιχείο είναι ο ρόμβος. Οι granulometries και antigranulometries είναι οι μορφολογικές που εξετάσαμε παραπάνω



Σχήμα 4.15: Η παραπάνω εικόνα έχει παραμέτρους $I = 4, J = 5, v = [0.085, 0.075, 0.065, 0.045, 0.05, 0.065, 0.07, 0.07, 0.075, 0.07], a = [1, 1, 1, 1, 1, 1, 1, 1, 1, 1]$ και $G = 1$. Δομικό στοιχείο και πάλι ο μοναδιαίος ρόμβος

Παρατηρούμε ότι πράγματι τα αποτελέσματα των συνθέσεων με αυτό το μοντέλο μας δίνουν εικόνες με ισχυρή γεωμετρική δομή. Επίσης οι Sivakumar & Goutsias στο [67] παρουσιάζουν έναν πολύ έξυπνο και πολύ γρήγορο τρόπο εκτίμησης των παραμέτρων του μοντέλου τους ενώ παρουσιάζουν και ιδιαίτερα επιτυχημένα αποτελέσματα κατηγοριοποίησης εννέα υφών από το άλμπουμ του Brodatz. Το μοντέλο όμως έχει το ισχυρό μειονέκτημα ότι πρέπει κάθε φορά να κάνουμε επιλογή του κατάλληλου δομικού στοιχείου ανάλογα με την εικόνα που θέλουμε να παράγουμε καθώς πρέπει να επιλέξουμε και το είδος της granulometry που θέλουμε να χρησιμοποιήσουμε. Τέλος όπως θα δούμε και παρακάτω έχει περιορισμένο εύρος υφών που μπορεί να παράγει ενώ παρουσιάζει αδυναμία στη σύνθεση πολλών φυσικών υφών.

Κεφάλαιο 5

Μια νέα πρόταση μοντέλου βασισμένο σε MRF

Κατά τη μελέτη του μοντέλου του GRF με μορφολογικούς περιορισμούς των Sivakumar & Goutsias εκτιμήσαμε αρκετές απο τις αρετές του. Κατ' αρχήν με ένα τέτοιο μοντέλο θα μπορούσαμε να συλλάβουμε γεωμετρικές δομές που επαναλαμβάνονται και οικοδομούν πολλές φυσικές υφές. Άρα συνδυάζοντας αυτές τις δομές με την τυχειότητα ή τη στατιστική δομή που μας προσφέρουν τα MRF θα μπορούσαμε να δημιουργήσουμε υφές πολύ κοντά στις πραγματικές. Ένα βασικό επίσης πλεονέκτημα του μοντέλου είναι ο τρόπος με τον οποίο εκτιμώνται οι παράμετροί του. Αντί να χρησιμοποιήσουν δοκιμασμένες αλλά εξαιρετικά χρονοβόρες λύσεις εκτίμησης (π.χ. εκτίμηση μεγίστης πιθανοφάνειας) χρησιμοποίησαν τη δομή της συνάρτησης ενέργειας ώστε να δημιουργήσουν έναν ιδιαίτερα γρήγορο και αξιόπιστο τρόπο εκτίμησης. Βέβαια κάτι τέτοιο αποβαίνει σε βάρος της ελεύθερης επιλογής όλων των παραμέτρων αλλά πρέπει να πούμε ότι κάτι τέτοιο δεν περιορίζει σημαντικά το εύρος των εικόνων που μπορούν να παραχθούν. Επίσης το αποτέλεσμα του πειράματος κατηγοριοποίησης υφών σε εννέα κλάσεις ήταν εντυπωσιακό και άξιο προσοχής κάτι που σημαίνει ότι ο τρόπος εκτίμησης των παραμέτρων είναι αξιόπιστος αλλά και ότι οι ίδιες οι παράμετροι αλλάζουν σημαντικά από υφή σε υφή και αποτελούν να λέγαμε την ταυτότητα κάθεμιας.

Βέβαια δεν θα μπορούσαμε να παραβλέψουμε και κάποια σημαντικά μειονεκτήματα. Το δομικό στοιχείο που εισάγεται στο μοντέλο αυτό όπως και οι granulometries και οι antigranulometries επηρεάζουν κατά πολύ την εκτίμηση παραμέτρων αλλά κυρίως τη σύνθεση υφών. Άρα για να συνθέσουμε ρεαλιστικά μια υφή πρέπει να βρούμε έναν τρόπο ώστε να επιλέξουμε το κατάλληλο δομικό στοιχείο κάτι το οποίο δεν είναι εύκολο. Ένα " λάθος " δομικό στοιχείο θα μας δώσει αποτελέσματα πολύ μακριά από τα επιθυμητά. Από την άλλη εστιάζοντας πάλι στη σύνθεση υφής το μοντέλο δεν λαμβάνει υπόψιν του τα χρώματα της ει-

κόνας (στην κλίμακα του γκρι) ούτε τα χρώματα των δομικών στοιχείων. Είναι δυνατόν λοιπόν να έχουμε επιλέξει σωστά το δομικό στοιχείο όμως η εικόνα που θα προκύψει από τη σύνθεση να μην είναι ρεαλιστική καθώς η χρωματική κατανομή (ιστόγραμμα) θα είναι τελείως διαφορετική από αυτήν της πραγματικής υφής. Αυτά κυρίως τα προβλήματα προσπαθήσαμε να λύσουμε προτείνοντας διάφορες νέες ιδέες. Στη συνέχεια θα τις περιγράψουμε αναλυτικά.

5.1 Θεωρητικό υπόβαθρο

Στο συγκεκριμένο μέρος θα παρουσιάσουμε αναλυτικά τον τρόπο με τον οποίο οδηγηθήκαμε στη συνάρτηση ενέργειας του μοντέλου μας. Πριν από αυτό όμως καλό θα είναι να δώσουμε κάποιους ορισμούς που θα μας χρειαστούν στη συνέχεια.

Ορισμός 5.1. Ένα μερικώς διατεταγμένο σύνολο (P, \leq) είναι ένα μη κενό σύνολο P στο οποίο μπορεί να οριστεί μια δυαδική σχέση για τα στοιχεία του, η οποία ονομάζεται μερική διάταξη, συμβολίζεται με \leq και ικανοποιεί τις παρακάτω ιδιότητες για όλα τα $x, y, z \in P$:

1. $x \leq x$ (ανακλαστική)
2. Αν $(x \leq y) \wedge (y \leq z) \Rightarrow x \leq z$ (μεταβατική)
3. Αν $(x \leq y) \wedge (y \leq x) \Rightarrow x = y$ (αντισυμμετρική)

Ένα τέτοιο σύνολο θα λέγεται συνολικά διατεταγμένο εαν \leq έχει ακόμα την ιδιότητα:

$$x \leq y \quad \text{ή} \quad y \leq x \quad \text{για κάθε ζεύγος } (x, y) \in P^2 \text{ με } x \neq y$$

Ορισμός 5.2. Ένα μερικώς διατεταγμένο σύνολο P θα λέγεται πλέγμα (lattice) εάν για κάθε δύο στοιχεία $x, y \in P$ ορίζονται το *infimum* και το *supremum* αυτών.

Θεωρώντας τώρα ότι μια γκριζα εικόνα είναι υποσύνολο ενός πλέγματος E τότε μπορούμε να ορίσουμε τους μορφολογικούς τελεστές dilation, erosion, opening και closing ως προς ένα δομικό στοιχείο $B \in E$ ως εξής:

$$(X \oplus B)(x) = \bigvee_{y \in B} X(x - y) \quad (5.1)$$

$$(X \ominus B)(x) = \bigwedge_{y \in B} X(x + y) \quad (5.2)$$

$$(X \circ B)(x) = \bigvee_{z \in B} \bigwedge_{y \in B} X(x + y - z) \quad (5.3)$$

$$(X \bullet B)(x) = \bigwedge_{z \in B} \bigvee_{y \in B} X(x - y + z) \quad (5.4)$$

Με βάση τους τελεστές αυτούς λοιπόν μπορούμε να ορίσουμε τις συναρτήσεις granulometry και antigranulometry όπως και κάναμε στο προηγούμενο κεφάλαιο.

Η λογική κατασκευής της συνάρτησης ενέργειας του μοντέλου μας γίνεται όπως και στο μοντέλο των Sivakumar & Goutsias με βάση τις granulometries και antigranulometries. Όμως θα θέλαμε αφένός να εξαλείψουμε το πρόβλημα επιλογής του δομικού στοιχείου ,διαλέγοντας κάτι που θα είναι πιο γενικό ,ενώ θα θέλαμε επίσης να μπορούμε να συλλάβουμε τη διεύθυνση των πρωταρχικών στοιχείων που αποτελούν τις υφές που εξετάζουμε. Για το πρώτο πρόβλημα θεωρήσαμε κατάλληλο το δομικό στοιχείο της ευθείας το οποίο είναι πράγματι γενικό και μπορεί να μας δώσει όλα τα χαρακτηριστικά της εικόνας που εξετάζουμε. Εκείνο που σκεφτήκαμε επίσης είναι ότι για να ανακτήσουμε χαρακτηριστικά που βρίσκονται σε διαφορετικές διευθύνσεις και κλίμακες θα μπορούσαμε να “σαρώσουμε” την εικόνα μας με ευθείες μεταβλητού μήκους και γωνίας ως προς τον οριζόντιο άξονα.

Η έμπνευση για κάτι τέτοιο αποτέλεσε ο ορισμός του κατευθυνόμενου φάσματος προτύπων (oriented pattern spectrum) που υπάρχει στην εργασία Maragos [42]. Ο ορισμός αυτός φαίνεται παρακάτω:

Ορισμός 5.3. Έστω $f(x, y)$ μία διδιάστατη γκριζα εικόνα , $g_\theta(z)$ μια μονοδιάστατη συνάρτηση με διεύθυνση θ και $A(f) = \sum_{(x,y)} |f(x, y)|$. Τότε το διακριτό oriented pattern spectrum ορίζεται ως:

$$\begin{aligned} OPS_f(+r) &= A(\max_{\theta} \{f \circ r g_{\theta}(x, y)\}) - \max_{\theta} \{f \circ (r + 1) g_{\theta}(x, y)\}, \\ & \quad r \geq 0, \\ OPS_f(+r) &= A(\min_{\theta} \{f \bullet r g_{\theta}(x, y)\}) - \min_{\theta} \{f \circ (r - 1) g_{\theta}(x, y)\}, \\ & \quad r > 0 \end{aligned} \quad (5.5)$$

Ο ορισμός (5.3) όπως μπορούμε να δούμε δεν λαμβάνει υπόψιν του την κατανομή της συνάρτησης $g_\theta(z)$ στην εικόνα f σε όλες τις διευθύνσεις θ παρά βρίσκει την “έπικρατούσα γωνία” αυτήν δηλαδή τη γωνία για την οποία το opening της εικόνας με τη συνάρτηση $g_\theta(z)$ είναι μεγαλύτερο ενώ το closing θα είναι μικρότερο.

Εμείς στο σημείο αυτό θα προτείνουμε ένα διαφοροποιημένο oriented pattern spectrum στο οποίο η γωνία θ θα υπάρχει ως παράμετρος. Επίσης για να

αντιμετωπίσουμε το πρόβλημα του χρώματος αποφασίσαμε κατά τον υπολογισμό του pattern spectrum να φιλτράρουμε τις εικόνες που προκύπτουν από τα opening και closing με μια συστοιχία φίλτρων M_r τα οποία ορίζονται ως εξής:

$$M_r(f(x, y)) = \begin{cases} r & \text{εάν } f(x, y) = r \\ -\frac{1}{L-1} & \text{αλλιώς} \end{cases} \quad (5.6)$$

όπου $r = 0, 1, \dots, L-1$ και L είναι ο αριθμός των σταθμών του γκριζου που έχει η εικόνα f . Η τοποθέτηση της συστοιχίας φίλτρων αυτού θα μας δείξει την εξέλιξη και το μέγεθος ευθειών συγκεκριμένης κατεύθυνσης και χρώματος. Η τιμή $-\frac{1}{L-1}$ χρησιμοποιήθηκε διότι έτσι θα μπορούσαμε να παρακολουθήσουμε και την εξέλιξη του μαύρου το οποίο ως γνωστόν έχει την τιμή 0, ενώ θα διατηρούντο οι διαφορές μεταξύ σταθμών του γκριζου.

