



Επεξεργασία Εικόνων από Ασύρματη Ενδοσκόπηση

Σπουδαστές:
Γεωργία Βεκοπούλου (1508)
Μαρία Παπαδοπούλου (1606)

Επιβλέπων καθηγητής
Δρ. Αναστάσιος Ι. Μπαλουκτσής

15 Μαΐου 2014

Πίνακας περιεχομένων

1	Εισαγωγή	3
1.1	Γενικά	3
1.2	Φυσιολογία του εντέρου	4
1.3	Το πεπτικό έλκος	6
1.4	Ασύρματη ενδοσκόπηση (AE)	7
1.4.1	Δομή και λειτουργία της κάψουλας	7
1.4.2	Σύστημα ψηφιακής καταγραφής δεδομένων	8
1.4.3	Λογισμικό επεξεργασίας δεδομένων	9
1.4.4	Μελέτη των δεδομένων καταγραφής	9
1.4.5	Περιορισμοί της AE	10
1.5	Βιβλιογραφική έρευνα	11
1.6	Σύντομη περιγραφή της παρούσας εργασίας	12
2	Αναπαράσταση εικόνων στον υπολογιστή	13
2.1	Γενικά	13
2.2	Χρωματικοί χώροι	13
2.2.1	Χρωματικός χώρος RGB	14
2.2.2	Χρωματικός χώρος HSV	14
2.2.3	Χρωματικός χώρος YCbCr	15
3	Προεπεξεργασία των εικόνων	17
3.1	Μετασχηματισμός Curvelet	17
3.2	Ταχύς διακριτός μετασχηματισμός Curvelet	20
4	Εξαγωγή χαρακτηριστικών υφής	22
4.1	Τοπικά Δυαδικά Πρότυπα (Local Binary Patterns)	23
4.2	Γενίκευση του LBP	24
4.3	LBP ανεξάρτητα περιστροφής	26
4.4	Ομοιόμορφα ανεξάρτητα περιστροφής LBP	27
5	Ταξινόμηση	28
5.1	Μηχανές διανυσμάτων υποστήριξης	28
6	Η προτεινόμενη μεθοδολογία	31
6.1	Εξαγωγή χρωματικών καναλιών	31
6.2	Προεπεξεργασία των εικόνων	32
6.3	Εξαγωγή χαρακτηριστικών υφής	34

6.4	Κατασκευή διανύσματος χαρακτηριστικών.....	36
6.5	Ταξινόμηση.....	37
6.6	Δεδομένα.....	39
6.7	Υλοποίηση της μεθοδολογίας.....	41
7	Αποτελέσματα.....	43
7.1	Χρωματικός χώρος RGB.....	43
7.2	Χρωματικός χώρος HSV.....	46
7.3	Χρωματικός χώρος YCbCr.....	49
7.4	Σύγκριση των χρωματικών χώρων.....	52
8	Αναφορές.....	54
9	Παράρτημα.....	56
9.1	Κώδικας MATLAB.....	56
9.2	Διάγραμμα ροής με εικόνα.....	57

1 Εισαγωγή

1.1 Γενικά

Βιοϊατρική τεχνολογία είναι η επιστήμη η οποία εφαρμόζει τις αρχές, το σκεπτικό και τις μεθόδους της επιστήμης της βιολογίας σε συνδυασμό με τις μεθοδολογίες ανάλυσης και σύνθεσης των θετικών επιστημών (φυσική, μαθηματικά, χημεία, πληροφορική) με σκοπό την ανάλυση και επίλυση προβλημάτων στον τομέα της ιατρικής [1]. Η βιοϊατρική τεχνολογία δεν προορίζεται να αντικαταστήσει την επιστήμη της ιατρικής αλλά συνδέεται άμεσα με αυτήν. Είναι γεγονός ότι η αλματώδης ανάπτυξη και αύξηση της αποτελεσματικότητας της ιατρικής επιστήμης οφείλεται στην ανάπτυξη της βιοϊατρικής τεχνολογίας. Σκοπός της είναι να συνδράμει στην κατανόηση των βασικών αρχών λειτουργίας των βιολογικών συστημάτων και στην ανάπτυξη αποδοτικών τεχνολογιών, βασισμένων στη βιολογία, για να καλύψει ένα ευρύ φάσμα αναγκών. Οι ανάγκες αυτές μπορεί να αφορούν τους τομείς της διάγνωσης, πρόληψης και θεραπείας ασθενειών, την ανάπτυξη νέων υλικών, συσκευών και διαδικασιών ακόμα και την αντιμετώπιση ευρύτερων περιβαλλοντικών προβλημάτων.

Η βιοϊατρική τεχνολογία έχει αναμφισβήτητα αλλάξει το ιατρικό γίγνεσθαι και ενδέχεται να αλλάξει και το μέλλον της ιατρικής. Δυστυχώς μέχρι σήμερα σε πολλές περιπτώσεις η ιατρική είναι ανήμπορη να θεραπεύσει κάποιες ασθένειες καθώς έρχεται αντιμέτωπη με αυτές όταν βρίσκονται στο τελικό τους στάδιο. Αυτό μπορεί να αλλάξει μόνο με την έγκαιρη διάγνωση κάποιων παθήσεων. Το πρόβλημα αυτό όμως δεν είναι τόσο εύκολο διότι ασθένειες τέτοιου τύπου, όπως κακοήθεις νεοπλασίες, ίσως δώσουν συμπτώματα κατά το τελικό στάδιο. Επίσης δεν γίνεται λόγω κόστους αλλά και λόγω χρόνου και διαδικασίας να γίνονται εκτεταμένες εξετάσεις για όλο το σώμα σε σύντομα χρονικά διαστήματα μόνο και μόνο για πρόληψη.

Πιο συγκεκριμένα, θα γίνει αναφορά στα προβλήματα του γαστρεντερικού σωλήνα (ΓΣ). Οι ασθένειες του γαστρεντερικού σωλήνα είναι μια πολύ συνηθισμένη απειλή για την υγεία του ανθρώπου. Σύμφωνα με στατιστικές στις ΗΠΑ, κάθε χρόνο εισάγονται στα νοσοκομεία 60 με 70 εκατομμύρια ασθενείς που πάσχουν από τέτοιες ασθένειες, ενώ ξοδεύονται δισεκατομμύρια δολάρια για κόστη διάγνωσης και περίθαλψης. Ο καρκίνος του ΓΣ και ειδικότερα του παχέος εντέρου αποτελεί τον τρίτο καρκίνο σε συχνότητα για άντρες και γυναίκες και τη δεύτερη σε αιτία θανάτου από καρκίνο [2]. Κάθε χρόνο σε όλο τον κόσμο, ο καρκίνος του παχέος εντέρου διαγιγνώσκεται σχεδόν σε ένα εκατομμύριο ανθρώπους και ο αριθμός αυτός αυξάνεται με την πάροδο των χρόνων. Πέραν, όμως, των παθήσεων του παχέος εντέρου, το οποίο είναι σχετικά εύκολα προσπελάσιμο, υπάρχουν σημαντικές δυσλειτουργίες και του λεπτού εντέρου. Το λεπτό έντερο έχει χαρακτηριστεί ως το «μαύρο κουτί» του ΓΣ εξαιτίας του μεγάλου μήκους του και των πολλών περιελίξεων εντός της κοιλιακής κοιλότητας. Η ανεύρεση του αιτίου αιμορραγίας του λεπτού εντέρου πολλές φορές είναι δύσκολη, ενώ η διάγνωση παθήσεων, όπως της νόσου του Crohn, καλοηθών/κακοηθών όγκων και συνδρόμων δυσαπορρόφησης, και η διαγνωστική προσπέλαση ασθενών με χρόνια διαρροϊκό σύνδρομο συχνά καθυστερούν σημαντικά, με αποτέλεσμα επιπτώσεις στην κλινική πορεία και θεραπεία των ασθενών [3].

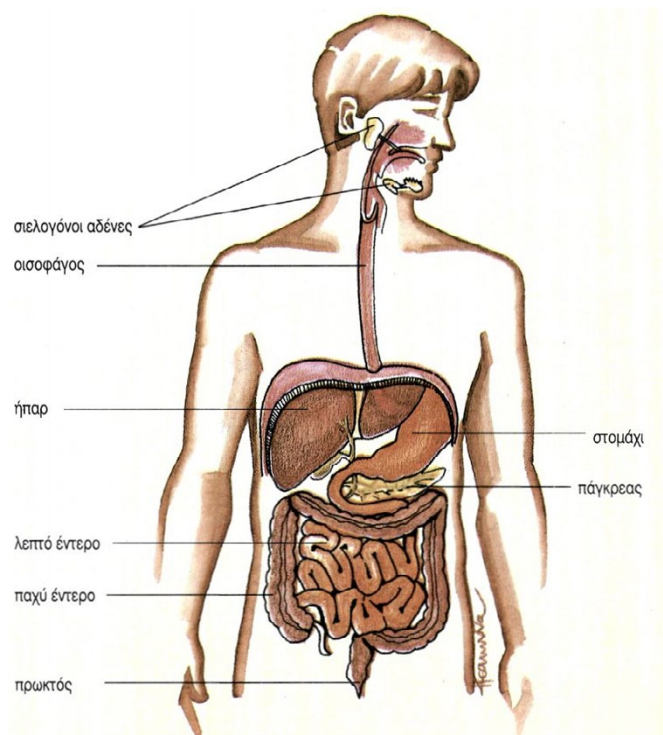
Εδώ και χρόνια η βασική μέθοδος διάγνωσης ασθενειών του ΓΣ ήταν η παραδοσιακή ενδοσκόπηση (με διάφορες τεχνικές), στην οποία ο γιατρός χειρίζεται ένα ευέλικτο ενδοσκόπιο για να εξετάσει το ανώτερο και κατώτερο άκρο του ΓΣ. Ωστόσο, οι συμβατικές

μέθοδοι είναι επεμβατικές, προκαλούν δυσφορία στον ασθενή ενώ παράλληλα έχουν περιορισμένη δυνατότητα απεικόνιση του αχανούς πεπτικού σωλήνα και κυρίως του λεπτού εντέρου. Τα τελευταία τριάντα χρόνια η ιατρική κοινότητα έγινε μάρτυρας μιας αλματώδους ανάπτυξης της τεχνολογίας στις ενδοσκοπήσεις του πεπτικού συστήματος. Η αντικατάσταση των ινοενδοσκοπίων από βιντεοσκόπια, η ανάπτυξη της ενδοσκοπικής υπερηχοτομογραφίας και της θεραπευτικής ενδοσκόπησης, η εφαρμογή της χρωμοενδοσκόπησης και η ανακάλυψη της βιντεοκάψουλας είναι η σημερινή εικόνα της σύγχρονης Γαστρεντερολογίας.

Σήμερα, η εφεύρεση της ασύρματης ενδοσκοπικής κάψουλας [4] αποτελεί ίσως τη σημαντικότερη εξέλιξη των τελευταίων ετών στον τομέα των απεικονιστικών μεθόδων. Η ενδοσκοπική κάψουλα [5] είναι μια νέα και επαναστατική μέθοδος ενδοσκοπικής προσέγγισης του πεπτικού σωλήνα που ανοίγει νέους ορίζοντες στην ενδοσκόπηση του πεπτικού συστήματος καθώς επιτρέπει μη επεμβατική, ανώδυνη και αποτελεσματική επισκόπηση όλου του μήκους του λεπτού εντέρου (από το δωδεκαδάκτυλο έως τον ειλέο) με τη λήψη λεπτομερών εικόνων από περιοχές όπου άλλες ενδοσκοπικές και ακτινολογικές μέθοδοι δεν επιτρέπουν [6]. Μπορεί να εφαρμοστεί σε παιδιά πολύ μικρής ηλικίας έως και σε επιβαρυσμένους ηλικιωμένους ασθενείς με καρδιο-αναπνευστικά προβλήματα. Δε θα ήταν υπερβολή ο ισχυρισμός ότι η ασύρματη ενδοσκόπηση είναι ίσως η πιο εύκολη και ανώδυνη εξέταση που μπορεί να υποστεί ένας ασθενής με προβλήματα του γαστρεντερικού.

1.2 Φυσιολογία του εντέρου

Το πεπτικό σύστημα είναι ένα από τα βασικότερα συστήματα του οργανισμού, τόσο του ανθρώπου, όσο και των ζώων το οποίο έχει ως ρόλο τις λειτουργίες διόδου, πέψης, και χρήσης της προσληφθείσας τροφής και αποβολής των αχώνευτων ουσιών. Το πεπτικό σύστημα αποτελείται από το γαστρεντερικό σωλήνα και από τους προσαρτημένους σ' αυτό αδένες, που είναι οι σιελογόνοι αδένες, το πάγκρεας και το ήπαρ [7] (Εικόνα 1).



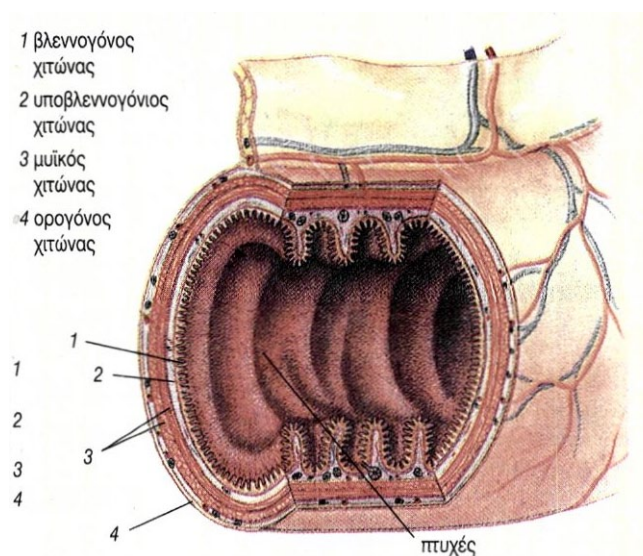
Εικόνα 1. Δομή του πεπτικού συστήματος [7].

Ο γαστρεντερικός ή πεπτικός σωλήνας είναι ένας ελικοειδής σωληνοειδής αγωγός με αφετηρία τη στοματική κοιλότητα και τερματισμό τον πρωκτό που έχει ως σκοπό της πρόσληψη τροφής, την αφομοίωσή της με σκοπό την παραγωγή ενέργειας και θρεπτικών συστατικών και τέλος, την απόρριψη των υπολειμμάτων. Σε έναν μέσο ενήλικα άνθρωπο ο ΓΣ έχει μήκος 8.5 - 9.0 μέτρα και χωρίζεται στον ανώτερο και στον κατώτερο ΓΣ. Ο ανώτερος ΓΣ αποτελείται από τα εξής:

- Στοματική κοιλότητα, από όπου γίνεται η εισαγωγή της τροφής
- Φάρυγγας ο οποίος αποτελεί την διόδο της στοματικής κοιλότητας προς τον οισοφάγο
- Οισοφάγος ο οποίος είναι ένας λεπτός μυώδης σωλήνας που οδηγεί την τροφή στο στομάχι
- Στομάχι το οποίο αποτελεί το κύριο όργανο επεξεργασίας της τροφής.

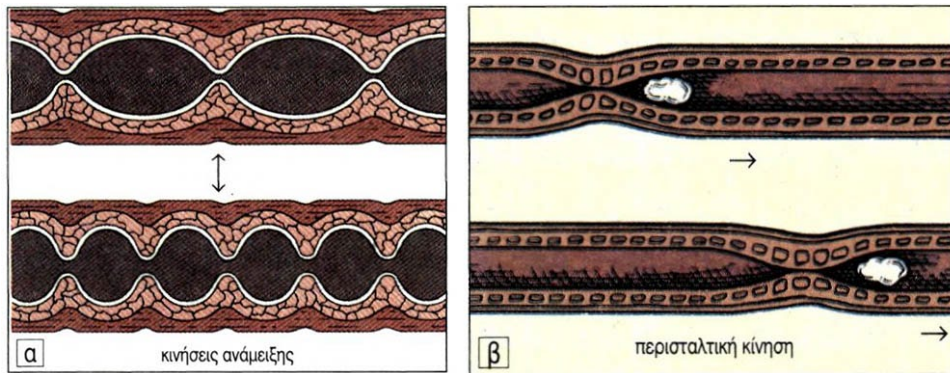
Ο κατώτερος ΓΣ αποτελείται από τα εξής:

- Δωδεκαδάκτυλος ο οποίος αποτελεί το πρώτο τμήμα του λεπτού εντέρου όπου λαμβάνει χώρα η χημική πέψη της τροφής.
- Λεπτό έντερο όπου γίνεται η απορρόφηση των θρεπτικών συστατικών της τροφής
- Παχύ έντερο όπου γίνεται απορρόφηση του νερού από την τροφή που δεν χωνεύτηκε στο λεπτό έντερο
- Πρωκτός από όπου γίνεται η απόρριψη των κοπράνων.



Εικόνα 2. Στιβάδες του γαστρεντερικού σωλήνα [7].

Το τοίχωμα του γαστρεντερικού σωλήνα αποτελείται από τέσσερις βασικές στιβάδες (βλεννογόνος χιτώνας, υποβλεννογόνιος χιτώνας, μυϊκός χιτώνας και ορογόνος χιτώνας) (Εικόνα 2) των οποίων η διάμετρος και η δομή παρουσιάζουν τοπικές διαφοροποιήσεις. Οι ανατομικές διαφοροποιήσεις των επιμέρους τμημάτων του γαστρεντερικού σωλήνα σχετίζονται με λειτουργικές διαφορές. Το κάθε τμήμα δηλαδή επιτελεί συγκεκριμένες λειτουργίες, που μπορεί να είναι κινητικές, εκκριτικές ή απορροφητικές. Με τις κινητικές λειτουργίες επιτυγχάνεται η ανάμειξη και προώθηση της τροφής κατά μήκος του γαστρεντερικού σωλήνα. Εμφανίζονται δύο κύριοι τύποι κινήσεων, οι κινήσεις ανάμειξης και οι κινήσεις προώθησης. Βασική προωθητική κίνηση είναι η περισταλτική κίνηση (Εικόνα 3).



Εικόνα 3. Κινήσεις του εντέρου. [α] κίνηση ανάμειξης [β] περισταλτική κίνηση.

1.3 Το πεπτικό έλκος

Έλκος ονομάζεται οποιαδήποτε εσωτερική ή εξωτερική πληγή αλλά ο όρος χρησιμοποιείται κυρίως για ανοιχτές πληγές στο ΓΣ. Το πεπτικό έλκος είναι μία χρόνια πάθηση του πεπτικού συστήματος και ιδιαίτερα του στομάχου και του δωδεκαδάκτυλου και πρόκειται για διάβρωση του βλεννογόνου χιτώνα. Ανάλογα με το βάθος της διάβρωσης (πληγής) καθορίζεται και η σοβαρότητα της πάθησης.

Οι κυριότεροι παράγοντες που συμβάλουν στην ανάπτυξη πεπτικού έλκους είναι ένα βακτήριο (*Helicobacter pylori*) και η διαταραχή της ισορροπίας κάποιων πεπτικών ενζύμων. Άλλα αίτια ανάπτυξης έλκους είναι ο ανθυγιεινός τρόπος ζωής (άγχος, αλκοόλ, κάπνισμα), η υπερβολική χρήση φαρμάκων (κυρίως μη στεροειδή αντιφλεγμονώδη) και άλλα νοσήματα (όπως ενδοκρινοπάθειες, νόσοι του ήπατος, πνευμονοπάθειες).

Το πεπτικό έλκος είναι μία από τις πιο συχνές παθήσεις αφού προσβάλλει περίπου το 25% των ανδρών και το 15% των γυναικών όπως αποδεικνύεται από ουλές που βρίσκονται στο στομάχι και το δωδεκαδάκτυλο σε νεκροτομικό υλικό [8]. Εν τούτοις μόνο το 5-10% των ατόμων εκδηλώνει συμπτώματα της νόσου του. Το δωδεκαδακτυλικό έλκος είναι 2-3 φορές πιο συχνό από το έλκος του στομάχου και 2-3 φορές πιο συχνό στους άνδρες απ' ό τι στις γυναίκες. Το έλκος στομάχου παρουσιάζεται με την ίδια σχεδόν συχνότητα και στα δύο φύλα. Αν και το έλκος μπορεί να εμφανισθεί σε οποιαδήποτε ηλικία, εν τούτοις τα συμπτώματα συνήθως αρχίζουν μεταξύ 20-40 ετών για τα δωδεκαδακτυλικά έλκη και 40-70 για τα έλκη στομάχου. Στα παιδιά, πιστεύεται, ότι το έλκος σπάνια εμφανίζεται αλλά πρόσφατες μελέτες έδειξαν ότι τα δωδεκαδακτυλικά έλκη δεν είναι σπάνια στα νεαρά άτομα. Τα συμπτώματα του πεπτικού έλκους περιλαμβάνουν:

- Πόνο της κοιλιακής χώρας, που συνδέεται στενά με τα γεύματα του ασθενή
- Φούσκωμα
- Ναυτία, τάση για εμετό
- Απώλεια όρεξης και βάρους, κ.α.

