Πολυτεχνείο Κρήτης

Τμήμα Ηλεκτρονικών Μηχανικών και Μηχανικών Υπολογιστών

Εκτίμηση ποιότητας μαρμάρινων επιγραφών

Θεοδώρα Γκορόγια

Σεπτέμβριος 2002

Επιτροπή:

καθηγητής Μιχάλης Ζερβάκης (επιβλέπων)
αναπληρωτής καθηγητής Κωνσταντίνος Καλαϊτζάκης
επίκουρος καθηγητής Ευριπίδης Πετράκης

Θα ήθελα να εκφράσω τις θερμές μου ευχαριστίες στον επιβλέποντα καθηγητή μου κ. Μιχάλη Ζερβάκη για την πολύτιμη καθοδήγησή και την άψογη συνεργασία του καθώς και στους γονείς και τους φίλους μου για την στήριξη που μου προσέφεραν.

Περιεχόμενα

Πρόλογος

Στην εργασία αυτή επιχειρείται η μελέτη εικόνων που έχουν ανακτηθεί από μαρμάρινες επιγραφές με μεθόδους επεξεργασίας εικόνας. Πιο συγκεκριμένα υλοποιείται ένα αυτοματοποιημένο σύστημα που εκτιμά την πληροφορία που περιέχεται σε τέτοιου είδους επιφάνειες, είτε αυτή είναι απλά επιφάνεια του μαρμάρου με ή χωρίς διάβρωση, είτε είναι χαρακτήρας της επιγραφής. Στόχος της εργασίας είναι να εξετάσει τη δυνατότητα μιας τέτοιας εκτίμησης με μεθόδους image processing και να συγκρίνει την αποδοτικότητα τέτοιων αλγορίθμων.

Ο αλγόριθμος που υλοποιείται βασίζεται κατά κύριο λόγο στην μελέτη και την εξαγωγή χαρακτηριστικών (feature extraction) που προσδιορίζουν επαρκώς την ποιότητα των διαφόρων υποπεριοχών των εικόνων. Επίσης τα χαρακτηριστικά αυτά αποτελούν τη βάση για το διαχωρισμό των διαφόρων περιοχών σε κατηγορίες με κοινή συμπεριφορά και συντελούν στην λεπτομερή εξέταση και εκτίμηση περιοχών που εμφανίζουν συγκεκριμένες ιδιότητες (περιοδικότητα, τυχαιότητα, κατευθυντικότητα, ομαλότητα). Αναφερόμαστε σε χωρικά και συχνοτικά χαρακτηριστικά που συντελούν στη σφαιρική εξέταση των διαφόρων περιοχών των εικόνων, σε περιοδικά χαρακτηριστικά που ανιχνεύουν το μέγεθος της περιοδικότητας που εμφανίζεται σε περιοχές με μεγάλη διάβρωση καθώς και σε κατευθυντικά χαρακτηριστικά που προσδιορίζουν τη δομή και την κατευθυντικότητα των περιοχών.

1. Εισαγωγή

1.1 Καθορισμός του προβλήματος

Η αρχαιολογική και ιστορική σημασία των αρχαίων επιγραφών είναι ανεκτίμητη αφού με αυτόν τον τρόπο έχει μείνει αποτυπωμένος στο χρόνο ο γραπτός λόγος των προγόνων μας. Συνεχής είναι η προσπάθεια μελέτης αλλά και συντήρησης αυτής της πολιτιστικής κληρονομιάς από αρχαιολόγους και συντηρητές αρχαιοτήτων. Δεν είναι λίγα τα προβλήματα που αντιμετωπίζουν στην προσπάθειά τους αυτή αν λάβουμε υπόψη μας ότι οι περισσότερες από αυτές τις επιγραφές είναι χαραγμένες σε μάρμαρο ή πέτρα, υλικά που η αρχική τους μορφή με το πέρασμα του χρόνου αλλοιώθηκε. Η έκθεση τους σε εξωγενείς παράγοντες, όπως νερό και αέρας, συντέλεσε στη διάβρωση και στην εν μέρει καταστροφή της επιφάνειάς τους.

Στην εργασία αυτή γίνεται μία προσπάθεια μελέτης εικόνων μαρμάρινων επιγραφών με μεθόδους image processing. Πιο συγκεκριμένα υλοποιείται ένα αυτοματοποιημένο σύστημα που εκτιμά την πληροφορία που περιέχεται σε τέτοιου είδους επιφάνειες, είτε αυτή είναι απλά επιφάνεια του μαρμάρου με ή χωρίς διάβρωση, είτε είναι χαρακτήρας της επιγραφής.

Στόχος της εργασίας είναι να εξετάσει τη δυνατότητα μιας τέτοια εκτίμησης με μεθόδους ψηφιακής επεξεργασίας εικόνας και να συγκρίνει την αποδοτικότητα τέτοιων αλγορίθμων. Έτσι η εργασία εστιάζεται στην επεξεργασία των δεδομένων παρά σε μεθόδους πιστής αποτύπωσης της εικόνας της επιγραφής. Για το σκοπό αυτό χρησιμοποιήθηκε μία απλή κάμερα (ψηφιακή) με flash, χωρίς να δίνεται ιδιαίτερη σημασία στην τοποθέτηση της κάμερας. Ένα τέτοιο σύστημα έχει πολύ χαμηλό κόστος, μπορεί να χρησιμοποιηθεί από τον οποιονδήποτε και σε ένα μεγάλο εύρος εφαρμογών, πέραν αυτής που εξετάζουμε.

Αυτό που πρέπει επίσης να επισημάνουμε είναι ότι στόχο δεν αποτελεί η αναγνώριση των χαρακτήρων αλλά ο διαχωρισμός των διάφορων ομαλών υποπεριοχών της εικόνας που περιέχουν χαρακτήρες από άλλες με πολύ θόρυβο που είτε περιέχουν γράμμα είτε όχι και η εκτίμηση της ποιότητά τους.

Για να οριστεί καλύτερα ο όρος «ποιότητα» θα πρέπει να λάβουμε υπόψη μας ότι τα γράμματα χαρακτηρίζονται από μία ποικιλία ως προς τη μορφή τους. Συναντάμε λοιπόν χαρακτήρες με πολύ λεπτή δομή, άλλους περισσότερο ευκρινείς με πιο ισχυρή δομή, «σπασμένους» χαρακτήρες και τέλος χαρακτήρες με έντονη εκτεταμένη δομή. Η μορφή με την οποία παρουσιάζεται το κάθε γράμμα οφείλεται στην ανομοιομορφία του μαρμάρου στο οποίο είναι χαραγμένα τα γράμματα. Στη συνολική επιφάνεια του μαρμάρου συναντάμε κάποιες σχετικά ομαλές περιοχές, περιοχές με περιοδική (πορώδη) διάβρωση, άλλες με γραμμική διάβρωση με τη μορφή επιμηκών δομών καθώς επίσης σπασίματα και πολύ έντονες τυχαίες γραμμώσεις («χαρακιές»). Γενικά, μία υποπεριοχή θεωρείται ότι έχει *αποδεκτή ποιότητα* όταν περιέχει χαμηλό ποσοστό θορύβου και μικρό ποσοστό διάβρωσης και όταν το γράμμα που τυχόν υπάρχει σε αυτήν την περιοχή είναι ευκρινές.

Για να προσδιοριστεί η ποιότητα κάθε υποπεριοχής επιχειρείται η εξαγωγή διάφορων χαρακτηριστικών, τα οποία δίνουν την απαραίτητη πληροφορία για τη μορφή και τις ιδιότητες της υποπεριοχής. Ανάλογα με τα ιδιόμορφα χαρακτηριστικά της κάθε περιοχής χρησιμοποιούνται εκείνα τα χαρακτηριστικά υφής που θεωρήθηκαν ότι ήταν τα πιο κατάλληλα για την εκτίμηση της ποιότητάς της.

Παράγοντες που δυσχεραίνουν την προσπάθεια εκτίμησης είναι η ανομοιομορφία του φωτισμού της εικόνας καθώς και η κακή ανάλυσή της, στοιχεία που αλλοιώνουν κατά κάποιο τρόπο την πληροφορία που υπάρχει στην εικόνα και επηρεάζουν αρνητικά την επεξεργασία της. Και τα δύο αυτά προβλήματα προκύπτουν από το σύστημα ανάκτησης της εικόνας που ίσως αν ήταν καλύτερο (καλύτερη ψηφιακή camera) τα προβλήματα αυτά να μην ήταν τόσο έντονα.

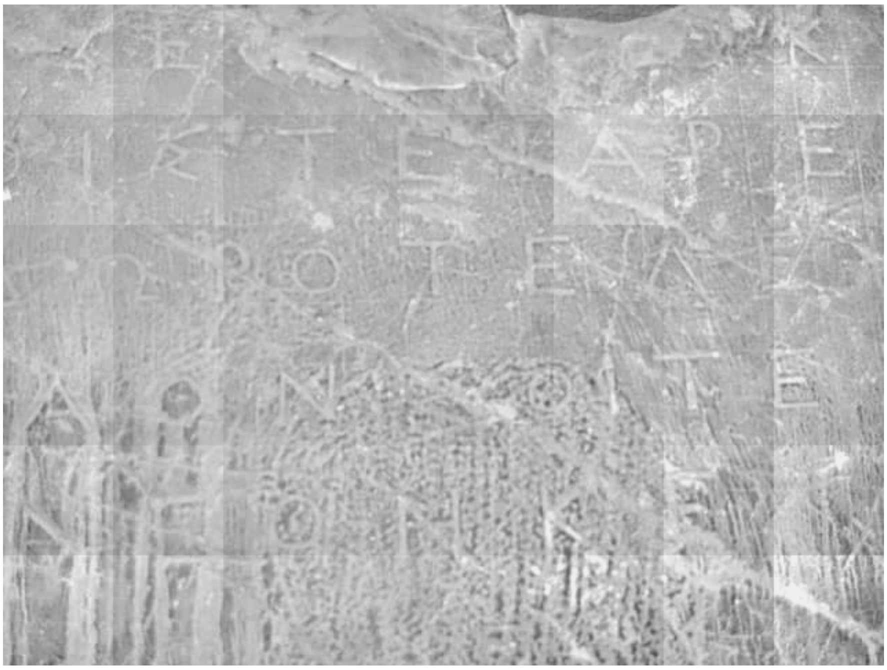
Η εργασία αυτή έχει την ακόλουθη δομή: Στην *Παράγραφο 1.2* γίνεται μία προεργασία στα πλαίσια της οποίας τμηματοποιείται η εικόνα σε υποπεριοχές στις οποίες εφαρμόζονται όλοι οι μέθοδοι που θα αναπτύξουμε στη συνέχεια και γίνεται μία προσπάθεια να εξομαλυνθούν ως ένα βαθμό τόσο τα προβλήματα του συστήματος ανάκτησης όσο και τα προβλήματα του ανομοιόμορφου background του μαρμάρου. Στο *2^{o} Κεφάλαιο* εξάγεται και μελετάται μία ομάδα χαρακτηριστικών οι ιδιότητες των οποίων προσδιορίζουν τη συμπεριφορά των διαφόρων υποπεριοχών και τα οποία αφενός μεν αποτελούν τη βάση για το διαχωρισμό των διαφόρων περιοχών σε κατηγορίες και αφετέρου συντελούν στην λεπτομερή εξέταση και εκτίμηση περιοχών που εμφανίζουν συγκεκριμένες ιδιότητες.
Στη συνέχεια στο *3^{o} Κεφάλαιο* παρουσιάζονται τα διάφορα στάδια του αλγορίθμου εκτίμησης της ποιότητας των υποπεριοχών. Ο αλγόριθμος χρησιμοποιεί τα χαρακτηριστικά του 2^{ου} Κεφαλαίου για να υλοποιηθεί. Στο 4^{o} και τελευταίο κεφάλαιο θα δούμε μερικά παραδείγματα εφαρμογής του αλγορίθμου στις διάφορες υποπεριοχές και θα παρουσιάσουμε το τελικό αποτέλεσμα στο οποίο καταλήγει η συγκεκριμένη εργασία.

1.2 Pre-processing

Η αρχική εικόνα (*Εικόνα 1*) χωρίζεται σε πολλά κομμάτια (segments) με σκοπό να απομονωθούν καλύτερα οι διάφορες περιοχές και η επεξεργασία τους να γίνει με μεγαλύτερη λεπτομέρεια. Η εικόνα διαιρείται με τέτοιον τρόπο έτσι ώστε στα διάφορα segments να περιέχονται κατά το δυνατόν ολόκληροι χαρακτήρες και το κάθε segment να περιέχει της ίδιας περίπου μορφής background. Για να επιτευχθεί αυτό η εικόνα διαιρείται σε 48 κομμάτια διαστάσεων 80x80. Άλλοι μέθοδοι τμηματοποίησης της εικόνας συζητούνται στις μελλοντικές επεκτάσεις.

Εικόνα 1: Μία από τις εικόνες της μαρμάρινης επιγραφής στην οποία βασίστηκε η εργασία. Στην εικόνα αυτή φαίνεται και πως ακριβώς έχει γίνει το segmentation.

Στη συνέχεια για να εξομαλυνθούν κάπως τα προβλήματα που αναφέρθηκαν προηγούμενα και αφορούν στην συνολική εμφάνιση της εικόνας, χρησιμοποιείται μία απλή μέθοδος σύμφωνα με την οποία πραγματοποιείται μία γραμμική κλιμάκωση (scaling) κάθε σημείου της εικόνας για να εξομαλυνθούν οι τιμές των διαφόρων υποπεριοχών και να βρίσκονται με αυτόν τον τρόπο στα ίδια επίπεδα. Υπολογίζεται η μέση τιμή μ_i των τιμών των εντάσεων του κάθε segment και ο μέσος όρος μ των μέσων τιμών και των 48 segments. Τέλος κάθε segment πολλαπλασιάζεται με το λόγο $\frac{\mu}{\mu_i}$ για να επιτευχθεί η ομοιομορφία που θέλουμε, φέρνοντας τη μέση ένταση της κάθε περιοχής στο ίδιο επίπεδο. Παρακάτω φαίνεται πως γίνεται η αρχική εικόνα μετά το scaling που εφαρμόζουμε.



Η εικόνα μετά το scaling

Ένας άλλος τρόπος για να αντιμετωπίσουμε το πρόβλημα του ανομοιόμορφου φωτισμού είναι η μέθοδος του *image leveling με* προσαρμόσιμες συναρτήσεις του background καθώς και η μέθοδος του *rank leveling* που προτείνονται από τον John C. Russ[1], μέθοδοι οι οποίες έχουν ως στόχο τους την «εξισορρόπηση» της αρχικής εικόνας.

Στην εργασία αυτή προτιμάται να εφαρμοστεί η μέθοδος του scaling διότι είναι μία εύκολη μέθοδος , με μικρό κόστος υλοποίησης. Επίσης ο τρόπος που επιλύεται το πρόβλημα -οι μέθοδοι που χρησιμοποιούνται- δεν επηρεάζεται σε μεγάλο βαθμό από την ποιότητα της αρχικής εικόνας.

2. Τύποι Χαρακτηριστικών Υφής

2.1 Γενική περιγραφή χαρακτηριστικών

Για τον καθορισμό της ποιότητας της υποπεριοχής της εικόνας σημαντικό ρόλο παίζουν η ομοιομορφία του background και η ευκρίνεια του γράμματος. Επειδή οι διάφορες υποπεριοχές της εικόνας εμφανίζουν μία ποικιλία ως προς τη μορφή του background (περιοδικότητα, τυχαιότητα, κατευθυντικότητα, ομαλότητα) αλλά και την ευκρίνεια του γράμματος (γράμματα με ικανοποιητική ευκρίνεια, θολά γράμματα, σπασμένοι χαρακτήρες λόγω διάβρωσης) υπεισέρχεται η ανάγκη να μελετηθούν κάποια χαρακτηριστικά υφής για να προσδιοριστούν με αυτόν τον τρόπο οι ιδιότητες και η συμπεριφορά των υποπεριοχών.

Η προσέγγιση του προβλήματος αρχικά γίνεται μελετώντας μερικά **ολικά χαρακτηριστικά** με σκοπό την υλοποίηση μιας πρωταρχικής κατηγοριοποίησης των υποπεριοχών. Γίνεται αναφορά σε *χ ω ρ ι κ ά* χαρακτηριστικά και πιο συγκεκριμένα στην **αυτοσυσχέτιση** της εικόνας η οποία παρέχει πολύτιμη πληροφορία για την συνολική ποιότητα της υποπεριοχής. Στη συνέχεια μελετώνται χαρακτηριστικά που εξάγονται στο επίπεδο της *σ υ χ ν ό τ η τ α ς* και προσδιορίζουν το ενεργειακό περιεχόμενο των υποπεριοχών.

Επόμενο βήμα είναι να εξεταστούν και να εξαχθούν χαρακτηριστικά που παρέχουν μία πιο λεπτομερή πληροφορία για την ποιότητα κάποιων υποπεριοχών οι οποίες εμφανίζουν συγκεκριμένες ιδιότητες (π.χ. περιοδικότητα ή κατευθυντικότητα). Συνεπώς εξετάζονται **περιοδικά** χαρακτηριστικά προσεγγίζοντας την έννοια της **περιοδικότητας** που εμφανίζεται σε πολλές υποπεριοχές και τέλος αναλύονται κάποια **κατευθυντικά** χαρακτηριστικά που προσδιορίζουν τη δομή της εικόνας.

2.2 Ολικά Χαρακτηριστικά

2.2.1 Αυτοσυσχέτιση (Autocorrelation)

Η **αυτοσυσχέτιση** (autocorrelation) μίας εικόνας διαστάσεων MxN ορίζεται:

$$\mathrm{A}(k,l) = \sum_x \sum_y f^*(x,y) \cdot f(x-k, y-l),$$

όπου k = 0,1,2,…M-1, l = 0,1,2,…,N-1

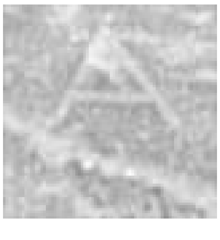
Για τον υπολογισμό της αυτοσυσχέτισης εξάγεται ο αντίστροφος μετασχηματισμός Fourier του φάσματος ισχύος της εικόνας:

$$\mathrm{A(k,l)} = \mathrm{IDFT}\ (|F(u,v|^2)$$

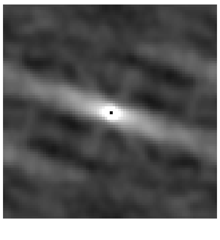
με φάσμα ισχύος $|F(u,v)|^2 = F(u,v) \cdot F^*(u,v)$

Η αυτοσυσχέτιση μιας εικόνας ουσιαστικά αποτελεί ένα σάρωμα (φιλτράρισμα) της εικόνας με τον εαυτό της. Όταν οι δύο όμοιες εικόνες ταυτιστούν (στο σημείο x=0, y=0) τότε έχουμε τις μέγιστες τιμές έντασης στην εικόνα και τη μέγιστη πληροφορία που αυτή περιέχει.

Στο *Σχήμα 1*(ii) παρουσιάζεται το autocorrelation μιας υποπεριοχής της εικόνας (*Σχήμα 1*(i))



(i)



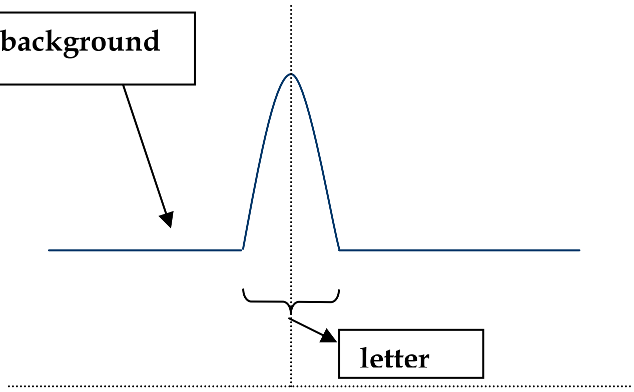
(ii)

Σχήμα 1

Από ότι παρατηρείται γύρω από το κέντρο της εικόνας της αυτοσυσχέτισης εμφανίζεται μία δομή σε μορφή «Χ» η οποία οφείλεται στην κατευθυντικότητα του γράμματος καθώς και έντονη μία διαγώνια δομή που οφείλεται στην «χαρακιά» της υποπεριοχής.