Οπότε μπορούμε τώρα να ορίσουμε το διαφοροποιημένο oriented pattern spectrum ως εξής:

$$\begin{aligned} p_f(s, \theta, r) &= \frac{1}{LW} A(M_r(f \circ s g_\theta(x, y)) - M_r(f \circ (s+1) g_\theta(x, y))), \\ & \quad s \geq 0, \\ p_f(s, \theta, r) &= \frac{1}{LW} A(M_r(f \bullet |s| g_\theta(x, y)) - M_r(f \bullet |s| - 1) g_\theta(x, y)), \\ & \quad s < 0 \end{aligned} \quad (5.7)$$

όπου W είναι ο αριθμός των pixels της εικόνας.

Με βάση τον παραπάνω ορισμό η συνάρτηση ενέργειας του μοντέλου μας θα είναι:

$$\begin{aligned} U(X) &= \sum_{r=0}^{L-1} \sum_{\theta=0}^{N-1} \sum_{i=0}^I a_{i\theta r} \left| A(\gamma_{i\theta r}(X) - \gamma_{(i+1)\theta r}(X)) - v_{i\theta r} LW \right|^p + \\ &+ \sum_{r=0}^{L-1} \sum_{\theta=0}^{N-1} \sum_{j=1}^J a_{-j\theta r} \left| A(\phi_{j\theta r}(X) - \phi_{(j-1)\theta r}(X)) - v_{-j\theta r} LW \right|^p \end{aligned} \quad (5.8)$$

όπου $\gamma_{i\theta r}(X) = M_r(X \circ iL \frac{\pi\theta}{N})$, $\phi_{j\theta r}(X) = M_r(X \bullet |j|L \frac{\pi\theta}{N})$ και L_θ είναι η μοναδιαία ευθεία προσανατολισμένη κατά γωνία θ .

5.2 Μια πρόταση αλγορίθμου σύνθεσης

Η παραπάνω συνάρτηση ενέργειας παρόλο που μπορεί να ανακτήσει πολλά από τα χαρακτηριστικά της υψής που επεξεργαζόμαστε αποδεικνύεται στην πράξη ότι η δειγματοληψία του τελικού μοντέλου που προκύπτει είναι μια εξαιρετικά χρονοβόρος υπολογιστική διαδικασία. Δοκιμάσαμε αρκετούς γνωστούς

αλγορίθμους (Metropolis ,Gibbs Sampler κ.α) οι οποίοι αποδείχτηκαν αρκετά αργοί στη σύγκλιση. Το πρόβλημα εντοπίζεται κυρίως σε υφές με έντονη κατεύθυνση που αποτελούνται από μεγάλα πρωταρχικά στοιχεία. Σε τέτοιες περιπτώσεις φαίνεται πιο αποδοτική η εισαγωγή προτύπων του δομικού στοιχείου σε διάφορες κλίμακες ,χρώματα και προσανατολισμούς παρά η αλλαγή ενός pixel σε κάθε κύκλο. Παρόλα αυτά για να έχουμε μικρό αριθμό απορρίψεων σε εικόνες που προκύπτουν από αυτή τη διαδικασία είναι εξαιρετικά σημαντική η προσεκτική επιλογή των παραμέτρων του δομικού στοιχείου που εισάγουμε κάθε φορά. Παρακάτω προτείνουμε έναν αλγόριθμο ο οποίος σε κάθε επανάληψη επιλέγει το χρώμα ,το μήκος και τη διεύθυνση του δομικού στοιχείου που εισάγει μέσω κάποιων συναρτήσεων πιθανότητας. Από πειράματα που έχουμε κάνει φάνηκε ότι ο συγκεκριμένος αλγόριθμος έχει πολύ μικρότερο λόγο απόρριψης από τους γνωστούς αλγόριθμους που αναφέραμε παραπάνω. Παρόλα αυτά ο χρόνος της κάθε επανάληψης παραμένει αρκετά μεγάλος κάτι το οποίο δε μας επέτρεψε τη σύνθεση εικόνων υψηλής ανάλυσης. Κάτι τέτοιο όμως αφορά την αποτελεσματικότητα των υπολογισμών των πολλαπλών opening και closing που προκύπτουν κατά τη διαδικασία και όχι την αποτελεσματική εκλογή των παραμέτρων του δομικού στοιχείου που εισάγουμε κάθε φορά. Τα βήματα του αλγορίθμου συνοψίζονται παρακάτω σε ψευδοκώδικα:

Αλγόριθμος Σύνθεσης

1. Ξεκινάμε με μια τυχαία εικόνα X η οποία όμως δεν περιέχει καθόλου στάθμες του γκρι που δεν εμφανίζονται καθόλου στο ιστόγραμμα της αρξικής (όχι ιδιαίτερα κρίσιμο απλώς επιταχύνει τη διαδικασία και η συνάρτηση ενέργειας ξεκινά από πιο μικρή τιμή).
2. Υπολογίζουμε τις ποσότητες της αρχικής εικόνας για κάθε $i \in [0, I], j \in [1, J], \theta \in [0, N - 1], r \in [0, L - 1]$
3. Διαλέγουμε ένα σημείο w τυχαία από όλα τα σημεία της εικόνας μας.
4. Διαλέγουμε μια στάθμη του γκρι με συνάρτηση πιθανότητας το ιστόγραμμα του πραγματικού δείγματος υφής.
5. Με βάση τη στάθμη r που επιλέξαμε στο προηγούμενο βήμα επιλέγουμε τώρα ένα μήκος s με ίση πιθανότητα από όλες τις τιμές $1 \cdots I$ αφού έχουμε αφαιρέσει εκείνες τις τιμές i για τις οποίες ισχύει

$$\sum_{\theta=0}^{N-1} (v_{i\theta r} + v_{-i\theta r}) = 0.$$

Αυτό σημαίνει ότι αφαιρούμε εκείνα τα μήκη για τα οποία το συγκεκριμένο χρώμα δεν εμφανίζεται σε καμία γωνία.

6. Έχοντας πλέον επιλέξει τη στάθμη r και το μήκος s επιλέγουμε μια γωνία θ με συνάρτηση πιθανότητας

$$P(\theta) = \frac{\frac{1}{\sum_{m=1}^s (v_{m\theta r} + v_{-m\theta r})}}{\sum_{\theta=0}^{N-1} \frac{1}{\sum_{m=1}^s (v_{m\theta r} + v_{-m\theta r})}}.$$

(Επειδή πολλές φορές προκύπτουν κάποια αθροίσματα ίσα με 0 για να μην έχουμε πρόβλημα στους αντίστοιχους παρονομαστές τα αφαιρούμε τελείως και βέβαια αφαιρούμε τη δυνατότητα επιλογής των συγκεκριμένων γωνιών).

7. Στη συνέχεια θεωρούμε εικόνα Y η οποία είναι ίδια με την εικόνα X στα σημεία $X \setminus (rL_\theta)_w$ ενώ για τα σημεία αυτά έχουμε ότι $X(p) = r, p \in (rL_\theta)_w$.
8. Υπολογίζουμε τις πιθανότητες των εικόνων X και Y ως

$$P(X) = \frac{1}{Z} \exp[-U(X)]$$

καθώς και το λόγο $b = \frac{P(Y)}{P(X)}$.

9. Διαλέγουμε έναν πραγματικό αριθμό l από το διάστημα $[0, 1]$ θεωρώντας ομοιόμορφη κατανομή.
10. Εάν $l < b$ τότε $X = Y$ και $\gamma_{i\theta r}(X) = \gamma_{i\theta r}(Y)$ και $\phi_{j\theta r}(X) = \phi_{j\theta r}(Y)$
11. Εάν $U(X) < \delta LW$ (όπου δ ένας μικρός θετικός πραγματικός π.χ. $\delta = 0.01$ τότε σταματάμε αλλιώς ξαναγυρίζουμε στο βήμα 3.

Πρέπει να πούμε ότι όσες συνθέσεις θα παρουσιάσουμε παρακάτω έγιναν με τον συγκεκριμένο αλγόριθμο. Συνεπώς επειδή τον τρέξαμε αρκετές φορές έχει αποδειχτεί ιδιαίτερα αξιόπιστος ενώ δεν έχει παρουσιαστεί ποτέ πρόβλημα στην επιλογή των παραμέτρων r, θ και s με βάση τις συναρτήσεις πιθανότητας που πρέπει να ικανοποιούν. Για τον αριθμό δ δώσαμε παραπάνω ένα παράδειγμα επιλογής της τιμής του όμως καλό θα είναι να επιλέγεται κάτι μεγαλύτερο κατά μία τάξη μεγέθους καθώς από προσομοιώσεις παρατηρήσαμε ότι η εικόνα που συνθέτουμε έχει ήδη συγχλίνει μέχρι εκείνο το σημείο ενώ στη συνέχεια έχουμε πολύ μεγάλο λόγο απόρριψης.

5.3 Εκτίμηση παραμέτρων

Η αποτελεσματική εκτίμηση των παραμέτρων του μοντέλου μας αλλά και κάθε μοντέλου είναι κάτι πάρα πολύ σημαντικό. Με τον τρόπο αυτό μπορεί κανείς να εξερευνήσει τις δυνατότητες και τις αδυναμίες του προσπαθώντας να εκτιμήσει παραμέτρους από πραγματικές υφές και κατόπιν να κάνει μια σύνθεση και να συγκρίνει τα αποτελέσματα.

Πολλά μοντέλα από αυτά που παρουσιάσαμε σε προηγούμενο κεφάλαιο ταλαιπωρούνται είτε από μη επαρκείς μεθόδους εκτίμησης παραμέτρων που υιοθετούν καθώς και από εξαιρετικά χρονοβόρες υπολογιστικές διαδικασίες. Το μοντέλο του μορφολογικά περιορισμένου GRF των Sivakumar & Goutsias πραγματικά ξεχωρίζει σ' αυτόν τον τομέα. Υποθέτοντας ότι $p = 2$ απέδειξαν ότι μπορούν να εκτιμήσουν τις παραμέτρους του μοντέλου τους με τη βοήθεια του pattern spectrum. Συγκεκριμένα απέδειξαν τη σχέση

$$\hat{v}_s(X) \cong p_X(s), \quad \text{για κάθε } s : -J \leq s \leq I \quad \text{έτσι ώστε } a_s = 1 \quad (5.9)$$

όπου $p_X(s)$ είναι το pattern spectrum. Η σχέση (5.9) μας επιτρέπει το γρήγορο και αποτελεσματικό υπολογισμό των παραμέτρων v_s χωρίς να χρειαστεί να καταφύγουμε σε μεθόδους χρονοβόρες όπως εκτίμηση μεγίστης πιθανοφάνειας. Μάλιστα αποδεικνύεται ότι για $p = 2$ και σε χαμηλές θερμοκρασίες είναι η εκτίμηση είναι σχεδόν ίδια με αυτήν του βέλτιστου εκτιμητή μεγίστης πιθανοφάνειας.

Όπως έχουμε αναφέρει και προηγουμένως ο τρόπος αυτός της εκτίμησης των παραμέτρων ήταν ένα από τα σημαντικά πλεονεκτήματα τη μοντέλου που μας έκαναν να το μελετήσουμε περαιτέρω και να προτείνουμε μια βελτίωσή του. Πρέπει όμως να αποδείξουμε ότι με τη νέα συνάρτηση ενέργειας που έχουμε εισάγει μπορούμε να εκτιμήσουμε τις παραμέτρους μας με έναν αντίστοιχο τρόπο.