Στην πλειονότητα των περιπτώσεων τα έλκη από μόνα τους ή μετά από θεραπεία επουλώνονται, αλλά σπανιότερα είναι δυνατόν να παρουσιάσουν σοβαρές επιπλοκές, όπως γαστρορραγία, διάτρηση ή πυλωρική στένωση. Από τις παραπάνω, σοβαρότερη είναι η διάτρηση που μπορεί να οδηγήσει ακόμα και στο θάνατο. Συνεπώς, η έγκαιρη διάγνωση του προβλήματος και η θεραπεία του κρίνεται απαραίτητη.

1.4 Ασύρματη ενδοσκόπηση (ΑΕ)

Η ενδοσκοπική απεικόνιση του λεπτού εντέρου, αποτελούσε ανέκαθεν ένα δύσκολο πρόβλημα για τους γαστρεντερολόγους, καθώς οι παλαιές μέθοδοι διερεύνησής του με τη χρήση ακτινοσκοπικού υγρού και ακτινολογικό έλεγχο, δεν επαρκούσαν για την πλήρη και λεπτομερή εξέτασή του. Το μήκος του λεπτού εντέρου (περίπου 8 μέτρα) και η ελικοειδής ανατομική του θέση μέσα στην κοιλιακή χώρα, καθιστούσαν δυνατή την ενδοσκοπική απεικόνιση μόνο του δωδεκαδακτύλου (του πρώτου τμήματος του λεπτού εντέρου, μετά το στομάχι) και του τελικού ειλεού (του τελευταίου τμήματος του λεπτού εντέρου, στην ένωση του με το παχύ) με τα υπάρχοντα συμβατά ενδοσκοπικά μέσα.

Το 2000 όμως, μια καινούργια τεχνική, αυτή της ασύρματης ενδοσκόπησης με κάψουλα (ΑΕ), εισήχθη στην ευρεία κλινική πρακτική, ανοίγοντας καινούργιους δρόμους στο τομέα της γαστρεντερολογίας. Το διαγνωστικό σύστημα της ΑΕ αναπτύχθηκε από την εταιρία Given Imaging Ltd, το οποίο έγινε γνωστό στο ευρύ κοινό αρχικά με την ονομασία M2A (η οποία προέρχεται από το Mouth to Anus) και στη συνέχεια με τη ονομασία Pillcam SB™ (Εικόνα 4) είναι ένα πρωτοποριακό διαγνωστικό μέσο, μη επεμβατικό, εύκολο στην εκτέλεσή του και παρέχει βελτιωμένο επίπεδο απεικόνισης για την πρόιμη ανίχνευση και διάγνωση παθήσεων σε όλο το μήκος του λεπτού εντέρου. Το σύστημα της ΑΕ αποτελείται από 3 βασικά μέρη: την κάψουλα, το σύστημα ψηφιακής καταγραφής δεδομένων και το λογισμικό επεξεργασίας δεδομένων.



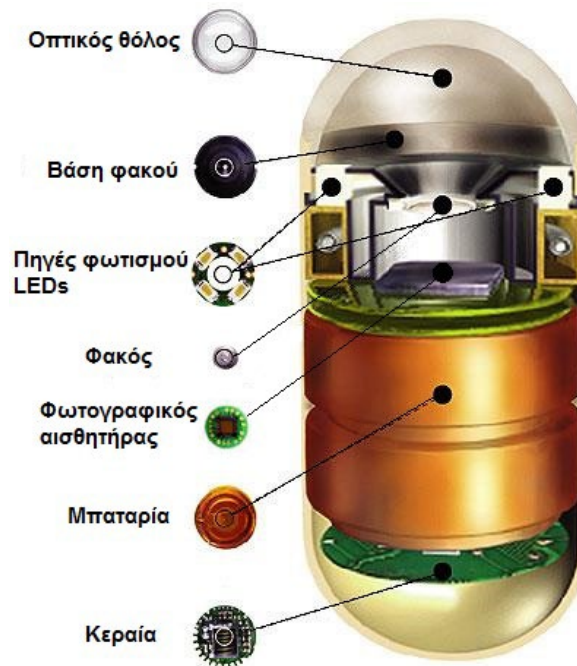
Εικόνα 4. Η πρώτη ασύρματη ενδοσκοπική κάψουλα M2A και Pillcam SB™.

1.4.1 Δομή και λειτουργία της κάψουλας

Η ασύρματη κάψουλα έχει διαστάσεις 11mm διάμετρο και 27mm μήκος σε σχήμα μακρόστενου χαπιού ή μεγάλης βιταμίνης, ζυγίζει μόλις 3.7gr και αποτελείται από (Σχήμα 5):

- Πλαστικό περίβλημα με διαφανές πρόσθιο τοίχωμα
- Οπτικό θόλο (αυτοκαθαριζόμενος από το βλεννογόνο του εντέρου) με σχήμα που εμποδίζει την ανάκλαση του φωτός από τις οπτικές πηγές
- Σύστημα ασφαιρικών φακών μικρής εστιακής απόστασης
- Πηγές φωτισμού τύπου LED
- Ψηφιακό αισθητήρα καταγραφής εικόνων (τύπου CMOS)
- Μπαταρίες

- Αναμεταδότη ASIC και κεραία εκπομπής RF.



Εικόνα 5. Δομή της κάψουλας Pillcam SB™.

Η κάψουλα είναι μιας χρήσεως και καταπίνεται από τον ασθενή με τη βοήθεια μικρής ποσότητας νερού. Η κίνησή της γίνεται με τη φυσιολογική περισταλτική κίνηση του εντέρου χωρίς να απαιτείται καταστολή του ασθενή ή εμφύσηση αέρα που αυξάνει την ενδοαυλική πίεση και προξενεί δυσφορία και πόνο. Κατά τη διάρκεια της κίνησής της, η κάψουλα λαμβάνει φωτογραφίες του τοιχώματος του ΓΣ με συχνότητα 2 εικόνες το δευτερόλεπτο. Η καταγραφή ξεκινάει με το που αποκολληθεί η κάψουλα από τη συσκευασία της και έχει συνολική διάρκεια 8 ώρες περίπου, όσο είναι δηλαδή και η διάρκεια της μπαταρίας. Προφανώς, αυτές τις ώρες ο ασθενής δεν χρειάζεται να παραμένει στο ιατρείο αλλά συνεχίζει κανονικά τις δραστηριότητές του. Οι εικόνες μεταδίδονται μέσω ραδιοκυμάτων σε ειδικούς αισθητήρες που είναι προσκολλημένοι πάνω στο σώμα του ασθενή και συνδέονται με το σύστημα καταγραφής όπου γίνεται η αποθήκευση των εικόνων. Τέλος, η κάψουλα αποβάλλεται με τα κόπρανα συνήθως εντός διαστήματος 24-72 ωρών.

1.4.2 Σύστημα ψηφιακής καταγραφής δεδομένων

Οι εικόνες που καταγράφονται κατά τη διάρκεια του ταξιδιού της κάψουλας δεν αποθηκεύονται εντός της κάψουλας αλλά αποστέλλονται ασύρματα και διαμέσου 8 ειδικών αισθητήρων (SensorArray™) (Εικόνα 6) μεταφέρονται και καταγράφονται σε ειδικό καταγραφέα δεδομένων (DataRecorder) τύπου walkman (Εικόνα 6). Η ειδική διάταξη των αισθητήρων τοποθετείται στην πρόσθια κοιλιακή χώρα του ασθενή. Στη συνέχεια, οι αισθητήρες συνδέονται με τον καταγραφέα δεδομένων και αυτός με τη σειρά του με ειδική μπαταρία τροφοδοσίας (διάρκειας 8 - 9 ωρών). Ο καταγραφέας και η μπαταρία τοποθετούνται πάνω σε μια ειδική ζώνη, γύρω από τη μέση των ασθενών. Μετά το τέλος της εξέτασης, ο καταγραφέας συνδέεται με τον ηλεκτρονικό υπολογιστή ο οποίος με τη βοήθεια

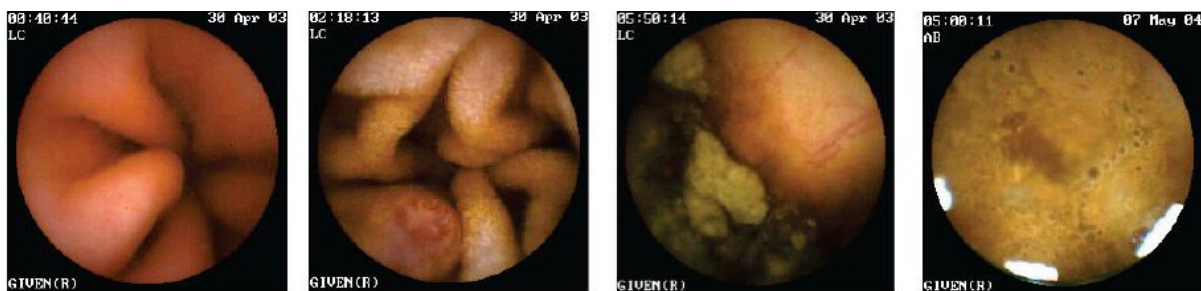
ειδικού λογισμικού δίνει τη δυνατότητα στον ιατρό να αναλύσει και να αναπαράγει τις εικόνες με τη μορφή video.



Εικόνα 6. Σύστημα αισθητήρων (αριστερά) και καταγραφέας δεδομένων ενσωματωμένος στην ειδική ζώνη (δεξιά).

1.4.3 Λογισμικό επεξεργασίας δεδομένων

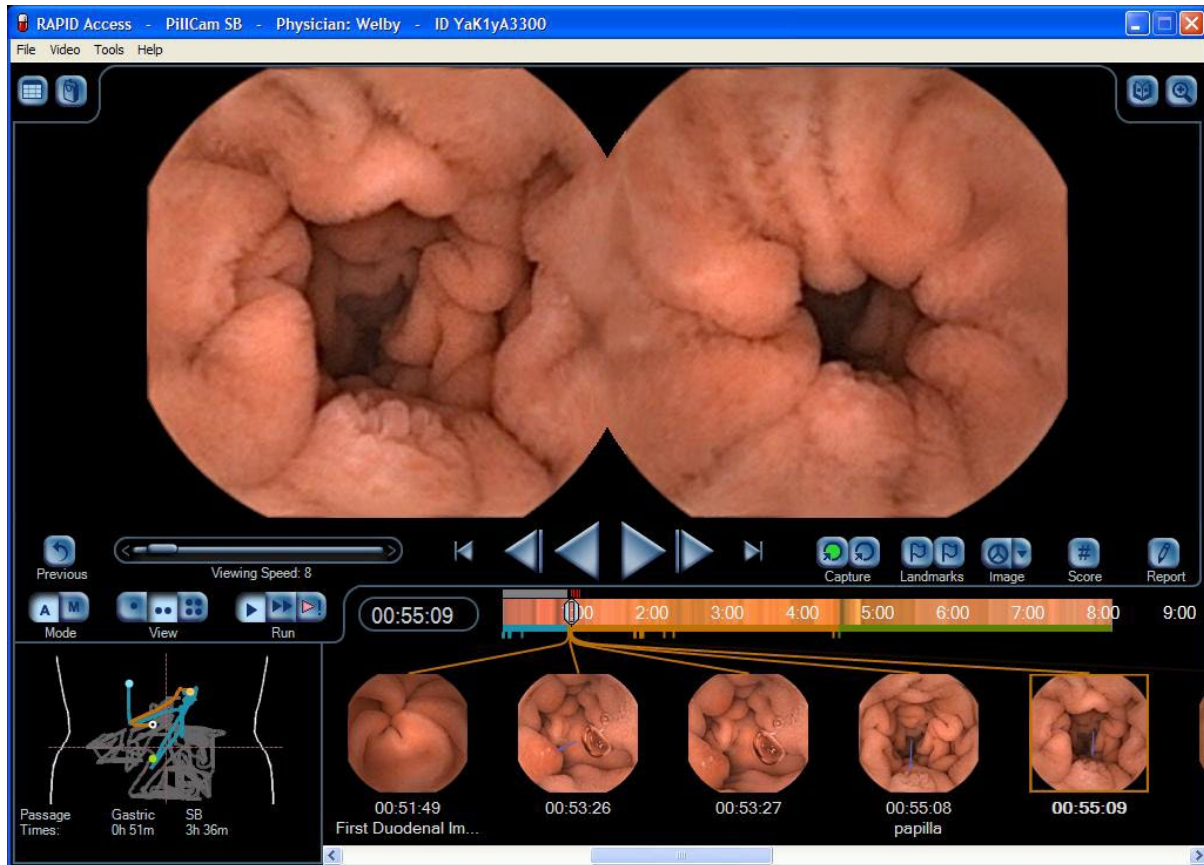
Το ειδικό λογισμικό που συνοδεύει το σύστημα ΑΕ της εταιρίας Given ονομάζεται RAPID και δίνει τη δυνατότητα μεταφόρτωσης των εικόνων από τον καταγραφέα στον ηλεκτρονικό υπολογιστή, επισκόπησης αυτών με τη μορφή video μεταβαλλόμενης ταχύτητας και αποθήκευσης-εξαγωγής εικόνων με τυχόν ευρήματα για περαιτέρω επεξεργασία σε δεύτερο χρόνο. Το υλικό αποτελείται από 55000-60000 εικόνες υψηλής ευκρίνειας (320x320 pixel). Ένα παράδειγμα από τέτοιες εικόνες δίνεται στην Εικόνα 7. Το ειδικό λογισμικό έχει επίσης την ικανότητα του κατά προσέγγιση προσδιορισμού της θέσης της κάψουλας στο λεπτό έντερο καθώς και τη δυνατότητα ανίχνευσης ενεργού αιμορραγίας στη διαδρομή της κάψουλας στον πεπτικό σωλήνα. Στην Εικόνα 8 φαίνεται το παράθυρο του λογισμικού RAPID.



Εικόνα 7. Παράδειγμα εικόνων από σύστημα ΑΕ της εταιρίας Given.

1.4.4 Μελέτη των δεδομένων καταγραφής

Η ανάγνωση του video της εξέτασης θα πρέπει να γίνεται από ιατρούς με αρκετή εμπειρία στην εκτίμηση ενδοσκοπικών εικόνων. Ο χρόνος που απαιτείται για τη μελέτη των δεδομένων καταγραφής είναι αρκετός και πιθανά αποτελεί έναν από τους λόγους περιορισμού της ευρείας χρήσης της. Πιο συγκεκριμένα, οι 55000-60000 εικόνες που καταγράφονται απαιτούν από 45 λεπτά έως 2.5 ώρες για την προσεκτική εξέτασή τους ανάλογα με την ταχύτητα απεικόνισής η οποία ποικίλει από 2 έως 25 εικόνες το δευτερόλεπτο. Μια μέση ταχύτητα απεικόνισης με συμβιβασμό ανάμεσα στη συνολική διάρκεια και την αποτελεσματικότητα και ασφάλεια στην εξαγωγή συμπεράσματος



Εικόνα 8. Το λογισμικό RAPID της εταιρίας Given.

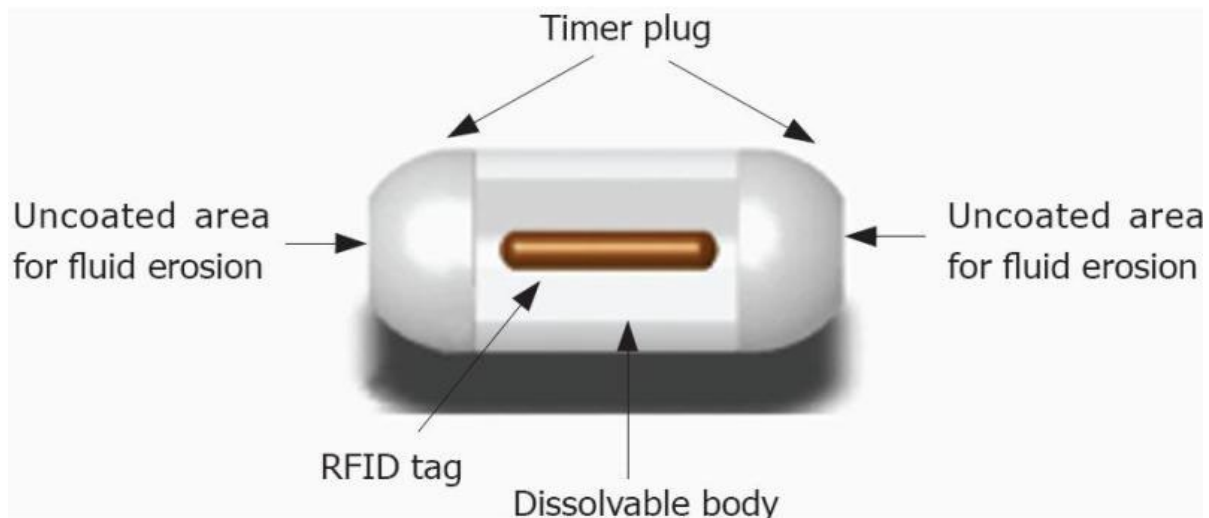
θεωρείται η ταχύτητα των 10 εικόνων το δευτερόλεπτο. Η ταχύτητα αυτή μπορεί να ακούγεται μικρή αλλά πρέπει να τονιστεί ότι η μελέτη των εικόνων προϋποθέτει τη συνεχή εγρήγορση και εκτεταμένη συγκέντρωση του ιατρού καθώς ένα παθολογικό εύρημα μπορεί να εμφανίζεται μόνο σε μία εικόνα ή τα χαρακτηριστικά του (χρώμα, μέγεθος, σχήμα) να είναι τέτοια που να μην επιτρέπουν την ξεκάθαρη αξιολόγησή του. Η αναπόσπαστη προσοχή είναι ένα αρκετά δύσκολο έργο για τον ιατρό δεδομένου των συνεχόμενων οπτικών ερεθισμάτων και της μεγάλης διάρκειας της εξέτασης.

1.4.5 Περιορισμοί της ΑΕ

Ο κυριότερος κίνδυνος από τη χρήση της ΑΕ είναι η κατακράτηση της κάψουλας. Κατακράτηση κάψουλας ορίζεται η παραμονή της κάψουλας στο πεπτικό σύστημα για τουλάχιστον 2 εβδομάδες. Σε τέτοια περίπτωση η κατακράτηση θεωρείται μόνιμη και απαιτείται χειρουργική επέμβαση για την αφαίρεσή της. Από τη μέχρι σήμερα εμπειρία, το ποσοστό μη διόδου της κάψουλας παγκοσμίως κυμαίνεται γύρω στο 1%. Στενώσεις του λεπτού εντέρου σε ασθενείς με νόσο Crohn, στενώσεις μετά από λήψη μη στεροειδών αντιφλεγμονωδών φαρμάκων, μετακτινικές και μετεγχειρητικές στενώσεις, αποτελούν τις κυριότερες αιτίες που οδήγησαν στη μη αποβολή της κάψουλας. Άλλες αντενδείξεις της μεθόδου πέραν της απόφραξης ή στένωσης του πεπτικού σωλήνα είναι:

- Διαταραχές κατάποσης
- Εγκυμοσύνη
- Παρουσία βηματοδότη ή άλλων βιοιατρικών συσκευών

Η διενέργεια αξονικής τομογραφίας και εντερόκλυσης πριν τη χορήγηση της κάψουλας εξασφαλίζουν σε σημαντικό βαθμό τη βατότητα του πεπτικού σωλήνα. Τέτοιες εξετάσεις όμως δεν είναι πρακτικές και έχουν σημαντικό κόστος. Για το λόγο αυτό δημιουργήθηκε από την εταιρία Given ένα σύστημα ελέγχου της βατότητας του πεπτικού συστήματος. Το "M2A™ Patency System" (Εικόνα 9) έχει ως σκοπό τη διαπίστωση στενώσεων εντός του πεπτικού σωλήνα που θα εμπόδιζαν τη δίοδο της κάψουλας. Στην ουσία πρόκειται για μια κάψουλα ίδιων ακριβώς διαστάσεων με την ενδοσκοπική κάψουλα με τη διαφορά ότι είναι κατασκευασμένη από υλικά που αποδομούνται πλήρως από τα γαστρεντερικά υγρά μετά την πάροδο μιας εβδομάδας. Η κάψουλα βατότητας έχει πολύ καλά αποτελέσματα και χρησιμοποιείται σχεδόν συστηματικά σε ασθενείς με νόσο του Crohn εξασφαλίζοντας την ασφαλή διενέργεια της ΑΕ.