Το autocorrelation υπολογίζεται για όλες της υποπεριοχές της εικόνας. Η τιμή της αυτοσυσχέτισης θεωρείται μηδενική στο κέντρο έτσι ώστε οι μεγάλες συνιστώσες της έντασης που εμφανίζονται εκεί να αγνοούνται και να μην επηρεάζουν τα αποτελέσματα.

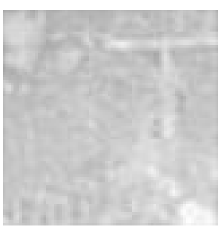
Παρατηρείται ότι ανάλογα με την ποιότητα και τη δομή της αρχικής εικόνας προκύπτει διαφορετική μορφή autocorrelation. Στο συγκεκριμένο πρόβλημά μας, η ιδανική μορφή που αναμένεται, φαίνεται στο παρακάτω σχήμα:

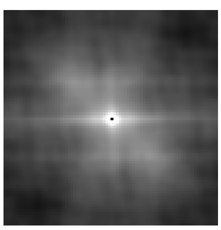


Σχήμα 2

Πιο συγκεκριμένα αναμένονται μεγάλες τιμές autocorrelation στο κέντρο που μειώνονται ομαλά γύρω από αυτό, οι οποίες οφείλονται στο γράμμα που υπάρχει στην υποπεριοχή (στη μικρή περιοχή που αυτό καταλαμβάνει). Επίσης αναμένεται ομοιόμορφο background στις υπόλοιπες περιοχές, με περίπου σταθερό autocorrelation καθώς απομακρυνόμαστε από το κέντρο της συνάρτησης. Συνεπώς με βάση την αυτοσυσχέτιση, επιχειρείται η ακόλουθη ταξινόμηση των υποπεριοχών:

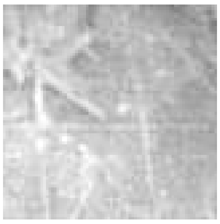
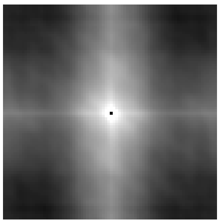
- Ο*μαλές* περιοχές που περιέχουν *εμφανές* γράμμα και καταλαμβάνουν μεγάλο εύρος στην autocorrelation μορφή τους αλλά παρουσιάζουν μεγάλες τιμές κυρίως γύρω από το κέντρο. Τέτοιου είδους περιοχές προσεγγίζουν την ιδανική μορφή του autocorrelation που περιγράφθηκε. Ένα παράδειγμα μιας τέτοιας υποπεριοχής και τη μορφή της αυτοσυσχέτισή της φαίνεται ακολούθως:



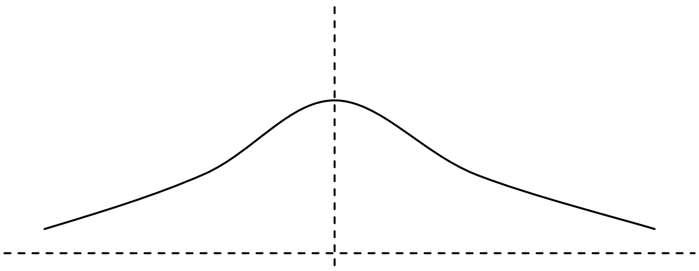


Υπάρχουν όμως άλλες περιοχές που η μορφή της αυτοσυσχέτισής τους αποκλίνει από την επιθυμητή όπως είναι οι παρακάτω περιπτώσεις:

- Περιοχές που περιέχουν μη ευδιάκριτο γράμμα αλλά εμφανίζουν κάποιες γραμμώσεις και μία ανομοιομορφία ως προς το φωτισμό τους, τείνουν να καταλαμβάνουν μεγάλη, με smooth μορφή, έκταση (large extent) στην autocorrelation μορφή τους:

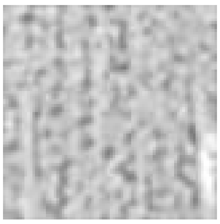
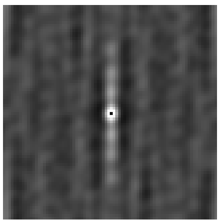



Η autocorrelation μορφή στον κατακόρυφο άξονα θα φαίνεται όπως στο παρακάτω σχήμα:

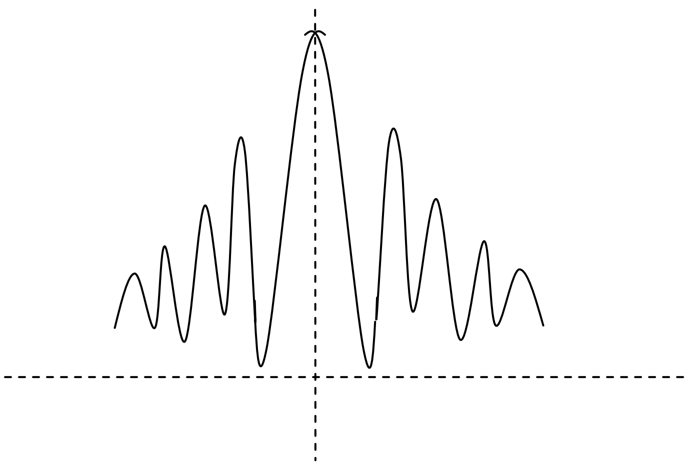


Παρατηρούμε ότι το άνοιγμα της καμπύλης είναι αρκετά μεγάλο και αυτό οφείλεται στο γεγονός ότι σε αυτόν τον άξονα η κυριαρχούσα δομή της εικόνας δεν αλλάζει, λόγω κυρίως των κατακόρυφων γραμμώσεων.

- *Περιοδικές (periodic)* εικόνες εμφανίζουν μεγάλη έκταση στο autocorrelation με περιοδική μορφή.

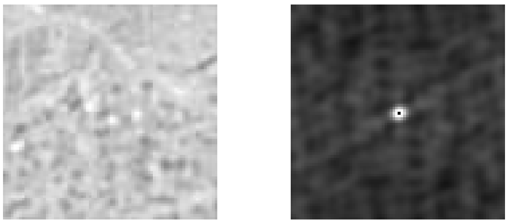



Τέτοιες περιοχές αναμένεται να έχουν την παρακάτω μορφή :

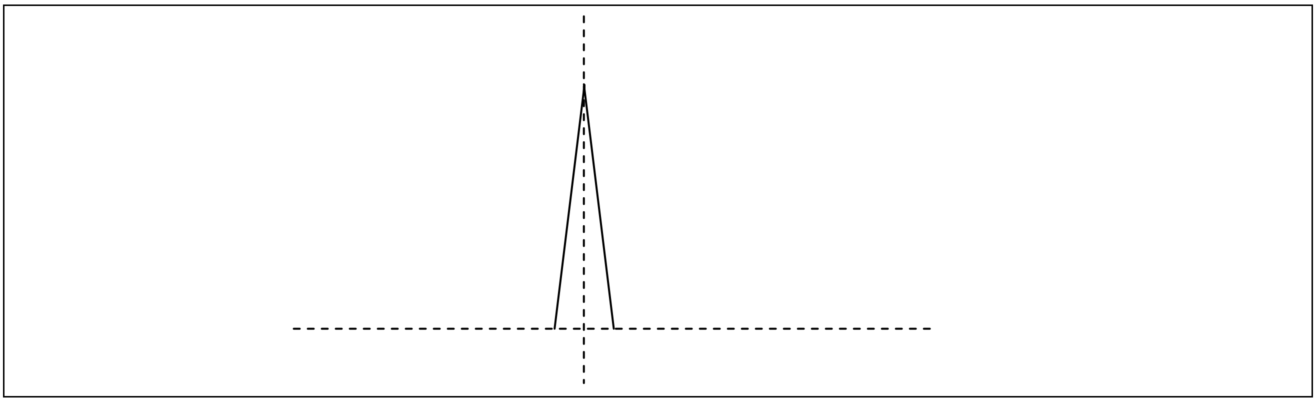


Γενικά, δεν είναι επιθυμητή μία τόσο έντονη περιοδική δομή γιατί αυτή προέρχεται από μεγάλη διάβρωση της περιοχής και που υποσκιάζει την ύπαρξη γράμματος.

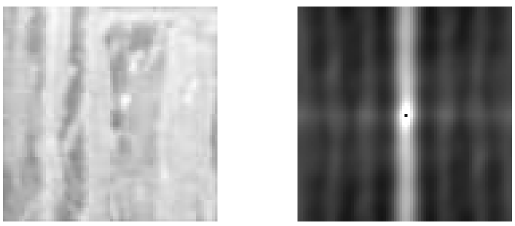
- *Random* εικόνες καταλαμβάνουν μικρή έκταση στο autocorrelation και έχουν μόνο πολύ μεγάλες τιμές στο κέντρο.



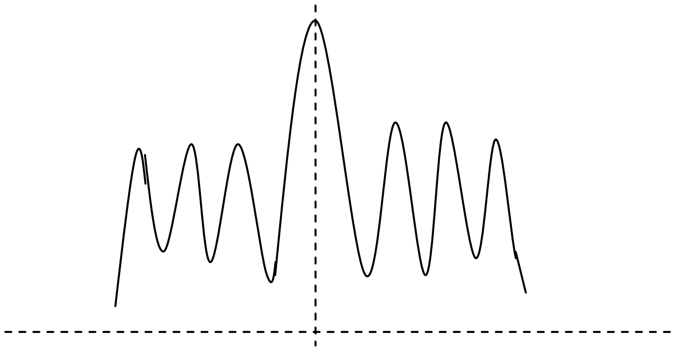
Τέτοιες περιοχές προσδοκάται να έχουν την παρακάτω μορφή :



- *Περιοδικές-Κατευθυντικές (periodic-directional)* εικόνες παρουσιάζουν κατευθυντικές και μεγάλου εύρους δομές στο autocorrelation:

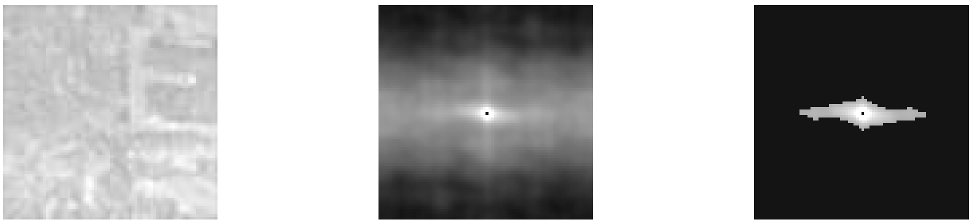


Η μορφή της αυτοσυσχέτισης στον ένα άξονα θα έχει πάλι περιοδική μορφή όπως ακόλουθα, αλλά στον κάθετο άξονα η αυτοσυσχέτιση αναμένεται να είναι περίπου σταθερή:



Για να απομονωθεί καλύτερα το πρόβλημα ορίστηκε μία μέθοδος που διαχωρίζει με πιο ευέλικτο τρόπο τις διάφορες περιοχές που εξετάζονται και βασίζεται στην μορφή της αυτοσυσχέτισης.

Από τα χαρακτηριστικά που μπορούμε να εξάγουμε από την αυτοσυσχέτιση, επιλέγουμε αυτά που βασίζονται στο "**New Wold Model**" των **Liu** και **Picard** [2],[3]. Από την εικόνα της αυτοσυσχέτισης εξάγεται μία μικρή περιοχή που διατηρεί ένα μεγάλο μέρος της συνολικής ενέργειας της αυτοσυσχέτισης (η ενέργεια ορίζεται με το άθροισμα των μέτρων των τιμών της εικόνας αυτοσυσχέτισης) και την κατευθυντικότητα της. Διατηρεί ουσιαστικά τη σημαντική πληροφορία που περιέχει η συνάρτηση αυτοσυσχέτισης στο κέντρο της. Αυτή η μικρή περιοχή αναφέρεται στη συνέχεια με τον όρο «περιοχή συγκέντρωσης». Η μορφή της περιοχής συγκέντρωσης για μία υποπεριοχή φαίνεται στο *Σχήμα 5.*



(i) (ii) (iii)

Σχήμα 5: (i) μία υποπεριοχή της εικόνας , (ii) η αυτοσυσχέτισή της, (iii) περιοχή συγκέντρωσης

Αναλυτικά η επεξεργασία, στην δική μας εργασία, η επιλογή της «περιοχής συγκέντρωσης» (*Σχήμα 4*(iii)) γίνεται ακολουθώντας τα παρακάτω βήματα:

- Εφαρμόζεται *thresholding* επιλέγοντας μόνο ένα ποσοστό του εύρους τιμών των τιμών αυτοσυσχέτισης (πειραματικά threshold=50%)

- *Connected-Component Labeling*

Η εικόνα της αυτοσυσχέτισης σαρώνεται και τοποθετούνται ετικέτες (labels) σε όλες εκείνες τις συνδεδεμένες περιοχές που πληρούν το 8-connectivity (κοιτάζοντας δηλαδή στην 8x8 γειτονιά του κάθε pixel αν έστω και ένα pixel έχει την τιμή του label θεωρείται connected) αλλά με τιμές μεγαλύτερες του threshold που προαναφέρθηκε. Αποτέλεσμα σε αυτή τη φάση είναι να δημιουργηθούν διακριτές περιοχές (components) καθεμιά με διαφορετική ετικέτα. Επιλέγεται όμως εκείνη την περιοχή γύρω από το κεντρικό pixel που πληρεί το 8- connectivity.

Η περιοχή συγκέντρωσης που εξήχθη με τον παραπάνω τρόπο θα αποτελέσει ένα χρήσιμό εργαλείο για την επίτευξη δύο πολύ σημαντικών στόχων:

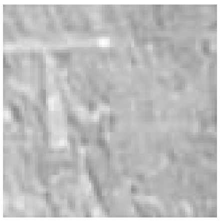
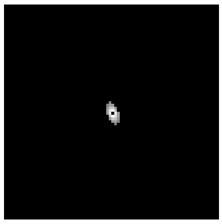
1. Για την κατηγοριοποίηση των υποπεριοχών σε σύνολα με κοινά χαρακτηριστικά και για την περαιτέρω περισσότερο λεπτομερή τους εξέταση.
2. Για την ανίχνευση περιοδικών χαρακτηριστικών που θα αναπτυχθούν σε επόμενο βήμα.

Αναλύοντας τον πρώτο στόχο εξετάσαμε πειραματικά τις περιοχές συγκέντρωσης ενός μεγάλου αριθμού υποεικόνων ενδιαφέροντος. Επιχειρώντας μία πρώτη κατηγοριοποίηση με βάση την περιοχή συγκέντρωσης , προτείνεται ο διαχωρισμός των περιοχών που αναλύεται παρακάτω:

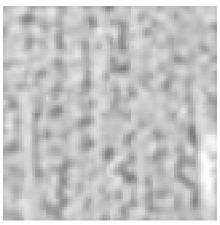
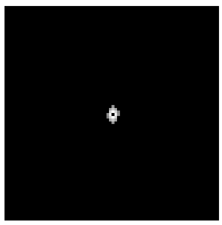
- *Υποπεριοχές τύπου 1*

Εικόνες με μικρή περιοχή συγκέντρωσης (αριθμός pixels περιοχής συγκέντρωσης<=105):

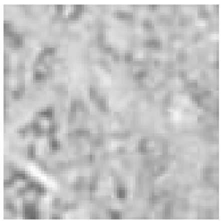
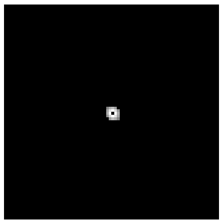
Περιέχουν ευκρινές γράμμα σε σχετικά ομαλό background π.χ:

ή είναι εικόνες με περιοδική μορφή που οφείλεται σε πορώδη διάβρωση π.χ.

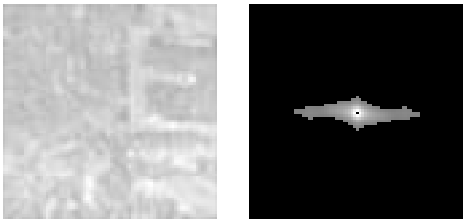



και τέλος είναι εικόνες με τυχαία δομή που οφείλεται στην ύπαρξη ακανόνιστου θορύβου π.χ.

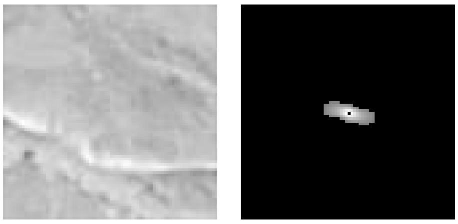



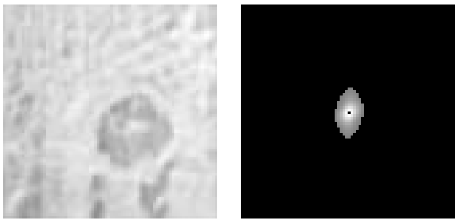
- *Υποπεριοχές τύπου 2*

Εικόνες με σχετικά μεγάλη περιοχή συγκέντρωσης (105<αριθμός pixels περιοχής συγκέντρωσης<=265): Είναι εικόνες με "θολή" δομή ή γράμμα σε ομαλό(smooth) background π.χ.



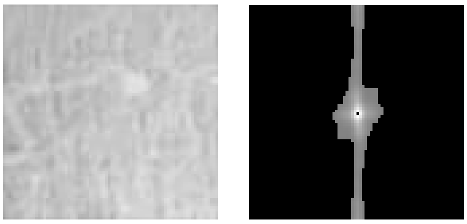
ή εικόνες με μεγάλου εύρους ("wide") δομή ή γράμμα π.χ.



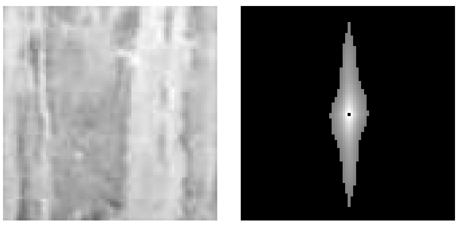


- *Υποπεριοχές τύπου 3*

Εικόνες με πολύ μεγάλη περιοχή συγκέντρωσης (αριθμός pixels περιοχής συγκέντρωσης>265): Είναι γενικά οι εικόνες με πολύ ομοιόμορφο background χωρίς κάποια συγκεκριμένη δομή ή γράμμα, όπως



ή εικόνες με πολύ έντονες επιμήκεις δομές που οφείλονται σε γραμμικής μορφής διάβρωση, όπως



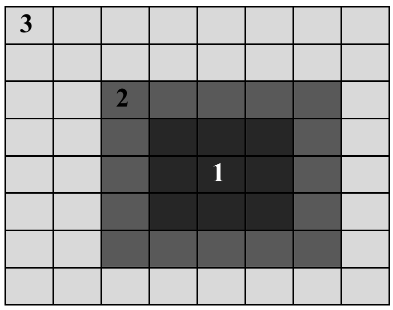
2.2.2 Χαρακτηριστικά που εξάγονται στο επίπεδο της συχνότητας

Η κατανομή της ενέργειας στο επίπεδο της συχνότητας μπορεί να προσδιορίσει την ποιότητα της υφής μιας εικόνας [5]. Χρησιμοποιείται ο διακριτός μετασχηματισμός Fourier $F(u,v)$ μιας εικόνας διαστάσεων $M \times N$ που ορίζεται από τη σχέση:

$$F(u,v) = \frac{1}{M \times N} \sum_{x=0}^{M} \sum_{y=0}^{N} exp(-2\pi iux / \mathrm{M}) \cdot exp(-2\pi iuy / N) \cdot f(x,y)$$

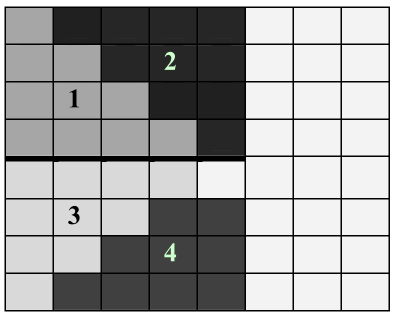
όπου u και v είναι οι διακριτές χωρικές συχνότητες, x και y δηλώνουν τις θέσεις των pixels .