Υποθέτουμε λοιπόν ότι θέλουμε να εκτιμήσουμε τις παραμέτρους \hat{v} με την εκτιμήτρια μεγίστης πιθανοφάνειας. Προφανώς οι παράμετροι αυτοί θα πρέπει να ικανοποιούν τις παρακάτω εξισώσεις:

$$A(\gamma_{i\theta r}(X) - \gamma_{(i+1)\theta r}(X)) = E_{\pi_{\hat{v}}} [A(\gamma_{i\theta r}(X) - \gamma_{(i+1)\theta r}(X))] \quad (5.10)$$

και

$$A(\phi_{j\theta r}(X) - \phi_{(j-1)\theta r}(X)) = E_{\pi_{\hat{v}}} [A(\phi_{j\theta r}(X) - \phi_{(j-1)\theta r}(X))] \quad (5.11)$$

Από τις παραπάνω εξισώσεις συμπεραίνουμε ότι η διαφοροποιημένη χωρική πυκνότητα η οποία είναι:

$$s_x(s, \theta, r | \hat{v}) = \frac{1}{LW} \begin{cases} E_{\pi_{\hat{v}}} [A(\gamma_{s\theta r}(X) - \gamma_{(s+1)\theta r}(X))] & \text{για } s \geq 0 \\ E_{\pi_{\hat{v}}} [A(\phi_{s\theta r}(X) - \phi_{(s-1)\theta r}(X))] & \text{για } s \leq -1 \end{cases} \quad (5.12)$$

είναι ίση με το διαφοροποιημένο pattern spectrum όπως το ορίσαμε στην εξίσωση 5.7. Έχουμε δηλαδή

$$s_X(s, \theta, r|\hat{v}) = p_X(s, \theta, r) \quad (5.13)$$

Επίσης αποδεικνύεται ότι

$$\hat{v}_{s,\theta,r}(X) = \lim_{T \rightarrow 0^+} s_X(s, \theta, r|\hat{v}). \quad (5.14)$$

Συνδυάζοντας λοιπόν τις 5.13 και 5.14 έχουμε ότι

$$\hat{v}_{s,\theta,r}(X) = p_X(s, \theta, r) \quad (5.15)$$

Πρέπει να παρατηρήσουμε ότι ουσιαστικά η απόδειξή μας δεν διαφοροποιήθηκε από αυτήν που προτείνουν οι Sivakumar και Goutsias για το δικό τους μοντέλο. Αυτό συμβαίνει διότι στην πραγματικότητα η επέκταση του μοντέλου με την προσθήκη των μεταβλητών θ και r δεν επηρεάζει το pattern spectrum ούτε το size density καθώς καμία από αυτές δεν υπεισέρχεται άμεσα στον υπολογισμό τους. Το μόνο που κάνουμε είναι να υπολογίζουμε το pattern spectrum με τον ίδιο τρόπο που έκαναν οι Sivakumar και Goutsias για όλες τις δυνατές τιμές των μεταβλητών αυτών.

Οι σκέψεις αυτές μας οδήγησαν στον παρακάτω αλγόριθμο οποίος δεν διαφέρει ιδιαίτερα από τον αλγόριθμο που προτείνουν οι Sivakumar και Goutsias.

Αλγόριθμος εκτίμησης παραμέτρων

1. Έστω X η εικόνα για την οποία θα εκτιμήσουμε τις παραμέτρους. Πρώτα υπολογίζουμε το διαφοροποιημένο pattern spectrum $p_X(s, \theta, r)$ για $-J \leq s \leq I, 0 \leq r \leq L-1, 0 \leq \theta \leq N-1$ όπου $I = J+1 = \max pX, pY$ όπου pX, pY είναι οι διαστάσεις της εικόνας X .
2. Επιλέγουμε εκείνο το I που ικανοποιεί κάποιο κριτήριο ελαχιστότητας π.χ.:

$$\sum_{\theta=0}^{N-1} \sum_{r=0}^{L-1} p_X(s, \theta, r) < \varepsilon$$

για $s > I$.

3. Τέλος κατασκευάζουμε το μοντέλο με παραμέτρους $p = 2, a_{ijk} = 1$ για κάθε

$$-J \leq i \leq I, \quad 0 \leq j \leq N-1, \quad 0 \leq k \leq L-1$$

και $v_{s,\theta,r} = p_X(s, \theta, r)$

5.4 Classification

Οι παράμετροι του μοντέλου μας που υπολογίζονται με τον παραπάνω τρόπο μπορούν να χρησιμοποιηθούν και για κατηγοριοποίηση υφής (texture classification). Ας υποθέσουμε για παράδειγμα ότι θέλουμε να κατατάξουμε μια δεδομένη υφή X σε μία από K διαφορετικές κλάσεις $C_k, k = 1, 2, \dots, K$. Κάθε κλάση C_k χαρακτηρίζεται από τη συνάρτηση ενέργειας $U_k(X)$ του μοντέλου μας με παραμέτρους $v^{(k)}$. Υποθέτοντας μια πιθανότητα κλάσης $Pr[C_k]$ η οποία δίνεται από τη σχέση

$$Pr[C_k] = \frac{Z_k}{\sum_{m=1}^K Z_m} \quad (5.16)$$

όπου Z_k είναι η partition function που αντιστοιχεί στην συνάρτηση ενέργειας της κλάσης k . Τότε θα έχουμε

$$\begin{aligned} Pr[C_k|X] &= \frac{Pr[X|C_k]Pr[C_k]}{\sum_{l=1}^K Pr[X|C_l]Pr[C_l]} = \\ &= \frac{\frac{1}{Z_k} \exp\{-\frac{1}{T}V_k(X)\} \frac{Z_k}{K}}{\sum_{m=1}^K Z_m} = \\ &= \frac{\sum_{l=1}^K \frac{1}{Z_l} \exp\{-\frac{1}{T}V_l(X)\} \frac{Z_l}{K}}{\sum_{m=1}^K Z_m} = \\ &= \frac{\exp\{-\frac{1}{T}V_k(X)\}}{\sum_{l=1}^K \exp\{-\frac{1}{T}V_l(X)\}} \end{aligned} \quad (5.17)$$

Άρα από την παραπάνω σχέση συμπεραίνουμε ότι ο βέλτιστος Bayesian ταξινομητής θα είναι:

$$\text{Διάλεξε το } X \in C_k \text{ αν } k = \arg \max_m \{Pr[C_m|X]\} = \arg \min_m \{V_m(X)\}$$

Με βάση την παραπάνω εξίσωση υλοποιούμε τον παρακάτω αλγόριθμο κατηγοριοποίησης:

Αλγόριθμος κατηγοριοποίησης

1. Θεωρούμε ότι μας δίνονται K εικόνες έστω $X_k, 1 \leq k \leq K$ οι οποίες αποτελούν τις διαφορετικές υφές. Με βάση τον αλγόριθμο εκτίμησης

παραμέτρων εκτιμούμε τις παραμέτρους $v_k(s, \theta, r)$ για κάθε κλάση. Το παραπάνω βήμα αποτελεί το βήμα εκμάθησης.

2. Στη συνέχεια αν μας δοθεί μια υφή X την κατατάσσουμε στην κλάση C_m για την οποία ισχύει:

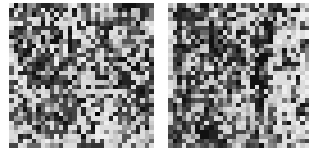
$$m = \arg \min_m \left\{ \sum_{r=0}^{L-1} \sum_{\theta=0}^{N-1} \sum_{s=0}^I (A(\gamma_{s\theta r}(X)) - \gamma_{(s+1)\theta r}(X)) - v_m(s, \theta, r)LW)^2 + \sum_{r=0}^{L-1} \sum_{\theta=0}^{N-1} \sum_{s=1}^J (A(\phi_{s\theta r}(X)) - \phi_{(s-1)\theta r}(X)) - v_m(-s, \theta, r)LW)^2 \right\} \quad (5.18)$$

Το παραπάνω βήμα είναι το βήμα ταξινόμησης.

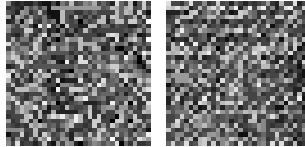
5.5 Αποτελέσματα πειραμάτων

5.5.1 Αποτελέσματα σύνθεσης

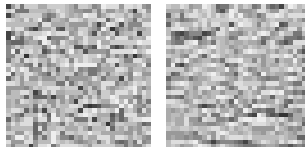
Οι συνθέσεις υφής που θα δούμε παρακάτω έγιναν όλες με το μοντέλο μας με τη βοήθεια του αλγόριθμου σύνθεσης που παρουσιάσαμε παραπάνω. Πρέπει να πούμε ότι λόγω της εξαιρετικά χρονοβόρας διαδικασίας της σύνθεσης δεν μπορέσαμε να συνθέσουμε εικόνες μεγαλύτερων διαστάσεων από 64×64 . Στη συνέχεια παραθέτουμε κάποιες συνθέσεις από το αλμπουμ του Brodatz [49]. Για να μπορέσουμε να εκτιμήσουμε παραμέτρους είμαστε υποχρεωμένοι να κάνουμε scale down τις πραγματικές εικόνες υφής ώστε να φτάσουν τις διαστάσεις 32×32 . Επίσης σε όλες τις εικόνες έχουμε χρησιμοποιήσει 16 στάθμες του γκρι ενώ οι διαστάσεις τους είναι 32×32 . Σε κάθε σχήμα αριστερά βρίσκεται η συνθετική υφή ενώ δεξιά η πραγματική που δόθηκε ως είσοδος στο σύστημά μας ώστε να εκτιμηθούν οι παράμετροι.



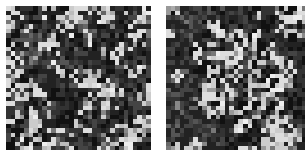
Σχήμα 5.1: Η παραπάνω υφή απεικονίζει το ανάγλυφο κορμού δέντρου. Το scale down εδώ δεν έχει διαφυλάξει την άγρια υφή σε όλη της τη διάσταση. Παρόλα αυτά παρατηρούμε ότι η σύνθεση πλησιάζει αρκετά.



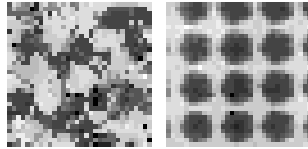
Σχήμα 5.2: Παρατηρούμε ότι η πραγματική εικόνα έχει μια έντονη κατεύθυνση περίπου 45° . Αυτό ελαφρά περνά και στη σύνθεση. Για καλύτερα αποτελέσματα ίσως χρειαζόμαστε μικρότερο δ .



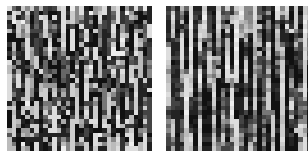
Σχήμα 5.3: Η πιο πετυχημένη μας σύνθεση. Το scale down δεν επηρεάζει ιδιαίτερα την αρχική δομή της αρχικής εικόνας ενώ η απαλή οριζόντια κατεύθυνση απεικονίζεται πολύ καλά.



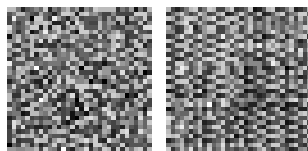
Σχήμα 5.4: Η συγκεκριμένη υφή απεικόνιζε ένα κέντημα με λουλούδια. Το θέμα (δηλαδή τα λουλούδια) ήταν ήδη σε μεγάλη κλίμακα στην φωτογραφία. Επίσης ένα τέτοιο σχέδιο ξεφεύγει αρκετά από τις απλοϊκές παραδοχές που έχουμε κάνει για τη δομή της υφής εισόδου. Συνεπώς τα αποτελέσματα είναι αναμενόμενο να μην είναι τόσο καλά.