Εικόνα 9. Κάψουλα βατότητας "M2A Patency Capsule".

1.5 Βιβλιογραφική έρευνα

Όπως γίνεται αντιληπτό από τα παραπάνω, υπάρχει επιτακτική ανάγκη να βρεθεί μια λύση στη χρονοβόρα και επισφαλή διαδικασία της χειροκίνητης εξέτασης των χιλιάδων εικόνων, τόσο για να διευκολυνθεί ο γιατρός, όσο και για να επιταχυνθεί η διαδικασία και να μειωθεί το κόστος της. Επίσης, μπορεί να υπάρχουν ανωμαλίες που δεν ανιχνεύονται από το γυμνό μάτι εξαιτίας το υ μεγέθο η, το υ χρώματο ς ή της υφής τους. Τα παραπάνω μειονεκτήματα έχουν κινητοποιήσει ερευνητικές δραστηριότητες για εύρεση λύσεων και την ανάπτυξη ολοκληρωμένων συστημάτων υποβοηθούμενης ή αυτόματης διάγνωσης. Οι περιοχές ερευνητικού ενδιαφέροντος περιλαμβάνουν αυτόματη κατάτμηση των εικόνων, αποθορυβοποίηση και βελτιστοποίηση των δεδομένων, αυτόματη ανίχνευση αιμορραγίας και αυτόματη ανίχνευση ανωμαλιών (όπως όγκος, έλκος, πολύποδας κτλ).

Η εταιρία Given, όπως αναφέρθηκε και παραπάνω, έχει ενσωματώσει στο λογισμικό της (RAPID) τη δυνατότητα της αυτοματοποιημένης ανίχνευσης ενεργούς αιμορραγίας και της θέσης της. Παρ' όλα αυτά όμως η ευαισθησία και η ειδικότητα του συστήματος δεν ξεπερνά το πολύ μικρό 21.5% και 41.8% αντίστοιχα. Με την ανίχνευση ενεργού αιμορραγίας ασχολήθηκαν και οι εργασίες [9] και [10] χρησιμοποιώντας τον μετασχηματισμό φάσματος χρώματος (colour spectrum) και χρωματικά χαρακτηριστικά υφής (chromaticity moments), αντίστοιχα. Επίσης, οι Sainju και Wahid [11] πρότειναν αλγόριθμο που εξάγει

χαρακτηριστικά χρώματος υπολογίζοντας τη μέση τιμή, την τυπική απόκλιση, τη λοξότητα και την ενέργεια από το ιστόγραμμα των τριών χρωματικών συνιστωσών RGB με σκοπό την ανίχνευση περιοχών αιμορραγίας.

Ο Cauvin [12], [13] δημιούργησε έναν έξυπνο άτλαντα με ενδοσκοπικά ευρήματα που χρησιμοποιούνται σαν αναφορά για αυτοματοποιημένη διάγνωση. Η μέθοδός τους βασίζεται σε ένα μέτρο ομοιότητας που εκτιμά τον βαθμό ομοιότητας μεταξύ του υπο εξέταση δείγματος και των εικόνων του άτλαντα. Στην εργασία του Καρκάνη [14] έγινε προσπάθεια αναγνώρισης όγκων χρησιμοποιώντας χαρακτηριστικά από τα χρωματικά κυματίδια (Color Wavelets). Οι Coimbra και Cunha [15] εκμεταλλεύτηκαν τα χαρακτηριστικά του προτύπου MPEG-7 για την αναγνώριση παθολογικών χαρακτηριστικών ενώ ο Xiaoying Liu [16] χρησιμοποίησαν μια μέθοδο που βασίζεται στη συνδιακύμανση στατιστικών χαρακτηριστικών δεύτερης τάξης (color wavelet covariance) για τον εντοπισμό περιοχών έλκους. Οι Καραγύρης και Μπουρμπάκης [17] ανέπτυξαν μέθοδο για την αποκοπή και αναγνώριση περιοχών έλκους χρησιμοποιώντας φίλτρα Log-Gabor. Ο Vieira [18] πρότεινε μια προσέγγιση βασισμένη στη μέθοδο Maximum A Posteriori για την αναγνώριση περιοχών με υποψία παθογόνων ευρημάτων συγκρίνοντας τις τιμές των εικονοστοιχείων της εικόνας και από τα τρία χρωματικά κανάλια ταυτόχρονα. Στη εργασία [19] ο Vilarino ασχολήθηκε με την ανίχνευση και απόρριψη εικόνων άχρηστων για την εξέταση εξαιτίας της απεικόνισης εντερικών υγρών χρησιμοποιώντας φίλτρα Gabor και αλγόριθμους εύρεσης φυσαλίδων. Επίσης, οι Τσέβας και Ιακωβίδης [20], [21] ασχολήθηκαν με την μείωση του όγκου των εικόνων που απαιτείται να εξετάσει ο ιατρός χρησιμοποιώντας τεχνικές αναγνώρισης και ομαδοποίησης όμοιων εικόνων. Ο Dhandra [22] ασχολήθηκε με την αναγνώριση ανωμαλιών σε ενδοσκοπικές εικόνες χρησιμοποιώντας τον αλγόριθμο κατάτμησης Watershed και τεχνικές εξομάλυνσης και φιλτραρίσματος στο χρωματικό χώρο Hue-Saturation-Intensity.

1.6 Σύντομη περιγραφή της παρούσας εργασίας

Σκοπός τη παρούσας πτυχιακής εργασίας είναι η επεξεργασία εικόνων του λεπτού εντέρου που προέρχονται από ασύρματη ενδοσκόπηση ώστε να επιτευχθεί αυτόματη αναγνώριση περιοχών που εμφανίζουν έλκος. Η επεξεργασία των εικόνων θα βασιστεί στη μεθοδολογία που προτείνεται στην εργασία των B. Li και M.Q.H. Meng [23] για την αναγνώριση ελκώδους ιστού από αντίστοιχες εικόνες ενώ παράλληλα γίνεται προσπάθεια περαιτέρω διερεύνησης και βελτίωσης. Στην εργασία αυτή οι εικόνες επεξεργάζονται στο χρωματικό χώρο YCbCr ενώ στην παρούσα πτυχιακή θα διερευνηθούν και οι χώροι HSV και RGB. Αρχικά, στις εικόνες εφαρμόζεται ο μετασχηματισμός Curvelet για να διευθυνθεί, στη συνέχεια, η εξαγωγή χαρακτηριστικών υφής χρησιμοποιώντας τη θεωρία των Τοπικών Δυαδικών Προτύπων (Local Binary Patterns). Τέλος, οι εικόνες ταξινομούνται σε φυσιολογικές και εικόνες με έλκος χρησιμοποιώντας τις Μηχανές Διανυσμάτων Υποστήριξης (Support Vector Machines) ενώ σαν επέκταση η παρούσα πτυχιακή θα εξετάσει πως επηρεάζεται το αποτέλεσμα της ταξινόμησης ανάλογα με τις παραμέτρους του ταξινομητή.

2 Αναπαράσταση εικόνων στον υπολογιστή

2.1 Γενικά

Η εικόνα είναι μια από τις πιο σημαντικές πηγές πληροφορίας την οποία χρησιμοποίησαν οι άνθρωποι από τα πρώτα χρόνια της ύπαρξής τους. Στη σημερινή εποχή η ψηφιακή εικόνα έχει ενταχθεί σχεδόν σε κάθε πτυχή της καθημερινότητας μας. Στην ιατρική το μεγαλύτερο ποσοστό των διαγνώσεων βασίζονται στην παρατήρηση και εξέταση εικόνων που καταγράφονται από κατάλληλες συσκευές οι οποίες δεν είναι τίποτα άλλο από εξειδικευμένες φωτογραφικές μηχανές. Φυσικά, οι διαφορές τους είναι τεράστιες αλλά ο απώτερος στόχος όλων είναι η απεικόνιση κάποιου μέρους του σώματος.

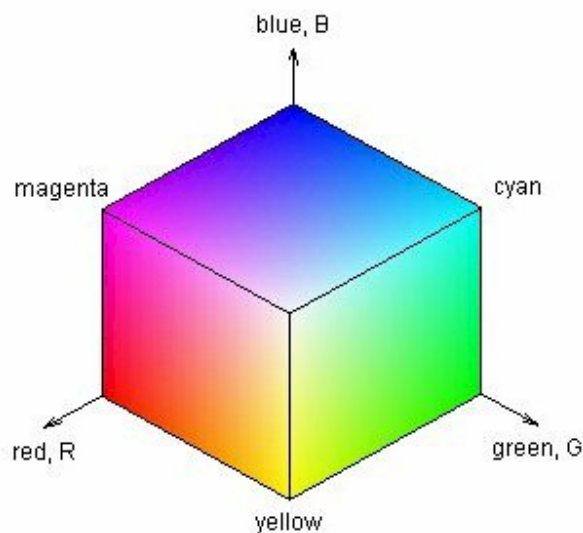
Μια πραγματική εικόνα μεταφέρεται στον ψηφιακό κόσμο με τη μορφή διακριτού σήματος που έχει τη μορφή ψηφιακών πινάκων. Μια ψηφιακή εικόνα μπορεί να είναι δυαδική, μονοχρωματική αποχρώσεων του γκρι ή έγχρωμη. Χαρακτηριστικό στοιχείο μιας ψηφιακής εικόνας είναι το πλήθος των εικονοστοιχείων (pixel) τα οποία την απαρτίζουν. Το κάθε εικονοστοιχείο (i,j) μιας εικόνας $I(i,j)$ με $i=1,\dots,M$ και $j=1,\dots,N$ διαστάσεων $M \times N$ εκφράζει την ένταση του φωτός που έχει καταγραφεί στο συγκεκριμένο σημείο της εικόνας. Με άλλα λόγια, η τιμή του $I(i,j)$ είναι ανάλογη της φωτεινότητας της εικόνας στο εικονοστοιχείο (i,j) . Κατή τη χρήση και αποθήκευση εικόνων στον υπολογιστή κάθε τέτοιο εικονοστοιχείο αποθηκεύεται σε μια θέση ενός πίνακα δύο διαστάσεων (πίνακας I) όπου το μέγεθος του πίνακα ισούται με το μέγεθος της εικόνας. Μια μονοχρωματική εικόνα αποτελείται από έναν πίνακα $I(i,j)$ που εκφράζει το πόσο φωτεινό ή σκοτεινό είναι ένα εικονοστοιχείο στη κλίμακα του γκρι. Σε αντίθεση, μια έγχρωμη εικόνα μπορεί να θεωρηθεί ότι συντίθεται από τρεις επιμέρους εικόνες κάθε μια από τις οποίες αντιστοιχεί σε μια χρωματική συνιστώσα. Επομένως, το χρώμα κάθε εικονοστοιχείου αναπαριστάται από ένα διάνυσμα τριών τιμών και προκύπτει από το συνδυασμό τριών χρωματικών συνιστωσών, $I_{\text{color}}(i,j) = [I_a(i,j), I_b(i,j), I_c(i,j)]$. Για παράδειγμα, στο χρωματικό χώρο RGB το κάθε χρώμα συντίθεται από τα χρώματα κόκκινο (Red), πράσινο (Green) και μπλε (Blue) με αποτέλεσμα κάθε εικονοστοιχείο να αποτελείται από τρεις τιμές (ένα διάνυσμα $[R \ G \ B]$), μία για κάθε βασικό χρώμα, οι οποίες εκφράζουν την ένταση (το ποσοστό) του κάθε χρώματος που υπάρχει στο συγκεκριμένο σημείο. Επομένως, μπορούμε να συμπεράνουμε ότι ο πίνακας στον οποίο αποθηκεύεται μια εικόνα στον υπολογιστή δεν είναι δύο αλλά τριών διαστάσεων της μορφής $M \times N \times 3$, όπου 3 είναι οι συνιστώσες του κάθε εικονοστοιχείου. Άρα, ο δισδιάστατος πίνακας I_a αναφέρεται στην πρώτη (κόκκινη) συνιστώσα, ο I_b αναφέρεται στη δεύτερη (πράσινη) συνιστώσα και ο I_c στην τρίτη (μπλε) συνιστώσα.

2.2 Χρωματικοί χώροι

Ο σκοπός της χρήσης χρωματικών χώρων είναι να διευκολύνεται ο ορισμός των χρωμάτων στα πλαίσια της τυποποίησης [24]. Στην ουσία, ο χρωματικός χώρος ή αλλιώς χρωματικό μοντέλο είναι ένας τρισδιάστατος χώρος όπου κάθε χρώμα αναπαρίσταται από ένα σημείο. Τα περισσότερα χρωματικά μοντέλα είναι προσαρμοσμένα στα φυσικά εξαρτήματα (π.χ. οθόνες, εκτυπωτές, αισθητήρες ψηφιακών μηχανών) ή σε εφαρμογές που είναι επιθυμητή η διαχείριση των χρωμάτων (π.χ. graphic animation). Στην πρώτη κατηγορία ανήκει ο χρωματικός χώρος RGB (Red-Green-Blue) ενώ στη δεύτερη ανήκουν οι χρωματικοί χώροι HSV (Hue-Saturation-Value) και YCbCr [24].

2.2.1 Χρωματικός χώρος RGB

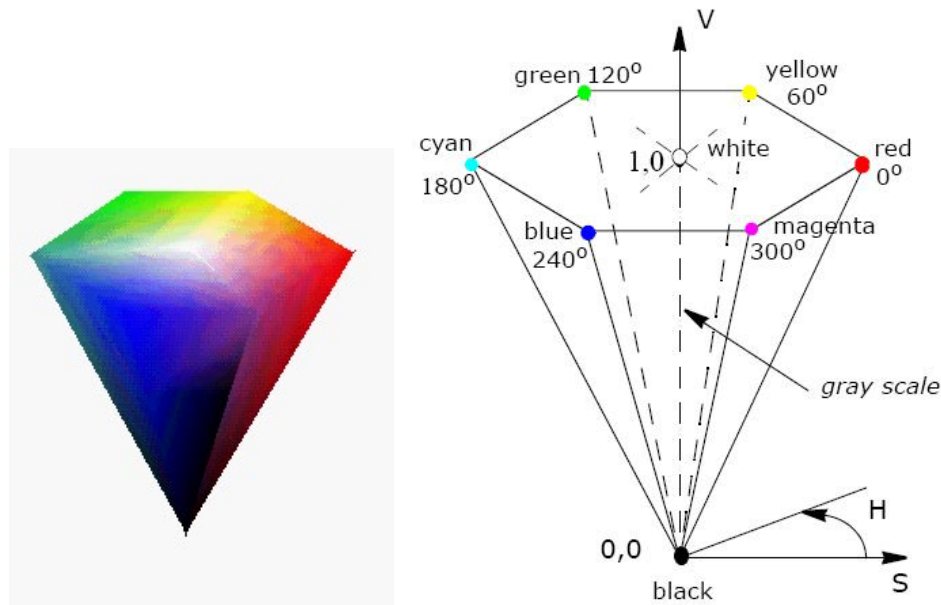
Ο RGB είναι ίσως ο πιο διαδεδομένος χρωματικός χώρος για την αναπαράσταση ψηφιακών εικόνων καθώς χρησιμοποιείται στις οθόνες και στους αισθητήρες καταγραφής ψηφιακών εικόνων (CCD, CMOS). Ο RGB είναι ο εγγενής χρωματικός χώρος στον οποίο καταγράφονται οι εικόνες από τους κοινούς φωτογραφικούς αισθητήρες που χρησιμοποιούνται στην ασύρματη ενδοσκοπική κάψουλα αλλά και σε κάθε είδους φωτογραφική μηχανή. Ο RGB είναι ένας "αθροιστικός" χώρος στον οποίο κάθε χρώμα προκύπτει από το συνδυασμό των τριών προκαθορισμένων χρωμάτων (κόκκινο, πράσινο, μπλε). Η ένταση του φωτός είναι που καθορίζει την αντίληψη των χρωμάτων. Μηδενική ένταση σημαίνει ότι κάθε χρώμα γίνεται αντιληπτό ως μαύρο, ενώ πλήρης ένταση οδηγεί σε αντίληψη λευκού χρώματος. Στο RGB οι διαφορετικές εντάσεις των τριών χρωμάτων παράγουν την απόχρωση του κάθε χρώματος. Στην Εικόνα 10 φαίνεται η χρωματική κατανομή των χρωμάτων σε κύβο RGB. Παρά τα πλεονεκτήματα που προσφέρει ο RGB (απλότητα) δεν παύει να είναι ένας χρωματικός χώρος προσανατολισμένος στο υλικό (hardware-oriented) ο οποίος δε συνάδει με τον τρόπο που αντιλαμβάνεται ο ανθρώπινος εγκέφαλος το χρώμα. Παρά το γεγονός ότι οι RGB κάμερες είναι εμπνευσμένες από το σύστημα της ανθρώπινης όρασης και εμφανίζουν πολλές ομοιότητες με αυτό, κανείς δε θα χρησιμοποιούσε τα ποσοστά του κόκκινου, του πράσινου και του μπλε για να περιγράψει ένα χρώμα.



Εικόνα 10. Ο χρωματικός χώρος RGB.

2.2.2 Χρωματικός χώρος HSV

Ο χρωματικός χώρος HSV επινοήθηκε το 1978 από τον Smith [25] και είναι ένας από τους χρωματικούς χώρους που χρησιμοποιείται από τους ανθρώπους για την επιλογή χρωμάτων (π.χ. μελάνια, μπογιές κτλ) από ένα δίσκο χρωμάτων ή μια παλέτα καθώς ο συγκεκριμένος χρωματικός χώρος είναι πιο κοντά στην ανθρώπινη αντίληψη του χρώματος. Το σύστημα συντεταγμένων στον HSV είναι κυκλικό και τα χρώματα βρίσκονται μέσα σε έναν εξάγωνο



Εικόνα 11. Ο χρωματικό χώρος HSV.