Από την ακτινοειδή κατανομή του φάσματος ισχύος του Fourier Μετασχηματισμού $|F|^2$ εξάγονται τα πρώτα χαρακτηριστικά που έχουν τη μορφή **rings** και αποτελούν ένα κριτήριο για το βαθμό coarseness που υπάρχει στην εικόνα. Μία «τραχιά» υφή θα έχει τις μεγάλες τιμές του φάσματος ισχύος γύρω από το κέντρο ενώ μία πιο ομαλή εικόνα θα τις έχει διασκορπισμένες παντού. Για αυτό το λόγο η κάθε υποπεριοχή διαιρείται σε 8x8 image segments και ύστερα υπολογίζεται η μέση τιμή των τιμών του φάσματος ισχύος σε διάφορα rings που σχηματίζονται γύρω από το κέντρο :



Σχήμα 3: **Rings**

Παρόμοια, η γωνιακή κατανομή των τιμών του power spectrum είναι ένα κριτήριο για την κατευθυντικότητα της υφής της εικόνας. Για αυτό το λόγο υπολογίζεται η μέση τιμή των τιμών του φάσματος ισχύος σε διάφορα **wedges** που σχηματίζονται με αυτόν τον τρόπο:



Σχήμα 4: **Wedges**

Από ότι παρατηρείται στο *Σχήμα 4* μόνο το ένα ημιεπίπεδο χρησιμοποιείται για τα wedges εξαιτίας του μεγάλου βαθμού συμμετρίας του φάσματος. Η (0,0) συχνότητα δεν συμπεριλαμβάνεται στους υπολογισμούς των wedges εφόσον η τιμή της ήταν αισθητά μεγαλύτερη από άλλες τιμές του φάσματος και δεν θα έπρεπε να επηρεάσει την συνολική τιμή που εξάγεται από το συγκεκριμένο wedge.

2.3 Περιοδικότητα της συνάρτησης Αυτοσυσχέτισης

Ένα άλλο χαρακτηριστικό πολλών περιοχών που περιέχουν διάβρωση μαρμάρου είναι η περιοδικότητα. Με τη μελέτη τέτοιων χαρακτηριστικών, ανιχνεύεται το μέγεθος της περιοδικότητας που εμφανίζεται σε κάθε υποπεριοχή. Η περιοδικότητα μιας εικόνας μεταφέρεται και στη συνάρτηση αυτοσυσχέτισής της. Στην εργασία αυτή επιχειρούμε ανίχνευση περιοδικών χαρακτηριστικών στη συνάρτηση αυτοσυσχέτισης η οποία , λόγω της εξομάλυνσης που επιτυγχάνει σαν smoothing operator, δεν περιέχει πολύ απότομες μεταβολές που οφείλονται για παράδειγμα σε θόρυβο.

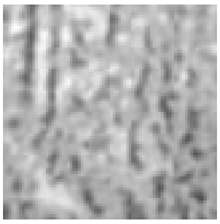
Στην προσπάθεια να επιτευχθεί αυτός ο στόχος υπολογίζεται ο Διακριτός Μετασχηματισμός Fourier (DFT) στην οριζόντια και κατακόρυφη κατεύθυνση του autocorrelation της κάθε εικόνας. Όμως ο απευθείας DFT του autocorrelation εμφανίζει απότομες εναλλαγές στις αρμονικές συχνότητες και έτσι δεν επιτρέπει τον εύκολο και γρήγορο προσδιορισμό τους.

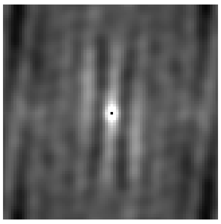
Ακολούθως υπολογίστηκε ο DFT πάνω στο αθροιστικό autocorrelation σε οριζόντια και κατακόρυφη κατεύθυνση. Δηλαδή αθροίζονται οι στήλες και οι γραμμές του autocorrelation κάθε εικόνας και εφαρμόζεται πάνω στο άθροισμά τους ο μετασχηματισμός Fourier. Με αυτόν τον τρόπο όμως διαπιστώνεται ότι η συνάρτηση αυτοσυσχέτισης χάνει τα τοπικά της χαρακτηριστικά γεγονός που συνεπάγεται απώλεια πληροφορίας.

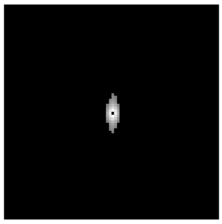
Πιο αποδοτικά αποτελέσματα εξάγονται κάνοντας ένα ειδικό φιλτράρισμα στο autocorrelation function. Συνοπτικά, η περιοδικότητα εξετάζεται σε μία ομαλοποιημένη μορφή της αυτοσυσχέτισης η οποία δημιουργείται φιλτράροντας την αρχική αυτοσυσχέτιση με ένα χαμηλοπερατό πυρήνα που διατηρεί τα κατευθυντικά χαρακτηριστικά της. Η διαδικασία που ακολουθείται και αποτελεί ένα από τα πιο χρήσιμα εργαλεία της εργασίας είναι η εξής :

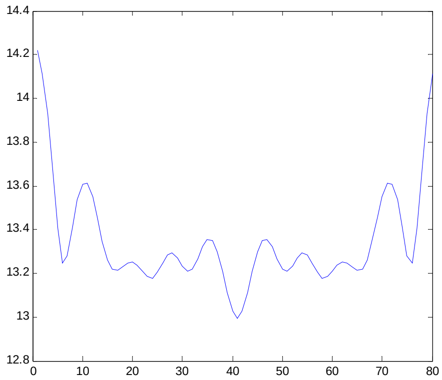
Ύστερα από τον προσδιορισμό της περιοχής συγκέντρωσης (§2.2.1), η εικόνα αυτοσυσχέτισης φιλτράρεται με την περιοχή συγκέντρωσης, παίρνοντας το **cross-correlation** μεταξύ των δύο αυτών εικόνων. Η περιοχή συγκέντρωσης παρουσιάζει την ιδιότητα να διατηρεί την κεντρική δομή της συνάρτησης αυτοσυσχέτισης με αποτέλεσμα να αποτελεί φίλτρο το οποίο επιτελεί προσαρμοστική κατευθυντική εξομάλυνση στη συνάρτηση αυτοσυσχέτισης. Ουσιαστικά, η λειτουργία αυτού του φίλτρου διατηρεί τη θέση και τη συχνότητα επανάληψης της κεντρικής δομής στην εικόνα της αυτοσυσχέτισης. Με αυτόν τον τρόπο λοιπόν είναι δυνατόν να ανιχνευθεί η περιοδικότητα που μπορεί να εμφανίζεται σε διάφορες υποπεριοχές. Προς αυτήν την κατεύθυνση εφαρμόζεται DFT στην οριζόντια και κατακόρυφη κατεύθυνση της φιλτραρισμένης εικόνας συσχέτισης, όπου υπολογίζεται το «συχνοτικό» περιεχόμενο της.

Για παράδειγμα στο *Σχήμα 6* φαίνονται το autocorrelation 6(ii) μίας υποπεριοχής 6(i), η μικρή περιοχή ενέργειας 6(iii) και στη συνέχεια το cross-correlation στο πεδίο του χώρου στην οριζόντια κατεύθυνση 6(iv) και το αντίστοιχο συχνοτικό περιεχόμενο 6(v) αφού εφαρμόστηκε δηλαδή DFT στο correlation του 6(iv). Τα γραφήματα 6(vi) και 6(vii) αναπαριστάνουν το cross-correlation στην κατακόρυφη κατεύθυνση για το πεδίο του χώρου και της συχνότητας αντίστοιχα.

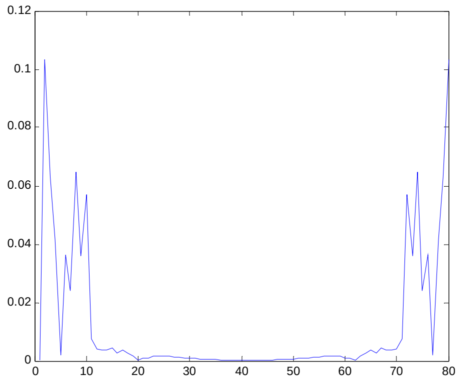

(i)


(ii)

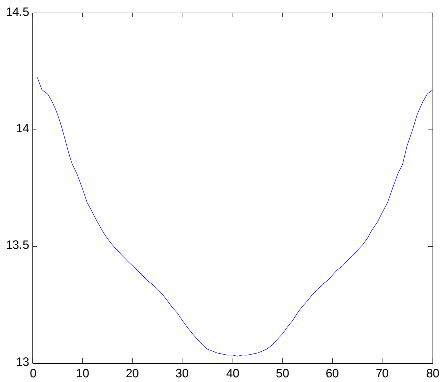

(iii)



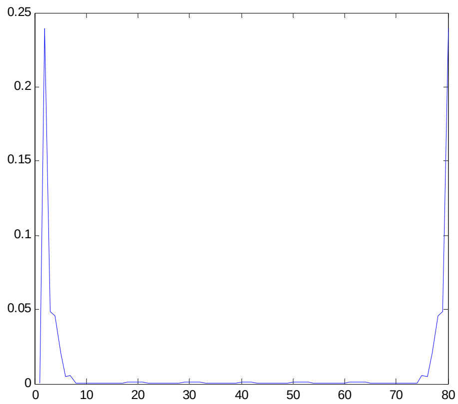
(iv)



(v)



(vi)

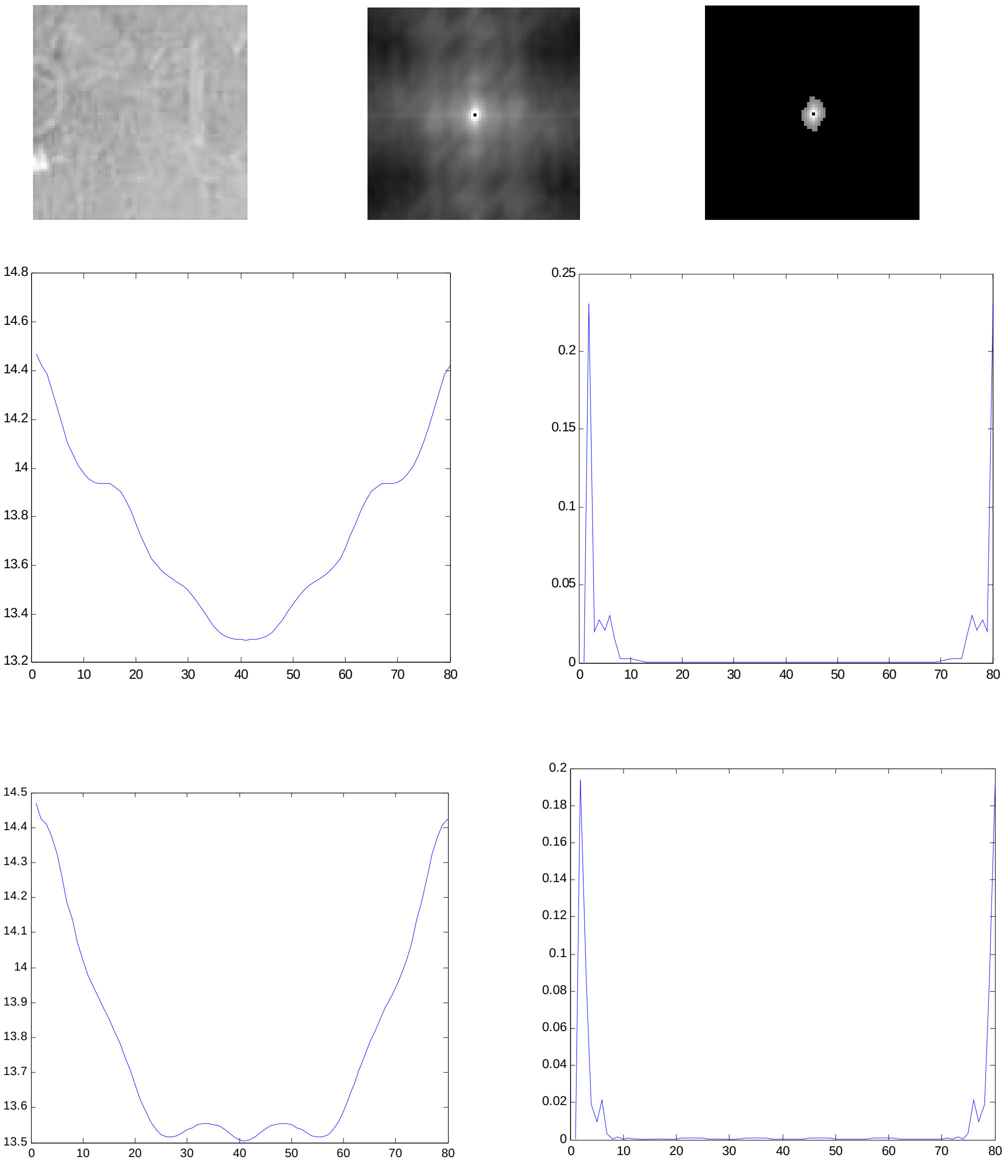


(vii)

Σχήμα 6

Από ότι παρατηρείται η συγκεκριμένη εικόνα εμφανίζει μεγάλη περιοδικότητα στις αρμονικές συχνότητες στην οριζόντια κατεύθυνση.

Στο *Σχήμα 7* παρουσιάζεται μία εικόνα που δεν εμφανίζει μεγάλη περιοδικότητα στη μορφή της και μάλιστα περιέχει κάποιους «αμυδρούς» χαρακτήρες. Στο σχήμα αυτό φαίνονται το autocorrelation της εικόνας, η μικρή περιοχή ενέργειας και τα γραφήματα του cross-correlation στο πεδίο του χρόνου και στο πεδίο της συχνότητας στους άξονες x και y αντίστοιχα.



Σχήμα 7

2.4 Κατευθυντικά Χαρακτηριστικά (Directional Features)

2.4.1 Tamura Features-Αριθμός σημείων ακμών

Για την ανίχνευση του πλήθους των ακμών και την κατευθυντικότητά τους χρησιμοποιούνται χαρακτηριστικά τα οποία έχουν την ιδιότητα να εντοπίζουν τον αριθμό σημείων που ανήκουν σε ακμές διαφόρων κατευθύνσεων. Πιο συγκεκριμένα στην αναφορά [6] προτείνεται μία γρήγορη μέθοδος για ανάλυση της υφής, βασισμένη σε δύο μόνο σαρώσεις της εικόνας κατά την οριζόντια και την κατακόρυφη κατεύθυνση. Το μέτρο |**ΔG**| των ακμών του texture της εικόνας και η γωνιακή κατεύθυνση θ της ακμής ορίζονται ως εξής:

$$|\Delta G| = (|\Delta_H| + |\Delta_V|)/2 \qquad (1)$$

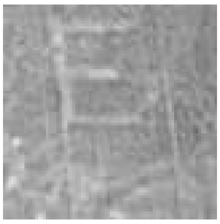
$$\theta = \tan^{-1}(\Delta_V / \Delta_H) + \frac{\pi}{2} \qquad (2)$$

όπου $\mathbf{\Delta_H}$ και Δ_V είναι οι εικόνες που προκύπτουν αν φιλτραριστούν με τις δυο παρακάτω μάσκες αντίστοιχα:

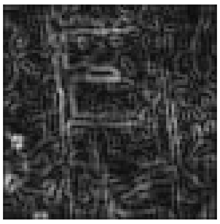
$$V = \begin{matrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{matrix} \qquad\qquad H = \begin{matrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{matrix}$$

Για την ακρίβεια η μάσκα V τελεί κατακόρυφη ανίχνευση ακμών ενώ η H οριζόντια ανίχνευση ακμών στις διάφορες υποπεριοχές.
Στη συγκεκριμένη εργασία η διαδικασία ανίχνευσης ακμών γίνεται ως εξής: Σαρώνονται συγχρόνως οι κατευθυντικές εικόνες $\mathbf{\Delta_H}$ και Δ_V βρίσκοντας για κάθε pixel της εικόνας το αντίστοιχο μέτρο |**ΔG**| και την κατευθυντικότητά του από τις σχέσεις (1) και (2). Γίνεται quantization της γωνιακής κατεύθυνσης θ σε τέσσερις κατευθύνσεις $0°$, $90°$, $45°$ και $135°$ και για την κάθε κατεύθυνση υπολογίζεται ο αριθμός των pixels του **ΔG** που αντιστοιχούν σε αυτήν την κατεύθυνση και έχουν μέτρο πάνω από ένα threshold (τιμή έντασης > 77), το οποίο θεωρούμε ως ένδειξη ύπαρξης ακμής.

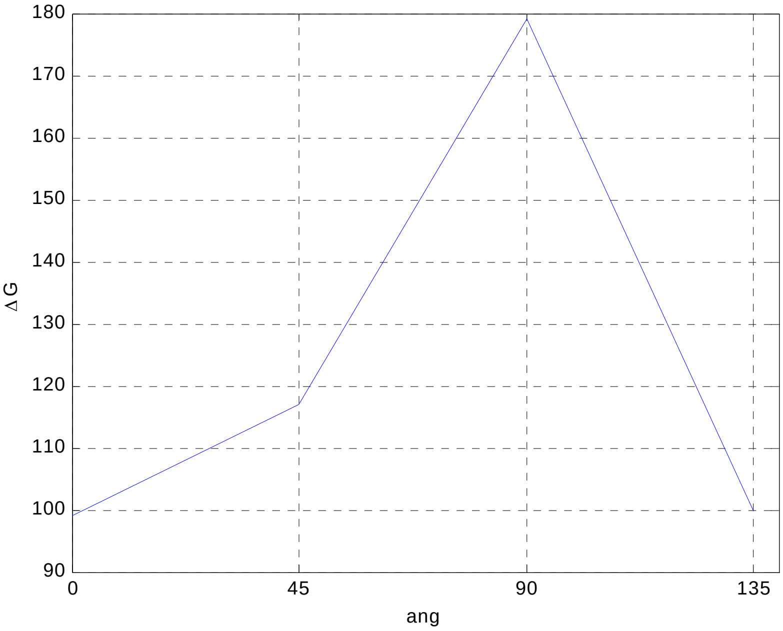
Στο *Σχήμα 8* παραθέτουμε κάποια παραδείγματα υποπεριοχών στις οποίες έχει γίνει ανίχνευση ακμών και υπολογισμός του αριθμού των pixels που ανήκουν σε ακμές διαφόρων κατευθύνσεων. Αναλυτικά, στα σχήματα *8* (ii) και *8*(v) φαίνεται η εικόνα ΔG (edge detected image) που προκύπτει αν εφαρμόσουμε τη σχέση (1) σε όλα τα pixels των κατευθυντικών εικόνων Δ_V και $\mathbf{\Delta_H}$. Στα σχήματα *8*(iii) και 8(vi) παρουσιάζονται τα γραφήματά στα οποία φαίνεται το πλήθος των σημείων του ΔG που βρίσκονται πάνω στις ακμές που αντιστοιχούν στις κατευθύνσεις $0°$, $90°$, $45°$ και $135°$. Παρατηρείται ότι στην πρώτη εικόνα επικρατούν οι κάθετες ακμές εξαιτίας της έντονης κάθετης ακμής του γράμματος καθώς και κάποιων κάθετων γραμμώσεων της εικόνας.