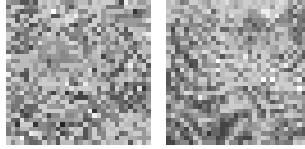
Σχήμα 5.5: Η παραπάνω υφή αποτελείται από επαναλαμβανόμενα μαύρα κυκλικά μοτίβα πλαισιωμένα από ένα λευκό περίγραμμα. Εδώ ο αλγόριθμος σύνθεσης δεν κατάφερε να συγκλίνει στην τιμή του δ που του είχαμε ορίσει. Άρα ενδεχομένως με άλλο τρόπο σύνθεσης (ίσως εισάγοντας άλλο δομικό στοιχείο) να μπορούσαμε να πάρουμε καλύτερα αποτελέσματα καθώς η ενέργεια της εικόνας που παρουσιάζουμε είναι αρκετά υψηλή.



Σχήμα 5.6: Στην πρωτότυπη εικόνα παρατηρούμε έντονη κάθετη κατεύθυνση από πρωταρχικά στοιχεία που είναι ευθείες σε διάφορες στάθμες του γκρι. Η σύνθεση έχει συλλάβει πολύ καλά τόσο την κατεύθυνση όσο και τη χρωματική αλληλουχία. Ίσως θα έπρεπε να τη συνθέσουμε με μεγαλύτερες παραμέτρους I, J όμως κάτι τέτοιο δεν κατέστη δυνατό καθώς επιβραδύνεται σημαντικά ο αλγόριθμος σύνθεσης.



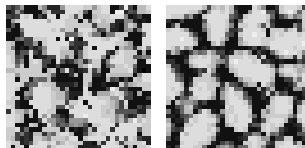
Σχήμα 5.7: Δύσκολη εικόνα καθώς έχουμε περισσότερες των μία κατευθύνσεων σε διάφορες κλίμακες. Η σύνθεση όμως πλησιάζει αρκετά. Ίσως λίγο μικρότερο δ να διόρθωνε κάποιες ατέλειες.



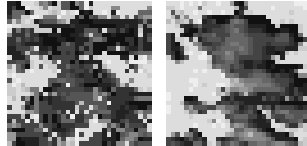
Σχήμα 5.8: Η πραγματική εικόνα απεικονίζει ένα κομμάτι ξύλου. Όσο μπορούμε να διακρίνουμε (και εδώ το scale down δεν είναι τόσο καλό) οι γραμμές του είναι αρκετά πολύπλοκες ενώ αλλάζουν συνεχώς κατεύθυνση. Μια τέτοια κατάσταση είναι αρκετά δύσκολο να απεικονιστεί στη σύνθεση. Πράγματι η εικόνα που προέκυψε από τη σύνθεση δεν έχει διατηρήσει τη συνέχεια των γραμμών αλλά τις έχει διατάξει με έναν πιο τυχαίο τρόπο. Παρόλα αυτά θα μπορούσαμε να πούμε ότι η σύνθεση δεν απέχει πολύ από την υφή του ξύλου.



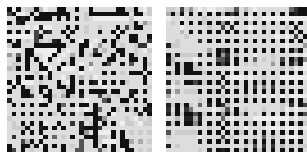
Σχήμα 5.9: Άλλη μια ιδιαίτερα επιτυχημένη σύνθεση. Εδώ φαίνεται πόσο σημαντική είναι η σάρωση σε όλες τις γωνίες διότι μόνο έτσι το μοντέλο μας θα μπορούσε να καταλάβει ότι παρόλο που είχε τη δυνατότητα ,κατά τη σύνθεση έπρεπε να επιλέξει ευθείες το πολύ μήκους 2.



Σχήμα 5.10: Εδώ έχουμε άλλη μια δύσκολη εικόνα η οποία απεικονίζει λευκά φύλλα σε μαύρο φόντο. Ο αλγόριθμος σύνθεσης πάλι δεν κατάφερε να συγχλί-νει για προκαθορισμένη τιμή του δ . Παρόλα αυτά παρατηρούμε ότι ήδη έχουν δημιουργηθεί στην συντεθειμένη εικόνα τα σύνορα των φύλλων. Με λίγο διαφορετικό τρόπο σύνθεσης ενδεχομένως να είχαμε καλύτερα αποτελέσματα.



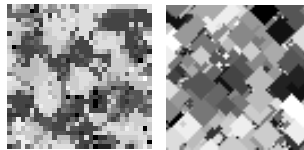
Σχήμα 5.11: Η συγκεκριμένη εικόνα απεικονίζει ένα σύννεφο στον ουρανό και θεωρείται εύκολη για ένα μοντέλο MRF. Πράγματι η σύνθεση είναι πολύ καλή αν και σε τέτοιου είδους συνθέσεις με μεγάλες συσσωματώσεις απαιτείται το φιλτράρισμα της παραγόμενης εικόνας με ένα median φίλτρο το οποίο κόβει διάφορες ασυνέχειες (ή κενά) που προκύπτουν κατά τη σύνθεση



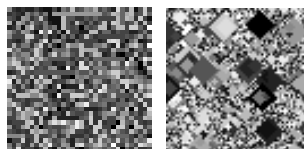
Σχήμα 5.12: Η πραγματική εικόνα απεικονίζει ένα ξύλινο ψαθί. Παρόλα αυτά το scale down είναι πάρα πολύ κακό και θυσιάζει αρκετή από τη δομή της. Όμως με βάση την είσοδο αυτή η σύνθεση κρίνεται αρκετά επιτυχημένη καθώς σε αρκετά σημεία διατηρείται η δομή πλέγματος.

Στη συνέχεια θα θέλαμε να κάνουμε μια σύγκριση με κάποιες από τις εικόνες που έχουμε συνθέσει με το δικό μας μοντέλο και αυτές που θα προέκυπταν από σύνθεση με το μοντέλο των Sivakumar και Goutsias.

Στις εικόνες 5.13, 5.14 και 5.15 μπορούμε να δούμε παράλληλες υλοποιήσεις με τα δύο μοντέλα. Στα αριστερά βρίσκονται οι υλοποιήσεις του δικού μας μοντέλου.



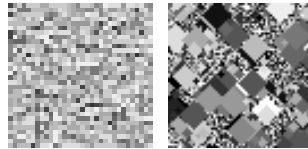
Σχήμα 5.13: Η πρώτη σύγκριση αφορά μια εικόνα η οποία δεν συντέθηκε με ιδιαίτερη επιτυχία από το δικό μας μοντέλο. Παρόλα αυτά αν και το μοντέλο των Sivakumar και Goutsias αντιλαμβάνεται ότι η εικόνα αποτελείται από μεγάλα συσσωματώματα δεν λαμβάνει υπόψη ούτε τις στάθμες του γκρι που πρέπει να έχει κάθε πρωταρχικό στοιχείο που συνθέτει την υφή ούτε και τη σχετική διάταξη αυτών στο επίπεδο της εικόνας.



Σχήμα 5.14: Η έντονη κατεύθυνση που έχουν τα πρωταρχικά στοιχεία της εικόνας υφής δεν απεικονίζονται στη δεξιά φωτογραφία. Επίσης πρέπει να παρατηρήσουμε ότι η σύγκλιση έγινε τόσο γρήγορα που δεν χρειάστηκε να αλλάξουν όλα τα pixel της προς σύνθεση εικόνας.

5.5.2 Αποτελέσματα classification

Στη συνέχεια προχωρήσαμε και σε ένα εκτεταμένο πείραμα ταξινόμησης υφών. Από το άλμπουμ του Brodatz επιλέχθηκαν 10 διαφορετικές εικόνες υφής. Οι εικόνες αυτές έχουν διαστάσεις 640×640 pixels. Κόψαμε το πρώτο 64×64 κομμάτι από κάθε εικόνα και το χρησιμοποιήσαμε για το βήμα εκμάθησης του αλγόριθμου κατηγοριοποίησης που παρουσιάσαμε προηγουμένως. Κατόπιν με τυχαίο τρόπο επιλέξαμε 16 διαφορετικά μη επικαλυπτόμενα κομμάτια 64×64 από κάθε εικόνα και τα χρησιμοποιήσαμε για το βήμα ταξινόμησης.



Σχήμα 5.15: Εδώ η διαφορά είναι εξαιρετικά εμφανής. Το μοντέλο μας πετυχαίνει υψηλό επίπεδο ρεαλισμού ενώ η σύνθεση με το μοντέλο των Sivakumar και Goutsias δεν απεικονίζεται η έμφαση στις μεσαίες στάθμες του γκρι που υπάρχει στην πραγματική εικόνα αλλά ούτε και η μικρή οριζόντια κατεύθυνση που παρατηρείται.

Τα αποτελέσματα είναι ιδιαίτερα ενθαρρυντικά και βρίσκονται στον πίνακα 5.1 με τη μορφή confusion matrix δηλαδή οριζόντια βλέπουμε τις πραγματικές εικόνες και κάθετα πού τις έχει κατατάξει ο αλγόριθμος. Οι εικόνες που χρησιμοποιήθηκαν για τη διαδικασία ταξινόμησης φαίνονται στα σχήματα 5.16, 5.17, 5.18 και 5.19.

Όπως βλέπουμε λοιπόν στον πίνακα 5.1 οι εικόνες D47, D9, D75 και D32 ταξινομούνται με απόλυτη επιτυχία. Στις εικόνες D18, D37, D10, D96, D56 έχουμε μία λάθος ταξινόμηση ενώ στην εικόνα D12 έχουμε δύο λάθος ταξινομήσεις. Πρέπει να πούμε ότι συνολικά το ποσοστό επιτυχίας βρίσκεται στο 95,27% κάτι που είναι ιδιαίτερα ενθαρρυντικό. Περαιτέρω βελτίωση των αποτελεσμάτων μπορεί να γίνει με αύξηση της πολυπλοκότητας του μοντέλου (αύξηση των συντελεστών I, J, θ, L) κάτι που κοστίζει όμως ιδιαίτερα τόσο σε υπολογιστικό χρόνο όσο και σε μνήμη. Επίσης πρέπει να αποφεύγονται λήψεις κομματιών υφής από τα άκρα των εικόνων καθώς πολλές φορές τόσο τα πρωταρχικά στοιχεία όσο και το κοντράστ αλλάζουν αρκετά και οδηγούν σε λάθος ταξινομήσεις.

5.6 Μετρήσεις πάνω στις εικόνες που προέκυψαν από σύνθεση

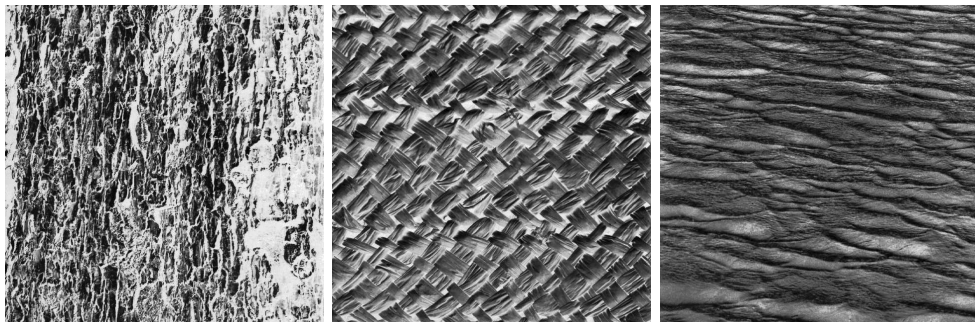
5.6.1 Μετρήσεις στατιστικής πρώτης τάξης

Παρόλο που σε περιπτώσεις σύνθεσης υφής το ανθρώπινο μάτι είναι ένας αρκετά αξιόπιστος τρόπος για να κριθούν τα αποτελέσματα ως προς το ρεαλισμό τους, θελήσαμε να πραγματοποιήσουμε μια σειρά από μετρήσεις οι οποίες να μας δείξουν πόσο κοντά είναι η εικόνα που έχουμε συνθέσει με την πραγματική υφή.