κόνο (Εικόνα 11). Οι παράμετροι που χρησιμοποιεί το HSV είναι η απόχρωση (Hue), ο κορεσμός (Saturation) ή αλλιώς χρωματική καθαρότητα που εκφράζει το βαθμό μίξης ενός καθαρού χρώματος με το λευκό και η τιμή (Value) που αναφέρεται στο βαθμό μείξης ενός καθαρού χρώματος με το μαύρο. Το H είναι η γωνία στο επίπεδο του εξαγώνου και παίρνει τιμές από 0° έως 360° , το V είναι ο κατακόρυφος άξονας και παίρνει τιμές από 0 έως 1 (όταν $V=0$ τότε το χρώμα είναι μαύρο, ενώ όταν $V=1$ τότε το χρώμα είναι άσπρο) ενώ το S είναι η απόσταση από τον κατακόρυφο άξονα V και παίρνει τιμές από 0 έως 1. Από μαθηματική σκοπιά οι παρακάτω σχέσεις εκφράζουν τη σχέση μεταξύ του χώρου HSV και RGB.

$$H_1 = \cos^{-1} \left\{ \frac{0.5[(R-G) + (R-B)]}{\sqrt{(R-G)^2 + (R-B)(G-R)}} \right\}$$

$$H = \begin{cases} H_1, & \text{εάν } B \leq G \\ 360^\circ - H_1, & \text{εάν } G < B \end{cases}$$

$$S = \frac{\max(R, G, B) - \min(R, G, B)}{\max(R, G, B)}$$

$$V = \frac{\max(R, G, B)}{255}$$

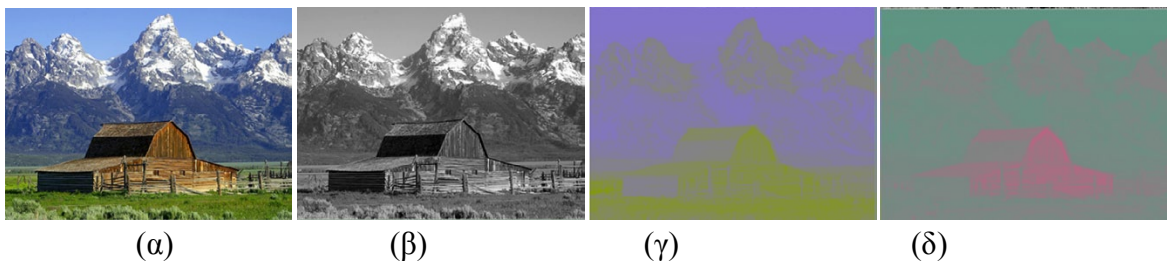
2.2.3 Χρωματικός χώρος YCbCr

Ο χρωματικός χώρος YCbCr δημιουργήθηκε για να καλύψει τις ανάγκες της ψηφιακής επεξεργασίας video καθώς είναι ένα μοντέλο το οποίο ανήκει στην οικογένεια των συστημάτων που χρησιμοποιούνται στην τηλεοπτική μετάδοση. Η δημιουργία του κάλυψε την ανάγκη προσθήκης της πληροφορίας του χρώματος στο μονοχρωματικό (κλίμακας του γκρι) σήμα. Στο YCbCr η συνιστώσα Y περιγράφει τη φωτεινότητα ενώ οι συνιστώσες Cb και Cr περιγράφουν τις πληροφορίες χρώματος. Πιο συγκεκριμένα, το Cb εκφράζει τη

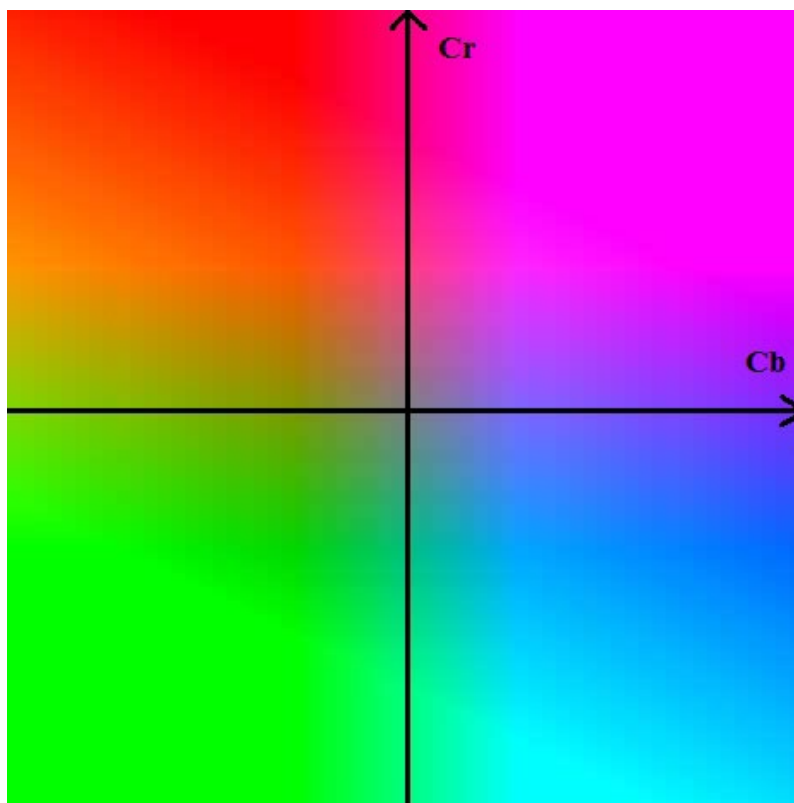
χρωματική διαφορά ως προς μπλε και το Cr εκφράζει τη χρωματική διαφορά ως προς το κόκκινο. Η μετατροπή του RGB χρωματικού χώρου στο YCbCr γίνεται ως εξής:

$$\begin{bmatrix} Y \\ C_b \\ C_r \end{bmatrix} = \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} + \begin{bmatrix} 65.481 & 128.553 & 24.966 \\ -37.797 & -74.203 & 112 \\ 112 & -93.786 & -18.214 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

Στην παραπάνω σχέση οι τιμές R, G, B θεωρούνται ότι βρίσκονται στο διάστημα [0,1] με αποτέλεσμα το Y να κυμαίνεται στο διάστημα [16,235] και τα Cb, Cr στο διάστημα [16,240]. Στην Εικόνα 12α φαίνεται μια RGB εικόνα αναλυμένη στις τρεις συνιστώσες του YCbCr ενώ στην Εικόνα 12β φαίνεται το επίπεδο CbCr για συγκεκριμένη τιμή στο Y.



Εικόνα 12α. Ο χρωματικές συνιστώσες του χώρου YCbCr: (α) Αρχική RGB εικόνα, (β) Y συνιστώσα, (γ) Cb συνιστώσα, (δ) Cr συνιστώσα.



Εικόνα 12β. Το επίπεδο CbCr για σταθερή τιμή Y=0.5.

3 Προεπεξεργασία των εικόνων

Είναι σύνθηρες φαινόμενο στον τομέα της Αναγνώρισης Προτύπων και της οπτικής αναγνώρισης αντικειμένων πριν την εξαγωγή των χαρακτηριστικών υφής, χρώματος, σχήματος, ακμών κλπ να προηγείται ένα στάδιο προεπεξεργασίας των δεδομένων. Το στάδιο της προεπεξεργασίας είναι σημαντικό ώστε τα εξαγόμενα χαρακτηριστικά να είναι πιο αποτελεσματικά και να περιέχουν όσο το δυνατόν πιο έγκυρες και ακριβείς πληροφορίες που να περιγράφουν επαρκώς και με αξιοπιστία τα δεδομένα. Το στάδιο της προεπεξεργασίας συνήθως αφορά την απόρριψη άχρηστων πληροφοριών (θόρυβος), τη βελτίωση της αναπαράστασης των δεδομένων (συνήθως μέσω μετασχηματισμών, όπως ο μετασχηματισμός Fourier, Wavelet κλπ.) και οποιαδήποτε άλλη κατάλληλη για τη συγκεκριμένη εφαρμογή επεξεργασία με τέτοιο τρόπο ώστε να διευκολύνονται τα επόμενα στάδια ανάλυσης που ακολουθούν.

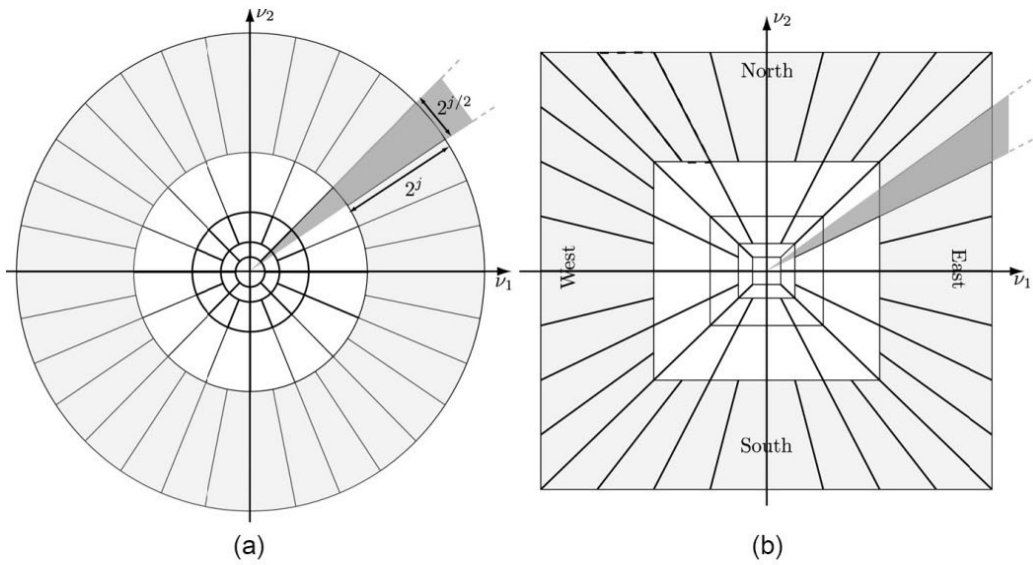
Στην παρούσα πτυχιακή εργασία χρησιμοποιείται ο διακριτός μετασχηματισμός Curvelet, καθώς έτσι προτείνεται στην εργασία [23] την οποία έχουμε ως οδηγό. Αυτός ο μετασχηματισμός ανήκει στην κατηγορία της πολυκλιμακωτής (multiscale) ανάλυσης εικόνας που έχει πλέον καθιερωθεί στην επιστημονική κοινότητα [24]. Η πολυκλιμακωτή ανάλυση βασίζεται στην τυπική θεωρία που επιτρέπει το χειρισμό και τον έλεγχο δομών της εικόνας σε διάφορες κλίμακες. Αυτό γίνεται με την αντιπροσώπευση της εικόνας από πολλές άλλες υπο-εικόνες που αντιστοιχούν σε διαφορετικές κλίμακες. Σκοπός της πολυκλιμακωτής ανάλυσης είναι η διάσπαση της εικόνας σε πολλαπλές κλίμακες ώστε να καταστεί δυνατή η ευκολότερη εξαγωγή των απαραίτητων πληροφοριών. Αυτό ακριβώς επιτυγχάνεται με τον μετασχηματισμό Curvelet.

3.1 Μετασχηματισμός Curvelet

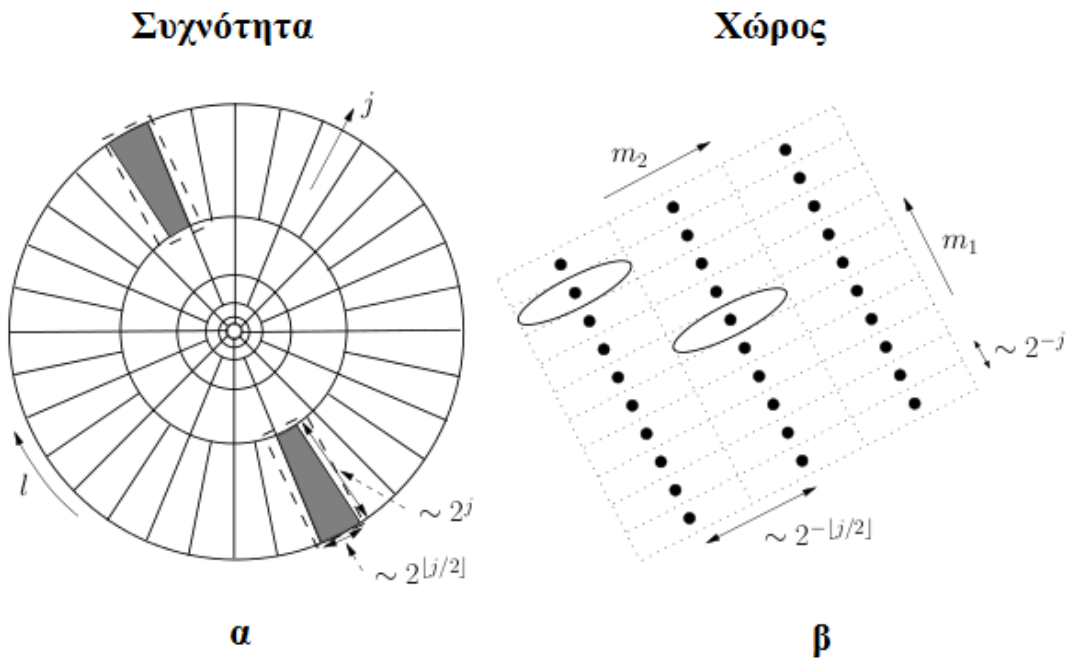
Ο μετασχηματισμός Curvelet αναπτύχθηκε από τους Candes και Donoho [26] το 2000 στην προσπάθειά τους να ξεπεράσουν τα προβλήματα που εμφάνιζε μέχρι τότε ο μετασχηματισμός Wavelet στη διερεύνηση και ανάλυση καμπυλόγραμμων δομών που εμφανίζονται σε εικόνες. Έτσι, στην εργασία [26] προτάθηκε η πρώτη γενιά του μετασχηματισμού Curvelet ο οποίος στην ουσία αποτελούσε επέκταση του μετασχηματισμού ridgelet σε μικρή κλίμακα, όπου ένα κομμάτι καμπύλης μπορεί να θεωρηθεί ευθεία. Παρά τα ενδιαφέροντα αποτελέσματα των Curvelet πρώτης γενιάς η αυξημένη μαθηματική και υπολογιστική πολυπλοκότητα οδήγησε στην ανάπτυξη του μετασχηματισμού curvelet δεύτερης γενιάς [27]. Σε αντίθεση με το μετασχηματισμό Curvelet πρώτης γενιάς, ο νέος μετασχηματισμός είναι πολύ πιο απλός στην υλοποίηση και απαλλάσσεται από τη χρήση των ridgelets, καθιστώντας τον γρηγορότερο.

Ο μετασχηματισμός curvelet χρησιμοποιεί ως συνάρτηση βάσης ένα μητρικό curvelet φ_j . Από αυτό το μητρικό curvelet προέρχονται όλα τα υπόλοιπα curvelets, που απαιτούνται για την αποσύνθεση των εικόνων, μετά από αλλαγή κλίμακας, προσανατολισμού και θέσης του μητρικού curvelet. Επομένως, ένα curvelet δεύτερης γενιάς ορίζεται σε κλίμακα 2^{-j} , προσανατολισμό θ_l και θέση $\mathbf{t}_k^{j,l} = R_{\theta_l}^{-1}(2^{-j}k, 2^{-j/2}l)$ από το μητρικό curvelet φ_j ως εξής

$$\varphi_{j,l,k}(\mathbf{t}) = \varphi_{j,l,k}(t_1, t_2) = \varphi_j\left(R_{\theta_l}(\mathbf{t} - \mathbf{t}_k^{j,l})\right),$$

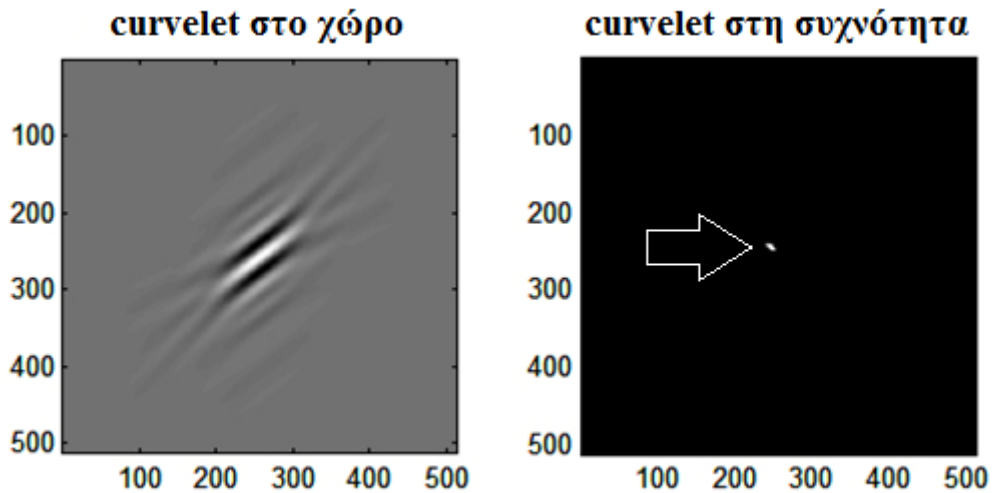


Εικόνα 13. (α) Διαμέριση curvelet στη συνεχή συχνότητα. Η γκριζα περιοχή αντιπροσωπεύει μια σφήνα που λαμβάνεται ως το γινόμενο του ακτινικού παραθύρου (ανοιχτό χρώμα) και του γωνιακού παραθύρου (σκοτεινό χρώμα). (β) Το καρτεσιανό πλέγμα στο χώρο που σχετίζεται με τη δομή του σχήματος (α) και υπακούει επίσης στην σχέση κλιμάκωσης.

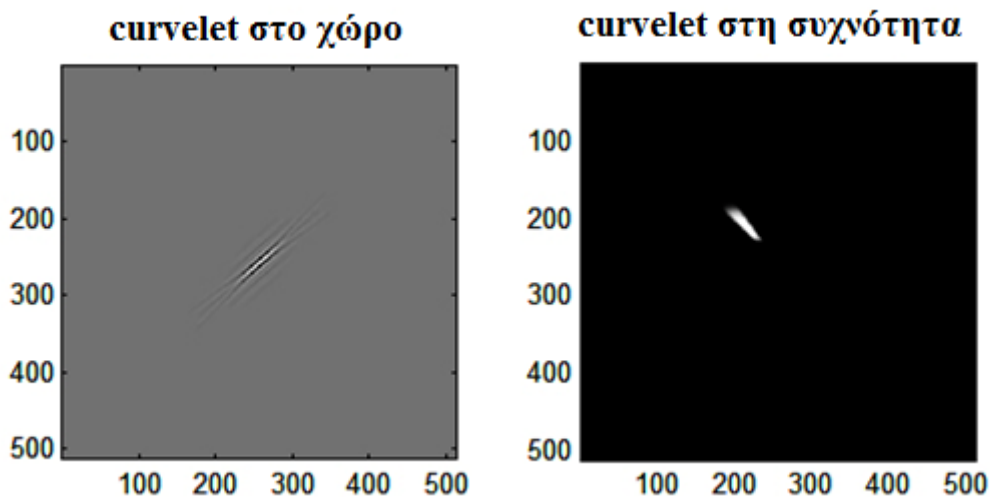


Εικόνα 14. Διαμέριση πεδίου συχνότητα (α) και πεδίου χώρου (β) για τον ορισμό των curvelet

όπου R_{θ_l} είναι η περιστροφή κατά θ_l rad και θ_l είναι η σειρά που ορίζεται από τη σχέση $\theta_l = 2\pi 2^{-\lfloor j/2 \rfloor} l$, με l τέτοιο ώστε $0 \leq \theta_l \leq 2\pi$. $k = (k_1, k_2) \in \mathbb{Z}^2$ είναι οι παράμετροι μετακίνησης στον χώρο. Η κυματομορφή φ_j ορίζεται μέσω του μετασχηματισμού Fourier $\widehat{\varphi}_j(v)$, εκφρασμένο σε πολικές συντεταγμένες (r, θ) στο πεδίο της συχνότητας:



Εικόνα 15. Παράδειγμα ενός curvelet τραχιάς κλίμακας στο χώρο (αριστερά) και στη συχνότητα (μετασχηματισμός Fourier) (δεξιά).



Εικόνα 16. Παράδειγμα ενός curvelet λεπτής κλίμακας στο χώρο (αριστερά) και στη συχνότητα (μετασχηματισμός Fourier) (δεξιά).

$$\widehat{\varphi}_j(r, \theta) = 2^{-\frac{3j}{4}} W(2^{-j}r) V\left(\frac{2^{\frac{j}{2}}|\theta|}{2\pi}\right).$$

Το πεδίο μη-μηδενικών τιμών της $\widehat{\varphi}_j$ είναι μια πολική παραβολική σφήνα που ορίζεται από το πεδίο τιμών των W, V που είναι τα ακτινικά και γωνιακά παράθυρα (ομαλά, θετικά και με πραγματικές τιμές), τα οποία εφαρμόζονται σε κάθε κατεύθυνση με μέγεθος παραθύρου που εξαρτάται από την κλίμακα. Ο τρόπος με τον οποίο διαμερίζεται η συχνότητα φαίνεται στην Εικόνα 13 ενώ στην Εικόνα 14 φαίνεται η διαμέριση του πεδίου της συχνότητας και του χώρου για ένα curvelet σε συγκεκριμένη κλίμακα, γωνία και θέση.

Στη συνεχή συχνότητα \mathbf{v} , οι συντελεστές curvelet της δισδιάστατης συνάρτησης $f(\mathbf{t})$ ορίζονται ως το εσωτερικό γινόμενο

$$\alpha_{j,l,k} := \langle f, \varphi_{j,l,k} \rangle = \int_{\mathbb{R}^2} f(\mathbf{v}) \widehat{\varphi}_j(\mathbf{R}_{\theta_l} \mathbf{v}) e^{it_k^{j,l} \cdot \mathbf{v}} d\mathbf{v}.$$

Τα curvelets που προκύπτουν έχουν μιγαδικές τιμές. Ωστόσο, εύκολα προκύπτουν curvelets με πραγματικές τιμές με χρήση της συμμετρικής εκδοχής $\widehat{\varphi}_j(r, \theta) + \widehat{\varphi}_j(r, \theta + \pi)$. Στις εικόνες 15 και 16 φαίνεται η μορφή ενός curvelet δεύτερης γενιάς στο χώρο και στη συχνότητα. Το πρώτο curvelet ανήκει σε μια τραχιά (coarse) κλίμακα, ενώ το δεύτερο ανήκει σε μια πιο λεπτή (fine) κλίμακα.

3.2 Ταχύς διακριτός μετασχηματισμός Curvelet

Για να χρησιμοποιηθεί ο μετασχηματισμός curvelet σε ψηφιακά δεδομένα, ο αλγόριθμος πρέπει προηγουμένως να προσαρμοστεί κατάλληλα. Υπάρχουν δυο ψηφιακές υλοποιήσεις του μετασχηματισμού που μπορούν να εφαρμοστούν στις δυο και στις τρεις διαστάσεις [28]. Αυτές είναι:

- Curvelets που βασίζονται σε άνισα διανεμημένους μετασχηματισμούς FFT (Unequally-spaced fast Fourier Transforms - USFFT).
- Curvelets που βασίζονται στο τύλιγμα (wrapping) ειδικά διαλεγμένων δειγμάτων Fourier.