(i)

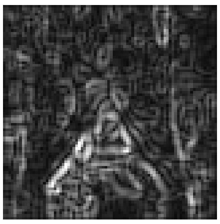


(ii)

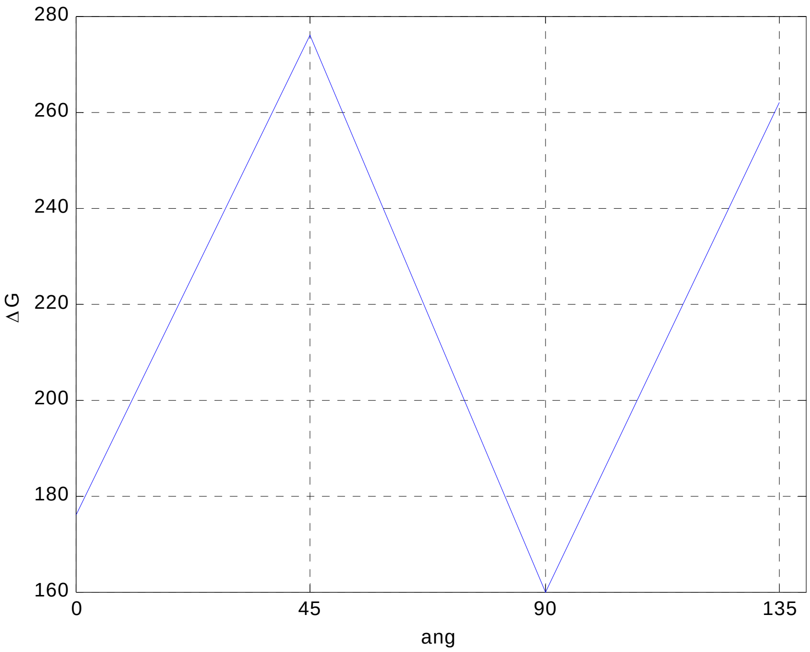


(iii)

(iv)



(v)



(vi)

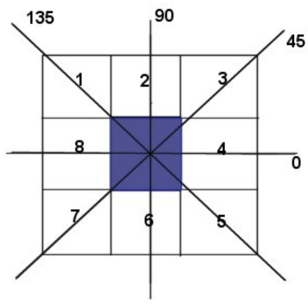
Σχήμα 8

Στη δεύτερη υποπεριοχή (*Σχήμα 8*(iv)) παρατηρείται μεγάλος αριθμός pixels του ΔG στις $45°$ και στις $135°$ μοίρες. Αυτό είναι επόμενο αφού στην εικόνα υπάρχει γράμμα με έντονες διαγώνιες ακμές καθώς και θόρυβος σε διαγώνια κατεύθυνση.

2.4.2 Χαρακτηριστικά Haralick

Τα χαρακτηριστικά Haralick είναι μεγέθη που μετρούν ποιοτικά χαρακτηριστικά των μεταβάσεων της έντασης σε μία κατεύθυνση. Σύμφωνα με τους Haralick, Shanmugam and Dinstein [7] η διαδικασία που χρησιμοποιείται για να εξαχθούν χαρακτηριστικά texture από μια εικόνα βασίζεται στην υπόθεση ότι η πληροφορία του texture μιας εικόνας περιέχεται στη συνολική χωρική σχέση που έχουν οι τιμές των pixels μεταξύ τους. Αυτή η πληροφορία για την υφή της εικόνας ορίζεται επαρκώς από ένα σύνολο **Gray-tone Spatial-Dependence(GTSD)** πινάκων που υπολογίζονται για διάφορες γωνιακές σχέσεις αλλά και για διάφορες αποστάσεις μεταξύ γειτονικών εικονοστοιχείων.

Πιο συγκεκριμένα για να γίνει σαφές τι είναι οι γωνιακές σχέσεις και αποστάσεις μεταξύ των pixels ας θεωρήσουμε ένα pixel που δεν βρίσκεται στην περιφέρεια της εικόνας. Το pixel (χρωματιστό κελί) με τους 8 κοντινότερους γείτονες του (που βρίσκονται σε απόσταση 1) καθώς και οι γωνιακές σχέσεις που έχει με τα pixels της γειτονιάς του φαίνονται στο *Σχήμα 9*:



4& 8, HORIZONTAL NEIGHBOURS
2 & 6, VERTICAL NEIGHBOURS
3 & 7, RIGHT DIAGONAL NEIGHBOURS
1 & 5, LEFT DIAGONAL NEIGHBOURS

Σχήμα 9

Ως Gray-Tone Spatial-Dependence Matrix (που καλείται και Gray-Level Co-occurrence Matrix) ορίζεται ο πίνακας των σχετικών συχνοτήτων p(i,j) με τις οποίες δυο γειτονικά pixels που απέχουν μεταξύ τους απόσταση d, το ένα με ένταση i και το άλλο με ένταση j, εμφανίζονται στην εικόνα.
Αυτός ο πίνακας λοιπόν που περιέχει gray-tone spatial-dependence συχνότητες είναι τόσο συνάρτηση της γωνιακής σχέσης των γειτονικών pixels όσο και της απόστασης που έχουν μεταξύ τους.

Παράδειγμα

Θεωρούμε την εικόνα (*Σχήμα 10*(i)) με διαστάσεις 4x4 με τέσσερα επίπεδα έντασης που κυμαίνονται με τιμές από 0 έως 3. Το *Σχήμα 10*(ii) δείχνει τη γενική μορφή του GTSD πίνακα όπου #(i,j) δηλώνει αριθμό μεταβάσεων από ένταση i σε j στην κατεύθυνση και γειτονιά που εξετάζουμε. Οι υπόλοιποι πίνακες του *Σχήματος 10*(iii) είναι οι τέσσερις πίνακες που υπολογίστηκαν οι gray-tone spatial-dependence συχνότητες για απόσταση 1 και γωνιακές σχέσεις $0°,90°,135°,45°$.

0	0	1	1
0	0	1	1
0	2	2	2
2	2	3	3

(i)

		Gray Tones			
		0	**1**	**2**	**3**
Gray Tones	**0**	#(0,0)	#(0,1)	#(0,2)	#(0,3)
	1	#(1,0)	#(1,1)	#(1,2)	#(1,3)
	2	#(2,0)	#(2,1)	#(2,2)	#(2,3)
	3	#(3,0)	#(3,1)	#(3,2)	#(3,3)

(ii)

4	2	1	0
2	4	0	0
1	0	6	1
0	0	1	2

$P_{Horizontal}(0°)$

6	0	2	0
0	4	2	0
2	2	2	2
0	0	2	0

$P_{Vertical}\ (90°)$

2	1	3	0
1	2	1	0
3	1	0	2
0	0	2	0

$P_{LeftDiagonal}(135°)$

2	1	3	0
1	2	1	0
3	1	0	2
0	0	2	0

$P_{RightDiagonal}(45°)$

(iii)

Σχήμα 10

Για παράδειγμα , το στοιχείο στη (2,1) θέση του $P_{Horizontal}$ δείχνει συνολικά πόσες φορές ένα pixel με τιμή έντασης 2 βρίσκεται αριστερά ή δεξιά, σε απόσταση 1, ενός γειτονικού pixel με τιμή έντασης 1, μέσα στην εικόνα. Τον αριθμό αυτόν τον υπολογίζουμε ελέγχοντας όλους τους συνδυασμούς των στοιχείων της εικόνας, π.χ. {(1,1),(1,2)}, {(1,2),(1,1)} , {(1,2),(1,3)}, {(1,3),(1,2)} κ.τ.λ. και αθροίζουμε πόσες φορές συναντήσαμε ζευγάρι τιμών (2,1) . Την ίδια διαδικασία ακολουθούμε για να υπολογίσουμε και τα υπόλοιπα στοιχεία του $P_{Horizontal}$. Για να υπολογίσουμε τον $P_{Vertical}$ αυτό που αλλάζει είναι ότι ελέγχουμε τους συνδυασμούς των ζευγαριών στον vertical άξονα π.χ. {(1,1),(2,1)}, {(2,1),(1,1)}... , τον $P_{LeftDiagonal}$ στις αριστερές διαγώνιους του πίνακα π.χ. {(1,1),(2,2)}, {(2,2),(1,1)}, {(2,1),(3,2)}, {(3,2),(2,1)}... και τον $P_{RightDiagonal}$ στις δεξιές διαγώνιους π.χ. {(1,2),(2,1)}, {(2,1),(1,2)}, {(3,1),(2,2)}, {(2,2),(3,1)}...

Αφού υπολογιστούν οι τέσσερις Gray-Tone Spatial-Dependence πίνακες, κανονικοποιούνται με τον εξής τρόπο: Σε κάθε πίνακα ξεχωριστά υπολογίζεται το άθροισμα όλων των στοιχείων του και ύστερα διαιρείται κάθε στοιχείο του με τον αριθμό αυτό.

Όπως αναφέρθηκε αρχικά, η πληροφορία για την υφή βρίσκεται στους Gray-Tone Spatial-Dependence Matrices. Σαν περιγραφικά μεγέθη της πληροφορίας αυτής εξάγουμε τρία χαρακτηριστικά υφής που είναι τα εξής: **Angular Second Moment, Contrast και Correlation** [7], [8].

1. **GTSD Angular Second Moment (ASM):**

$$f1 = \sum_{i=1} \sum_{j=1} (p(i,j))^2 \qquad (1)$$

με p(i,j) κάθε στοιχείο του gray-tone spatial-dependence (GTSD) πίνακα.
Χρησιμοποιεί δηλαδή την ενέργεια κάθε στοιχείου του GTSD πίνακα για να υπολογιστεί.

Φυσική Σημασία

Αποτελεί μέτρο ομοιογένειας των μεταβολών της έντασης σε κατευθυντικές γειτονιές της εικόνας. Σε μια ομοιογενή εικόνα εμφανίζονται λίγες επικρατούσες τονικές μεταβάσεις με αποτέλεσμα ο GTSD πίνακας να περιέχει λίγα στοιχεία με μεγάλο μέτρο. Αντίθετα σε μία μη ομοιογενή εικόνα υπάρχουν πολλές διαφορετικές μεταβάσεις στην ένταση των εικονοστοιχείων οπότε ο Gray-Level Co-occurrence πίνακας περιέχει μεγαλύτερο αριθμό στοιχείων με μικρά όμως μέτρα. Έτσι σε αυτή την περίπτωση το ASM χαρακτηριστικό (που είναι το άθροισμα των τετραγώνων των στοιχείων του GTSD πίνακα) είναι σχετικά μικρό.

2. **GTSD Contrast:**

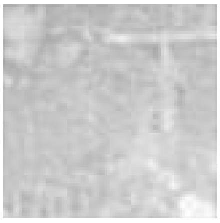
$$f2 = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i,j)(i-j)^2 \qquad (2)$$

όπου N_g = αριθμός των διακριτών gray levels και i, j οι τιμές των διακριτών gray levels της εικόνας.

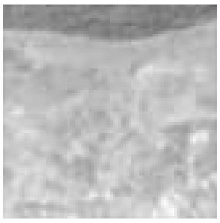
Φυσική Σημασία

Δηλώνει την ομοιομορφία γειτονικών καταστάσεων (Similarity of States). Αποτελεί μέτρο της συχνότητας των μεταβάσεων έντασης της εικόνας και του μεγέθους των μεταβάσεων αυτών.

Όταν σε μια εικόνα υπάρχει ομοιομορφία στις τιμές των στοιχείων της, δηλαδή δεν υπάρχουν πολλές και μεγάλες διαφοροποιήσεις στην ένταση (όπως στην *εικόνα 2*) ή όταν έχουμε μεγάλες συγκεντρωμένες περιοχές έντασης με λίγες απότομες μεταβάσεις από περιοχή σε περιοχή (*εικόνα 3*), ο GTSD πίνακας θα έχει τις περισσότερες μεγάλες τιμές του πάνω (ή κοντά) στη διαγώνιο όπου η διαφορά (i – j) είναι μηδέν (ή πολύ μικρή) και διάσπαρτα λίγες μέτριες τιμές. Παρατηρώντας τη σχέση (2) οι όροι p(i,j) και (i – j) θα παίρνουν μικρές τιμές οπότε η εικόνα θα έχει μικρό contrast.

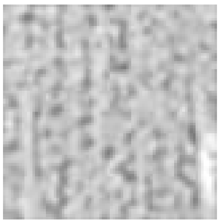


εικόνα 2

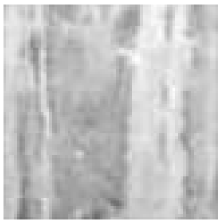


εικόνα 3

Για πολλές μικρές περιοχές όμοιας έντασης αλλά διάσπαρτες στην έκταση της εικόνας (όπως στην *εικόνα 4*) δηλαδή περιοχές όμοιων αλλά πολλών μεταβάσεων, ο GTSD πίνακας θα έχει μακριά από τη διαγώνιό του διάσπαρτα λίγες αλλά μεγάλες τιμές (μεγάλο p(i , j)) που τείνουν να αυξάνουν το contrast. Το contrast αυξάνεται επίσης όταν μία εικόνα εμφανίζει συγκεντρωμένες περιοχές έντασης αλλά περιέχει πολλές και απότομες μεταβάσεις από τη μία περιοχή στην άλλη (π.χ. *εικόνα 5*).



εικόνα 4



εικόνα 5

Ποιοτικά επομένως το contrast αυξάνει είτε λόγω μεγάλων διαφοροποιήσεων της έντασης (μεγάλο $(i-j)^2$), είτε λόγω πολλών μεταβάσεων (μεγάλο p(i,j) εκτός διαγωνίου).

3. GTSD Correlation

$$f3 = \frac{\left(\sum_i \sum_j i * j * p(i,j)\right) - \mu_x * \mu_y}{\sigma_x * \sigma_y} \qquad (3)$$

με μ_x, μ_y η μέση τιμή και σ_x , σ_y το standard deviation του p_x , p_y πίνακα αντίστοιχα.

Ο p_x πίνακας προκύπτει αν αθροιστούν οι γραμμές του GTSD : $p_x(i) = \sum_{j=1}^{N_g} P(i,j)$

Αντίστοιχα, ο p_y προκύπτει αν αθροιστούν οι στήλες του GTSD : $p_y(j) = \sum_{i=1}^{N_g} P(i,j)$

Επίσης $\mu_x = \sum_{i=0}^{N_g-1} i * p_x(i)$, $\qquad \mu_y = \sum_{j=0}^{N_g-1} j * p_x(j)$

και $\sigma_x = \sqrt{\sum_{i=0}^{N_g-1} (i-\mu_x)^2 * p_x(i)}$, $\sigma_y = \sqrt{\sum_{j=0}^{N_g-1} (j-\mu_y)^2 * p_y(j)}$

Φ υ σ ι κ ή Σ η μ α σ ί α

Γενικά μεγάλο correlation ενός σήματος δηλώνει ότι υπάρχει μία προβλέψιμη και γραμμική σχέση μεταξύ των δειγμάτων του. Αν για παράδειγμα υπάρχει μία πολύ μεγάλη τιμή συσχέτισης μεταξύ ενός pixel αναφοράς και του γειτονικού του pixel, που εκφράζεται με τη σχέση n=2r+2, όπου n είναι η τιμή του γειτονικού pixel και r του σημείου αναφοράς, αν r=1 τότε το γειτονικό pixel αναμένεται να έχει τιμή n=4, αν το r=4 τότε n=10. Μεγάλο correlation λοιπόν σημαίνει μεγάλη προβλεψιμότητα για τη σχέση που έχουν τα pixels μεταξύ τους. Σε αυτό το σημείο γίνεται αναφορά σε κάποιες **ιδιότητες** του correlation :

- Ένα αντικείμενο από μόνο του εμφανίζει μεγαλύτερο correlation από ότι δύο γειτονικά αντικείμενα μεταξύ τους
- Τα pixels είναι περισσότερα συσχετισμένα με αυτά που βρίσκονται στη γειτονιά τους από ότι με αυτά τα pixels που είναι πιο απομακρυσμένα.
- Όσο πιο μεγάλο παράθυρο χρησιμοποιείται για να υπολογιστεί το correlation τόσο μικραίνει η τιμή του correlation σε αυτό το παράθυρο.

Γενικότερα στο επίπεδο της εικόνας το correlation εκφράζει την ομοιότητα στον τρόπο μεταβολής της έντασης σε μία κατεύθυνση. Πιο συγκεκριμένα εκφράζει τη συχνότητα αυξομειώσεων της έντασης σε μία κατεύθυνση και δεν εξαρτάται τόσο από τις τιμές των εντάσεων μεμονωμένα. Σταθερή ή γραμμικά μεταβαλλόμενη ένταση δίνει μεγάλο correlation , ενώ τακτικές περιοδικές ή τυχαίες μεταβολές έντασης καταλήγουν σε μικρό ή ακόμα και αρνητικό correlation. Εμβαθύνοντας στην έννοια του correlation, παίρνουμε την ακραία περίπτωση (*εικόνα 6*) μιας εικόνας που αν τη σαρώσουμε κάθετα συναντάμε την ίδια τιμή έντασης ενώ αν τη σαρώσουμε οριζόντια η τιμή της έντασης κάθε φορά θα μεταβάλλεται γραμμικά κατά μία τιμή (από 0 έως 255 τιμή έντασης). Ο GTSD πίνακας στην πρώτη περίπτωση θα έχει όλα τα στοιχεία του πάνω στη διαγώνιο και στα υπόλοιπα σημεία μηδέν. Αν N είναι το άθροισμα των στοιχείων του GTSD τότε πάνω στη διαγώνιο θα έχουμε στοιχεία-πιθανότητες 1/N. Αν τώρα σαρώσουμε την εικόνα οριζόντια οι πιθανότητες 1/N δε θα βρίσκονται πάνω στη διαγώνιο αλλά στην πιο κοντινή παράλληλο της $(\mathrm{i} - \mathrm{j} = 1)$.

εικόνα 6

Όταν τα στοιχεία του GTSD βρίσκονται στη διαγώνιο, το correlation σύμφωνα με τη σχέση (3) γίνεται:

$$f3 = \frac{\left(\frac{\sum i \cdot i}{N}\right) - \frac{\sum i}{N} \cdot \frac{\sum i}{N}}{\sigma_x \cdot \sigma_y} = \frac{\frac{N\sum i^2 - (\sum i)^2}{N^2}}{\sigma_x \cdot \sigma_y}$$

Παρατηρούμε ότι ο αριθμητής είναι αρκετά μεγάλος άρα το correlation θα είναι αυξημένο. Αλλά και στην περίπτωση που τα στοιχεία βρίσκονται στην off-διαγώνιο, το correlation όμοια διατηρεί μεγάλες τιμές.
Για την συγκεκριμένη εικόνα το correlation υπολογισμένο βάση του horizontal ($0°$) GTSD θα είναι 0.999 ενώ με βάση τον vertical ($90°$) GTSD είναι 1.

Αν πάρουμε την άκρως αντίθετη περίπτωση όπου η εικόνα εμφανίζει μεγάλη τυχαιότητα στις τιμές τις, στον GTSD πίνακα οι τιμές είναι ισοκατανεμημένες παντού. Οι πιθανότητες p_x, p_y θα είναι τώρα είναι ανεξάρτητες οπότε :

$$p_{xy}(i,j) = p_x(i) \cdot p_y(j)$$

και η σχέση (3) γίνεται :

$$f3 = \frac{\sum i \cdot p_x(i) \cdot \sum j \cdot p_y(j) - \mu_x \cdot \mu_y}{\sigma_x \cdot \sigma_y} = \frac{\mu_x \cdot \mu_y - \mu_x \cdot \mu_y}{\sigma_x \cdot \sigma_y} = 0$$

οπότε το correlation θα είναι μηδέν.