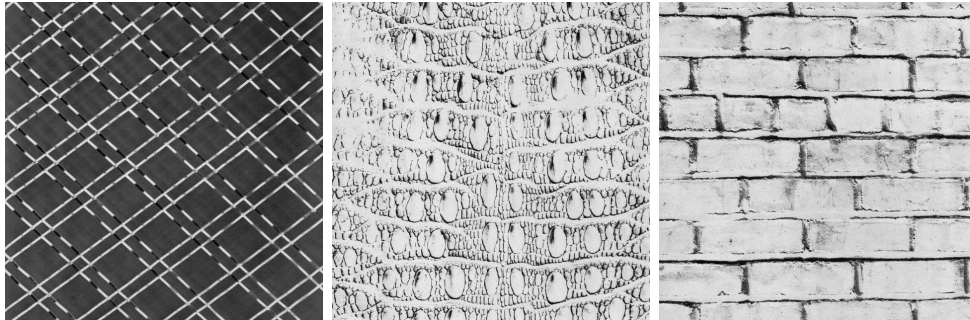
Στο κεφάλαιο 2 είχαμε αναφερθεί στον ορισμό από τον Julesz της στατιστικής πρώτης και δεύτερης τάξης. Ένα καλό κριτήριο για αξιολόγηση στατιστικής

	D12	D18	D37	D47	D10	D96	D56	D9	D75	D32
D12	14	2	0	0	0	0	0	0	0	0
D18	0	15	1	0	0	0	0	0	0	0
D37	0	1	15	0	0	0	0	0	0	0
D47	0	0	0	16	0	0	0	0	0	0
D10	0	0	0	0	15	1	0	0	0	0
D96	0	0	0	0	1	15	0	0	0	0
D56	1	0	0	0	0	0	15	0	0	0
D9	0	0	0	0	0	0	0	16	0	0
D75	0	0	0	0	0	0	0	0	16	0
D32	0	0	0	0	0	0	0	0	0	16

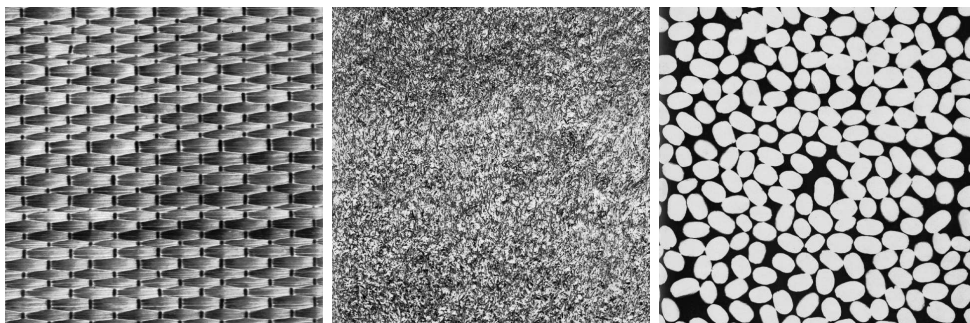
Πίνακας 5.1: Αποτελέσματα κατηγοριοποίησης



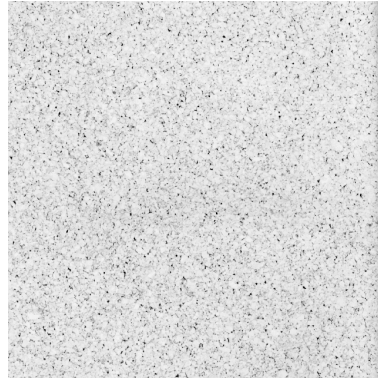
Σχήμα 5.16: Οι εικόνες D12,D18,D37.



Σχήμα 5.17: Οι εικόνες D47,D10,D96.



Σχήμα 5.18: Οι εικόνες D56,D9,D75.



Σχήμα 5.19: Η εικόνα D32.

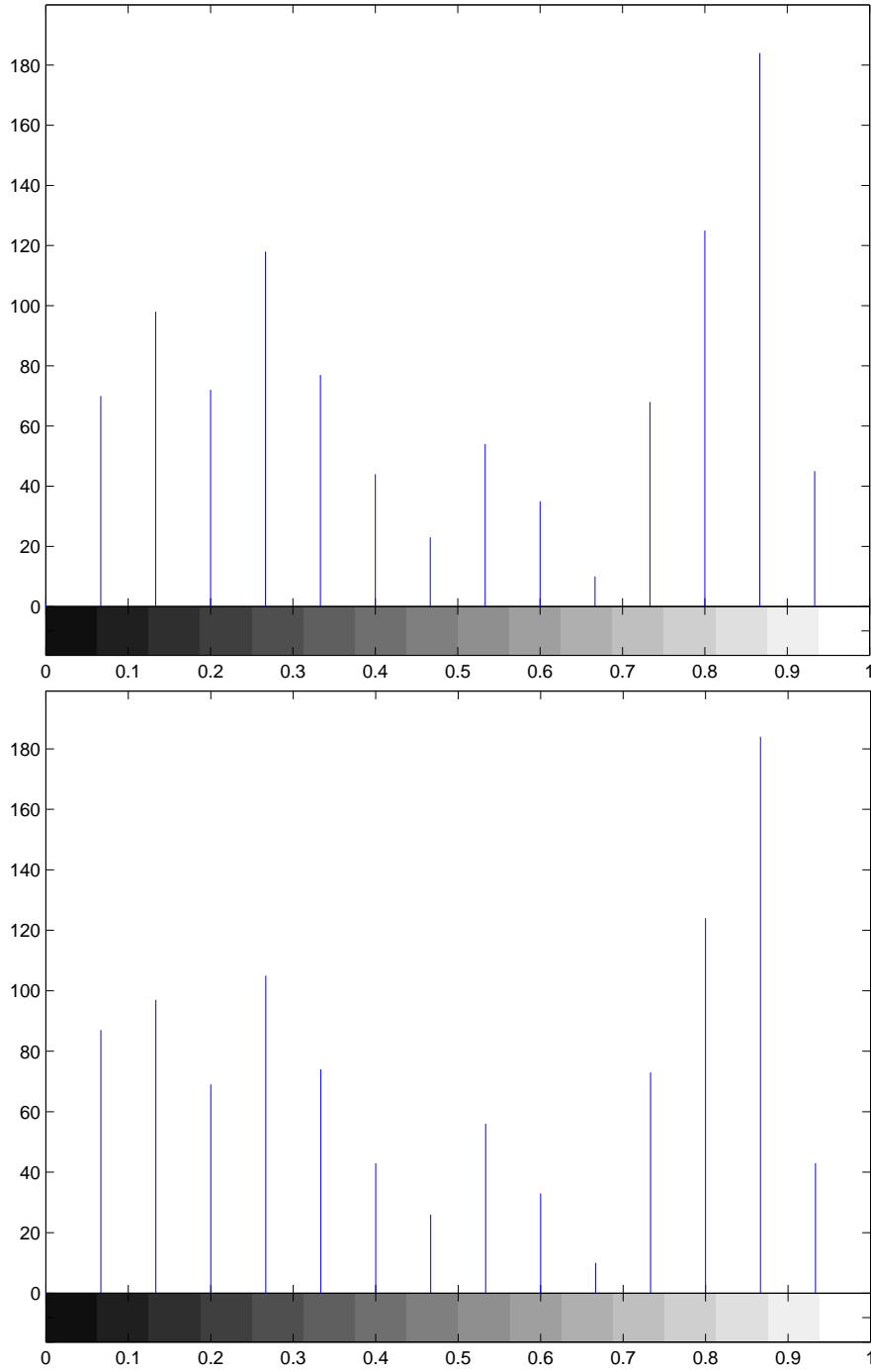
πρώτης τάξης (first order statistics) είναι το ιστογράμμα (histogramm) μιας εικόνας, το οποίο είναι ένα διάγραμμα που μας δείχνει την κατανομή των pixels της εικόνας στις στάθμες του γκριζου τις οποίες περιλαμβάνει. Στη συνέχεια στα σχήματα 5.20, 5.21, 5.22, 5.23 παραθέτουμε τα ιστογράμματα κάποιων από τις εικόνες που έχουμε συνθέσει. Σε κάθε σχήμα επάνω βρίσκεται το ιστογράμμα της εικόνας που έχει προκύψει από σύνθεση ενώ το δεύτερο της πραγματικής που χρησιμοποιήθηκε σαν δείγμα.

Από τα ιστογράμματα αυτά μπορούμε να δούμε ότι ο περιορισμός για το χρώμα που έχουμε θέσει στο μοντέλο αποδίδει πολύ καλά καθώς η κατανομές πλησιάζουν πάρα πολύ. Κάτι τέτοιο βέβαια ήταν ήδη ορατό από την αντιπαράβολή συνθετικής και πραγματικής υφής. Σημαντικό είναι επίσης ότι η σωστή κατανομή των σταθμών του γκρι στην εικόνα δεν εμπεριέχει κανέναν συμβιβασμό ως προς τη γενικότερη δομή της.

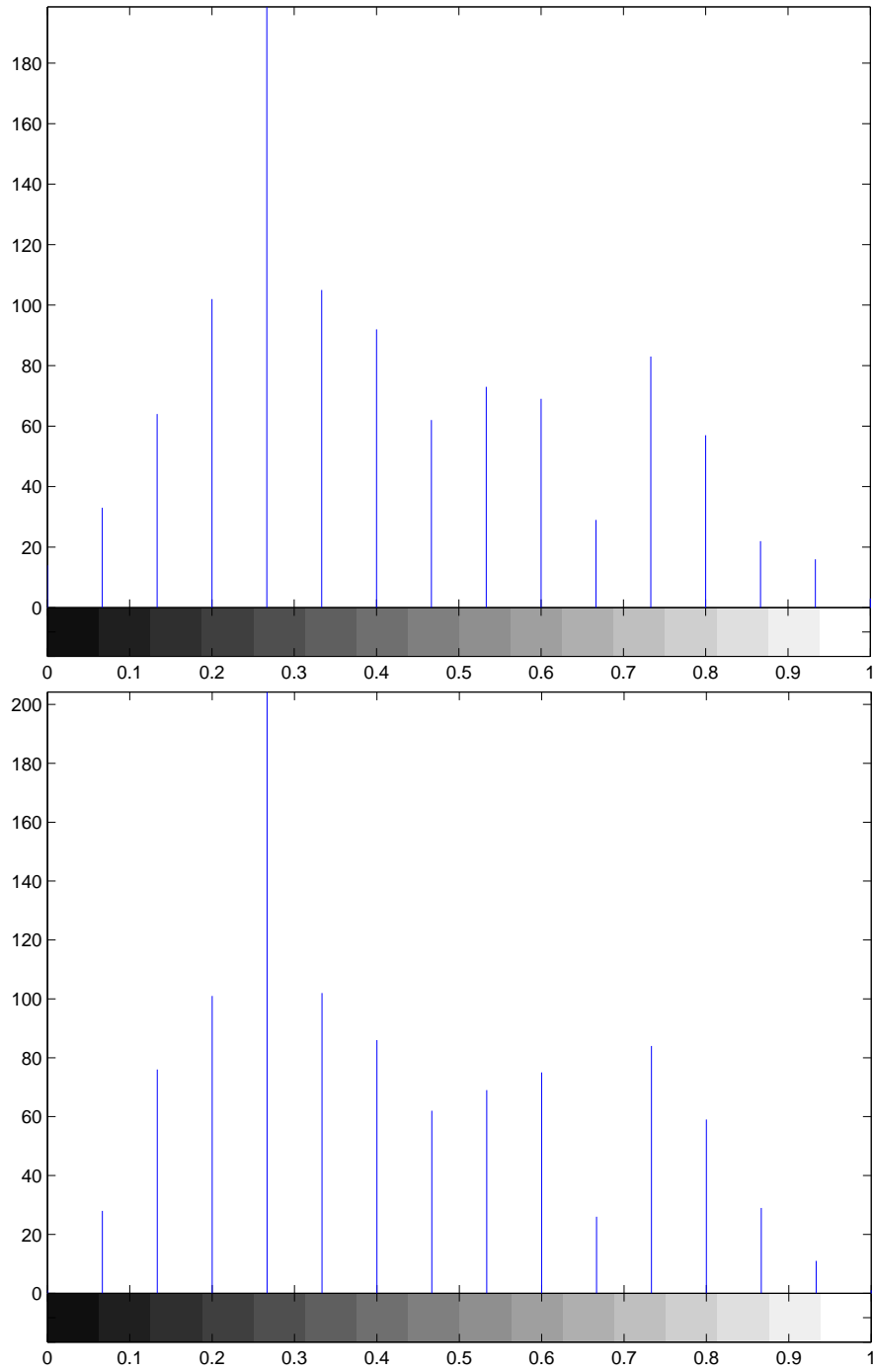
5.6.2 Μετρήσεις στατιστικής δεύτερης τάξης

Αφού είδαμε ότι οι μετρήσεις στατιστικής πρώτης τάξης ήταν ιδιαίτερα ενδιαφέρουσες αποφασίσαμε να προχωρήσουμε σε μετρήσεις στατιστικής δεύτερης τάξης (second order statistics). Ένας εύκολος τρόπος για να κάνουμε κάτι τέτοιο είναι να συγκρίνουμε τα *περιοδογράμματα* της πραγματικής υφής και αυτής που έχει προκύψει από σύνθεση.