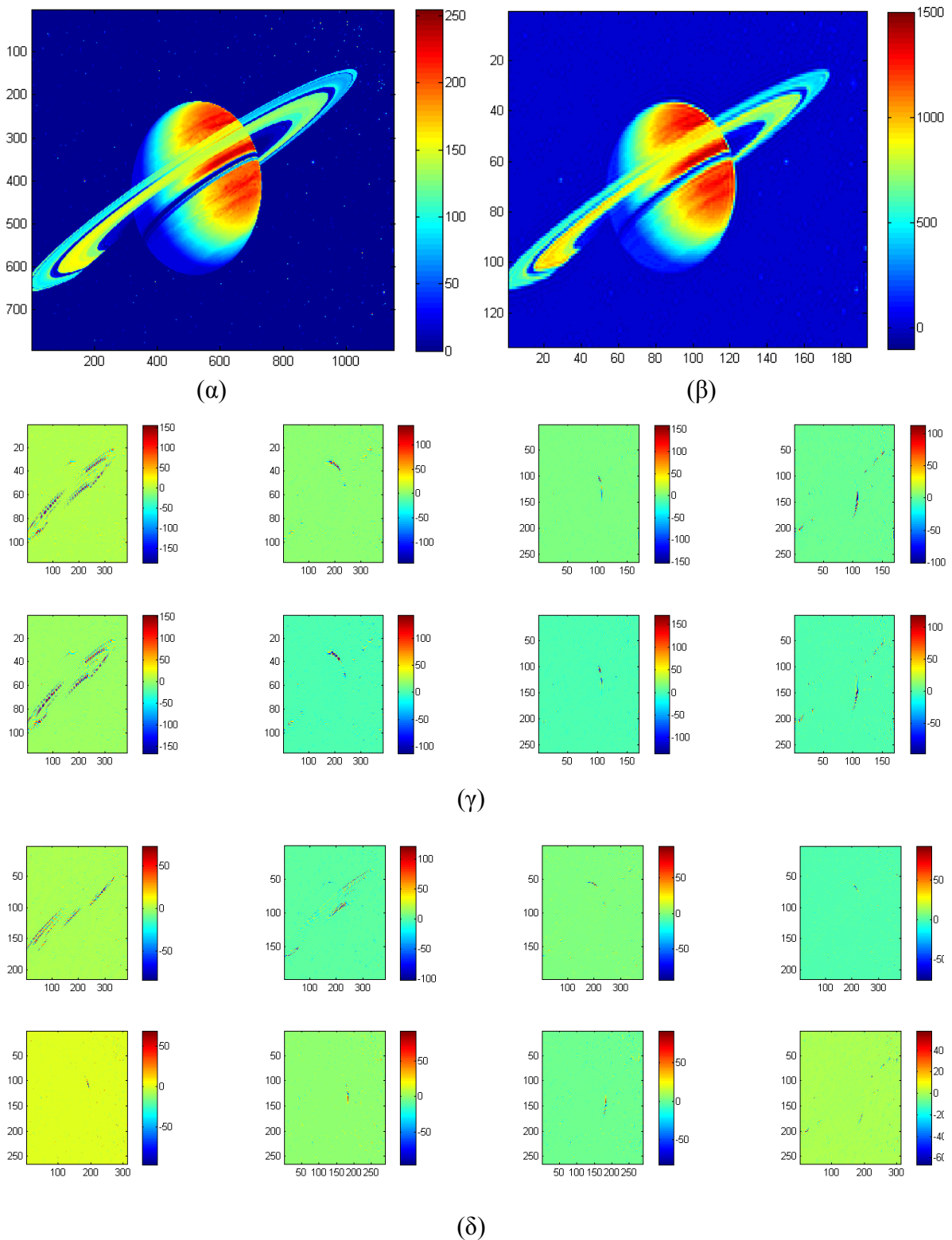
Οι δύο υλοποιήσεις ουσιαστικά διαφέρουν στην επιλογή του χωρικού πλέγματος που χρησιμοποιείται για να μετακινήσει τα curvelet σε κάθε κλίμακα και γωνία. Υλοποιήσεις και των δυο αλγορίθμων είναι διαθέσιμες στη μορφή του πακέτου λογισμού CurveLab (Matlab, C++) που διατίθεται δωρεάν για ερευνητικούς σκοπούς (www.curvelet.org).

Επομένως, δημιουργείται ένα, αντίστοιχα με το αναλογικό, ψηφιακό curvelet της μορφής $\varphi_{j,l,k}^D$ όπου οι μεταβλητές j, l, k αντιστοιχούν στην κλίμακα, στη γωνία και στη θέση. Εφαρμόζοντας ένα τέτοιο curvelet πάνω σε εικόνα $Im[t_1, t_2]$, $0 \leq t_1, t_2 \leq N$ δημιουργείται ένας συντελεστής $c^D(j, l, k)$ για τη συγκεκριμένη κλίμακα, γωνία, θέση και ορίζεται ως

$$c^D(j, l, k) = \sum_{0 \leq t_1, t_2 < n} f[t_1, t_2] \overline{\varphi_{j,l,k}^D[t_1, t_2]},$$

όπου κάθε $\varphi_{j,l,k}^D$ είναι η ψηφιακή κυματομορφή curvelet (ή αλλιώς συνάρτηση βάσης). Όλοι αυτοί οι συντελεστές για διάφορες κλίμακες, γωνίες και θέσεις αποτελούν το μετασχηματισμό curvelet μιας εικόνας. Ο αριθμός των κλιμάκων καθώς και το πλήθος των γωνιών σε κάθε κλίμακα επιλέγεται από το χρήστη ενώ το πλήθος των θέσεων καθορίζεται αυτόματα από το μέγεθος της εικόνας. Όλοι οι συντελεστές (που αντιστοιχούν σε όλες τις θέσεις) για συγκεκριμένη κλίμακα και γωνία συνιστούν μια υπο-εικόνα ή υπο-μπάντα.

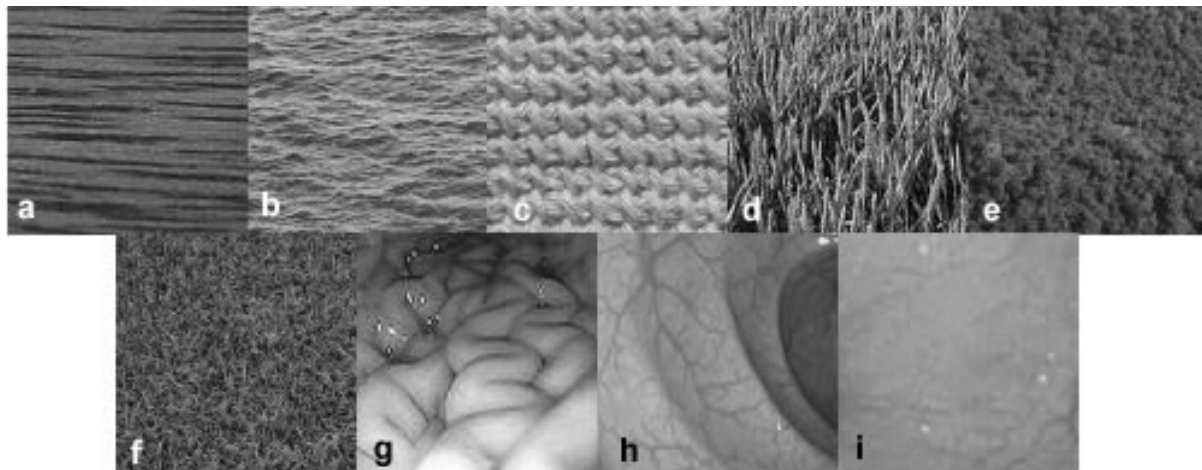
Στην Εικόνα 17 φαίνεται η ανάλυση μιας εικόνας σε 3 κλίμακες, με 8 γωνίες στη δεύτερη κλίμακα και 16 γωνίες στην τρίτη (εάν και απεικονίζονται μόνο οι 8 λόγω χώρου). Η πρώτη κλίμακα έχει πάντα μόνο μία γωνία.



Εικόνα 17. Παράδειγμα ανάλυσης εικόνας με Curvelet transform. (α) Αρχική εικόνα, (β) Υπο-εικόνα της κλίμακας 1 (1 γωνία), (γ) Υπο-εικόνες της κλίμακας 2 (8 γωνίες), (δ) Υπο-εικόνες της κλίμακας 3 (16 γωνίες αλλά απεικονίζονται μόνο οι 8 πρώτες)

4 Εξαγωγή χαρακτηριστικών υφής

Η υφή μιας εικόνας είναι μια έννοια αρκετά αφηρημένη καθώς δεν υπάρχει ένας αυστηρώς ορισμός αλλά οι διάφοροι ερευνητές χρησιμοποιούν διάφορους ορισμούς ανάλογα με τις συγκεκριμένες εφαρμογές που τους απασχολούν. Κάποιος, λοιπόν, μπορεί να πει ότι υφή είναι τα ιδιαίτερα χαρακτηριστικά της επιφάνειας ενός υλικού και η αίσθηση που δημιουργείται από την επαφή μαζί του. Η υφή γίνεται εύκολα αντιληπτή από τον άνθρωπο και θεωρείται πλούσια πηγή πληροφορίας, τόσο για τη φύση όσο και για το τρισδιάστατο σχήμα των αντικειμένων. Υφή, επίσης, μπορεί να θεωρηθεί και το λεπτομερές μοτίβο που επαναλαμβάνεται αρκετές φορές για να καλύψει μια εικόνα (Εικόνα 18). Γενικά, ένας



Εικόνα 18. Παράδειγμα εικόνων με διαφορετική υφή: a) ξύλο, b) νερό, c) ύφασμα, d) βλαστοί καλαμποκιών, e) δάσος, f) γρασίδι, g) στομάχι, h) παχύ έντερο, i) οισοφάγος.

διαδεδομένος ορισμός για την υφή είναι αυτός που την ορίζει ως τη χωρική διακύμανση της έντασης των εικονοστοιχείων. Υφή είναι πολύπλοκα οπτικά πρότυπα που αποτελούνται από οντότητες ή υπο-πρότυπα που έχουν χαρακτηριστική φωτεινότητα, χρώμα, κλίση, μέγεθος, κτλ. Έτσι, η υφή μπορεί να θεωρηθεί ως «ομαδοποίηση ομοιότητας» σε μια εικόνα. Οι ιδιότητες των τοπικών υπο-προτύπων προκαλούν την αντίληψη της ομοιομορφίας, της πυκνότητας, της τραχύτητας, της λεπτότητας, της απαλότητας κτλ. της υφής ως σύνολο. Στην ανάλυση υφής υπάρχουν τέσσερα σημαντικά ζητήματα [29]:

- Εξαγωγή χαρακτηριστικών: για την περιγραφή της υφής
- Διαχωρισμός υφής: για την τμηματοποίηση μιας εικόνας σε περιοχές όπου η καθεμιά χαρακτηρίζεται από ομογενή υφή
- Ταξινόμηση υφής: για να προσδιοριστεί σε ποιά κατηγορία (από ένα συγκεκριμένο πλήθος διαφορετικών κατηγοριών) ανήκει μια περιοχή με ομογενή υφή
- Ανακατασκευή σχήματος από υφή: για την ανακατασκευή 3D επιφανειακής γεωμετρίας.

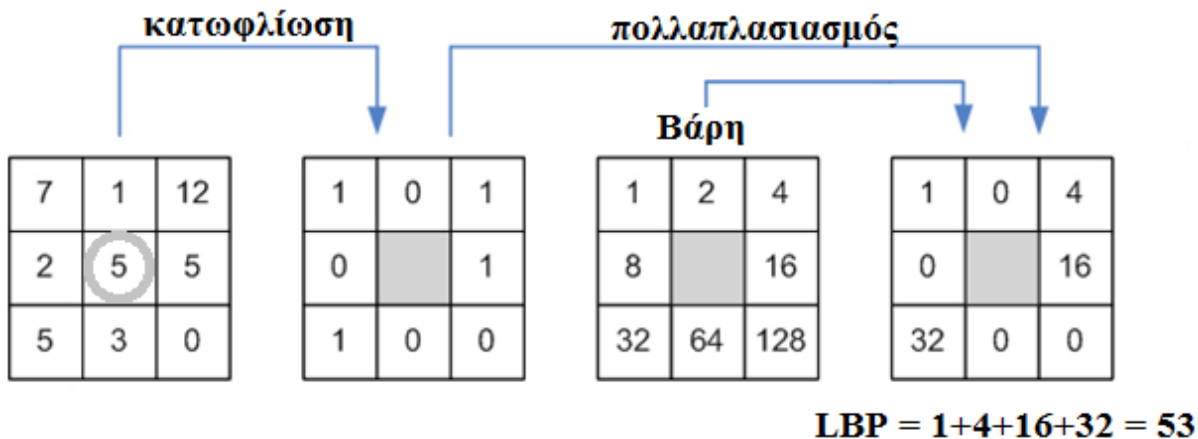
Γενικά, ως "χαρακτηριστικό" μπορεί να θεωρηθεί οποιοδήποτε μετρήσιμο μέγεθος που εξάγεται από μια εικόνα [24]. Τα χαρακτηριστικά χρησιμοποιούνται για να περιγράψουν αντικείμενα που περιέχονται στις εικόνες και πρέπει να επιλέγονται κατάλληλα και ανάλογα με τις ιδιαιτερότητες της κάθε εφαρμογής. Ουσιαστικά, η περιγραφή των αντικειμένων με χαρακτηριστικά απεικονίζει τα αντικείμενα στο χώρο των χαρακτηριστικών, με αποτέλεσμα,

η αναγνώρισή τους να ισοδυναμεί με τη μέτρηση της ομοιότητας μεταξύ των χαρακτηριστικών των αντικειμένων. Τα χαρακτηριστικά, γενικά, μπορούν να κατηγοριοποιηθούν ως χαρακτηριστικά χώρου και χαρακτηριστικά από μετασχηματισμό. Τα χαρακτηριστικά χώρου μπορούμε να τα διακρίνουμε σε γεωμετρικά χαρακτηριστικά, στατιστικά χαρακτηριστικά και χαρακτηριστικά υφής [24]. Τα γεωμετρικά χαρακτηριστικά περιγράφουν την υφή με καλά ορισμένα θεμελιώδη στοιχεία (primitives) και μια ιεραρχία της χωρικής κατανομής αυτών των στοιχείων. Η επιλογή ενός θεμελιώδους στοιχείου και η πιθανότητα το επιλεγμένο στοιχείο να τοποθετηθεί σε μια συγκεκριμένη θέση μπορεί να είναι συνάρτηση της θέσης ή των στοιχείων κοντά στη θέση. Το πλεονέκτημα της δομικής προσέγγισης είναι ότι παρέχει καλή συμβολική περιγραφή της εικόνας. Ωστόσο, το χαρακτηριστικό αυτό είναι πιο χρήσιμο σε εφαρμογές σύνθεσης παρά ανάλυσης. Τα στατιστικά χαρακτηριστικά, σε αντίθεση με τα γεωμετρικά, περιγράφουν την υφή μέσα από τις στατιστικές ιδιότητες της κατανομής των φωτεινοτήτων των εικονοστοιχείων μέσα στην εικόνα. Τέλος τα χαρακτηριστικά υφής προσπαθούν να περιγράψουν την υφή από επαναλαμβανόμενες δομές που οφείλονται σε διακυμάνσεις των φωτεινοτήτων της εικόνας, τα οποία για τη συγκεκριμένη ευκρίνεια της εικόνας, είναι τόσο λεπτά που δεν μπορούν να διακριθούν ως ξεχωριστά αντικείμενα.

4.1 Τοπικά Δυαδικά Πρότυπα (Local Binary Patterns)

Τα Τοπικά Δυαδικά Πρότυπα ή πιο γνωστά Local Binary Patterns (LBP) προτάθηκαν το 1996 από τους Ojala κ.α. [30] και έκτοτε έχουν χρησιμοποιηθεί ευρύτατα σε πολλές εφαρμογές ως ισχυρά και αποτελεσματικά χαρακτηριστικά υφής χάριν της διακριτικής τους ικανότητας και της χαμηλής υπολογιστικής πολυπλοκότητας. Τα LBP ουσιαστικά αποτυπώνουν και κωδικοποιούν τοπικά χαρακτηριστικά των εικόνων, τη γειτονιά, δηλαδή, γύρω από κάθε εικονοστοιχείο, και η κατανομή τους χαρακτηρίζει τις εικόνες. Η βασική ιδέα για την ανάπτυξη των LBP ήταν ότι η υφή σε δυσδιάστατες επιφάνειες μπορεί να περιγραφεί από δύο συμπληρωματικά μέτρα: τοπικά χωρικά πρότυπα (υπολογίζονται από τα LBP) και αντίθεση στην κλίμακα του γκρι (συνήθως χρησιμοποιείται συμπληρωματικά με τα LBP).

Στην αρχική τους μορφή τα LBP υπολογίζονται ως εξής (Εικόνα 19): Σε μια γειτονιά 3×3 pixels, η οποία ονομάζεται και μάσκα, η φωτεινότητα του κεντρικού εικονοστοιχείου λαμβάνεται ως τοπικό κατώφλι με αποτέλεσμα τα 8 γειτονικά εικονοστοιχεία να πάρουν την τιμή 1 ή 0 ανάλογα εάν η φωτεινότητά τους είναι μεγαλύτερη (ή ίση) ή μικρότερη από την τιμή του κατωφλίου (δηλαδή τη φωτεινότητα του κεντρικού εικονοστοιχείου). Όπως φαίνεται και στην Εικόνα 19, το κατώφλι (κεντρικό εικονοστοιχείο) έχει τιμή 5. Όλα τα εικονοστοιχεία που έχουν φωτεινότητα μεγαλύτερη ή ίση από 5 παίρνουν την τιμή 1 ενώ όλα τα υπόλοιπα την τιμή 0. Στη συνέχεια, η τιμή κάθε εικονοστοιχείου πολλαπλασιάζεται με ένα βάρος (δύναμη του 2) ανάλογα με τη θέση του στη 3×3 γειτονιά. Για παράδειγμα, το πάνω αριστερά εικονοστοιχείο (σε σχέση με το κεντρικό) έχει βάρος $2^0=1$, το δεξιό του έχει βάρος $2^1=2$, ενώ το κάτω δεξιά έχει βάρος $2^7=128$. Τέλος, οι τιμές των οκτώ εικονοστοιχείων αθροίζονται ώστε να προκύψει η τιμή του LBP για τη συγκεκριμένη περιοχή η οποία είναι και η τιμή για το συγκεκριμένο πρότυπο (ή μονάδα υφής). Αυτή η τιμή προφανώς είναι ακέραιη, πραγματική στο διάστημα $[0, 255]$ αφού από 8 γειτονικά εικονοστοιχεία μπορούν να προκύψουν $2^8=256$ διαφορετικοί συνδυασμοί. Με άλλα λόγια, ένα σύνολο από 256 διαφορετικά πρότυπα μπορεί να προκύψει από μια γειτονιά 3×3 . Εφαρμόζοντας την



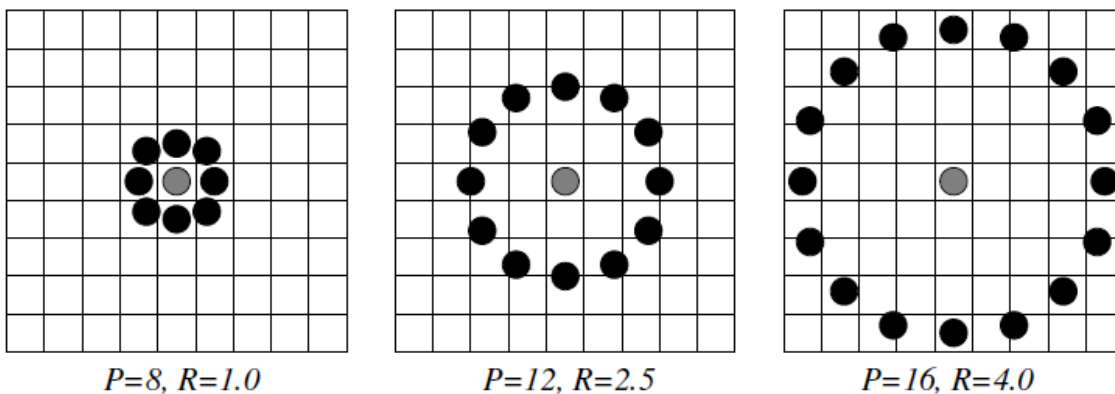
Εικόνα 19. Υπολογισμός LBP χαρακτηριστικών σε γειτονιά 3x3 pixel.