Συμπερασματικά, όταν τα στοιχεία της εικόνας είναι γενικά ομοιόμορφα κατανεμημένα ή μεταβάλλονται ομοιόμορφα σε κάποια κατευθυντικά μοτίβα (patterns) της εικόνας, τα στοιχεία του GTSD πίνακα περιορίζονται γύρω από τη διαγώνιο ή κάποια παράλληλο της διαγωνίου με αποτέλεσμα το correlation να αυξάνει. Αντίθετα τακτικές ή τυχαία εναλλασσόμενες δομές στην εικόνα οδηγούν σε μικρή τιμή correlation (ή ακόμα και αρνητική). Η τυχαιότητα στη δομή της εικόνας τείνει να μειώνει το correlation.

2.5 Ιδιότητες των Χαρακτηριστικών

Σε αυτό το στάδιο συζητάται η χρησιμότητα των τύπων χαρακτηριστικών που αναφέραμε έως αυτό το σημείο.

Έγινε αναφορά σε κάποια ολικά χαρακτηριστικά που είτε είναι χωρικά (**αυτοσυσχέτιση**) είτε συχνοτικού περιεχομένου (**Rings–Wedges**). Η πληροφορία που παρέχουν αυτά τα χαρακτηριστικά είναι γενική και καλύπτει όλο το φάσμα των υποπεριοχών της εικόνας .

- **Αυτοσυσχέτιση**

Τα χαρακτηριστικά αυτά όμως μπορούν να επιτελέσουν μία πρωταρχική κατηγοριοποίηση με βάση τη μορφή της αυτοσυσχέτισης και πιο συγκεκριμένα την ενεργεια της περιοχής συγκέντρωσης (§2.2.1), εφόσον παρατηρείται μία διαφοροποίηση της από υποπεριοχή σε υποπεριοχή. Αυτή η διαφορετική μορφή αυτοσυσχέτισης όπως έχει ήδη αναφερθεί προκύπτει από το γεγονός ότι οι διάφορες υποπεριοχές της εικόνας εμφανίζουν μία ποικιλία ως προς τη μορφή του background (περιοδικότητα, τυχαιότητα, κατευθυντικότητα, ομαλότητα) αλλά και την ευκρίνεια του γράμματος (γράμματα με ικανοποιητική ευκρίνεια, θολά γράμματα, σπασμένοι χαρακτήρες λόγω διάβρωσης).

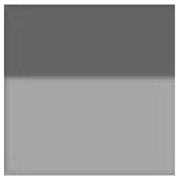
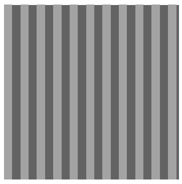
- **Περιοδικά Χαρακτηριστικά**

Όσον αφορά στα περιοδικά χαρακτηριστικά της §2.2.1, η χρησιμότητά τους επικεντρώνεται σε τύπους υποεικόνων που εμφανίζουν έντονα αδρά χαρακτηριστικά, όπως αυτές του τύπου 1 (§2.2.1). Σε αυτήν την περίπτωση, περιοδικότητα σημαίνει έντονη διάβρωση και πρέπει να απορριφθεί. Η χρήση αυτών των χαρακτηριστικών επικεντρώνεται στην περίπτωση που η επιφάνεια της περιοχής συγκέντρωσης είναι σχετικά μικρή έτσι ώστε να είναι δυνατή, φιλτράροντας τη συνάρτηση αυτοσυσχέτισης αυτών των περιοχών, η ανίχνευση περιοδικών χαρακτηριστικών.

- **Κατευθυντικά Χαρακτηριστικά**

Για τα κατευθυντικά χαρακτηριστικά, αυτό που πρέπει να επισημανθεί είναι ότι εξαιτίας της φυσικής τους σημασίας δεν έχουν όλα την ικανότητα να διαχωρίσουν όλους τους πιθανούς τύπους υποπεριοχών. Από την παρακάτω ανάλυση γίνεται φανερό ότι η χρήση των χαρακτηριστικών αυτών είναι πιο ουσιαστική στις εικόνες τύπου 2. Πιο αναλυτικά μπορούμε να εξάγουμε τα παρακάτω συμπεράσματα:

Η τιμή του κατευθυντικού χαρακτηριστικού Angular Second Moment (§2.4.2) δεν εξαρτάται κατά βάση από τις εντάσεις των εικονοστοιχείων αλλά από την ποικιλία και τη συχνότητα των γειτονικών καταστάσεων (*state transition density*). Για παράδειγμα, συγκρίνουμε την τιμή του ASM στην οριζόντια κατεύθυνση των δύο εικόνων που ακολουθούν. Και οι δύο εικόνες έχουν δύο ισοκατανεμημένες περιοχές έντασης.: η πρώτη παραμένει σταθερή στην οριζόντια κατεύθυνση, ενώ η δεύτερη αλλάζει τιμή ανά ένα pixel.

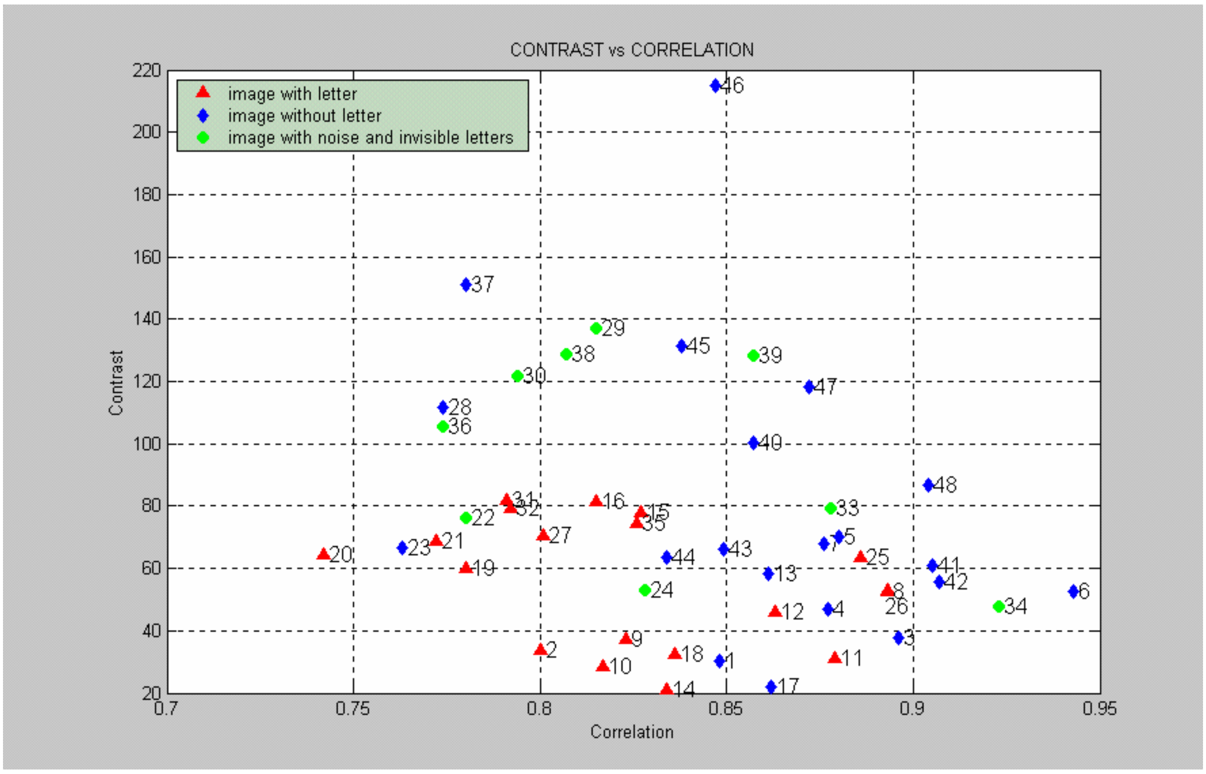
 

Σε αυτές τις εικόνες δεν επιτυγχάνεται κανένας διαχωρισμός μιας και η τιμή του ASM είναι η ίδια. Αυτό συμβαίνει διότι στην πρώτη εικόνα έχουμε σταθερή κατάσταση σε κάθε γραμμή αλλά συνολικά δύο καταστάσεις με την ίδια συχνότητα. Δύο καταστάσεις με την ίδια συχνότητα έχουμε και στη δεύτερη περίπτωση άσχετα αν εκεί έχουμε αλλαγή κατάστασης ανά 1 pixel. Συνεπώς το ASM δεν μπορεί να διαχωρίσει πληροφορία έντασης, αλλά μόνο πληροφορία καταστάσεων (states και state transition). Αυτό που παρατηρούμε επίσης είναι ότι το ASM αλλά και το Correlation δεν εξαρτώνται από τη διαφοροποίηση εντάσεων, στοιχείο που είναι πολύ σημαντικό για την ανίχνευση ακμών και γενικότερα δομών μέσα στην εικόνα. Το contrast αντίθετα, λαμβάνει υπόψη του τη συχνότητα των μεταβάσεων αλλά και τις διαφοροποιήσεις της έντασης πράγμα που το καθιστά σημαντικό κατευθυντικό χαρακτηριστικό.

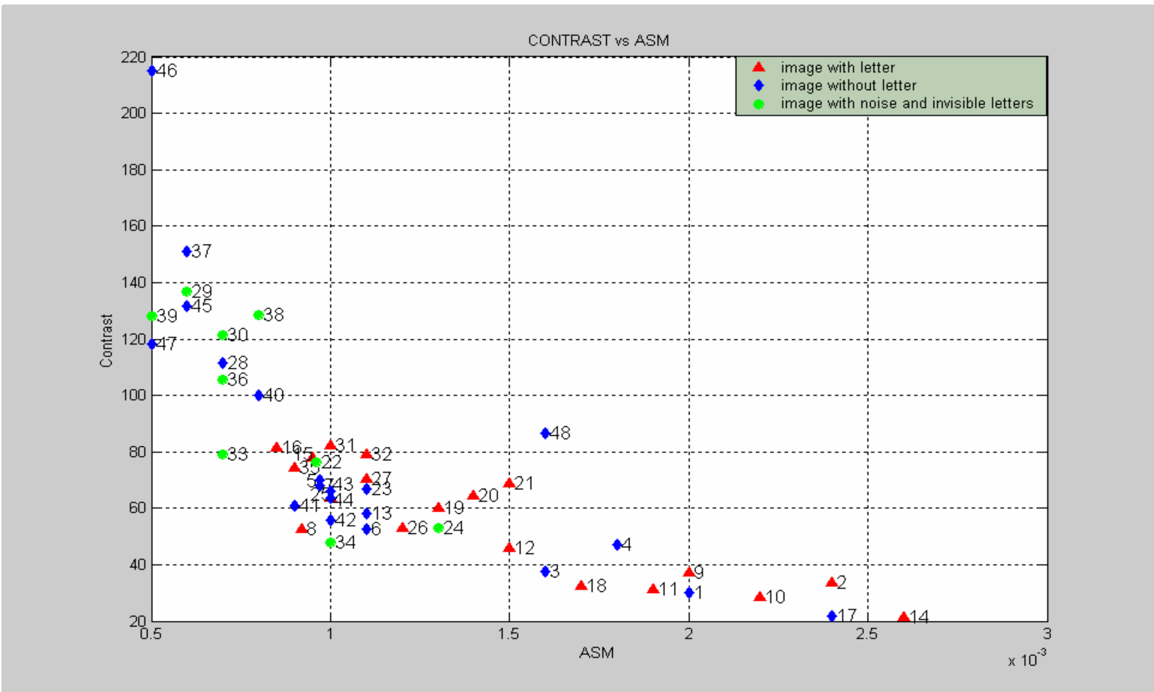
Έτσι λοιπόν τα χαρακτηριστικά Haralick έχουν μικρή αποτελεσματικότητα σε περιπτώσεις πολλών τυχαίων και απότομων μεταβάσεων στην περιοχή της υποεικόνας, μιας και εξαρτώνται περισσότερο από τον αριθμό των μεταβάσεων της περιοχής και όχι από τη δομή της. Μεγαλύτερη χρησιμότητα εντοπίζεται σε περιπτώσεις λιγότερων μεταβάσεων όπου μπορούν να διαχωρίσουν περιπτώσεις μεταβάσεων συγκεντρωμένων γύρω από τη διαγώνιο του πίνακα P από περιπτώσεις λίγων αλλά μεγάλων μεταβάσεων έντασης (μεγάλο contrast) που εντοπίζονται μακριά από τη διαγώνιο στον πίνακα P. Με βάση τα παραπάνω, η χρήση των χαρακτηριστικών Haralick είναι προτιμητέα σε υποεικόνες τύπου 2 (§2.2.1) παρά σε αυτές του τύπου 1.

Ενισχύοντας περισσότερο τις παραπάνω παρατηρήσεις, εξετάζεται η διαχωριστική ικανότητα των χαρακτηριστικών Haralick στο σύνολο των υποεικόνων. Τα τρία χαρακτηριστικά ASM, Contrast και Correlation αφού υπολογίζονται και στις τέσσερις διευθύνσεις $0°$, $90°$, $45°$ και $135°$ στη συνέχεια υπολογίζεται ο μέσος όρος των 4 τιμών του κάθε χαρακτηριστικού. Οι υποεικόνες κατατάσσονται ποιοτικά, με οπτική παρατήρηση σε τρεις κατηγορίες : Εικόνες με γράμμα, εικόνες χωρίς γράμμα και εικόνες με δυσδιάκριτο γράμμα που συμβολίζονται με κόκκινο τρίγωνο, μπλε ρόμβο και πράσινο κύκλο αντίστοιχα στα διαγράμματα που ακολουθούν.

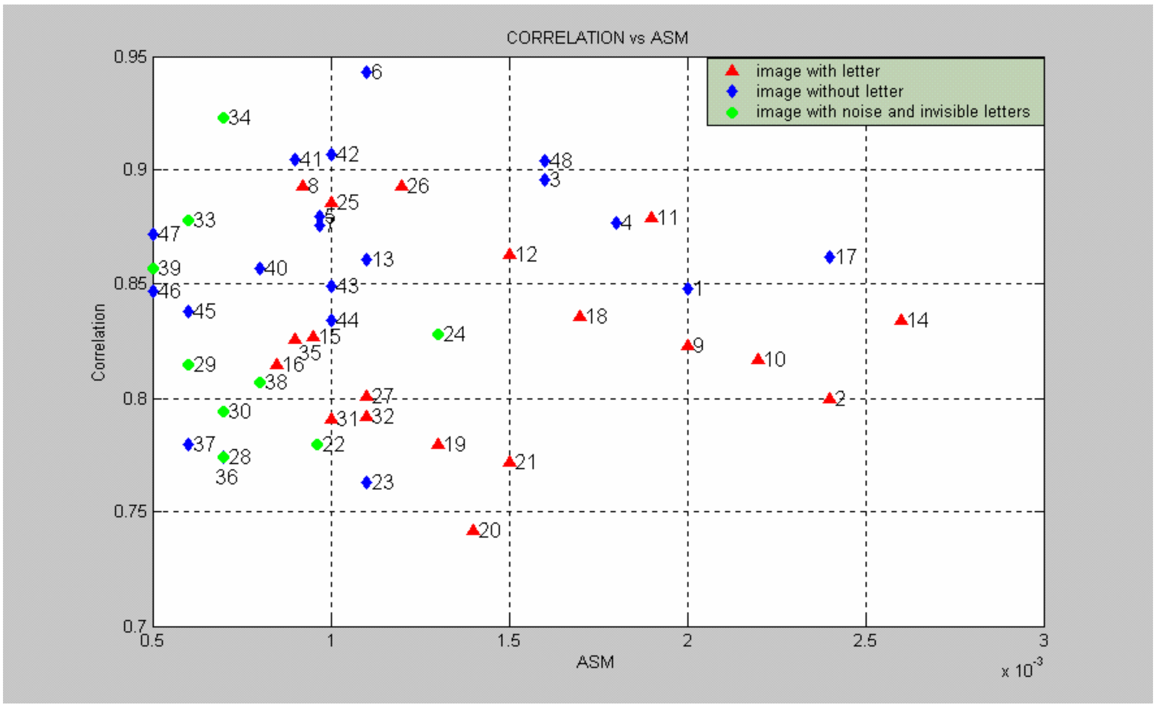
Οι τιμές των συνδυασμών contrast versus correlation, contrast versus ASM και correlation versus ASM παρουσιάζονται στα διαγράμματα 1,2 ,3 αντίστοιχα.



Διάγραμμα 1



Διάγραμμα 2



Διάγραμμα 3

Αμέσως παρατηρείται ότι υπάρχουν πολύ λίγα περιθώρια για σωστό διαχωρισμό του τύπου υποπεριοχής από τέτοιους συνδυασμούς χαρακτηριστικών. Πιο συγκεκριμένα, όλοι οι τύποι των υποπεριοχών εντοπίζονται διάσπαρτα στα διαγράμματα αυτά. Έτσι η εφαρμογή των χαρακτηριστικών Haralick σε όλο το φάσμα των υποπεριοχών δεν μπορεί να δώσει χρήσιμες πληροφορίες για κατάταξη. Στη συνέχεια επιχειρείται η εξέταση της συσχέτισης αυτών των χαρακτηριστικών για το σύνολο των υποεικόνων.

Παρατηρείται γενικά ότι δεν υπάρχει συσχέτιση μεταξύ contrast και correlation διότι για το ίδιο correlation δύο εικόνες μπορούν να έχουν τελείως διαφορετικό contrast. Αυτό εξηγείται διότι οι εικόνες αυτές μπορεί να έχουν την ίδια δομή (ίδιο correlation) αλλά με διαφορετικές εντάσεις, στοιχείο που επηρεάζει την τιμή του contrast. Επίσης δύο εικόνες με ίδια τιμή contrast μπορεί να έχουν τελείως διαφορετική δομή (συγκεντρωμένες ή τυχαία διάσπαρτες περιοχές) πράγμα που οδηγεί σε διαφορετική τιμή correlation.

Μεγαλύτερη συσχέτιση εμφανίζεται στο *Διάγραμμα 2* μεταξύ του contrast και του ASM. Αυτή η συσχέτιση ισχύει για τον συγκεκριμένο φωτισμό της εικόνας, αλλά δεν μπορεί να τεκμηριωθεί γενικά, μιας και ίδια δομή υποεικόνας (σταθερό ASM) κάτω από διαφορετικές συνθήκες φωτισμού μπορεί να εμφανίζεται με τελείως διαφορετικές εντάσεις (εντελώς διαφορετικό contrast). Ίδια συμπεράσματα εξάγονται και από το *Διάγραμμα 3* μεταξύ correlation και ASM.