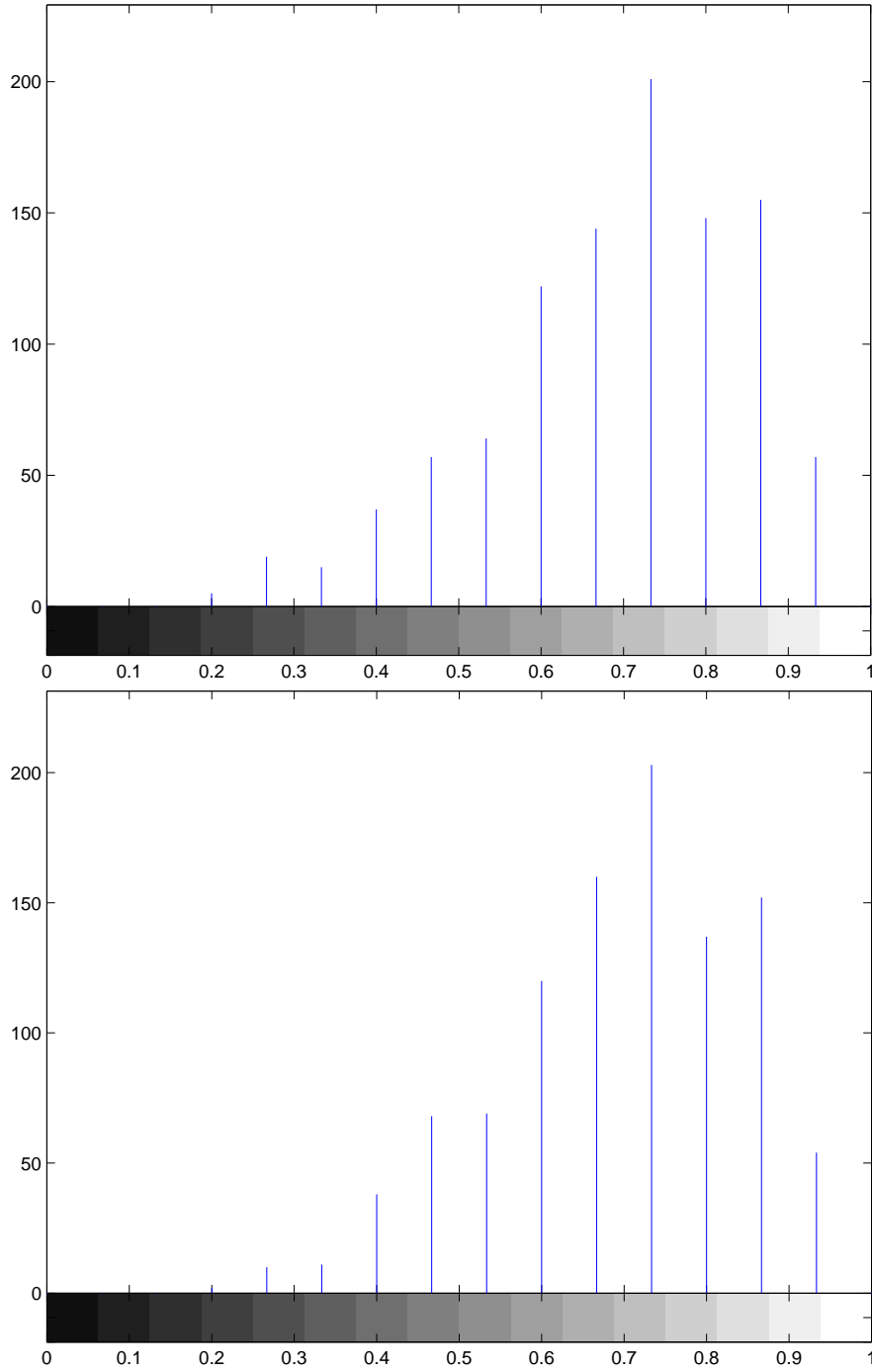
Το περιοδογράμμα είναι μία εκτίμηση του μη ντετερμινιστικού φάσματος ισχύος (power spectrum) μιας εικόνας. Συγκεκριμένα έστω εικόνα I διαστάσεων $x \times y$ η οποία διαιρείται σε N πιθανώς επικαλυπτόμενες εικόνες $I_m, 1 \leq m \leq N$ διαστάσεων $a \times b$ όπου $a < x, b < y$. Τότε το περιοδογράμ-



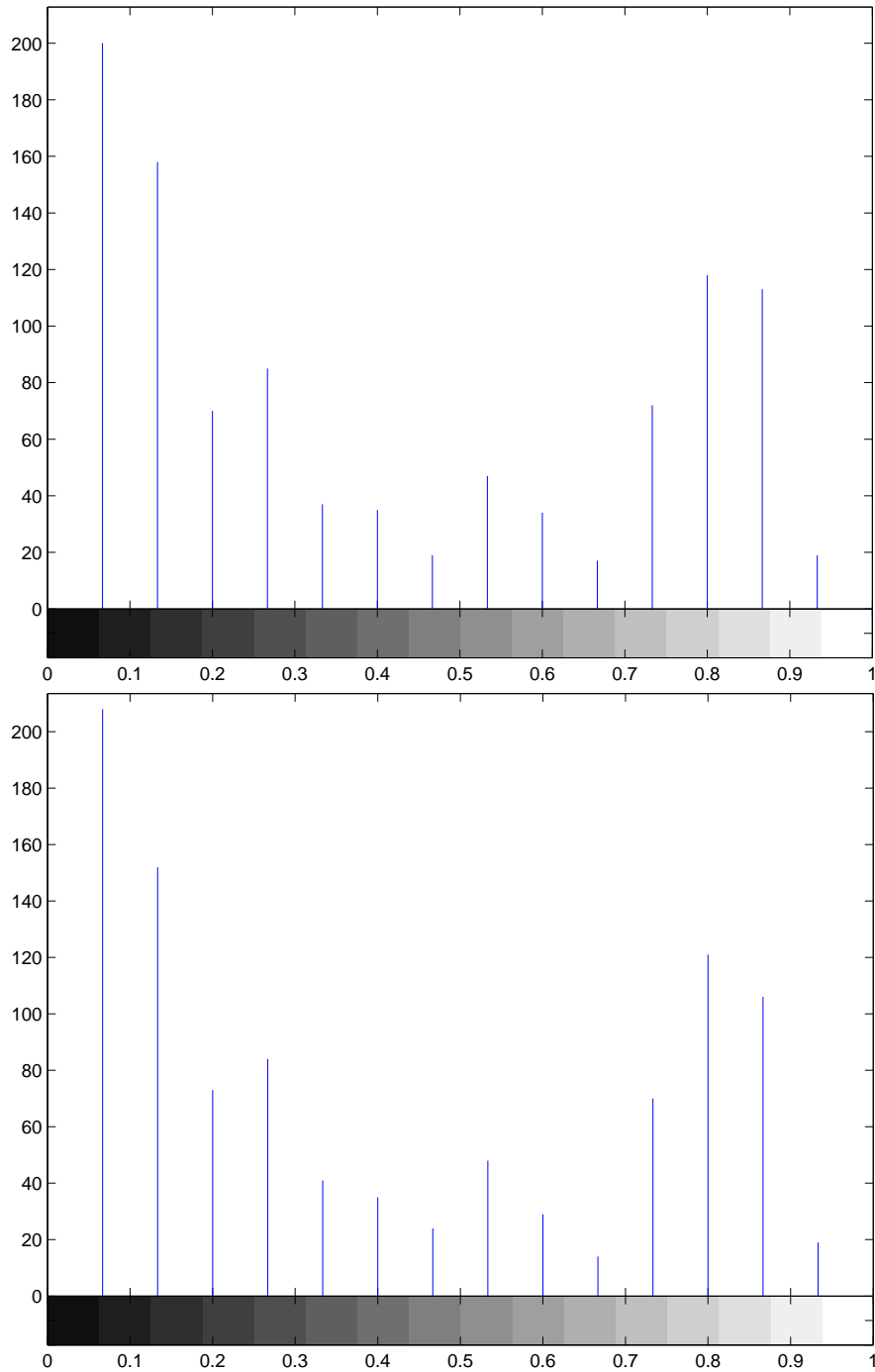
Σχήμα 5.20: Σύγκριση ιστογραμμάτων της συντεθειμένης και της πραγματικής εικόνας D12.



Σχήμα 5.21: Σύγκριση ιστογραμμάτων της συντεθειμένης και της πραγματικής εικόνας D18.



Σχήμα 5.22: Σύγκριση ιστογραμμάτων της συντεθειμένης και της πραγματικής εικόνας D37.



Σχήμα 5.23: Σύγκριση ιστογραμμάτων της συντεθειμένης και της πραγματικής εικόνας D51.

5.6 Μετρήσεις πάνω στις εικόνες που προέκυψαν από σύνθεση 101

Εικόνα	Επικάλυψη 25%	Επικάλυψη 50%
D12	2.01%	0.67%
D18	0.35%	2.36%
D37	1.75%	1.77%
D42	6.79%	13.65%
D48	0.39%	1.65%
D51	2.90%	3.26%
D56	1.52%	0.77%
D71	0.24%	2.7%
D75	2.78%	0.61%
D88	2.28%	2.86%
D91	30.07%	31.46%
D101	6.13%	6.33%

Πίνακας 5.2: Αποτελέσματα μετρήσεων second order statistics (Ποσοστιαία μεταβολή του περιοδογράμματος πραγματικής από συντεθειμένη εικόνα υπολογισμένη με βάση την ευκλείδεια νόρμα ενώ block size ίσο με 8×8 και $N = 8 \times 8$)

μα της εικόνας I θα είναι:

$$per(I) = \frac{\sum_{m=1}^N |\mathfrak{F}(I_m)|^2}{N} \quad (5.19)$$

όπου $\mathfrak{F}(I)$ είναι ο διδιάστατος μετασχηματισμός Fourier της εικόνας I .