παραπάνω διαδικασία σε όλη την εικόνα μπορούμε να σχηματίσουμε το ιστόγραμμα των LBP τιμών το οποίο μπορεί να ληφθεί ως το διάνυσμα χαρακτηριστικών της συγκεκριμένης εικόνας.

Αξίζει να αναφερθεί ότι τα LBP χαρακτηριστικά είναι ανθεκτικά σε μονοτονικές αλλαγές της κλίμακας του γκρι που οφείλονται π.χ. σε διαφορετικές συνθήκες φωτισμού. Δηλαδή, εάν οι αρχικές τιμές φωτεινότητας αυξηθούν όλες κατά μια σταθερή τιμή τότε οι τιμές των LBP δεν μεταβάλλονται. Αντίθετα, πρέπει να τονισθεί ότι η σειρά ανάθεσης των βαρών στα εικονοστοιχεία δεν υπακούει κάποιον κανόνα και είναι στην ευχέρεια του χρήστη. Αυτό αποτελεί μειονέκτημα καθώς αλλαγή στη σειρά από τα βάρη συνεπάγεται αλλαγή στην τιμή του LBP.

4.2 Γενίκευση του LBP

Το 2002 οι Ojala κ.α. πρότειναν μια γενίκευση των LBP [31] ώστε να αποτελούν έναν απλό αλλά αποτελεσματικό τρόπο πολυδιακριτικής ανάλυσης. Σύμφωνα με τη γενίκευση αυτή, ο καθορισμός των τιμών των LBP δε γίνεται σε μια γειτονιά 3x3 αλλά πάνω στην περιφέρεια ενός κύκλου (P,R). R είναι η ακτίνα του κύκλου από το κεντρικό εικονοστοιχείο και P είναι το πλήθος των γειτονικών σημείων που κατανέμονται ομοιόμορφα στην περιφέρεια του κύκλου. Τρία χαρακτηριστικά παραδείγματα από γειτονιές διαφορετικής ακτίνας και πλήθους γειτόνων δίνονται στην Εικόνα 20 όπου οι συντεταγμένες



Εικόνα 20. Κυκλικές συμμετρικές γειτονιές για διάφορες τιμές P και R.

των γειτονικών σημείων (x_p, y_p) προκύπτουν από τις σχέσεις

$$x_p = x_c + R \cdot \cos\left(\frac{2\pi p}{P}\right)$$

$$y_p = y_c - R \cdot \sin\left(\frac{2\pi p}{P}\right)$$

όπου (x_c, y_c) είναι οι συντεταγμένες του κεντρικού εικονοστοιχείου. Όταν οι συντεταγμένες των γειτόνων δε συμπίπτουν ακριβώς με το κέντρο ενός εικονοστοιχείου τότε οι τιμές των φωτεινότητων προκύπτουν από παρεμβολή.

Ο υπολογισμός των νέων τιμών των τοπικών χαρακτηριστικών LBP μπορεί να εκφραστεί από την παρακάτω σχέση.

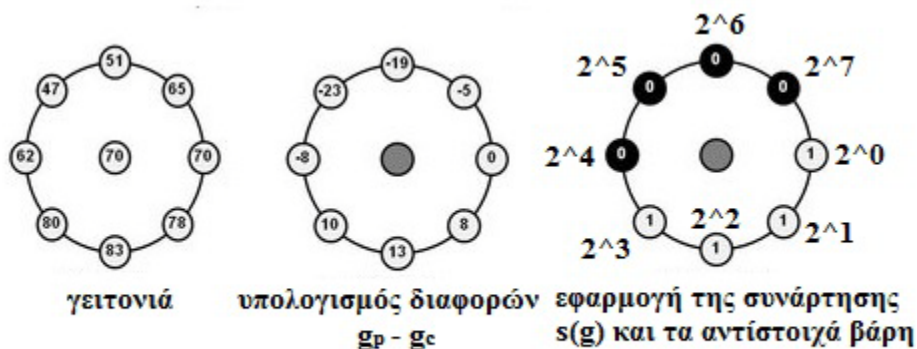
$$LBP_{P,R} = \sum_{p=0}^{P-1} s(g_p - g_c) 2^p$$

όπου g_c είναι η φωτεινότητα του κεντρικού εικονοστοιχείου, g_p είναι οι φωτεινότητες των γειτονικών εικονοστοιχείων που κατανέμονται συμμετρικά στην περιφέρεια του κύκλου και συνήθως το $p=0$ αντιστοιχεί στο εικονοστοιχείο που βρίσκεται δεξιά από το κεντρικό, και

$$s(g) = \begin{cases} 1 & \text{αν } g \geq 0 \\ 0 & \text{αν } g < 0 \end{cases}$$

η οποία εκφράζει την καταφλίωση όπως περιγράφηκε στην προηγούμενη παράγραφο. Στην Εικόνα 21 φαίνεται ο υπολογισμός του $LBP_{P,R}$ για γειτονιά $(8,R)$.

Είναι φανερό ότι για συγκεκριμένα P και R εξάγονται 2^P διαφορετικές τιμές χαρακτηριστικών $LBP^{P,R}$ που αντιστοιχούν στους 2^P διαφορετικούς δυαδικούς αριθμούς που προκύπτουν. Με άλλα λόγια, ένα σύνολο από 2^P διαφορετικά πρότυπα μπορεί να προκύψει. Υπογραμμίζεται ότι ο υπολογισμός των $LBP^{P,R}$ γίνεται μόνο για εικονοστοιχεία που απέχουν R ή περισσότερο από τα άκρα της εικόνας. Τα υπόλοιπα εικονοστοιχεία που δεν μπορούν να έχουν ένα πλήρες κύκλο R γύρω τους αμελούνται. Να σημειωθεί ότι το $LBP_{8,1}$ είναι



$$LBP = 1*1 + 1*2 + 1*4 + 1*8 + 0*16 + 0*32 + 0*64 + 0*128 = 15$$

Εικόνα 21. Υπολογισμός $LBP_{8,R}$.

παρόμοιο με το αρχικό LBP που προκύπτει από γειτονιά 3×3 με δύο όμως διαφορές: τα εικονοστοιχεία στο $LBP_{8,1}$ σχηματίζουν μια κυκλική αλυσίδα και οι τιμές των φωτεινότητων των διαγώνιων εικονοστοιχείων υπολογίζονται με παρεμβολή. Και οι δυο αυτές αλλαγές είναι απαραίτητες ώστε να έχουμε κυκλικά συμμετρική γειτονιά που δύναται να οδηγήσει

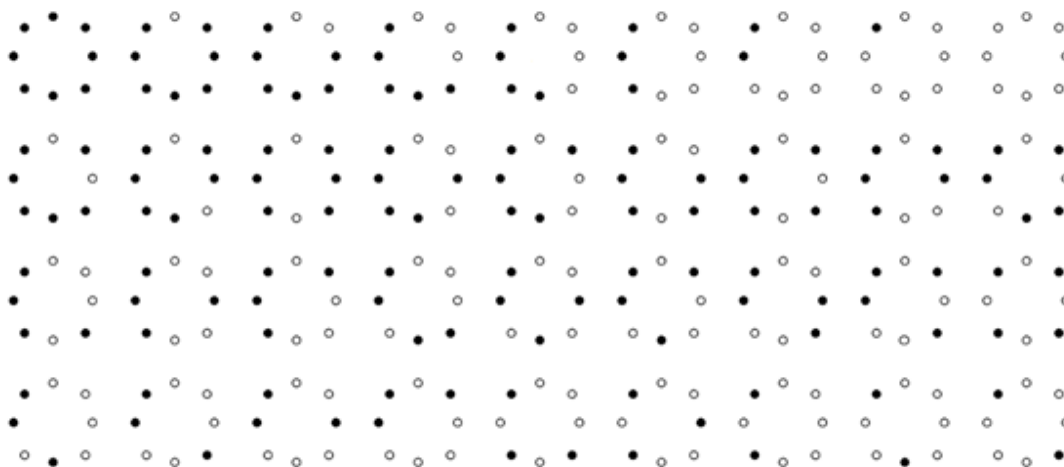
στην εξαγωγή LBP χαρακτηριστικών ανεξάρτητων περιστροφής τα οποία θα περιγραφούν στην επόμενη παράγραφο.

4.3 LBP ανεξάρτητα περιστροφής

Ένα σημαντικό μειονέκτημα των LBP όπως παρουσιάστηκαν μέχρι στιγμής είναι η μη ανθεκτικότητά τους στην περιστροφή της εικόνας ή στην σειρά με την οποία ορίζονται τα βάρη. Με άλλα λόγια, εάν θεωρήσουμε σταθερή την αρχική θέση από τα βάρη (π.χ. $p=0$ πάντα στο εικονοστοιχείο δεξιά του κεντρικού) τότε όταν η εικόνα περιστραφεί, οι τιμές της έντασης g_p θα μετακινηθούν αντίστοιχα κατά μήκος του κύκλου γύρω από το g_0 . Αφού το g_0 περιέχει πάντα την τιμή της έντασης του στοιχείου $(0, R)$ στα δεξιά του g_c , η περιστροφή ενός συγκεκριμένου δυαδικού προτύπου θα οδηγήσει σε διαφορετική τιμή $LBP_{p,R}$. Το ίδιο ισχύει αντίστοιχα όταν θεωρήσουμε σταθερή την εικόνα και αλλάξουμε τη σειρά από τα βάρη (π.χ. $p=0$ να αντιστοιχεί στο εικονοστοιχείο πάνω από το κεντρικό). Φυσικά, αυτό το μειονέκτημα δεν ισχύει για πρότυπα που αποτελούνται αποκλειστικά από 1 ή αποκλειστικά από 0, αφού όπως και να περιστραφούν το τελικό αποτέλεσμα θα είναι το ίδιο. Για να απαλειφθεί το πρόβλημα που προκύπτει σε περίπτωση περιστροφής της εικόνας ή διαφορετικής σειράς των βαρών και να δοθεί μια μοναδική ταυτότητα σε κάθε δυαδικό πρότυπο (ανεξάρτητα της κατεύθυνσης) ορίζονται τα LBP ανεξάρτητα περιστροφής ως εξής:

$$LBP_{p,R}^i = \min\{ROR(LBP_{p,R}, i) \mid i = 0, 1, \dots, P - 1\}$$

όπου η συνάρτηση $ROR(x, i)$ πραγματοποιεί κυκλική ολίσθηση του δυαδικού αριθμού x (P ψηφίων) προς τα δεξιά i φορές. Με λίγα λόγια, ο κώδικας για ανεξαρτησία περιστροφής παράγεται από ολίσθηση του δυαδικού αριθμού που παράγεται από τα κλασικά LBP μέχρι την εύρεση της ελάχιστης τιμής του. Αυτό σημαίνει ότι όλες οι 2^P διαφορετικές τιμές ενός συγκεκριμένου μοτίβου αντιστοιχίζονται σε μια τιμή αναφοράς, την ελάχιστη. Με τον τρόπο αυτό ομαδοποιούνται όλες οι «περιστρεμμένες» εκδοχές ενός δυαδικού αριθμού σε μία, και από σχηματικής σκοπιάς ομαδοποιούνται πολλά πρότυπα σε λιγότερα πρότυπα ανεξάρτητα της περιστροφής, που αντιστοιχούν σε συγκεκριμένα μικρο-χαρακτηριστικά της εικόνας.



Εικόνα 22. Τα 36 μοναδικά ανεξάρτητα περιστροφής πρότυπα σε γειτονιά $(8,R)$. Οι μαύρες βούλες αντιστοιχούν σε 0 και οι άσπρες σε 1.

Σε μια γειτονιά $(8, R)$ τα $2^8 = 256$ διαφορετικά πρότυπα που προκύπτουν από τον απλό τελεστή LBP, μειώνονται σε 36 μοναδικά ανεξάρτητα περιστροφής πρότυπα με τη χρήση του τελεστή $LBP_{8,R}^i$ (Εικόνα 22).

4.4 Ομοιόμορφα ανεξάρτητα περιστροφής LBP

Πολύ ενθαρρυντικά τα αποτελέσματα από τα ανεξάρτητα περιστροφής LBP καθώς μειώνουν σημαντικά το πλήθος των διαφορετικών προτύπων άρα και το μέγεθος του ιστογράμματος. Παρόλα αυτά, στην πράξη διαπιστώθηκε ότι απαιτείται περαιτέρω μείωση του χαρακτηριστικού διανύσματος χωρίς όμως να μειώσουμε την αναλυτικότητα (αριθμός των γειτόνων). Έτσι λοιπόν παρατηρήθηκε ότι συγκεκριμένα τοπικά δυαδικά πρότυπα αποτελούν θεμελιώδεις ιδιότητες της υφής και καταλαμβάνουν τη συντριπτική πλειοψηφία των προτύπων που εμφανίζονται. Αυτά τα θεμελιώδη πρότυπα ονομάστηκαν "ομοιόμορφα" γιατί έχουν ως κοινό στοιχείο μια ομοιόμορφη κυκλική δομή. Ως ομοιόμορφη ορίζεται η δομή η οποία που περιέχει λίγες μεταβάσεις από 0 σε 1 και αντίστροφα.

Επομένως, η έννοια της ομοιομορφίας στην περίπτωση των δυαδικών προτύπων U ορίζεται ως το πλήθος των μεταβάσεων από 0 σε 1 και από 1 σε 0. Για παράδειγμα, το πρότυπο 11111111_2 έχει τιμή $U=0$ ενώ το πρότυπο 11111110_2 έχει $U=2$ (δεν πρέπει να ξεχνάμε ότι τα πρότυπα είναι κυκλικά, ανεξάρτητα περιστροφής, δηλαδή το 11111110_2 είναι το ίδιο με το 11101111_2). Γενικά, για να θεωρείται ένα πρότυπο ομοιόμορφο πρέπει να έχει τιμή U μικρότερη ή ίση με 2. Ο μαθηματικός ορισμός της ομοιομορφίας σε μια γειτονιά g δίνεται από τον παρακάτω τύπο.

$$U(G) = |s(g_{P-1} - g_c) - s(g_0 - g_c)| + \sum_{p=1}^{P-1} |s(g_p - g_c) - s(g_{p-1} - g_c)|,$$

Τα ομοιόμορφα ανεξάρτητα περιστροφής LBP ($LBP_{P,R}^i$) ορίζονται ως εξής:

$$LBP_{P,R}^{iu2} = \begin{cases} \sum_{p=0}^{P-1} s(g_p - g_c), & \text{εαν } U(LBP_{P,R}) \leq 2 \\ P + 1, & \text{σε άλλη περίπτωση} \end{cases}$$

Επομένως, το σύνολο των ομοιόμορφων ανεξάρτητων περιστροφής προτύπων που μπορεί να εμφανιστούν σε μια γειτονιά (P,R) είναι $P+1$. Στην Εικόνα 23 παρουσιάζονται τα ομοιόμορφα ανεξάρτητα περιστροφής πρότυπα μιας γειτονιάς $(8,R)$ και οι αντίστοιχες τιμές τους. Τα πρότυπα αυτά λειτουργούν ως δομικά συστατικά μικροδομών στις εικόνες όπως ένα φωτεινό σημείο (πρότυπο 0), ένα σκοτεινό σημείο ή επίπεδη περιοχή (πρότυπο 8) και ακμές ποικίλης καμπυλότητας (πρότυπα 1-7).



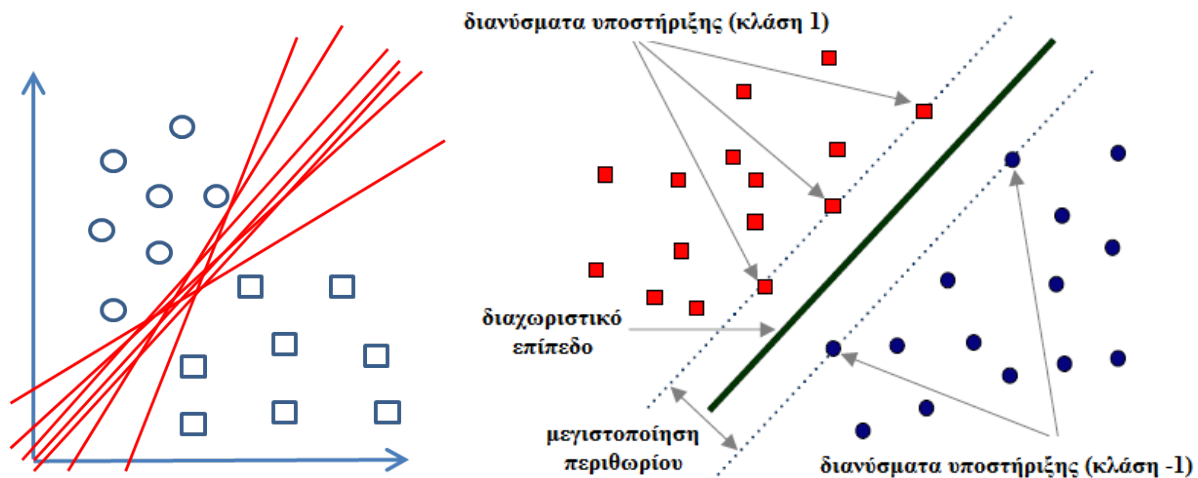
Εικόνα 23. Ομοιόμορφα ανεξάρτητα περιστροφής πρότυπα μιας γειτονιάς $(8,R)$.

5 Ταξινόμηση

Η ευκολία με την οποία αντιλαμβανόμαστε τα αντικείμενα γύρω μας είναι ουσιαστικά διαδικασίες αναγνώρισης προτύπων. Οι άνθρωποι εκπαιδεύουν τα κύτταρα του εγκεφάλου τους ώστε να μπορούν να επιλύουν προβλήματα τέτοιου είδους προβλήματα. Στην περιοχή της ψηφιακής επεξεργασίας εικόνων το πρόβλημα της αναγνώρισης προτύπων ανάγεται στην ταξινόμηση ενός συνόλου κατάλληλων χαρακτηριστικών σε κλάσεις με τη βοήθεια κατάλληλων ταξινομητών [24]. Το σύνολο των χαρακτηριστικών που χρησιμοποιείται σε ένα σύστημα ταξινόμησης ονομάζεται διάνυσμα χαρακτηριστικών ή πρότυπο. Οι ταξινομητές είναι ουσιαστικά αλγόριθμοι που χρησιμοποιούνται για να επιτύχουμε την ταξινόμηση των προτύπων ομοιογενείς από άποψη χαρακτηριστικών (όσο είναι αυτό δυνατό) κλάσεις. Υπάρχουν ταξινομητές διαφόρων ειδών και δυνατοτήτων, όπως στατιστικοί, νευρωνικοί, βασισμένοι σε ιεραρχικό αλγόριθμο, σαφούς και ασαφούς λογικής κλπ. Στη παρούσα πτυχιακή εργασία θα ασχοληθούμε με έναν ευρέως αποδεκτό ταξινομητή που ονομάζεται Μηχανές Διανυσμάτων Υποστήριξης (Support Vector Machines - SVM).

5.1 Μηχανές διανυσμάτων υποστήριξης

Οι μηχανές διανυσμάτων υποστήριξης [32], [33] (από εδώ και στο εξής θα αναφέρεται ως SVM) πρωτοεμφανίστηκαν το 1993 και βασίζονται στη στατιστική θεωρία μάθησης. Έχουν τύχει ευρείας αποδοχής και εφαρμόζονται σε πολλά προβλήματα ταξινόμησης. Ο SVM χωρίζει το χώρο όπου βρίσκονται τα δεδομένα εκπαίδευσης σε δύο κλάσεις με τη βοήθεια των διανυσμάτων υποστήριξης (Εικόνα 24). Ουσιαστικά ο SVM



Εικόνα 24. Διάφοροι τρόποι διαχωρισμού του επιπέδου σε δύο κλάσεις (αριστερά).