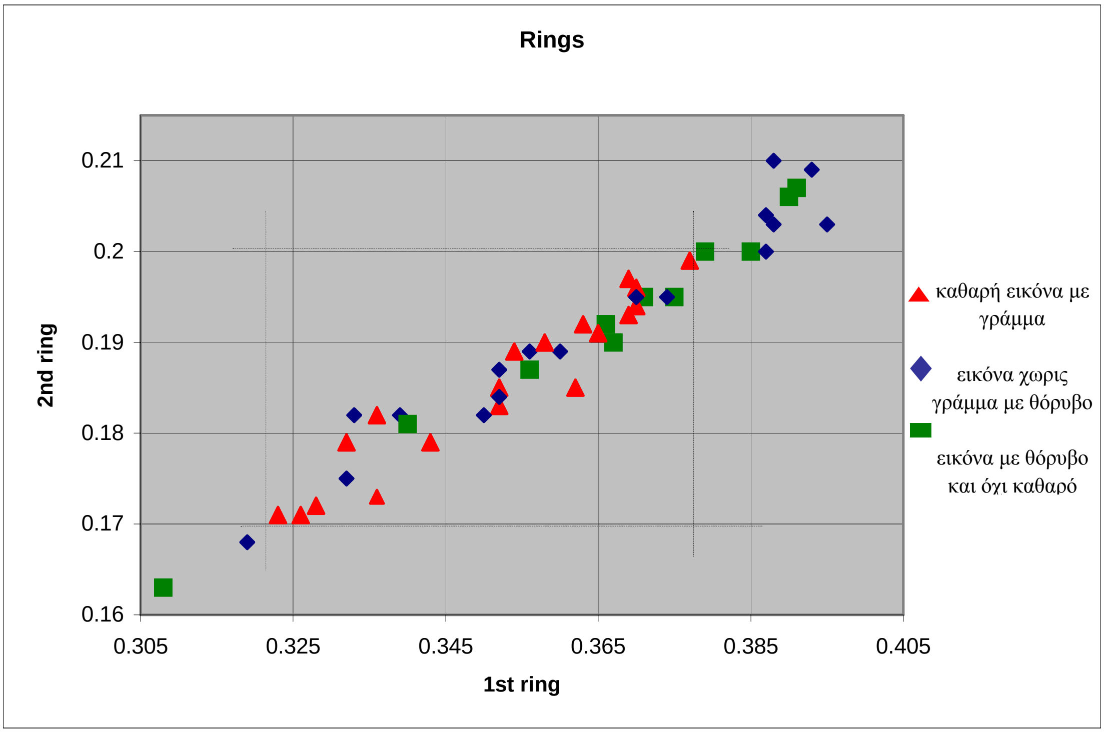
Γενικά τα 3 αυτά χαρακτηριστικά, παρουσιάζουν μικρό βαθμό συσχέτισης, έτσι ώστε το ένα δεν δίνει ασφαλή πληροφορία για το άλλο. Από τα τρία αυτά τα χαρακτηριστικά το contrast εμφανίζεται να αποτυπώνει την πιο χρήσιμη πληροφορία μιας και εμπεριέχει όχι μόνο πληροφορία σχετική με τις μεταβάσεις της έντασης αλλά και πληροφορία σχετικά με τις εντάσεις που εμφανίζονται στην κάθε υποεικόνα. Για τους παραπάνω λόγους από τα κατευθυντικά χαρακτηριστικά Haralick χρησιμοποιούμε το contrast και ειδικότερα μόνο στις υποπεριοχές τύπου 2.

Ένα άλλο σημαντικό χαρακτηριστικό για τις υποπεριοχές τύπου 2 είναι το μέγεθος ΔG (που αναλύσαμε στην §2.4.1) το οποίο αποτελεί μία ένδειξη του αριθμού των edges. Η χρησιμότητα του ΔG είναι να ξεχωρίζει περιπτώσεις εικόνων που περιέχουν έντονο θόρυβο από ομαλές εικόνες που περιέχουν κάποια δομή γράμματος. Αυτό επιτυγχάνεται βάση της παρατήρησης ότι μεγάλος αριθμός ΔG θα αντιστοιχεί σε εικόνα με πολλές ακμές θορύβου και γενικά άχρηστης πληροφορίας. Όταν η περιοχή περιέχει γράμμα σε ομαλό background, το ΔG προέρχεται από τις συγκεκριμένες ακμές του γράμματος και γενικά διατηρεί χαμηλότερες τιμές σε σχέση με την προηγούμενη περίπτωση. Ο συνδυασμός των κατευθυντικών χαρακτηριστικών contrast Haralick και ΔG αναλύεται στο επόμενο κεφάλαιο σε σχετίζεται με τον προτεινόμενο αλγόριθμο.

- **Rings–Wedges**

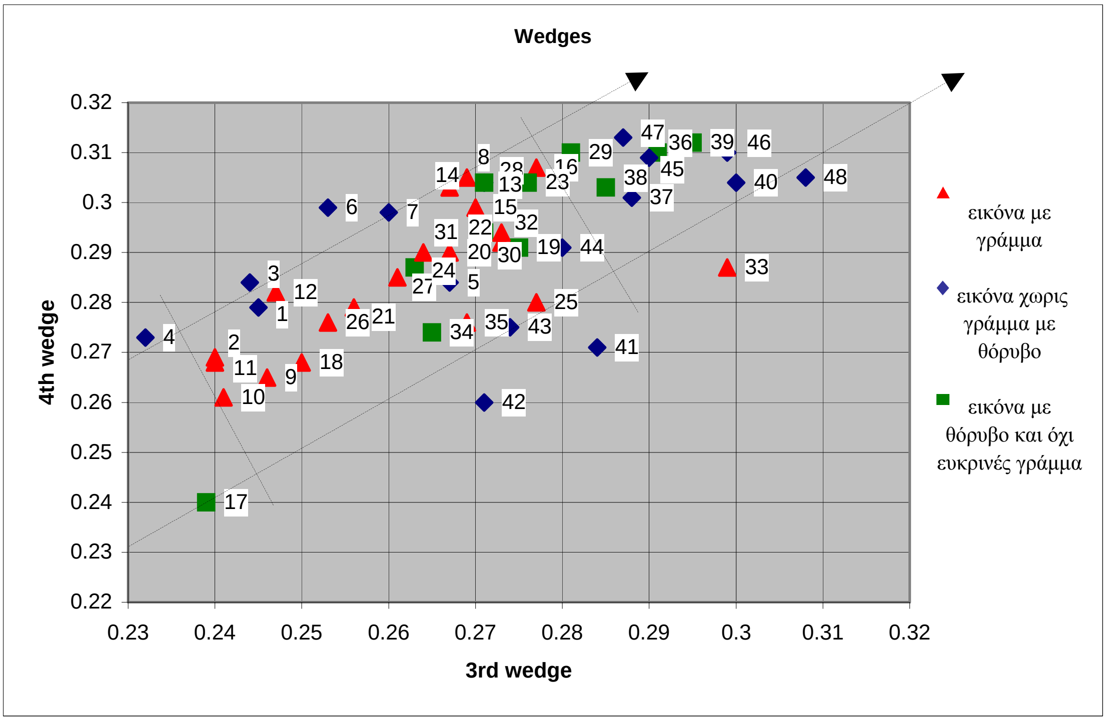
Κατά τη διάρκεια προσέγγισης του προβλήματος μελετώνται και χαρακτηριστικά στο πεδίο της συχνότητας , η περιγραφή των οποίων έγινε στην παράγραφο §2.2.2 . Ελέγχεται αν το ενεργειακό περιεχόμενο που περικλείεται είτε σε σχηματισμούς **rings** είτε σε κατευθυντικούς σχηματισμούς με τη μορφή **wedges** θα μπορούσε να δώσει χρήσιμη πληροφορία για τον διαχωρισμό κάποιων περιοχών και την εκτίμηση της ποιότητας τους.

Πειραματικά λοιπόν κάθε segment της εικόνας (διαστάσεων 80x80) χωρίζεται σε 100 τμήματα (διαστάσεων 8x8). Σε αυτά τα μικρά segments υπολογίζεται η ενέργεια του φάσματος ισχύος στα διάφορα rings και wedges που σχηματίζονται (§2.2.2). Ύστερα υπολογίζεται η μέση τιμή των ενεργειών των rings και των wedges για τα 100 μικρά segments της κάθε υποπεριοχής. Όσον αφορά στα rings επειδή γενικά η ενέργεια εμφανίζεται αρκετά μεγαλύτερη στο $1^{ο}$ ring από ότι στα υπόλοιπα ($2^{ο}$ και $3^{ο}$) συγκρίναμε την τιμή του $1^{ου}$ ring με την τιμή του $2^{ου}$ και $3^{ου}$ ring μαζί. Αυτή η συσχέτιση τιμών για όλες τις υποπεριοχές παρουσιάζεται στο *Διάγραμμα 6*. Από ότι παρατηρούμε η γραμμική διάρθρωση που εμφανίζουν οι τιμές των rings δεν μπορεί να διαχωρίσει καθόλου ικανοποιητικά τις διάφορες υποπεριοχές και έτσι δεν μπορεί να αποτελέσει κατάλληλο κριτήριο για την εκτίμηση ποιότητας των υποπεριοχών.



Διάγραμμα 6

Όσον αφορά στα wedges παρατηρήσαμε ότι η περισσότερη ενέργεια περικλείεται στο 3^{o} και 4^{o} wedge κάτι που οφείλεται στην κατευθυντικότητα των περιοχών. Η συσχέτιση της μέσης τιμής της ενέργειας των δύο wedges για κάθε υποπεριοχή φαίνεται στο *Διάγραμμα 7*.



Διάγραμμα 7

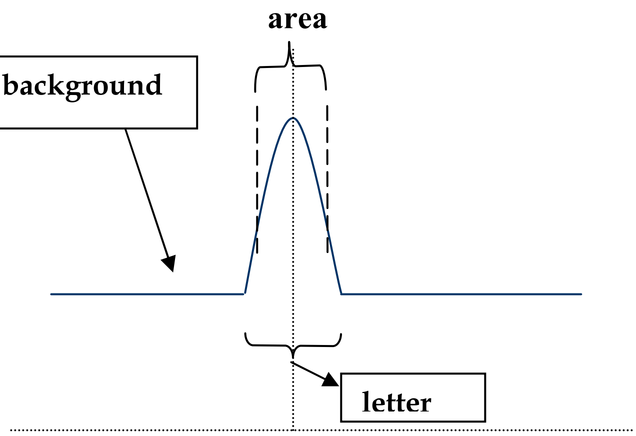
Με βάση το *Διάγραμμα 7* επιχειρείται μία κατάταξη περιοχών σε αποδεκτές ή μη, χρησιμοποιώντας τα γραμμικά όρια που φαίνονται στο διάγραμμα αυτό. Πιο συγκεκριμένα αν καταταχτούν ως αποδεκτές περιοχές εντός της ορθογώνιας περιοχής, το λάθος κατάταξης ή απόρριψης είναι σχετικά μικρό. Έτσι η ενέργεια των wedges θα μπορούσε να χρησιμοποιηθεί για χαρακτηρισμό και κατάταξη υποπεριοχών, με κάποιο σχήμα περιορισμού αυτής της ενέργειας. Γενικά δεν αποδεχόμαστε περιοχές που η ενέργεια τους στο 3^o αλλά και στο 4^o wedge είναι πολύ μεγάλη είτε πολύ μικρή. Στην πρώτη περίπτωση οι μεγάλες τιμές προέρχονται από έντονο θόρυβο που οφείλεται στην πορώδη αλλά και γραμμική διάβρωση των περιοχών και στη δεύτερη περίπτωση δεν θέλουμε πολύ μικρές τιμές διότι η περιοχή είναι πάρα πολύ ομαλή και δεν περιέχει χρήσιμη πληροφορία (γράμμα).

Ο *αλγόριθμος* που αναπτύχθηκε προκειμένου να επιτευχθεί ο διαχωρισμός των περιοχών και η εξαγωγή χαρακτηριστικών από κάθε κατηγορία περιοχών, παρουσιάζεται αναλυτικά στο Κεφάλαιο 3.

3. Αλγόριθμος

3.1 Στάδιο Ι - Βασικός Διαχωρισμός υποπεριοχών σε υποσύνολα

Στο πρώτο στάδιο λοιπόν γίνεται ένας πρώτος διαχωρισμός των υποπεριοχών, ταξινομώντας τις σε τρία υποσύνολα. Η κατάταξη αυτή έγινε με βάση την επιφάνεια της περιοχής συγκέντρωσης που κατασκευάσαμε στην παράγραφο §2.2.1. Με τον όρο επιφάνεια εννοούμε τον αριθμό των pixels που ανήκουν στη μικρή περιοχή. Από εδώ και στο εξής η επιφάνεια ατή θα καλείται «area». Αυτή η περιοχή συγκέντρωσης προσδιορίζει καλύτερα τη βασική δομή της αυτοσυσχέτισης. Αν κάνουμε μια αναδρομή σε αυτό που θεωρείται ιδανική μορφή της αυτοσυσχέτισης το "area", όπως έχει αναφερθεί, καταλαμβάνει τη βασική δομή του autocorrelation, ένα μέρος της καμπύλης όπως φαίνεται στο *Σχήμα 11.* Το μέγεθος του συνεπώς προδίδει σε τι είδους περιοχή βρισκόμαστε.

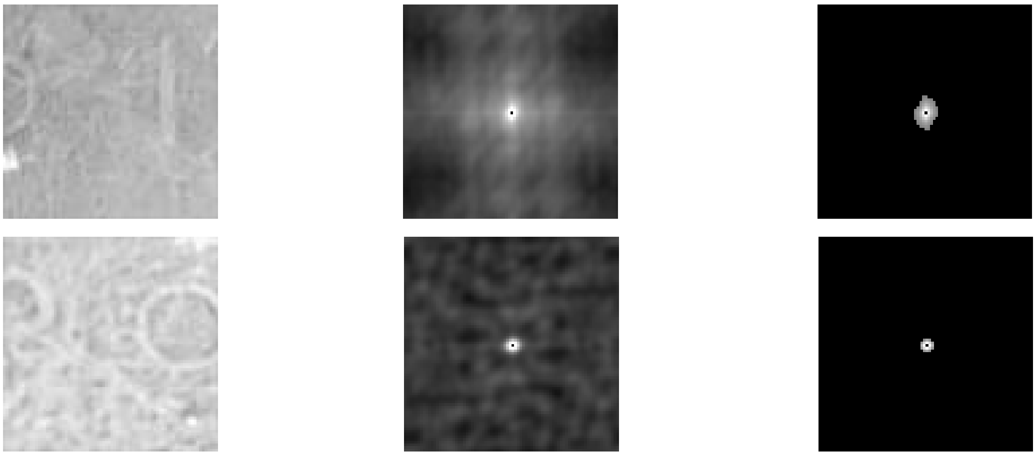


Σχήμα 11

Τα αριθμητικά όρια που τίθενται επιλέχτηκαν ύστερα από παρατήρηση των τιμών του area σε όλο το φάσμα των υποπεριοχών καθώς και της μορφής της αυτοσυσχέτισής τους. Παρατηρείται λοιπόν μία κοινή συμπεριφορά κάποιων υποπεριοχών που οδήγησε στην ομαδοποίηση τους σύμφωνα με τα κοινά τους χαρακτηριστικά:

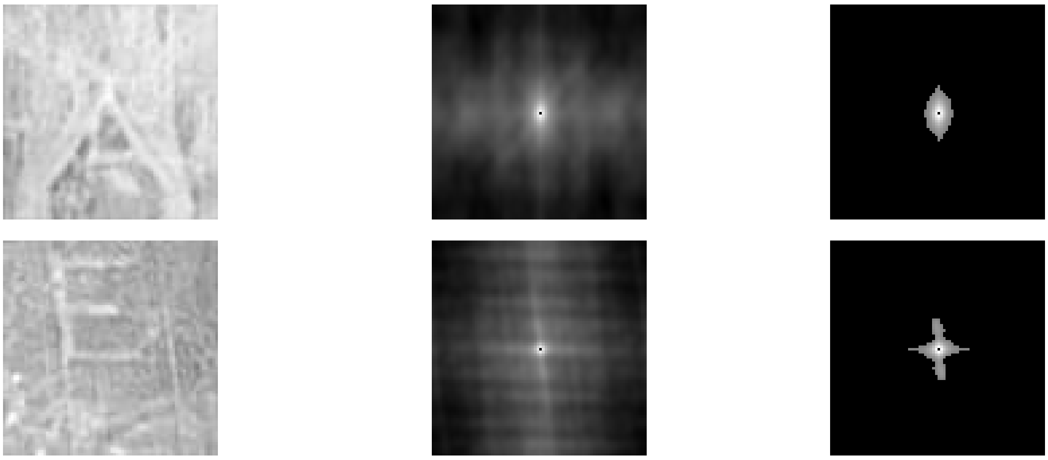
- **Τύπος 1**

Σε αυτήν την κατηγορία ανήκουν εικόνες με μικρό area. Μικρό area θεωρείται **area<=105.** Προκύπτει από εικόνες με μικρή "thin" δομή ή γράμμα σε σχετικά ομαλό background καθώς και από εικόνες με περιοδική ή τυχαία δομή (εξαιτίας της διάβρωσης του μαρμάρου). Ακολουθεί ένα παράδειγμα για μία σχετικά ομαλή εικόνα με πολύ «αχνούς» χαρακτήρες και ένα παράδειγμα από μία περιοδικά διαβρωμένη εικόνα



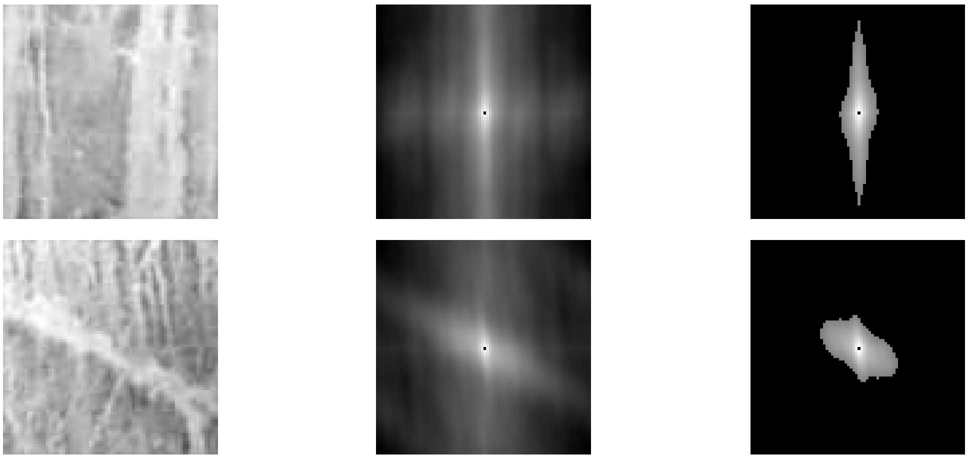
Τύπος 2

Εικόνες με ενδιάμεση τιμή area. Μετρίου μεγέθους area (**105<area<=265**) προκύπτει από εικόνες με “θολή” δομή ή γράμμα σε ομαλό background ή εικόνες με μεγάλου εύρους δομή ή θόρυβο. Το μέγεθος του area μεγαλώνει σε εικόνες τέτοιου είδους γιατί οι δομές «απλώνουν» ή γίνονται πιο δυσδιάκριτες και το background διατηρείται ομαλό. Παραδείγματα τέτοιων εικόνων:



Τύπος 3

Μεγάλο area (**area>265**) προκύπτει από εικόνες με ανομοιόμορφο background χωρίς κάποια συγκεκριμένη δομή και χωρίς φανερό γράμμα ή από εικόνες με πολύ έντονες επιμήκεις δομές που οφείλονται σε γραμμικής μορφής διάβρωση. Αυτές οι πολύ ισχυρές γραμμώσεις έχουν ως αποτέλεσμα μεγάλη έκταση στο autocorrelation άρα και πολύ μεγάλο area.



3.2 Στάδιο ΙΙ - Εξαγωγή χαρακτηριστικών σε κάθε υποσύνολο

Όπως έχει ήδη αναφερθεί στην κάθε κατηγορία υποπεριοχών εφαρμόζονται εκείνα τα χαρακτηριστικά που παρέχουν την κατάλληλη και χρήσιμη πληροφορία για την ποιότητα της υποπεριοχής.