Το κριτήριό μας για το αν δύο εικόνες η πραγματική I_{real} και αυτή που έχει προκύψει από σύνθεση I_{synth} έχουν παρόμοια στατιστική δεύτερης τάξης είναι η ποσοστιαία μεταβολή των δύο περιοδογραμμάτων τους υπολογισμένη με βάση την ευκλείδεια νόρμα. Συγκεκριμένα μετράμε το ποσοστό

$$K = \frac{\left(\sum_{i=1}^a \sum_{j=1}^b (per_{I_{synth}}(i, j) - per_{I_{real}}(i, j))^2 \right)^{1/2}}{\left(\sum_{i=1}^a \sum_{j=1}^b (per_{I_{real}}(i, j))^2 \right)^{1/2}} \times 100\%. \quad (5.20)$$

Τα αποτελέσματα φαίνονται στους πίνακες 5.2 και 5.3. Πρέπει να πούμε ότι η μέθοδος του περιοδογράμματος δίνει αξιόπιστα αποτελέσματα για μεγάλες εικόνες I και αντίστοιχα για μεγάλα τμήματα I_m . Εμείς λόγω της αδυναμίας μας να συνθέσουμε μεγάλες εικόνες προσπαθήσαμε να μετρήσουμε το περιοδογράμμα διαιρώντας τις εικόνες 32×32 που έχουμε φτιάξει είτε σε 25 εικόνες

Εικόνα	Επικάλυψη 0%	Επικάλυψη 50%
D12	1.43%	7.99%
D18	1.42%	5.06%
D37	0.25%	2.29%
D42	2.23%	15.08%
D48	4.98%	4.21%
D51	3.48%	5.68%
D56	1.03%	2.27%
D71	0.23%	4.09%
D75	2.69%	1.67%
D88	2.96%	4.86%
D91	32.5%	30.95%
D101	6.26%	6.71%

Πίνακας 5.3: Αποτελέσματα μετρήσεων second order statistics (Ποσοστιαία μεταβολή του περιοδογράμματος πραγματικής από συντεθειμένη εικόνα υπολογισμένη με βάση την ευκλείδεια νόρμα ενώ block size ίσο με 16×16 και $N = 16 \times 16$)

8×8 με επικάλυψη 25% είτε σε 49 εικόνες 8×8 με επικάλυψη 50%. Επίσης δοκιμάσαμε να υπολογίσουμε τις ίδιες ποσότητες με εικόνες 16×16 και επικάλυψη 0% και 50%. Είναι προφανές ότι μια εικόνα 8×8 δεν μπορεί να περιέχει σημαντική πληροφορία ιδιαίτερα όταν η αρχική εικόνα υψής αποτελείται από μεγάλα συσσωματώματα (όπως π.χ. η εικόνα *D91*). Γενικά παρατηρώντας τις τιμές του πίνακα 5.2 και τις αντίστοιχες υφές στις οποίες αναφέρονται βλέπουμε ότι έχουμε σχετικά μικρές αποκλίσεις στο περιοδογράμμα. Εξάιρεση αποτελούν κυρίως οι εικόνες *D42*, *D101* και *D91*. Για τις δύο πρώτες έχουμε να πούμε ότι η σύνθεση δεν ήταν η καλύτερη δυνατή κάτι που αποδεικνύεται και από την αδυναμία σύγκλισης του αλγορίθμου σύνθεσης για δεδομένο δ . Για την *D91* έχουμε να παρατηρήσουμε ότι αποτελείται από ένα μεγάλο συσσωματώμα ακανόνιστου σχήματος. Αυτό κατά τη σύνθεση προσδίδει μεγάλο βαθμό ελευθερίας καθώς οι συντελεστές του μοντέλου δεν έχουν ανιχνεύσει κάποια επικρατούσα κατεύθυνση ή κάποια συγκεκριμένη έστω τοπική διάταξη. Αυτό σημαίνει ότι η σύνθεση θα είναι μεν ρεαλιστική αλλά η στατιστική δεύτερης τάξης μπορεί να μην είναι ίδια καθώς δεν αντικατοπτρίζει πολλές από τις ιδιότητες της συγκεκριμένης υψής. Στον πίνακα 5.3 καθαρά για λόγους σύγκρισης παραθέτουμε τις ίδιες μετρήσεις αλλά με blocks διαστάσεων 16×16 και $N = 16 \times 16$. Εδώ επειδή διαλέξαμε μεγάλα blocks, για να ανεβάσουμε τον αριθμό τους από τα 4 (όπου τότε έχουμε 0% επικάλυψη) στα 9 χρησιμοποιούμε επικάλυψη 50% κάτι που εξηγεί τις μεγάλες διαφορές στους υπολογισμούς καθώς έχουμε με-

5.6 Μετρήσεις πάνω στις εικόνες που προέκυψαν από σύνθεση 103

γαλύτερη συσχέτιση ανάμεσα στα δείγματα που λαμβάνουμε. Γενικά όμως αν παρατηρήσουμε τη στήλη με 0% επικάλυψη θα δούμε ότι τηρείται η σχετική διαφορά των αποτελεσμάτων μεταξύ διαφορετικών εικόνων, σε σχέση με αυτή του προσηγούμενου πίνακα, με ελάχιστες αποκλίσεις.

Κεφάλαιο 6

Συμπεράσματα-Κατευθύνσεις για μελλοντική έρευνα

6.1 Συμπεράσματα

Στα πλαίσια αυτής της διπλωματικής εργασίας στην αρχή μελετήσαμε γενικά την έννοια της υψής και τα προβλήματα που σχετίζονται με αυτήν. Ερευνήσαμε τις μελέτες που έχουν γίνει μέχρι τώρα και αφορούν την ανθρώπινη αντίληψη της. Επίσης αναζητήσαμε τους πιο γνωστούς τρόπους με τους οποίους έγινε προσπάθεια να μοντελοποιηθεί η υφή. Από τα μοντέλα αυτά ξεχωρίσαμε τα Μαρκοβιανά τυχαία πεδία (MRF) τα οποία αποτελούν την πιο σύγχρονη άποψη για τη μοντελοποίηση πολλών μη ντετερμινιστικών φαινομένων. Ιδιαίτερα επικεντρωθήκαμε στην κατανόηση της βασικής τους θεωρίας καθώς και θέματα που σχετίζονται με αυτά όπως δειγματοληψία και εκτίμηση των παραμέτρων τους. Κατόπιν προχωρήσαμε σε υλοποιήσεις μοντέλων που βασίζονται σε αυτά και έχουν χρησιμοποιηθεί για ανάλυση υψής. Αξιολογώντας τα πλεονεκτήματα και τα μειονεκτήματα του κάθε μοντέλου επιλέξαμε ένα με στόχο να του προσθέσουμε κάποια χαρακτηριστικά τα οποία θα οδηγήσουν στη βελτίωσή του.

Συμπερασματικά η προσπάθεια βελτίωσης των αποτελεσμάτων σύνθεσης του μοντέλου στέφθηκε με επιτυχία. Χρησιμοποιήσαμε όσο το δυνατόν μεγαλύτερη γκάμα υψών για σύνθεση με στόχο να αναδείξουμε τόσο τις αρετές όσο και τις αδυναμίες του μοντέλου μας. Παρατηρήσαμε ότι σε κάθε περίπτωση το ιστόγραμμα της πραγματικής εικόνας υψής διατηρείται σε μεγάλο βαθμό και στη συνθετική εικόνα ενώ στις περισσότερες περιπτώσεις διατηρείται τόσο το μέγεθος όσο και η κατεύθυνση των πρωταρχικών στοιχείων που την απαρτίζουν. Πρόβλημα παρατηρήσαμε στη σύνθεση τέλειων γεωμετρικών σχημάτων (π.χ. κύκλων) που επαναλαμβάνονται περιοδικά εντός της εικόνας. Επίσης μικρό πρόβλημα παρατηρήσαμε στην απεικόνιση υψών που αποτελούνται από

γραμμές καμπύλες στις οποίες δηλαδή η κατεύθυνση αλλάζει συνεχώς. Τόσο η εκτίμηση παραμέτρων (που αξιολογείται από τα αποτελέσματα της σύνθεσης) όσο και τα αποτελέσματα του classification κρίνονται ως ιδιαίτερα επιτυχημένα. Απόδειξη της διατήρησης της χρωματικής κατανομής της πραγματικής εικόνας και στη συνθετική αποτελεί η απευθείας σύγκριση των ιστογραμμάτων τους (τα οποία αποτελούν τις μετρήσεις στατιστικής πρώτης τάξης) τα οποία είναι πανομοιότυπα. Μικρή απόκλιση παρατηρούμε σε στάθμες μικρότερης πιθανότητας. Τέλος η προσπάθεια μέτρησης στατιστικής δεύτερης τάξης επισκιάστηκε από την έλλειψη μεγάλου πλήθους δεδομένων (στην προκειμένη περίπτωση εικόνων μεγάλων διαστάσεων) τα οποία θα μας εξασφαλίσουν αφ' ενός στατιστική ανεξαρτησία και αφ' ετέρου εγκυρότητα των αποτελεσμάτων.

6.2 Κατευθύνσεις για μελλοντική έρευνα

Η πίεση του χρόνου καθώς και τα στενά πλαίσια της διπλωματικής εργασίας δεν μας επέτρεψαν την περαιτέρω διερεύνηση πολλών ακόμη θεμάτων τα οποία είναι ιδιαίτερα ενδιαφέροντα. Μερικά από αυτά τα οποία μας ενδιαφέρουν να τα μελετήσουμε στη συνέχεια είναι τα παρακάτω:

- Πολλές φορές αναφερθήκαμε στους περιορισμούς που μπαίνουν στη μελέτη των αποτελεσμάτων σύνθεσης του μοντέλου μας λόγω της ανεπάρκειας του αλγορίθμου σύνθεσης. Θεωρούμε επομένως καίρια την ανάπτυξη ενός ταχύ αλγόριθμου που θα επιτρέπει τη σύνθεση εικόνων υψηλής ανάλυσης εκμεταλλευόμενοι τη φύση του δομικού στοιχείου (ευθεία).
- Παρατηρήσαμε ότι για να έχουμε πολύ καλή απόδοση στην κατεύθυνση των πρωταρχικών στοιχείων θυσιάσαμε αρκετά την δυνατότητα απεικόνισης απλών γεωμετρικών σχημάτων τα οποία εμφανίζονται στην εικόνα με περιοδικό τρόπο. Για να λύσουμε κάτι τέτοιο σχεδιάζουμε τον εμπλουτισμό της συνάρτησης ενέργειας με όρους οι οποίοι μπορούν να συλλάβουν τέτοια χαρακτηριστικά. Μια σκέψη είναι η χρήση συνάρτησης αυτοσυσχέτισης.
- Θα μας ενδιέφερε να διερευνήσουμε την εισαγωγή στην συνάρτηση ενέργειας του μοντέλου μας ποσοτήτων οι οποίες έχουν παρόμοιες ιδιότητες με το pattern spectrum (άρα να μπορούν να εκτιμηθούν οι παράμετροί τους με αντίστοιχο τρόπο) προσδίδουν όμως διαφορετικές ιδιότητες στην παραγόμενη εικόνα. Μία από αυτές είναι και το power spectrum το οποίο σχεδιάζουμε να εισάγουμε στη συνάρτηση ενέργειας του μοντέλου μας κατα πάσα πιθανότητα υπολογισμένο στην έξοδο συστοιχίας φίλτρων.

- Ενδιαφέρον επίσης παρουσιάζει η εφαρμογή του παραπάνω μοντέλου σε φιλτραρισμένες εικόνες ή ακόμα καλύτερα σε έξοδο συστοιχίας φίλτρων. Έτσι δημιουργούμε τόσα μοντέλα όσα και τα φίλτρα και εκτιμούμε παραμέτρους για κάθε φίλτρο έτσι ώστε η σύνθεση να προκύψει ως το άθροισμα των συνθέσεων για κάθε φίλτρο.
- Ιδιαίτερο ενδιαφέρον παρουσιάζει η εφαρμογή του μοντέλου σε αποδιαμορφωμένες εικόνες υψής και συγκεκριμένα στο πλάτος αυτών το οποίο δεν παρουσιάζει πολλές διακυμάνσεις οπότε με μια εκτίμηση της συχνότητας να μπορούμε να έχουμε πιο ρεαλιστικά αποτελέσματα σε narrowband εικόνες και σε εικόνες που εμφανίζουν μικρή διαμόρφωση.
- Τέλος θα θέλαμε να πειραματιστούμε με φίλτρα sharpening και smoothing έτσι ώστε να έχουμε τη δυνατότητα να συλλάβουμε την έντονη κατεύθυνση η οποία δεν αφορά μια συγκεκριμένη μόνο στάθμη φωτεινότητας.