Σύμφωνα με την θεωρία των SVM ο διαχωρισμός γίνεται με βάση τα διανύσματα υποστήριξης ώστε να μεγιστοποιείται η απόσταση

προσπαθεί να βρει έναν γραμμικό διαχωριστή, δηλαδή ένα υπερεπίπεδο, που να διαχωρίζει τις δύο κατηγορίες με μέγιστο περιθώριο. Γενικά, η εξίσωση του υπερεπιπέδου διαχωρισμού θα είναι της ακόλουθης μορφής, όπου x_i , $i=1, \dots, N$ (N = πλήθος δεδομένων) το διάνυσμα χαρακτηριστικών (προς εκπαίδευση), w^T διάνυσμα κάθετο στο υπερεπίπεδο και b πραγματικός αριθμός.

$$w^T x_i + b = 0$$

Στόχος είναι λοιπόν να επιλεγούν κατάλληλα τα w και b ώστε να μεγιστοποιηθεί το περιθώριο, δηλαδή η απόσταση μεταξύ δύο παράλληλων εφαπτόμενων υπερεπιπέδων τα οποία διαχωρίζουν τα διανύσματα x σε δύο κλάσεις (1 και -1) (Εικόνα 24). Τα υπερεπίπεδα αυτά ικανοποιούν την εξίσωση

$$y_i(w^T x_i + b) - 1 = 0$$

όπου $y_i = \pm 1$ ανάλογα με την κλάση στην οποία ανήκει κάθε διάνυσμα x_i . Επομένως, η απόσταση μεταξύ των δύο υπερεπιπέδων είναι $2/\|w\|$. Η απόσταση αυτή είναι το περιθώριο του υπερεπιπέδου διαχωρισμού, που τοποθετείται στο μέσον της απόστασης των δύο εφαπτόμενων υπερεπιπέδων. Όπως αναφέρθηκε ήδη, ο στόχος του SVM είναι να βρει το υπερεπίπεδο διαχωρισμού με το μέγιστο περιθώριο. Οπότε, προκύπτει τα παρακάτω πρόβλημα βελτιστοποίησης:

$$\text{ελαχιστοποίησε } 0.5\|w\|^2$$

με περιορισμό

$$y_i(w^T x_i + b) \geq 1, i = 1, \dots, N$$

Με τη διαδικασία αυτή μεγιστοποιούμε το περιθώριο με παράλληλη ικανοποίηση του περιορισμού που προστάζει όλα τα διανύσματα εκπαίδευσης να βρίσκονται από τη μια ή την άλλη πλευρά του υπερεπιπέδου διαχωρισμού. Τα διανύσματα που βρίσκονται πάνω στο υπερεπίπεδο καλούνται διανύσματα υποστήριξης. Η διαδικασία που ακολουθείται ώστε να βρεθούν τα βέλτιστα w και b δε θα αναπτυχθεί καθώς ξεφεύγει από τα πλαίσια της εργασίας. Θα αναφέρουμε μόνο ότι οι λύσεις για τα w και b είναι

$$w = \sum_{i=1}^N a_i y_i x_i \quad \text{και} \quad b = y_k - w^T x_k,$$

για κάθε x_k τέτοιο ώστε $a_k \neq 0$, όπου a_k οι συντελεστές Lagrange και η ταξινόμηση στις δύο κλάσεις γίνεται ανάλογα με το πρόσημο της συνάρτησης $g(x_k) = w^T x_k + b$. Εάν $g(x_k) > 0$ το x_k ταξινομείται στην κλάση 1 αλλιώς στην κλάση -1.

Πρέπει να σημειωθεί όμως ότι ο περιορισμός που τέθηκε παραπάνω είναι αρκετά αυστηρός. Για παράδειγμα, ενδέχεται να μην είναι δυνατή η εύρεση γραμμικού διαχωριστή που να διαχωρίζει πλήρως τα παραδείγματα εκπαίδευσης ή ενδέχεται να προτιμούμε ένα υπερεπίπεδο διαχωρισμού που έχει μεγαλύτερο περιθώριο αλλά κατατάσσει λανθασμένα ή εντός του περιθωρίου κάποια παραδείγματα εκπαίδευσης από κάποιο άλλο που ικανοποιεί όλους τους περιορισμούς αλλά έχει μικρότερο περιθώριο. Για τους λόγους αυτούς είναι δυνατόν να χαλαρώσουν οι περιορισμοί με αποτέλεσμα να κατασκευαστεί ένα ανεκτικότερο πρόβλημα βελτιστοποίησης το οποίο ορίζεται ως εξής:

$$\text{ελαχιστοποίησε } 0.5\|w\|^2 + C \sum_{i=1}^N \xi_i^p$$

με περιορισμό

$$y_i(w^T x_i + b) \geq 1 - \xi_i, i = 1, \dots, N$$

όπου ξ_i είναι το σφάλμα για κάθε διάνυσμα εκπαίδευσης, C το κόστος (ανοχή) που δίνεται στο συνολικό σφάλμα και p μπορεί να πάρει την τιμή 1 ή 2. Η επιλογή της τιμής του C είναι σημαντική καθώς όσο μεγαλύτερη είναι η τιμή της, τόσο πιο αυστηρό είναι ο SVM στον καθορισμό του υπερεπιπέδου που μπορεί να διαχωρίσει επιτυχώς τα περισσότερα διανύσματα εκπαίδευσης. Μεγάλες τιμές του C μπορεί να οδηγήσουν στο φαινόμενο του υπερταϊριάσματος (overfitting) με αποτέλεσμα να μην μπορεί να γίνει γενίκευση όταν συμπεριληφθούν νέα δεδομένα. Αντίθετα, μικρές τιμές του C μπορεί να οδηγήσουν σε υποταΐριασμα (underfitting) που συνεπάγεται χαμηλά ποσοστά σωστής ταξινόμησης.

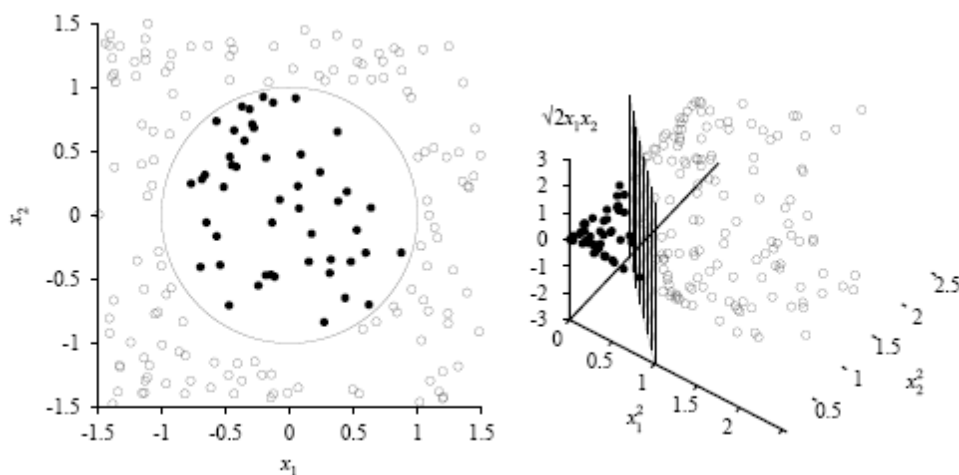
Σε πολλές εφαρμογές υπάρχει η περίπτωση τα διανύσματα εισόδου να μην μπορούν να διαχωριστούν με γραμμικό διαχωριστή (υπερεπίπεδο). Τότε μπορεί να εφαρμοστεί ένας μετασχηματισμός πυρήνα (kernel) ο οποίος μεταβιβάζει τα διανύσματα χαρακτηριστικών σε ένα χώρο περισσότερων διαστάσεων στον οποίο τα διανύσματα ενδεχόμενος μπορούν να διαχωριστούν γραμμικά (Εικόνα 25). Ο μετασχηματισμός πυρήνα ορίζεται ως

$$K(x, y) = \varphi(x)^T \varphi(y)$$

όπου $\varphi(x)$ είναι μια συνάρτηση που μετασχηματίζει ένα χώρο σε έναν άλλο μεγαλύτερης διάστασης από τον αρχικό. Με τον τρόπο αυτό εφαρμόζονται όσα αναφέρθηκαν παραπάνω με τη διαφορά ότι αντί για x_i χρησιμοποιείται το $\varphi(x_i)$. Οι πιο γνωστοί πυρήνες είναι

- γραμμικός
- πολυωνυμικός
- ακτινωτής βάσης (radial basis function)
- σιγμοειδής

Ο πιο διαδεδομένος είναι ο πυρήνας ακτινωτής βάσης $K(x, y) = \exp(-0.5\|x-y\|^2/\sigma^2)$ στον οποίο η παράμετρος σ καθορίζει το πλάτος της καμπάνας εκθετικής συνάρτησης K και επηρεάζει σε σημαντικό βαθμό τα αποτελέσματα της εκπαίδευσης και επομένως της ταξινόμησης.



Εικόνα 25. Παράδειγμα μετασχηματισμού πυρήνα που μεταφέρει τα διανύσματα χαρακτηριστικών από τις δύο διαστάσεις στις τρεις.

6 Η προτεινόμενη μεθοδολογία

Όπως αναφέρθηκε και στην εισαγωγή, στόχος της παρούσας πτυχιακής εργασίας είναι η επεξεργασία ενδοσκοπικών εικόνων που προέρχονται από ασύρματη ενδοσκόπηση με καινοτόμες τεχνικές επεξεργασίας σήματος ώστε να επιτευχθεί επιτυχημένη ταξινόμησή τους σε εικόνες που απεικονίζουν έλκος και σε εικόνες που απεικονίζουν φυσιολογικό ιστό. Η επεξεργασία των εικόνων θα βασιστεί στη μεθοδολογία που προτείνεται στην εργασία των B. Li και M.Q.H. Meng [23]. Σύμφωνα με την εργασία αυτή οι RGB ενδοσκοπικές εικόνες μεταφέρονται στον YCbCr χρωματικό χώρο προς επεξεργασία. Αρχικά εφαρμόζεται ο μετασχηματισμός Cuvvelet και στη συνέχεια εξάγονται τοπικά δυαδικά πρότυπα. Τέλος, η ταξινόμηση βασίζεται στις Μηχανές διανυσμάτων υποστήριξης. Ένα τυπικό σύστημα οπτικής αναγνώρισης αντικειμένων [24] απαρτίζεται από 4 στάδια: 1) προεπεξεργασία, 2) τμηματοποίηση, 3) εξαγωγή χαρακτηριστικών και 4) ταξινόμηση. Όπως γίνεται φανερό, στην παραπάνω μεθοδολογία αναγνώρισης ελκώδους ιστού απουσιάζει το στάδιο της τμηματοποίησης. Αυτό συμβαίνει διότι οι συγγραφείς δεν ενδιαφέρονται να εξετάσουν τεχνικές τμηματοποίησης αλλά επιθυμούν να επικεντρωθούν σε τεχνικές εξαγωγής χαρακτηριστικών υφής. Ως αποτέλεσμα, χρησιμοποιούν ενδοσκοπικές εικόνες εκ των προτέρων "κομμένες" στις περιοχές ενδιαφέροντος (ελκώδης ιστός, φυσιολογικός ιστός, πτυχώσεις, φυσαλίδες κλπ.).

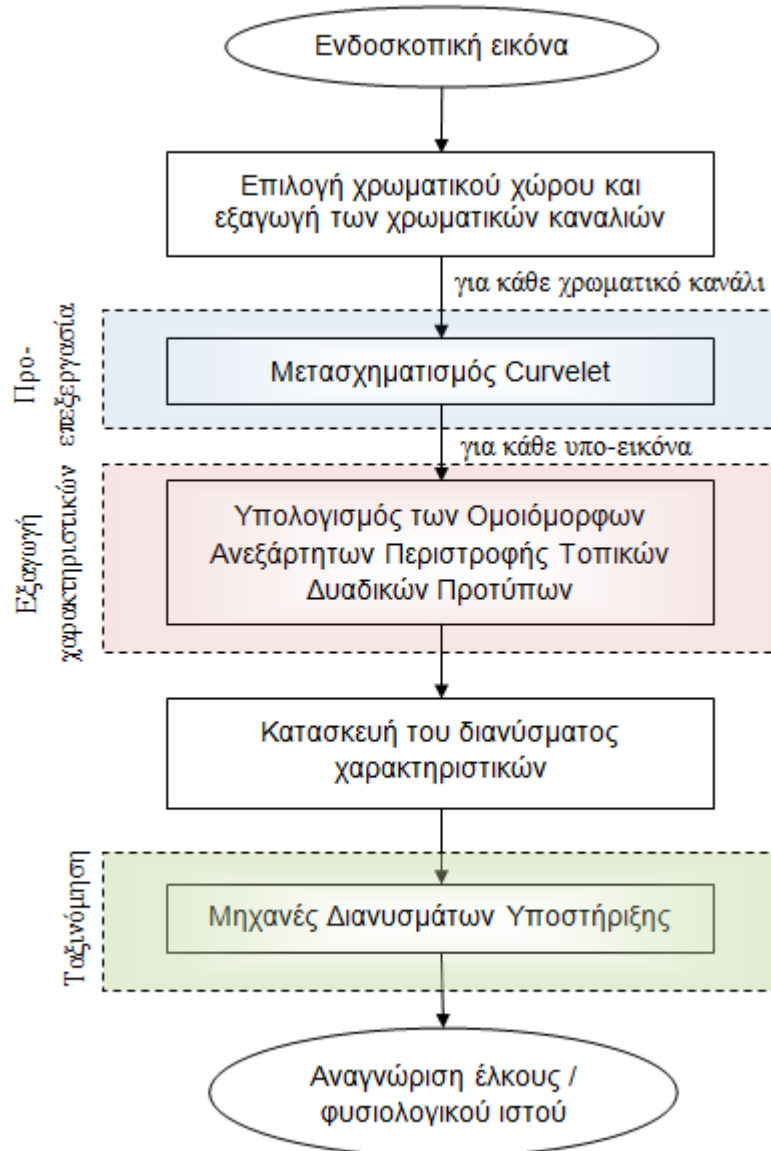
Στην παρούσα πτυχιακή εργασία πέρα από την ανάπτυξη της προτεινόμενης μεθοδολογίας γίνεται προσπάθεια για περαιτέρω διερεύνηση κ βελτίωση. Έτσι, λοιπόν, θα εξετάσουμε νέους χρωματικούς χώρους για την επεξεργασία των εικόνων ενώ θα γίνει προσπάθεια μελέτης της επιρροής των παραμέτρων του ταξινομητή στα αποτελέσματα επιτυχούς αναγνώρισης. Στην Εικόνα 26 φαίνεται το διάγραμμα ροής της μεθοδολογίας.

6.1 Εξαγωγή χρωματικών καναλιών

Οι ενδοσκοπικές εικόνες είναι εικόνες έγχρωμες και για το λόγο αυτό θα πρότεινε κάποιος να ασχοληθούμε με το πρόβλημα της χρωματικής υφής [23]. Παρά τις προσπάθειες που έχουν γίνει για την ανάλυση χρώματος και υφής ταυτόχρονα, η συγκεκριμένη επιστημονική περιοχή είναι ακόμη στα σπλάχνα της και τα αποτελέσματα είναι αναξιόπιστα. Συνεπώς, υπάρχουν διάφορες στρατηγικές αντιμετώπισης αυτού του θέματος. Η πιο δημοφιλής είναι η επεξεργασία του κάθε χρωματικού καναλιού ξεχωριστά [23] σαν να είχαμε τρεις μονοχρωματικές εικόνες κλίμακας του γκρι.

Επομένως, στο πρώτο στάδιο η εικόνα αποσυντίθεται στις τρεις χρωματικές συνιστώσες της. Παρόλο που η χρωματική πληροφορία της εικόνας δεν χρησιμοποιείται άμεσα ως χαρακτηριστικό ταξινόμησης, έχει διαπιστωθεί ότι διαφορετικοί χρωματικοί χώροι μπορούν να επηρεάσουν σημαντικά την απόδοση του συστήματος ταξινόμησης. Η εργασία [23] προτείνει η ανάλυση να γίνει στον χρωματικό χώρο YCbCr καθώς αναμένεται καλύτερη επίδοση καθώς υπάρχει διαχωρισμός στην πληροφορία χρώματος και φωτεινότητας. Εδώ θα εξετάσουμε επιπλέον τις δυνατότητες που προσφέρει τόσο ο χώρος RGB όσο και ο χώρος HSV.

Στην Εικόνα 27 φαίνεται ένα τυπικό παράδειγμα μιας φυσιολογικής εικόνας με σκοτεινή περιοχή και μιας εικόνας με έλκος αναλυμένες στις τρεις συνιστώσες Y, Cb, Cr. Μετά από παρατήρηση φαίνεται ότι η συνιστώσα Y διατηρεί σε μεγάλο βαθμό την

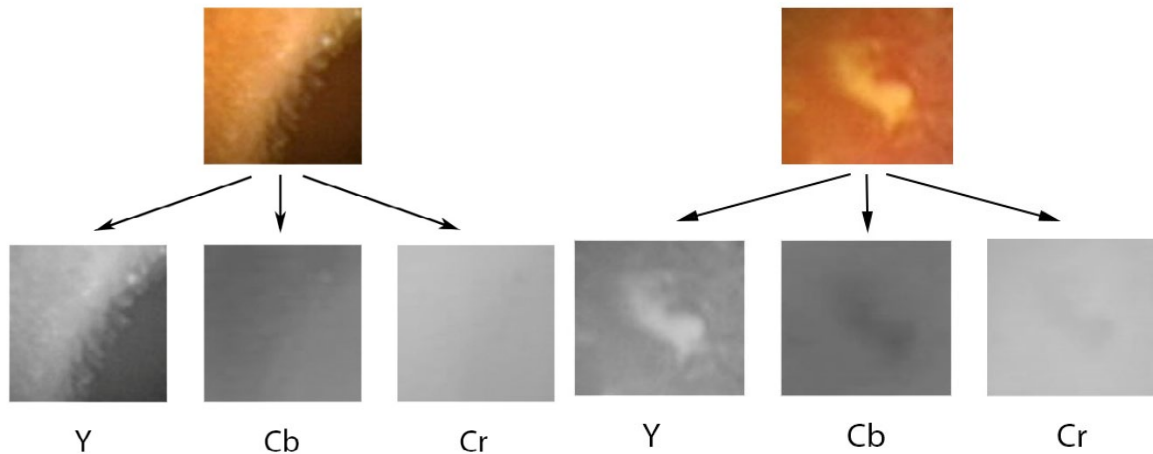


Εικόνα 26. Διάγραμμα ροής της μεθοδολογίας αναγνώρισης έλκους.

πληροφορία της αρχικής εικόνας. Ωστόσο αυτό δεν είναι πάντα επιθυμητό, όπως για παράδειγμα στην περίπτωση της φυσιολογικής εικόνας, όπου διατηρείται η έντονη αλλαγή στη φωτεινότητα. Αυτό πιθανόν να δυσκολέψει τη διαδικασία ταξινόμησης, γιατί η συνιστώσα Y της εικόνας με έλκος παρουσιάζει επίσης έντονη αλλαγή της φωτεινότητας. Αντίθετα, οι χρωματικές συνιστώσες C_b, C_r της φυσιολογικής εικόνας απαλλάσσονται σε μεγάλο βαθμό από την ανομοιογένεια στη φωτεινότητα που προκαλεί η σκοτεινή περιοχή, ενώ παράλληλα οι συνιστώσες C_b, C_r της εικόνας με έλκος διατηρεί μια σχετική εναλλαγή φωτεινότητας στην περιοχή όπου εντοπίζεται το έλκος.