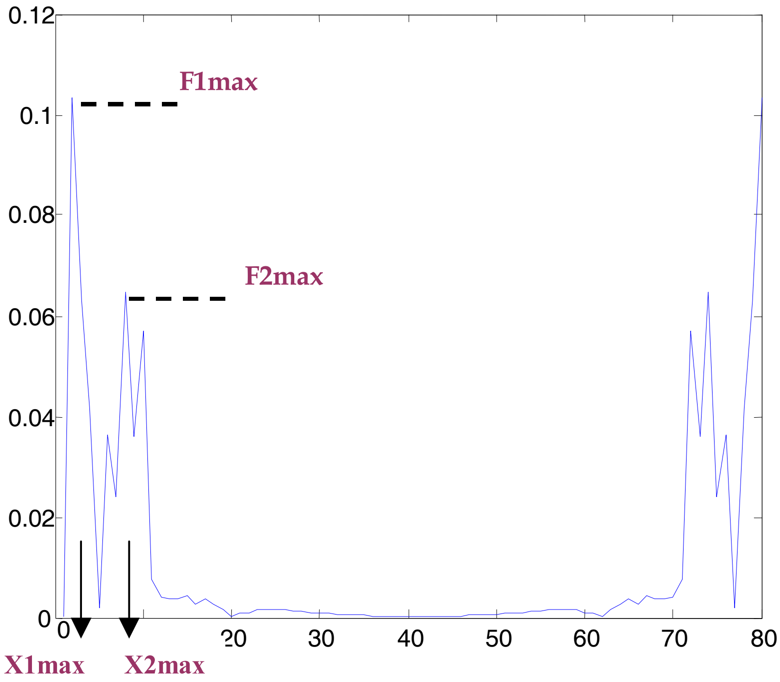
3.2.1 Υποπεριοχές Τύπου 1

Για τις υποπεριοχές **τύπου 1** (εικόνες με ευκρινές γράμμα σε σχετικά περιοδικό background, εικόνες με έντονη περιοδική μορφή, εικόνες με τυχαία δομή) εξαιτίας της περιοδικότητας που εμφανίζουν χρειάζεται μία μέθοδος ανίχνευσης της περιοδικότητας τους έτσι ώστε να διαχωριστούν περιοχές με πολύ έντονη περιοδικότητα (που απορρίπτονται) από περιοχές που περιέχουν γράμμα. Στην παράγραφο §2.3 αναλύεται ο τρόπος με τον οποίο ανιχνεύονται περιοδικά χαρακτηριστικά και ως αποτέλεσμα προκύπτουν κάποια γραφήματα στο επίπεδο την συχνότητας που αποτελούν ένδειξη της περιοδικότητας της κάθε υποπεριοχής.
Μελετήθηκε η μορφή των γραφημάτων προσπαθώντας με αυτόν τον τρόπο να τεθούν κάποια κριτήρια για την αποδοχή ή μη των υποπεριοχών βάση της περιοδικότητάς τους.
Γίνεται έλεγχος και σύγκριση μεταξύ των εξής μετρικών :

- ✓ Συχνότητα πρώτης αρμονικής (**x1max**). Είναι φανερό παρατηρώντας το *Σχήμα 11* ότι όσο το x1max μετατοπίζεται προς τα δεξιά στον άξονα x, δηλαδή αυξάνεται η συχνότητα, η υποπεριοχή θα έχει όλο και πιο έντονη περιοδική μορφή.

- ✓ Τιμή του πρώτου μέγιστου magnitude των αρμονικών της συχνότητας (**F1max**). Η τιμή αυτή υπολογίζεται παίρνοντας τη μεγαλύτερη τιμή των μεγίστων αρμονικών σε ένα συγκεκριμένο εύρος συχνοτήτων που επιλέξαμε (0-15). Το μέτρο αυτό δίνει τη μέγιστη περιοδική δομή του autocorrelation. Μικρή τιμή, κάτω από ένα όριο που επιλέγεται βάση παρατηρήσεων, σημαίνει πολύ μικρή περιοδικότητα.

- ✓ Λόγος των τιμών της πρώτης αρμονικής συχνότητας ως προς τη δεύτερη μέγιστη αρμονική συχνότητας (**Frequency_Ratio**=F1max/F2max). Το Frequency Ratio αποτελεί τη σύγκριση μεταξύ των δυο βασικών δομών της εικόνας. Αν η τιμή του μειώνεται πλησιάζοντας τη μονάδα, σημαίνει ότι η υποπεριοχή περιέχει έντονες περιοδικές δομές.



Σχήμα 12

Τα αριθμητικά κριτήρια που τίθενται παρουσιάζονται ακόλουθα :

- ✓ Αν το **x1max**>=4 τότε γίνεται απόρριψη της υποπεριοχής λόγω μεγάλης περιοδικότητας.

- ✓ Αν ισχύουν **F1max**>0.095 και **Frequency_Ratio**<=1.6 συγχρόνως τότε επίσης γίνεται απόρριψη της υποπεριοχής λόγω μεγάλης περιοδικότητας. Επίσης αν **F1max**<0.095 η πρώτη πιο ισχυρή αρμονική θεωρείται πολύ «ασθενής» στοιχείο που οδηγεί στο συμπέρασμα ότι ολόκληρη η υποπεριοχή είναι ομοιόμορφη , χωρίς κάποια επικρατούσα ισχυρή δομή.

Τα παραπάνω κριτήρια πρέπει να πληρούνται και για την οριζόντια αλλά και την κατακόρυφη κατεύθυνση συγχρόνως. Αρκεί δηλαδή μία υποπεριοχή να μην ικανοποιεί τους περιορισμούς για τη μία μόνο κατεύθυνση για να μη γίνει αποδεκτή. Αυτό είναι λογικό εφόσον μπορεί να εμφανίζει έντονη περιοδικότητα μόνο στη μία κατεύθυνση.

3.2.2 Υποπεριοχές Τύπου 2

Οι υποπεριοχές **τύπου 2** (εικόνες με "θολή" δομή ή γράμμα σε ομαλό (smooth) background, εικόνες με μεγάλου εύρους ("wide") δομή ή γράμμα σε ομαλό background) χρειάζονται κάποια κατευθυντικά μεγέθη που να μετρούν τις μεταβάσεις της έντασης σε μία κατεύθυνση και να πιστοποιήσουν με αυτόν τον τρόπο την ύπαρξη γράμματος.

Όπως αναφέρθηκε και στην παράγραφο 2.5 τα χαρακτηριστικά εκείνα που είναι κατάλληλα και χρήσιμα για την ανίχνευση κατευθυντικότητας είναι ο αριθμός pixels των ακμών ΔG και από τα χαρακτηριστικά του Haralick το contrast.

Αφού υπολογίζεται το ΔG στις 4 κατευθύνσεις: $0°$, $90°$, $45°$ και $135°$ ξεχωριστά (βλ. §2.4.1), ύστερα για κάθε υποπεριοχή αθροίζονται όλα τα ΔG για τις 4 κατευθύνσεις. Με αυτόν τον τρόπο δηλαδή υπολογίζεται ο αριθμός των pixels που ανήκουν στις ακμές των υποπεριοχών του τύπου 2 κατηγορίας. Επίσης υπολογίζεται το contrast για τις τέσσερις διευθύνσεις $0°$, $90°$, $45°$ και $135°$ αλλά η τιμή που χρησιμοποιείται στον αλγόριθμο είναι ο μέσος όρος των τεσσάρων τιμών του contrast για τις τέσσερις κατευθύνσεις .

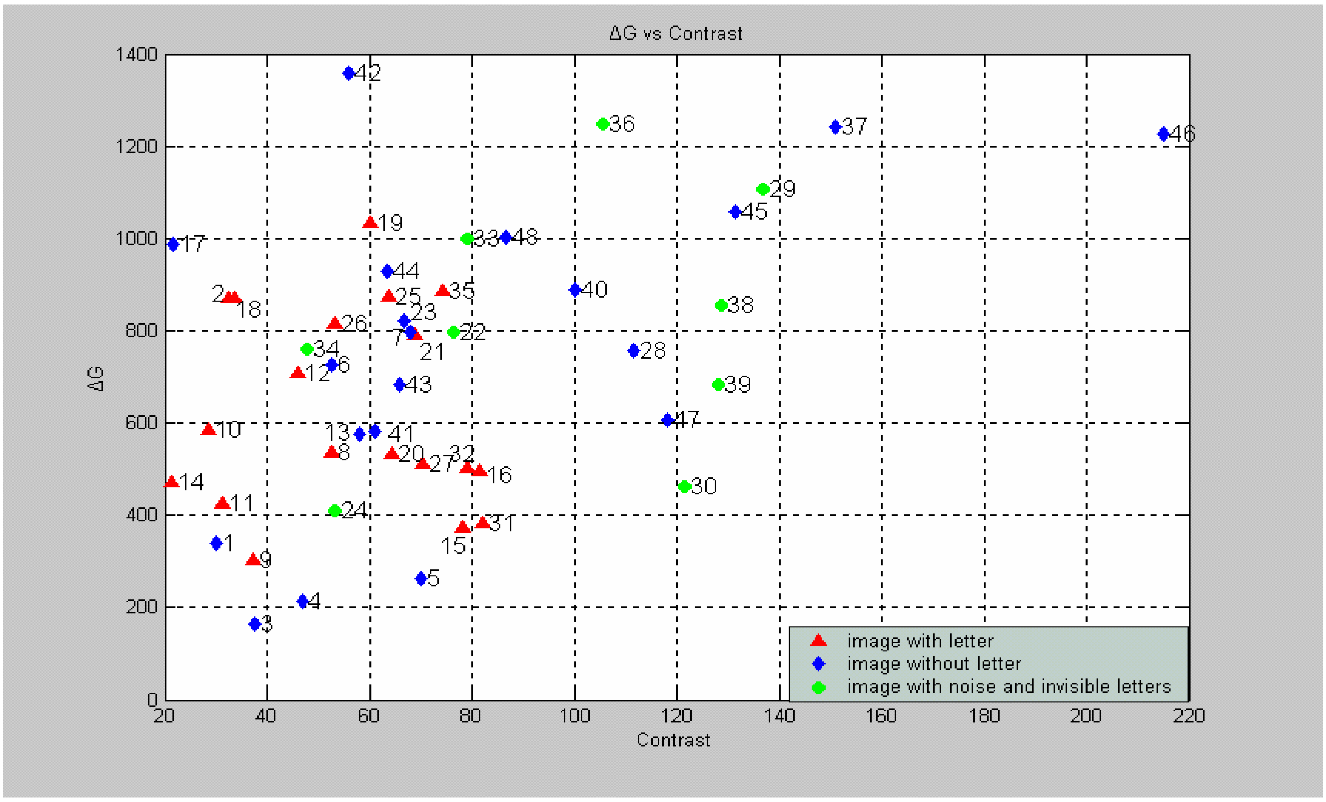
Ο συνδυασμός χαρακτηριστικών που περιέχει σημαντική πληροφορία είναι το μέγεθος ΔG των ακμών σε συνάρτηση με το contrast. Για παράδειγμα ο συνδυασμός μεγάλου contrast με μεγάλο αριθμό ΔG συνεπάγεται την ύπαρξη πολλών διάσπαρτων δομών στην εικόνα οι οποίες προέρχονται κυρίως από διάβρωση στην υπό εξέταση περιοχή. Επίσης συνδυασμός μικρού contrast με μικρή τιμή ΔG παραπέμπει σε περιοχές πολύ ομοιογενείς που πιθανότατα δεν υπάρχει γράμμα.

Βάσει αυτών των παρατηρήσεων τίθεται το κριτήριο για τις περιοχές αυτού του τύπου :

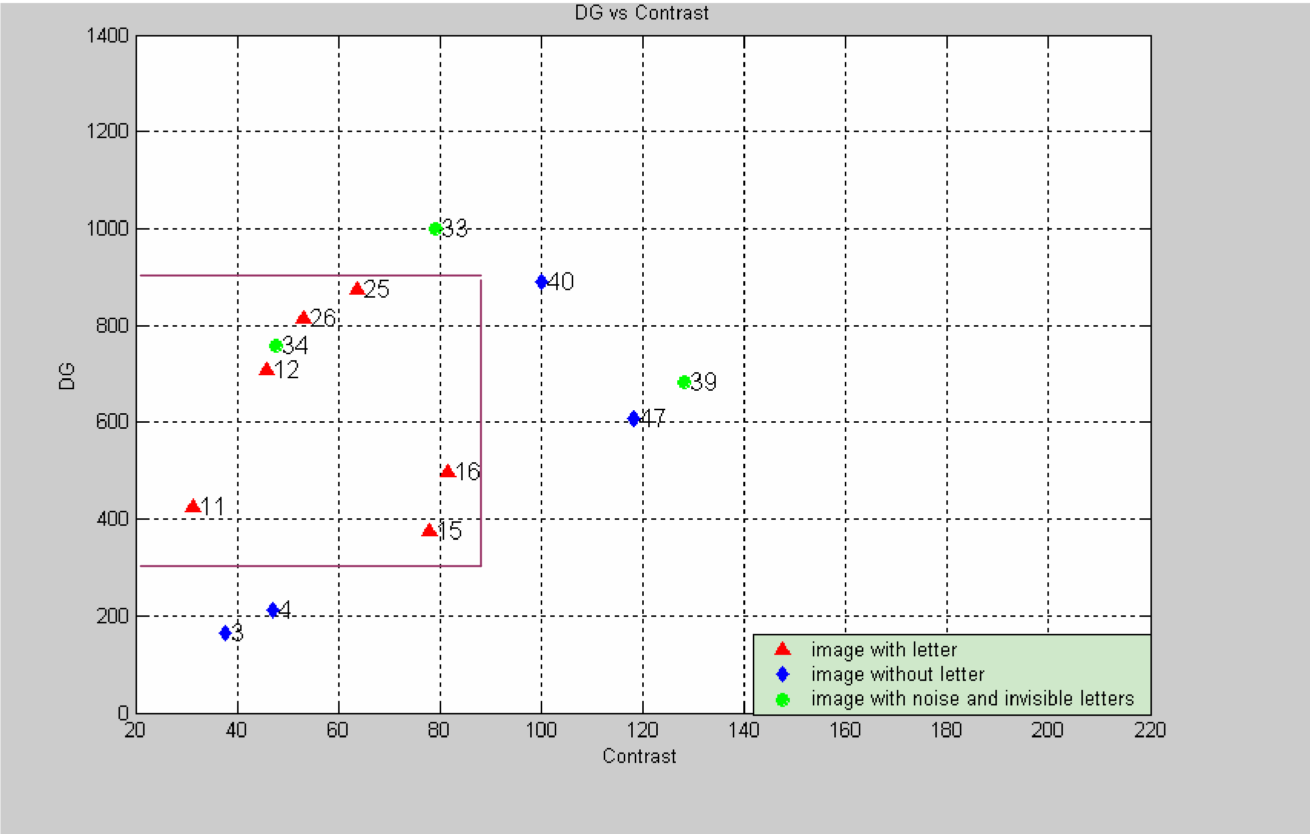
- ✓ Αν το contrast>85 και ταυτόχρονα ΔG<300 ή ΔG>900 τότε γίνεται απόρριψη της περιοχής.

Στο *Διάγραμμα 4* φαίνεται πως κυμαίνεται το ΔG σε σχέση με το Contrast για όλες τις υποπεριοχές . Παρατηρείται ένας αρκετά ευκρινής διαχωρισμό ανάμεσα στις διάφορες περιοχές που περιέχουν γράμμα από αυτές που δεν περιέχουν ή περιέχουν δυσδιάκριτο γράμμα σε θορυβώδες background.

Στο *Διάγραμμα 5* απομονώνονται οι περιοχές τύπου 2 και παρατηρείται πως κυμαίνεται το ΔG σε σχέση με το Contrast. Σύμφωνα με το κριτήριο που τέθηκε (contrast>85 και ταυτόχρονα ΔG<300 ή ΔG>900) παρατηρείται ότι επιτυχώς γίνονται αποδεκτές εικόνες με πολύτιμη πληροφορία και απορρίπτονται αυτές που είτε δεν έχουν γράμμα είτε είναι πολύ θορυβώδεις.



Διάγραμμα 4



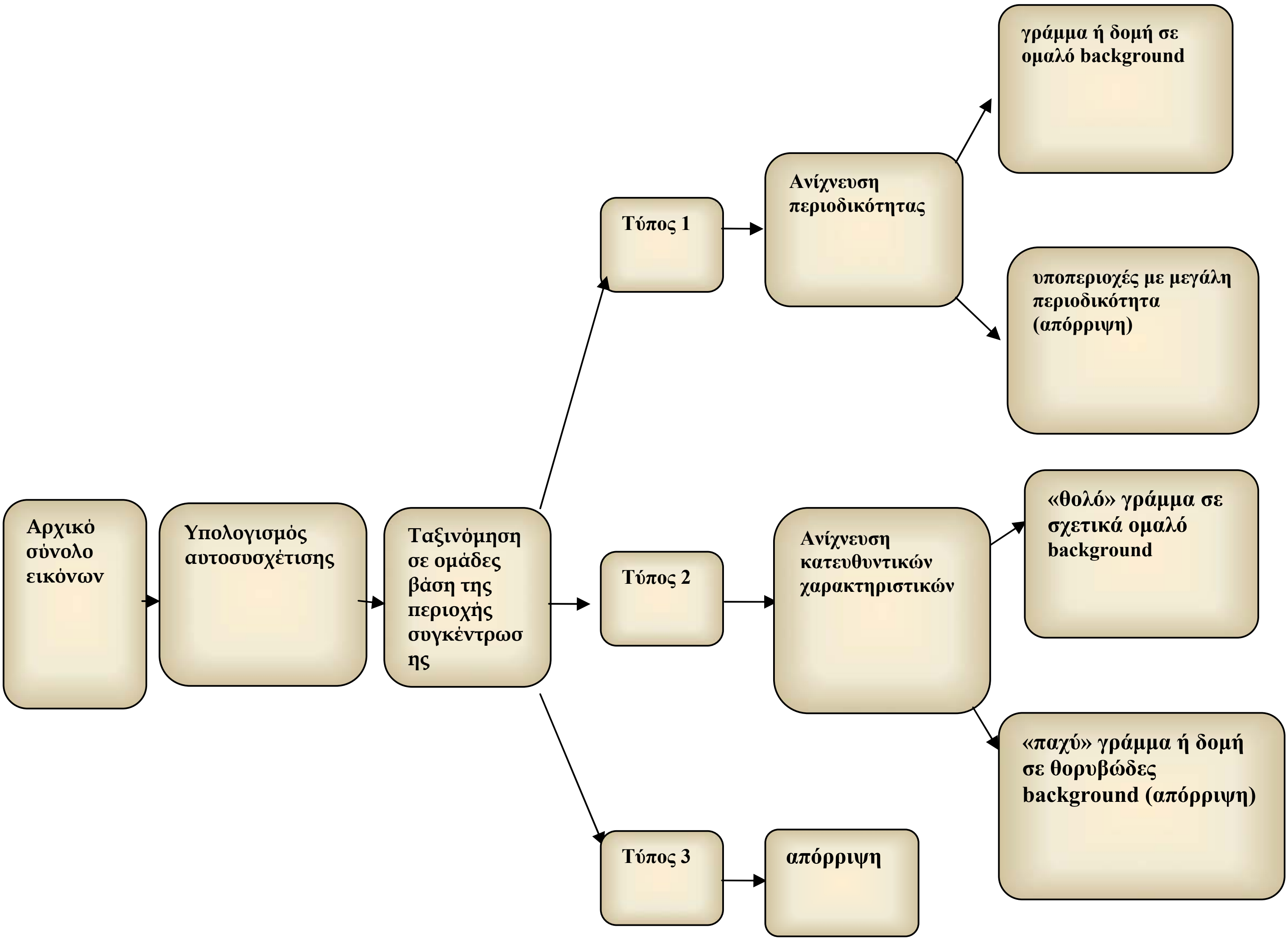
Διάγραμμα 5

3.2.3 Υποπεριοχές τύπου 3

Οι υποπεριοχές τύπου 3 κατηγορίας (εικόνες με πολύ ανομοιόμορφο background, εικόνες με πολύ έντονες επιμήκεις δομές) δεν γίνονται δεκτές εφόσον είναι περιοχές χωρίς κάποια δομή και εμφανίζουν μεγάλο βαθμό γραμμικής διάβρωσης.