Αυτά είναι κάποια από τα θέματα που θα απασχολήσουν την έρευνά μας στο μέλλον.

Βιβλιογραφία

- [1] Zobrist A.I. και W.B Thompson. Building a distance function for gestalt grouping. *IEEE Transactions on Computers*, C-4:718–728, 1985.
- [2] Mumford D. Lee A. και Huang J. Occlusion models for natural images: A statistical study of a scale-invariant dead leaves model. *International Journal of Computer Vision*, 41:35–59, 2001.
- [3] Rosenfeld A. και E. Troy. Visual texture analysis. Τεχνική Αναφορά υπ. αριθμ. 70-116, University of Maryland, College Park, 1970.
- [4] A. Sutter Beck J. και R. Ivry. Spatial frequency channels and perceptual grouping in texture segregation. *Computer Vision, Graphics, and Image Processing*, 37:299–325, 1987.
- [5] J. Besag. Spatial interaction and the statistical analysis on lattice systems. *Journal of Royal Statistical Society*, 36:192–236, 1974.
- [6] McCormick B.H. και S.N. Jayaramamurthy. Time series model for texture synthesis. *International Journal of Computer and Informatin Sciences*, 3:329–343, 1974.
- [7] Julesz B. Visual pattern discrimination. *IRE Trans. on Information Theory*, 8:84–92, 1962.
- [8] Julesz B. Experiments in the visual perception of texture. *Scientific American*, 232:34–43, 1975.
- [9] Julesz B. A theory of preattentive texture discrimination based on first-order statistics of textons. *Biological Cybernetics*, 41:131–138, 1981.
- [10] Keller J. M. S. Chen και R. M. Crownover. Texture description and segmentation through fractal geometry. *Computer Vision, Graphics, and Image Processing*, 45:150–166, 1989.

- [11] Fernand S. Cohen και David B. Cooper. Simple parallel hierarchical and relaxation algorithms for segmenting noncausal markovian random fields. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 9(2):195–219, 1987.
- [12] George R. Cross και Anil K. Jain. Markov random field texture models. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 5(1):149–163, 1983.
- [13] Therrien C. An estimation-theoretic approach to terrain image segmentation. *Computer Vision, Graphics and Image Processing*, 22:313–326, 1983.
- [14] Haluk Derin και Howard Elliot. Modeling and segmentation of noisy and textured images using gibbs random fields. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 9(1):39–55, 1987.
- [15] Havlicek J. Harding D. και A. Bovik. Multidimensional quasi-eigenfunction approximations and multicomponent am-fm models. *IEEE Transactions on Image Processing*, 9(2):227–242, 2000.
- [16] F.Romeo D.Mitra και A.Sangiovanni Vincentelli. Convergence and finite time behavior of simulated annealing. *Advances in Applied Probability*, 18:747–771, 1986.
- [17] Alexei Efros και Thomas Leung. Texture synthesis by non parametric sampling. Corfu ,Greece, September 1999. IEEE International Conference on Computer Vision.
- [18] Deutsch E.L και N.J. Belknap. Texture descriptions using neighborhood information. *Computer Graphics and Image Processing*, 1:145–168, 1972.
- [19] Darling E.M. και R.D. Joseph. Pattern recognition from satellite altitudes. *IEEE Transactions on Systems, Man and Cybenetics*, SMC-4:38–47, 1968.
- [20] Kokkinos I. G. Evangelopoulos και P. Maragos. Advances in texture analysis: Energy dominant component & multiple hypothesis testing. Int”l Conf. Image Processing (ICIP-2004), 2004.
- [21] S. Geman και D.Geman. Stochastic relaxation, gibbs distributions, and the bayesian resoration of images. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 6:721–741, 1984.

- [22] Winkler Gerhard. *Image Analysis, Random Fields and Markov Chain Monte Carlo Methods*. Springer, 2003.
- [23] B. Gidas. Nonstationary markov chains and convergence of the annealing algorithm. *Journal of Statistical Physics*, 39:73–131, 1985.
- [24] B. Gidas. *Topics in Contemporary Probability and its Applications*, κεφάλαιο Metropolis type Monte Carlo simulation algorithms and simulated annealing. Boca Raton, Florida, 1995.
- [25] Matheron G. *Random Sets and Integral Geometry*. Wiley, New York, 1975.
- [26] B. Hajek. Cooling schedules for optimal annealing. *Mathematics of Operation Research*, 13:311–329, 1988.
- [27] R.M Haralick. Statistical and structural approaches to texture. Στο *Proceedings of the IEEE*, αριθμός 67, 1979.
- [28] J. K. Hawkins. *Picture Processing and Psychopictorics*, κεφάλαιο Textural Properties for Pattern Recognition. Academic Press, New York,, 1969.
- [29] Besag J. P. Green D. Higdon και K. Mengersen. Baysian computation and stochastic systems. *Statistical Science*, 10(1):3–66, 1995.
- [30] Elfadel I. και W. Picard. Gibbs random fields,cooccurences and texture modeling. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(1):24–37, 1994.
- [31] Box J.E. και G.M. Jenkins. *Time Series Analysis*. Holden-Day, San Francisco, 1970.
- [32] Havlicek J. και A. Bovik. *Handbook of Image and Video Processing*, κεφάλαιο Image Modulation Models, σελίδες 305–316. Bovik Ed., 2000.
- [33] Goutsias John. Markov random fields:interacting particle systems for statistical image modeling and analysis. Τεχνική Αναφορά υπ. αριθμ. 96-01, The John Hopkins University, 1996.
- [34] Serra J. και G. Verchery. Mathematical morphology applied to fibre composite materials. *Film Science and Technology*, 6:141–158, 1973.
- [35] Sklansky J. Image segmentation and feature extraction. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-8:237–247, 1978.

- [36] Reed J.S. και S.N. Jayaramamurthy. Automatic generation of texture feature detectors. *IEEE Transactions on Computers*, C-21:803–812, 1972.
- [37] L. A. Shepp Julesz B., E. N. Gilbert και H. L. Frisch. Inability of humans to discriminate between visual textures that agree in second-order statistics - revisited. *Perception*, 2:391–405, 1973.
- [38] B. Julesz. Textons, the elements of texture perception, and their interactions. *Nature*, 290:91–97, 1981.
- [39] Laws. Rapid texture identification. Στο *Proceedings of the SPIE Conference on Image Processing for Missile Guidance*, 1980.
- [40] B. B. Mandelbrot. *The Fractal Geometry of Nature*. Freeman, San Francisco, 1983.
- [41] B.S. Manjunath και R. Chellappa. Unsupervised texture segmentation using markov random field models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(5):478–482, 1991.
- [42] P. Maragos. Pattern spectrum and multiscale shape representation. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 11(7):701–716, 1989.
- [43] Gurelli M. και Onural L. On a parameter estimation method for gibbs-markov random fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(4):424–430, 1994.
- [44] Galloway M.M. Texture analysis using gray level run lengths. *Computer Graphics and Image Processing*, 4:172–179, 1975.
- [45] H. Tamura S. Mori και Y. Yamawaki. Textural features corresponding to visual perception. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-8:460–473, 1978.
- [46] Tuceryan M. και A. K. Jain. Texture segmentation using voronoi polygons. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12:211–216, 1990.
- [47] Ahuja N. Dot pattern processing using voronoi neighborhoods. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 4:336–343, 1982.

- [48] R. Paget. Strong markov field model. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2002.
- [49] Brodatz P. *Textures: A Photographic Album For Artists and Designers*. Dover, New York, 1966.
- [50] Chen P.C και T. Pavlidis. Segmentation by texture using a co-occurrence matrix and a split-and-merge algorithm. *Computer Graphics and Image Processing*, 10:172–182, 1979.
- [51] A. Pentland. Fractal-based description of natural scenes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 9:661–674, 1984.
- [52] Bartel P.H. και G.L. Wied. *Mamalian Cells: Probes and Problems*, κεφάλαιο Extraction and Evaluation of Information from Digitized Cell Images, σελίδες 15–28. U.S NTIS Technical Information Center, Springfield VA, 1975.
- [53] Maragos P. *Ανάλυση εικόνων και όραση υπολογιστών*. Ε.Μ.Π., 2005.
- [54] Maragos P. και A. Bovik. Image demodulation using multidimensional energy separation. *Journal of the Optical Society of America*, 12(9):1867–1876, 1995.
- [55] Chellappa R. και S. Chatterjee. Classification of textures using gaussian markov random fields. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 33(4):959–963, 1985.
- [56] Chellappa R. και R.L. Kashyap. Digital image restoration using spatial interaction models. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 30(3):461–472, 1982.
- [57] Haralick R. και L. Shapiro. *Computer and Robot Vision*. Addison Wesley, 1992.
- [58] W. Richards και A. Polit. Texture matching. *Kybernetik*, 16:155–162, 1974.
- [59] Haralick R.M. A texture-context feature extraction algorithm for remotely sensed imagery. Στο *Proceedings of the IEEE Conference on Computer Graphics, Pattern Recognition and Data Structure*, 1971.
- [60] Haralick R.M. A textural transform for images. Στο *Proceedings of the IEEE Conference on Computer Graphics, Pattern Recognition and Data Structure*, 1975.

- [61] Sutton R.N και E.L. Hall. Texture measurements for automatic classification of pulmonary disease. *IEEE Transactions on Computers*, C-21:667–676, 1972.
- [62] P.E. Hart R.O. Duda και D.G. Stork. *Pattern Classification*. Wiley, 2001.
- [63] Paget R και Longstaff I.D. Texture synthesis via a noncausal nonparametric multiscale markov random field. *IEEE TRANSACTIONS ON IMAGE PROCESSING*, 7(6):1–9, 1998.
- [64] Voss R. *Scaling Phenomena in Disordered Systems*, κεφάλαιο Random fractals: Characterization and measurement. Plenum, New York, 1986.
- [65] Suhail S. Saquib και Charles A. Bouman. Ml parameter estimation for markov random fields with applications to bayesian tomography. *IEEE TRANSACTIONS ON IMAGE PROCESSING*, 7(7):1029–1044, 1998.
- [66] Zhu S.C. Embedding gestalt laws in markov random fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(11):1170–1187, 1999.
- [67] K. Sivakumar και J. Goutsias. Morphologically constrained grfs: Applications to texture synthesis and analysis. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 21(2):99–113, 1999.
- [68] Lakshmanan S. και H. Derin. Simultaneous parameter estimation and segmentation of gibbs random fields using simulated annealing. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(8):799–813, 1989.
- [69] Y. Wu S.Zhu και D. Mumford. Filters, random fields and maximum entropy: towards a unified theory for texture modeling. *Int.l J. Computer Vision*, 27(2):1–20, 1998.
- [70] Caelli T. *Visual Perception*. Pergamon Press, 1981.
- [71] Charles A. Bouman. Markov Random Fields and Stochastic Image Models. Available from: <http://dynamo.ecn.purdue.edu/~bouman>, χ.χ.
- [72] M. Tuceryan και A. K. Jain. *The Handbook of Pattern Recognition and Computer Vision*, κεφάλαιο 2.1, σελίδες 207–248. World Scientific Publishing Co., 1998.

-
- [73] G. Voronoi. Nouvelles applications des parametres continus a la theorie des formes quadratiques. deuxieme memoire: Recherches sur les paralleloedres primitifs,. *J. Reine Angew. Math*, 134:198–287, 1908.
- [74] Wu Y. και Liw S. Equivalence of julesz ensembles and frame models. *International Journal of Computer Vision*, 38:247–265, 2000.
- [75] Song Chun Zhu και David Mumford. Prior learning and gibbs reaction-diffusion. *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE*, 19(11):1236–1250, 1997.
- [76] S. W. Zucker και K. Kant. Multiple-level representations for texture discrimination. Στο *Proceedings of the IEEE Conference on Pattern Recognition and Image Processing*, σελίδες 609–614, Dallas, TX., 1981.