3.3 Συνολική δομή αλγορίθμού

Για να έχουμε μία καλύτερη οπτική εικόνα του αλγορίθμου και για να ακολουθούμε καλύτερα τη ροή του παραθέτουμε το block diagram του:



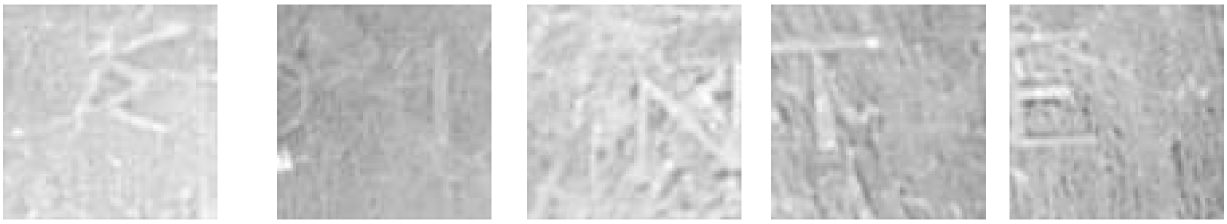
4. Εφαρμογές αλγορίθμου

4.1 Παραδείγματα από τα διάφορα στάδια εφαρμογής του αλγορίθμου

Σε αυτό το σημείο παρουσιάζονται κάποια αποτελέσματα της εφαρμογής του αλγορίθμου για τις τρεις κατηγορίες υποπεριοχών και κάποιες χαρακτηριστικές περιοχές που γίνονται δεκτές και άλλες που απορρίπτονται:

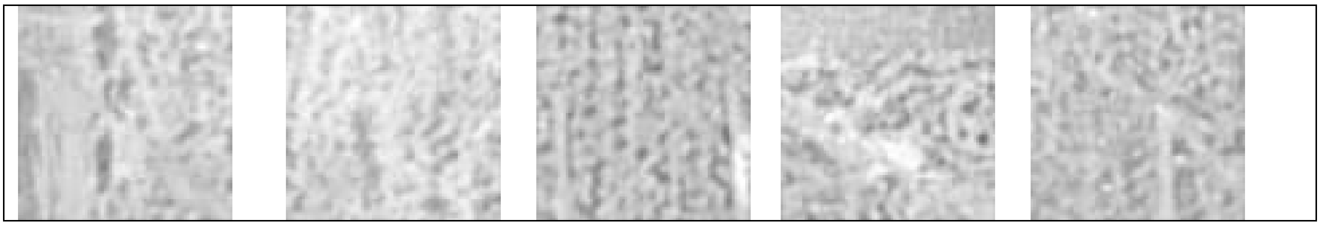
Τύπος 1

Κάποιες περιοχές που γίνονται δεκτές από αυτήν την κατηγορία παρουσιάζονται ακόλουθα:



Από ότι φαίνεται οι περιοχές που αποδέχεται ο αλγόριθμός είναι είτε ομαλές με «λεπτούς χαρακτήρες» είτε περιοδικές με ευκρινείς χαρακτήρες.

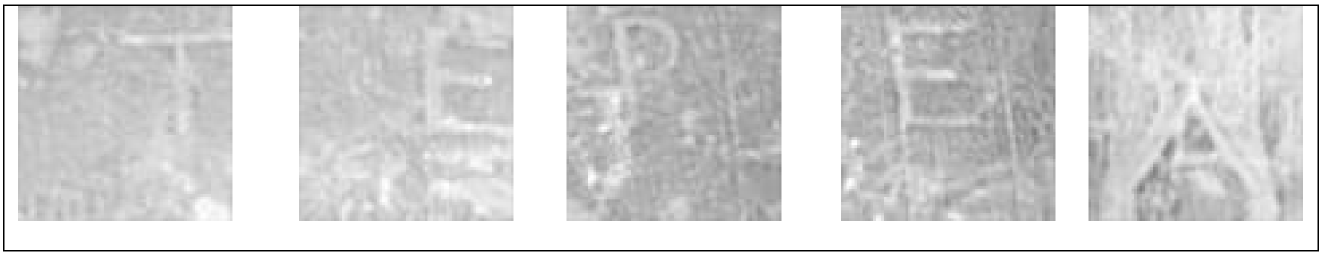
Περιοχές που απορρίπτονται:



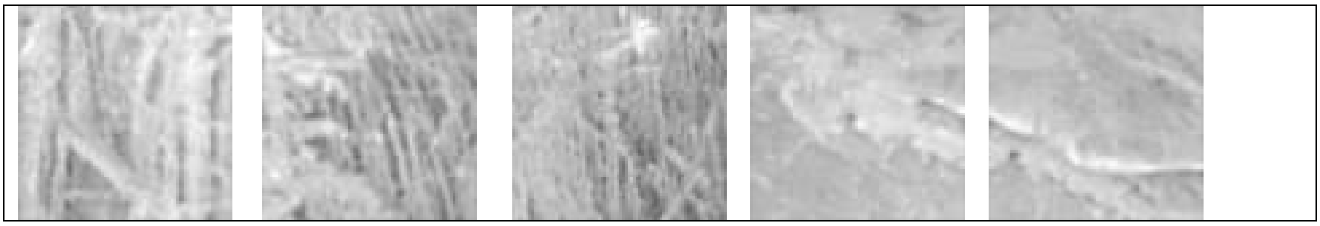
Είναι περιοχές με έντονη περιοδικότητα λόγω διάβρωσης και οι χαρακτήρες που τυχόν υπάρχουν στην περιοχή είναι δυσδιάκριτοι, οπότε ορθά ο αλγόριθμος τις απορρίπτει.

Τύπος 2

Κάποιες περιοχές που γίνονται δεκτές από αυτήν την κατηγορία είναι ομαλές με «θολά» βέβαια γράμματα αλλά ευκρινή:

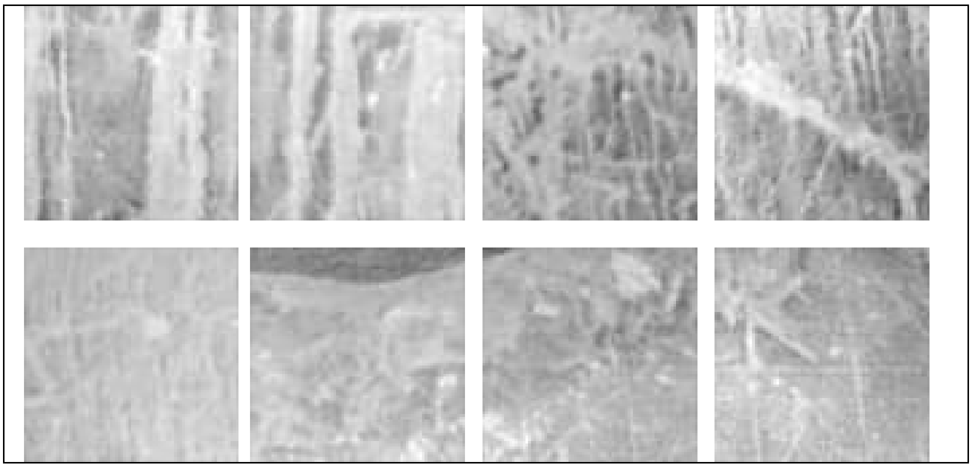


Περιοχές που απορρίπτονται είναι περιοχές θορυβώδεις με γραμμική διάβρωση ή περιοχές με έντονες ακμές που προέρχονται από εγκοπές του μαρμάρου:



Τύπος 3

Οι περιοχές αυτού του τύπου είναι όλες απορριπτέες. Είναι υποπεριοχές πολύ ομαλές ή με έντονη περιοδική διάβρωση ή πολύ ομαλές σε κατευθυντικές περιοχές και φαίνονται παρακάτω:



4.2 Τελικό Αποτέλεσμα

Το αποτέλεσμα του διαχωρισμού περιοχών προέκυψε ύστερα από τη διαδικασία κάποιων ελέγχων οι οποίοι εξηγήθηκαν σε προηγούμενα στάδια αναλυτικά. Συνολικά οι έλεγχοι αυτοί και η αποτελεσματικότητά τους φαίνονται συγκεντρωμένοι για όλες τις υποπεριοχές στον παρακάτω πίνακα. Από τους ελέγχους αυτούς ο τρίτος είναι ουσιαστικά ο έλεγχος στον οποίο βασίζεται ο αλγόριθμος για να διαχωρίσει τις διάφορες υποπεριοχές. Σε σύγκριση με τον προτεινόμενο αλγόριθμο (Test3) παρουσιάζουμε και τους ελέγχους που αφορούν στην ενέργεια που περικλείεται στους σχηματισμούς rings και wedges του φάσματος ισχύος των υποπεριοχών οι οποίοι πετυχαίνουν και αυτοί κάποιο διαχωρισμό των υποπεριοχών.

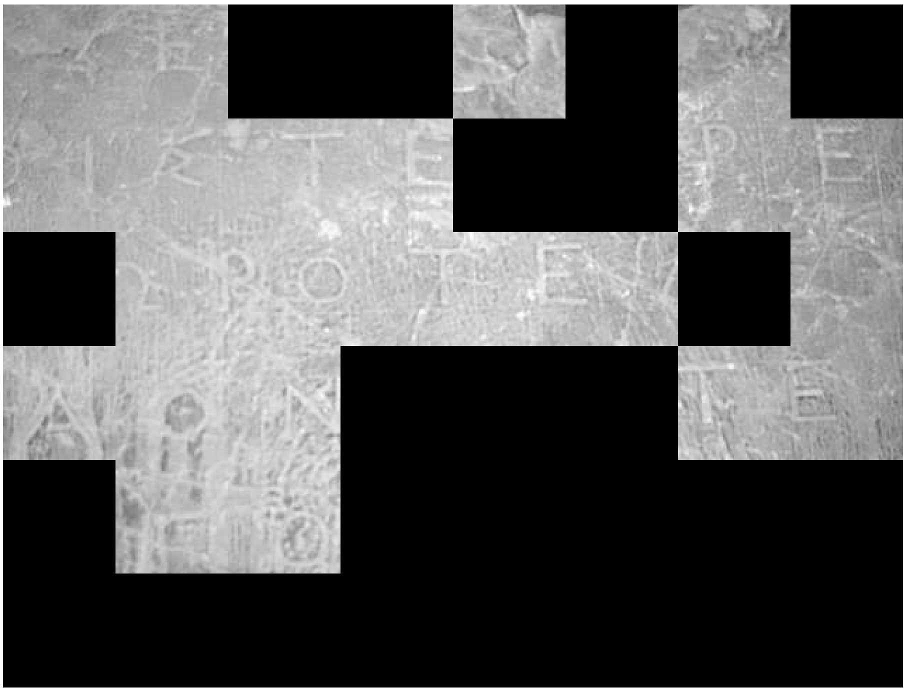
Image no and image type: ▲ Εικόνα με γράμμα ◆ Εικόνα χωρίς γράμμα ■ Εικόνα με θόρυβο και όχι ευκρινές ή ολόκληρο γράμμα	**Test1** Energy of Rings Απορρίπτουμε τις εικόνες που δεν ανήκουν σε μια επιλεγμένη περιοχή ενέργειας.	**Test2** Energy of Wedges Απορρίπτουμε τις εικόνες που δεν ανήκουν σε μια επιλεγμένη περιοχή ενέργειας.	**Test3** εξετάζουμε την τιμή του τύπος : **τύπος 1: value<=105** (if xmax>=4→ rejection) OR (If fmax>0.095 and FreqRatio<=1.6→ rejection) **τύπος 2: 105<value<265** (if contrast>85 ->rejection) AND (if ΔG>1100 ή ΔG<280→ rejection) **τύπος 3: value>265** →rejected
1. ◆			τύπος **1**
2. ▲			τύπος **1**
3. ◆		–	τύπος **2** –
4. ◆	–	–	τύπος **2** –
5. ◆			τύπος **1**
6. ◆		–	τύπος **3** –
7. ◆		–	τύπος **3** –
8. ▲			τύπος **3** –
9. ▲			τύπος **1**
10. ▲			τύπος **1**
11. ▲			τύπος **2**
12. ▲			τύπος **2**
13. ◆			τύπος **1** –
14. ▲			τύπος **1** –
15. ▲			τύπος **2**
16. ▲			τύπος **2**
17. ◆	–		τύπος **3** –
18. ▲			τύπος **1**
19. ▲			τύπος **1**
20. ▲			τύπος **1**
21. ▲			τύπος **1**

22.	■			τύπος **1**	
23.	◆			τύπος **1**	–
24.	■			τύπος **1**	
25.	▲			τύπος **2**	
26.	▲			τύπος **2**	
27.	▲			τύπος **1**	
28.	◆			τύπος **1**	–
29.	■	–	–	τύπος **1**	–
30.	■			τύπος **1**	–
31.	▲			τύπος **1**	
32.	▲			τύπος **1**	
33.	■		–	τύπος **2**	
34.	■			τύπος **2**	
35.	▲			τύπος **1**	
36.	■	–	–	τύπος **1**	–
37.	◆	–	–	τύπος **1**	–
38.	■		–	τύπος **1**	–
39.	■	–	–	τύπος **2**	–
40.	◆	–	–	τύπος **2**	–
41.	◆		–	τύπος **3**	–
42.	◆		–	τύπος **3**	–
43.	◆			τύπος **1**	–
44.	◆			τύπος **1**	–
45.	◆	–	–	τύπος **1**	–
46.	◆	–	–	τύπος **1**	–
47.	◆	–	–	τύπος **2**	–
48.	◆	–	–	τύπος **3**	–

Με έντονο χρώμα συμβολίζονται οι περιοχές που απορρίπτει ο αλγόριθμός στους διάφορους ελέγχους. Παρατηρείται ότι από τους ελέγχους που πραγματοποιούνται ο 3^{ος} έλεγχος είχε την καλύτερη απόδοση. Επιτελώντας λεπτομερή έλεγχο, επιτυχώς απορρίπτει περιοχές με πολύ κακή ποιότητα και δέχεται περιοχές που τα γράμματα είναι ευδιάκριτα και το background ομαλό. Το κριτήριο την ενέργειας των rings κατάφερε να απορρίψει λίγες μόνο περιοχές. Λίγο καλύτερα λειτουργεί το κριτήριο των wedges. Οι εικόνες στις οποίες ο αλγόριθμος λειτουργεί με σχεδόν 100% επιτυχία είναι οι υποπεριοχές τύπου 2 λόγω του ότι είναι πιο ομαλές περιοχές και πιο λίγες μέσα στη συνολική εικόνα. Το αποτέλεσμα για τις περιοχές τύπου 2 φαίνεται καλύτερα στο *Διάγραμμα 5* που παραθέσαμε στην §3.2.2 Οι περιοχές τύπου 1 είναι και οι περισσότερες που συναντώνται στο σύνολο των υποπεριοχών και πολλές από αυτές περιέχουν αρκετό θόρυβο και ανεπιθύμητα χαρακτηριστικά («εγκοπές» ή διάβρωση) που αλλοιώνουν κατά κάποιο βαθμό τα αποτελέσματα για αυτό και προέκυψαν κάποιες απώλειες στους ελέγχους.

Για να εκτιμηθεί καλύτερα το τελικό αποτέλεσμα δείχνουμε ξανά την αρχική εικόνα:

Η εικόνα που προκύπτει μετά την εφαρμογή του αλγορίθμου, με τις περιοχές που απορρίπτονται να είναι σκιασμένες, φαίνεται παρακάτω:



Η περιοχές της εικόνας που έχουν γίνει αποδεκτές ως επί το πλείστον είναι σχετικά ομαλές περιοχές που περιέχουν γράμματα τα οποία μάλιστα είναι ευκρινή σε ικανοποιητικό βαθμό. Φυσικά εμφανίζεται απώλεια κάποιων γραμμάτων που είτε είναι γράμματα στα οποία παρεμβάλλεται μία έντονη πορώδης ή γραμμική διάβρωση είτε γράμματα που ανήκουν σε υποπεριοχές με έντονες εγκοπές οι οποίες αλλοιώνουν το αποτέλεσμα. Κάποιες βελτιώσεις που επιτρέπουν ίσως τη μείωση του “false alarm” και του “false reject” rate του αλγορίθμου περιγράφονται στις μελλοντικές επεκτάσεις.

4.3 Μελλοντικές Επεκτάσεις

Αν και εξετάζει πολλές μικρές υποπεριοχές, η εφαρμογή ολόκληρου του αλγορίθμου αν και εφαρμόστηκε και σε άλλη μία εικόνα της μορφής:

πρέπει να τεκμηριωθεί σε μεγάλο πλήθος εικόνων, οι οποίες δεν είναι διαθέσιμες άμεσα. Η εφαρμογή αυτή θα δώσει τη δυνατότητα καλύτερης και πιστότερης επιλογής των thresholds του αλγορίθμου στο συγκεκριμένο πρόβλημα.

Στην εργασία μας η εφαρμογή του αλγορίθμου έγινε σε segments της εικόνας που ήταν διακριτά και συνεχόμενα (μη επικαλυπτόμενα). Μία διαφορετική προσέγγιση είναι να εφαρμοστεί ο αλγόριθμος σε διαφορετικής χωρικής θέσης υποπεριοχές που να είναι είτε επικαλυπτόμενες είτε συνεχόμενες αλλά διαφορετικά τμηματοποιημένες. Η προσέγγιση αυτή θα μπορούσε να εξετάσει λεπτομερέστερα τις διάφορες περιοχές ώστε ο αλγόριθμος να αποδεχτεί κάποια κομμάτια που με τη δική μας εργασία απορρίπτει.

Επιπρόσθετα, η εξέταση των περιοχών σε διαφορετική ανάλυση (resolution) θα μπορούσε να παράγει χρήσιμη πληροφορία. Ένας τρόπος για να εξετάσουμε εικόνες σε διαφορετικό resolution είναι η χρήση του *Wavelet Transform* ο οποίος επιτελεί ανάλυση (decomposition) της εικόνας σε πολλά επίπεδα (levels) το καθένα με διαφορετικό resolution. Με αυτόν τον τρόπο καθώς προχωράμε σε χαμηλότερα επίπεδα (σε βάθος της πυραμίδας) και εφαρμόζοντας τον αλγόριθμο, ίσως προσθέτοντας και κάποια άλλα χαρακτηριστικά, μπορούμε να αποκομίσουμε περισσότερη πληροφορία και να τη συνδυάσουμε κατάλληλα.

References

[1] The Image Processing Handbook, 3rd edition John C. Russ

[2] F. Liu and R.W. Picard. ‘ Periodicity, directionality and randomness: Wold features for image modeling and retrieval ‘ IEEE Trans. on PAMI, vol.18 , no. 7, , pp. 722-733, July 1996.

[3] A New Wold Ordering for Image Similarity, R. W. Picard and F. Liu, Proc. ICASSP , Adelaide, Australia, Vol. V, pp. 129-132, Apr 1994

[4] Digital Image Processing Rafael C. Gonzalez and Richard E. Woods.

[5] Marijke F. Augusteijn, Laura E. Clemens, and Kelly A. Saw, Performance Evaluation of Texture Measures for Ground Cover Identification in sattelite images by Means of a Neural Network Classifier, IEEE Transactions on Geoscience and remote sensing, Vol. 33, No 3, May 1995.

[6] H. Tamura, S. Mori, and T. Yamawaki. Texture features corresponding to visual perception. IEEE T. Sys., Man and Cyber., SMC-8(6): 460-473, 1978.

[7] R.M.Haralick, K.Shanmugam, I.Dinstein, Textural Features for Image Classification, IEEE Transactions on Systems, Man and Cybernetics, 1973, p.610-621.

[8] GLCM texture: a tutorial, Mryka Hall-Beyer, Ph.D. Department of Geography
University of Calgary Canada