6.2 Προεπεξεργασία των εικόνων

Στη συνέχεια, σε κάθε μια από τις χρωματικές συνιστώσες που προέκυψαν εφαρμόζεται ο γρήγορος μετασχηματισμός Curvelet με τη μέθοδο Wrapping (FDCT via Wrapping). Ο σκοπός του βήματος αυτού είναι να γίνει αποσύνθεση της εικόνας σε άλλες



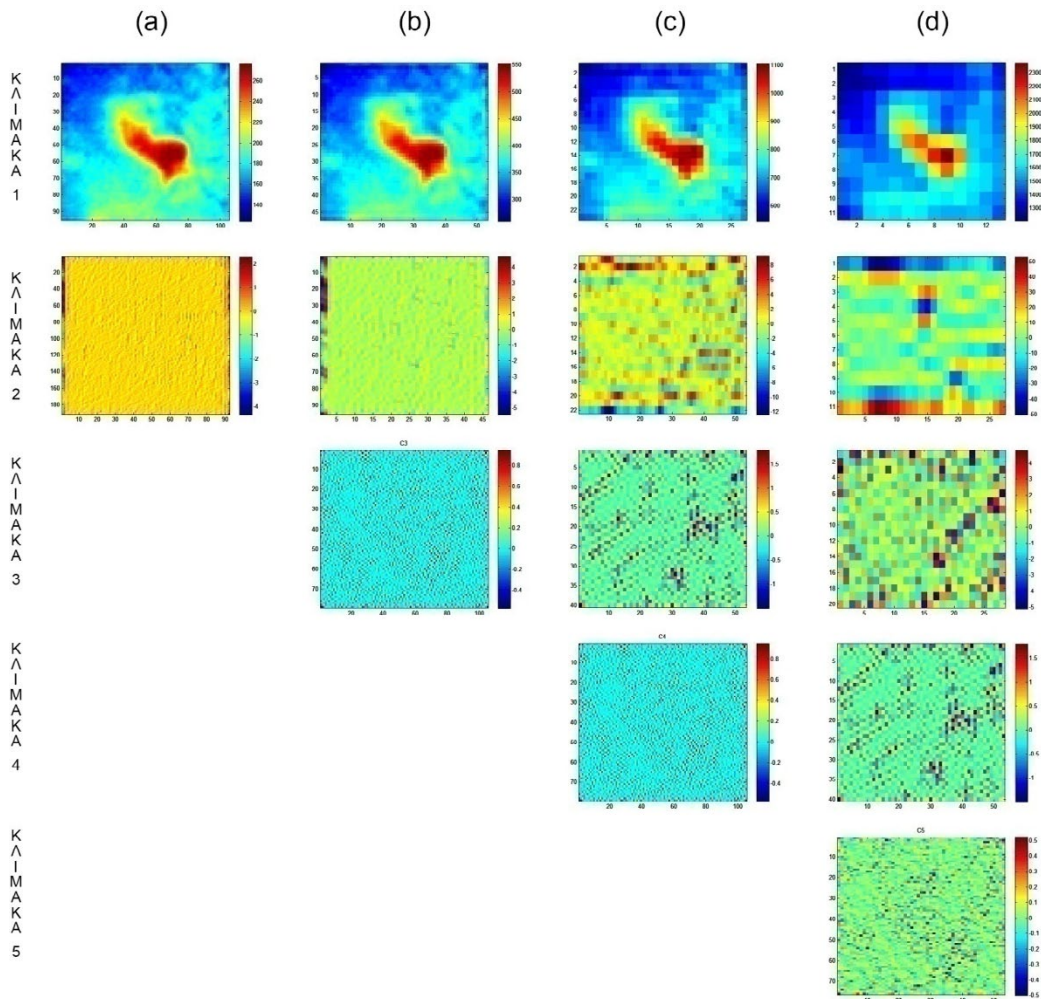
Εικόνα 27. Y, Cb, Cr συνιστώσες φυσιολογικής εικόνας (αριστερά) και εικόνας με έλκος (δεξιά)

μικρότερες με βάση κάποια κριτήρια κλίμακας και γωνίας ώστε να αναδειχθούν σημαντικές δομικές πληροφορίες και πληροφορίες υφής που εμπεριέχονται κρυμμένες κατά μήκος καμπυλόγραμμων σχημάτων και δομών μέσα στις εικόνες. Με άλλα λόγια, μετασχηματίζουμε την εικόνα σε άλλες μικρότερες υπο-εικόνες (στην πραγματικότητα δεν είναι εικόνες αλλά ομάδες συντελεστών) έτσι ώστε να γίνει πιο εύκολη και αποδοτική ή προσπάθεια εξαγωγής χαρακτηριστικών υφής που είναι απαραίτητα για την αυτοματοποιημένη διάκριση φυσιολογικού και παθολογικού ιστού. Για περισσότερες λεπτομέρειες για το μετασχηματισμό Curvelet και για τον τρόπο υπολογισμού των υπο-εικόνων σε διάφορες κλίμακες και γωνίες ο αναγνώστης παραπέμπεται στο τρίτο κεφάλαιο της παρούσας εργασίας αλλά και στα [26], [27], [28].

Ο μετασχηματισμός Curvelet απαιτεί από το χρήστη μεταξύ άλλων, τον καθορισμό δύο παραμέτρων:

- Αριθμός κλιμάκων αποσύνθεσης: Είναι ουσιαστικά το βάθος της αποσύνθεσης της εικόνας.
- Αριθμός γωνιών στη δεύτερη κλίμακα: Είναι ουσιαστικά το πλήθος των υποεικόνων που θα βρίσκονται στη δεύτερη κλίμακα αποσύνθεσης. Α αριθμός αυτός πρέπει να είναι τουλάχιστον 8 και πολλαπλάσιο του 4. Ο αριθμός γωνιών στις υπόλοιπες κλίμακες καθορίζεται αυτόματα.

Στην Εικόνα 28 φαίνεται η ανάλυση της συνιστώσας Y της εικόνας με έλκος (Εικόνα 27) για κλίμακες αποσύνθεσης από 2 έως 5 και αριθμό γωνιών 8. Για πρακτικούς λόγους, από κάθε κλίμακα αποσύνθεσης παρουσιάζεται μόνο η πρώτη γωνία, δηλαδή μια υπο-εικόνα. Όπως φαίνεται, καθώς το βάθος αποσύνθεσης αυξάνεται, τόσο μειώνεται η ανάλυση των εικόνων στις πρώτες κλίμακες. Από ένα σημείο και μετά η ανάλυση σε pixels των εικόνων στις πρώτες κλίμακες θα γίνει τόσο μικρή, που η πληροφορία υφής θα αρχίσει να χάνεται. Έτσι, συμπεραίνουμε ότι τίθεται ένα όριο στο μέγιστο βάθος αποσύνθεσης που έχει νόημα να χρησιμοποιήσουμε. . Για τις ενδοσκοπικές εικόνες που το μέγεθος τους είναι 256×256 pixels και ακόμα μικρότερο μετά την κατάτμηση των περιοχών ενδιαφέροντος, η ανάλυση σε 4 κλίμακες (εικόνα 5.6 (c)) φαίνεται να είναι ιδανική. Στην παρούσα εργασία θα

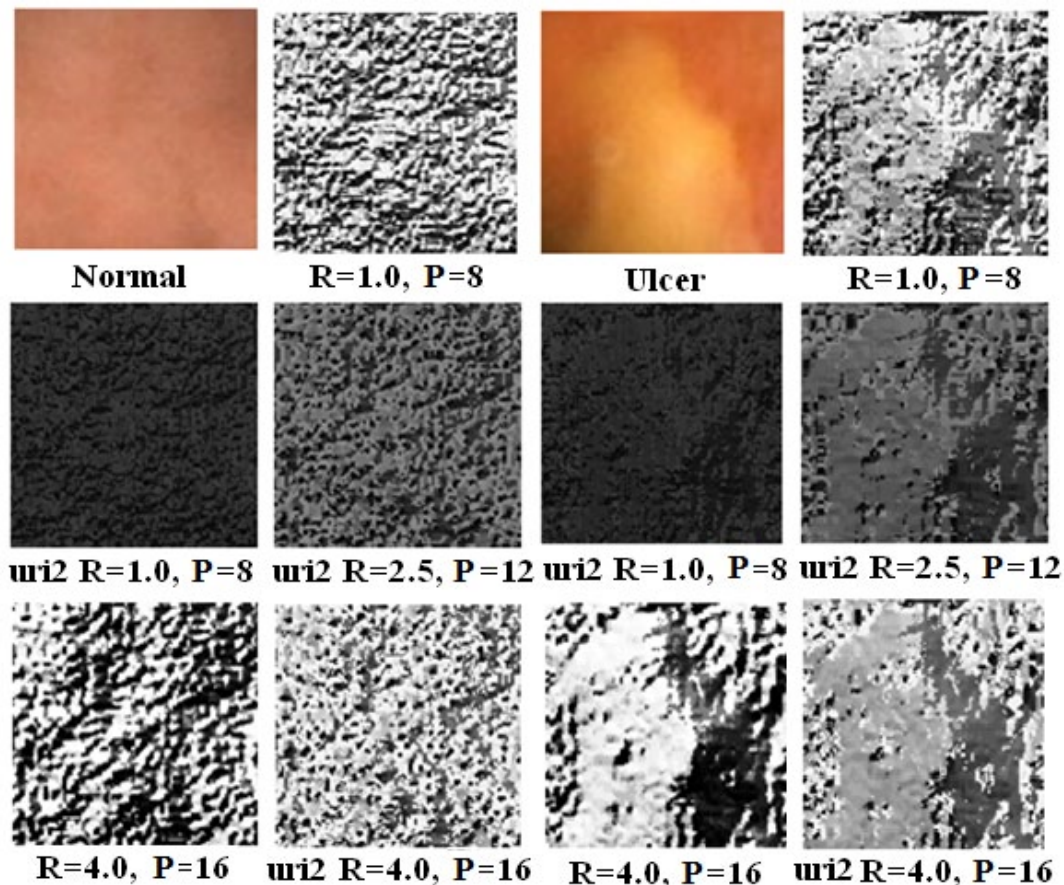


Εικόνα 28. FDCT εικόνας έλκους σε (α) 2 κλίμακες, (β) 3 κλίμακες, (γ) 4 κλίμακες, (δ) 5 κλίμακες και 8 γωνίες. Στην εικόνα απεικονίζεται μόνο η πρώτη γωνία για κάθε κλίμακα

χρησιμοποιήσουμε μόνο τις 2 πρώτες κλίμακες αποσύνθεσης και μόνο την πρώτη γωνία από κάθε κλίμακα.

6.3 Εξαγωγή χαρακτηριστικών υφής

Το επόμενο βήμα είναι η εξαγωγή των χαρακτηριστικών υφής. Αυτό επιτυγχάνεται με τη χρήση των ομοιόμορφων ανεξάρτητων περιστροφής τοπικών δυαδικών προτύπων. Οι ενδοσκοπικές εικόνες υπόκεινται σε διακυμάνσεις του φωτισμού που οφείλονται σε διάφορους λόγους όπως κίνηση της κάμερας, ανομοιόμορφες και περιορισμένες συνθήκες φωτισμού εντός του γαστρεντερικού σωλήνα, παρουσία πεπτικού υγρού και φυσαλίδων κ.α. Επομένως, πρέπει να λάβουμε υπόψη τις επιρροές των διαφοροποιήσεων του φωτισμού πάνω στην υφή των ενδοσκοπικών εικόνων καθώς τα χαρακτηριστικά της υφής δεν παραμένουν σταθερά σε τέτοιου είδους μεταβολές. Τα ομοιόμορφα ανεξάρτητα περιστροφής τοπικά δυαδικά πρότυπα επιδεικνύουν εύρωστη απόδοση σε διακυμάνσεις της φωτεινότητας. Επιπλέον, τα πρότυπα αυτά, μπορούν να διακρίνουν μικροδομές όπως φωτεινά και σκοτεινά σημεία, κάτι εξαιρετικά σημαντικό στην ανίχνευση παθήσεων από ενδοσκοπικές εικόνες καθώς πολλές παθήσεις, ανάμεσα σε αυτές και το έλκος, εμφανίζουν τοπικά χαρακτηριστικά μοτίβα. Για τους δύο λόγους που προαναφέρθηκαν γίνεται χρήση των ομοιόμορφων



Εικόνα 29. Απεικόνιση των LBP και των ομοιόμορφων ανεξάρτητων περιστροφής LBP (uir2) για φυσιολογική εικόνα (αριστερά) και εικόνα με έλκος (δεξιά).

ανεξάρτητων περιστροφής τοπικών δυαδικών προτύπων για την εξαγωγή χαρακτηριστικών υφής κατόπιν της εφαρμογής του μετασχηματισμού Curvelet. Στην Εικόνα 29 φαίνεται η ανάλυση δύο ενδοσκοπικών εικόνων (μια φυσιολογική και μια με έλκος) με τα τοπικά δυαδικά πρότυπα (LBP) και τα ομοιόμορφα ανεξάρτητα περιστροφής LBP για διάφορες τιμές των παραμέτρων P και R. Αρχικά να τονίσουμε ότι όσο μεγαλύτερος είναι ο αριθμός των γειτόνων P και η ακτίνα R τόσο καλύτερη η αναπαράσταση της υφής. Επίσης, από τη χρήση των απλών LBP φαίνεται ξεκάθαρα η ευαισθησία στην κατευθυντικότητα. Για παράδειγμα στην εικόνα έλκους R=4.0, P=16, η περιοχή έλκους απεικονίζεται σαν βουνό που φωτίζεται από τα αριστερά και η σκιά του πέφτει στα δεξιά. Αυτή η απεικόνιση θα άλλαζε και ενδεχομένως να μην ήταν τόσο ξεκάθαρη η υφή του έλκους εάν επιλέγαμε διαφορετική σειρά για τα βάρη ή εικόνα ήταν περιστρεμμένη. Αντίθετα, στην εικόνα έλκους uir2 R=4.0, P=16 η περιοχή του έλκους φαίνεται και πάλι ξεκάθαρα χωρίς όμως αυτή τη φορά να εμφανίζεται το πρόβλημα της κατευθυντικότητας. Επομένως, στην παρούσα εργασία επιλέχθηκαν οι τιμές R=4.0, P=16 για τον υπολογισμό των ομοιόμορφων ανεξάρτητων περιστροφής LBP τα οποία υπολογίζονται για κάθε υπο-εικόνα ξεχωριστά που προέκυψε από το προηγούμενο βήμα. Στη συνέχεια, υπολογίζεται το ιστόγραμμα των LBP (για κάθε υπο-εικόνα) το οποίο μπορεί να αποτελέσει το διάλυσμα χαρακτηριστικών.

6.4 Κατασκευή διανύσματος χαρακτηριστικών

Όπως αναφέρθηκε και προηγουμένως, τα ιστογράμματα των LBP κάθε υπο-εικόνας μπορούν να χρησιμοποιηθούν ως το διάνυσμα χαρακτηριστικών. Κάτι τέτοιο όμως θα οδηγούσε σε υπερβολικά μεγάλο σε διάσταση διάνυσμα το οποίο θα δυσκόλευε τον ταξινομητή στο να εκπαιδευτεί σωστά. Για το λόγο αυτό υπολογίζουμε 6 στατιστικά μεγέθη του ιστογράμματος των LBP που μπορούν να το περιγράψουν επαρκώς και με επιτυχία και χρησιμοποιούνται ευρέως σε τέτοιες εφαρμογές. Τα μεγέθη αυτά είναι:

- **Μέση Τιμή.** Η μέση τιμή των αριθμών x_1, x_2, \dots, x_v ορίζεται ως

$$M = \frac{1}{v} \sum_{i=1}^v x_i$$

- **Τυπική Απόκλιση.** Η τυπική απόκλιση των αριθμών x_1, x_2, \dots, x_v ορίζεται ως

$$TA = \sqrt{\frac{\sum_{i=1}^v (x_i - M)^2}{v}}$$

όπου M είναι η μέση τιμή.

- **Ενέργεια.** Η ενέργεια των αριθμών x_1, x_2, \dots, x_v ορίζεται ως

$$E = \sum_{i=1}^v x_i^2$$

- **Εντροπία.** Η εντροπία εκφράζει την τυχαιότητα και την πολυπλοκότητα του ιστογράμματος. Η εντροπία των αριθμών x_1, x_2, \dots, x_v ορίζεται ως

$$ENT = - \sum_{i=1}^v p_i \log_2 p_i$$

όπου p_i είναι η πιθανότητα εμφάνισης της τιμής x_i .

- **Συντελεστής Κύρτωσης.** Ο συντελεστής κύρτωσης εκφράζει την αιχμηρότητα ή την ύπαρξη έντονων κορυφών στο ιστόγραμμα και για τους αριθμούς x_1, x_2, \dots, x_v ορίζεται ως

$$K = \frac{1}{v * TA^4} \sum_{i=1}^v (x_i - M)^4$$

όπου M είναι η μέση τιμή και TA η τυπική απόκλιση.

- **Συντελεστής Λοξότητας ή Ασυμμετρίας.** Ο συντελεστής λοξότητας εκφράζει το κατά πόσο είναι ομοιόμορφα κατανομημένες οι τιμές του ιστογράμματος γύρω από τη μέση τιμή και για τους αριθμούς x_1, x_2, \dots, x_v ορίζεται ως

$$K = \frac{1}{v * TA^3} \sum_{i=1}^v (x_i - M)^3$$

όπου M είναι η μέση τιμή και TA η τυπική απόκλιση.

Επομένως, αντί να χρησιμοποιήσουμε όλες τις τιμές του ιστογράμματος χρησιμοποιούμε μόνο αυτές τις 6 τιμές οι οποίες αποτελούν τα χαρακτηριστικά (features) που θα χρησιμοποιήσουμε για την ταξινόμηση των εικόνων.

Υπενθυμίζουμε ότι τα 6 αυτά στατιστικά μεγέθη υπολογίζονται για κάθε υπο-εικόνα που προέκυψε από το δεύτερο βήμα. Επομένως έχουμε συνολικά (3 χρωμ. κανάλια) \times (2 υπο-εικόνες/κανάλι) \times (6 χαρακτηριστικά/υπο-εικόνα) = 36 χαρακτηριστικά, 12 για κάθε χρωματικό κανάλι. Με σκοπό τη διερεύνηση της αποτελεσματικότητας κάθε χρωματικού χώρου αλλά και κάθε χρωματικού καναλιού εξετάζουμε 4 διαφορετικά σενάρια διανύσματος χαρακτηριστικών για κάθε χρωματικό χώρο.

- Σενάριο 1: Το διάνυσμα χαρακτηριστικών περιέχει μόνο τα features από το 1ο χρωματικό κανάλι (π.χ. Κόκκινο)
- Σενάριο 2: Το διάνυσμα χαρακτηριστικών περιέχει μόνο τα features από το 2ο χρωματικό κανάλι (π.χ. Πράσινο)
- Σενάριο 3: Το διάνυσμα χαρακτηριστικών περιέχει μόνο τα features από το 3ο χρωματικό κανάλι (π.χ. Μπλε)
- Σενάριο 4: Το διάνυσμα χαρακτηριστικών περιέχει όλα τα features και από το 3 χρωματικά κανάλια (π.χ. Κόκκινο-Πράσινο-Μπλε)

6.5 Ταξινόμηση

Η διαδικασία της ταξινόμησης είναι ένα από τα πιο ζωτικής σημασίας στάδια του επιστημονικού πεδίου Αναγνώρισης Προτύπων. Τα χαρακτηριστικά που εξάγαμε στο προηγούμενο βήμα μας είναι ουσιαστικά άχρηστα εάν δεν υπάρχει κάποιος αλγόριθμος ταξινόμησης να τα χρησιμοποιήσει και αξιολογήσει. Ένας αλγόριθμος ταξινόμησης (ονομάζεται και ταξινομητής) δέχεται ως είσοδο ένα διάνυσμα χαρακτηριστικών και παρέχει στην έξοδο την απόφασή του σχετικά με το είδος ή την προέλευση του διανύσματος αυτού. Για να υπάρξει όμως αυτή η δυνατότητα της εκτίμησης πρέπει πρώτα ο ταξινομητής να εκπαιδευτεί πάνω στα διανύσματα χαρακτηριστικών και τις αντίστοιχες κατηγορίες στις οποίες ανήκουν. Επομένως, αρχικά δίνουμε σαν είσοδο στον ταξινομητή ένα πλήθος διανυσμάτων χαρακτηριστικών καθώς και την κατηγορία στην οποία ανήκει το κάθε διάνυσμα. Έχοντας αυτά τα στοιχεία ο ταξινομητής εκπαιδεύεται από μόνος του ώστε να αναγνωρίζει ποια χαρακτηριστικά αντιστοιχούν σε κάθε κατηγορία. Στη συνέχεια, πρέπει να ελέγξουμε εάν ο ταξινομητής εκπαιδεύτηκε σωστά. Για το λόγο αυτό, δίνουμε στην είσοδο διανύσματα χαρακτηριστικών (διαφορετικά από τα προηγούμενα που χρησιμοποιήθηκαν στην εκπαίδευση) των οποίων όμως ξέρουμε εκ των προτέρων την κατηγορία. Συλλέγουμε τα αποτελέσματα που έβγαλε ο ταξινομητής και τα συγκρίνουμε με τα πραγματικά. Με τον τρόπο αυτό υπολογίζουμε την απόδοση του ταξινομητή, την πιθανότητα δηλαδή να προβλέψει σωστά την κατηγορία στην οποία ανήκει ένα διάνυσμα χαρακτηριστικών. Η απόδοση και η αποτελεσματικότητα ενός ταξινομητή μετράται με τους δείκτες: ακρίβεια (accuracy), ευαισθησία (sensitivity) και ειδικότητα (specificity).

- Η ακρίβεια ορίζεται ως

$$Acc = \frac{TP + TF}{TP + TF + FP + FN}$$

- Η ευαισθησία ορίζεται ως

$$Sens = \frac{TP}{TP + FN}$$

- Η ειδικότητα ορίζεται ως

$$Spec = \frac{TN}{TN + FP}$$

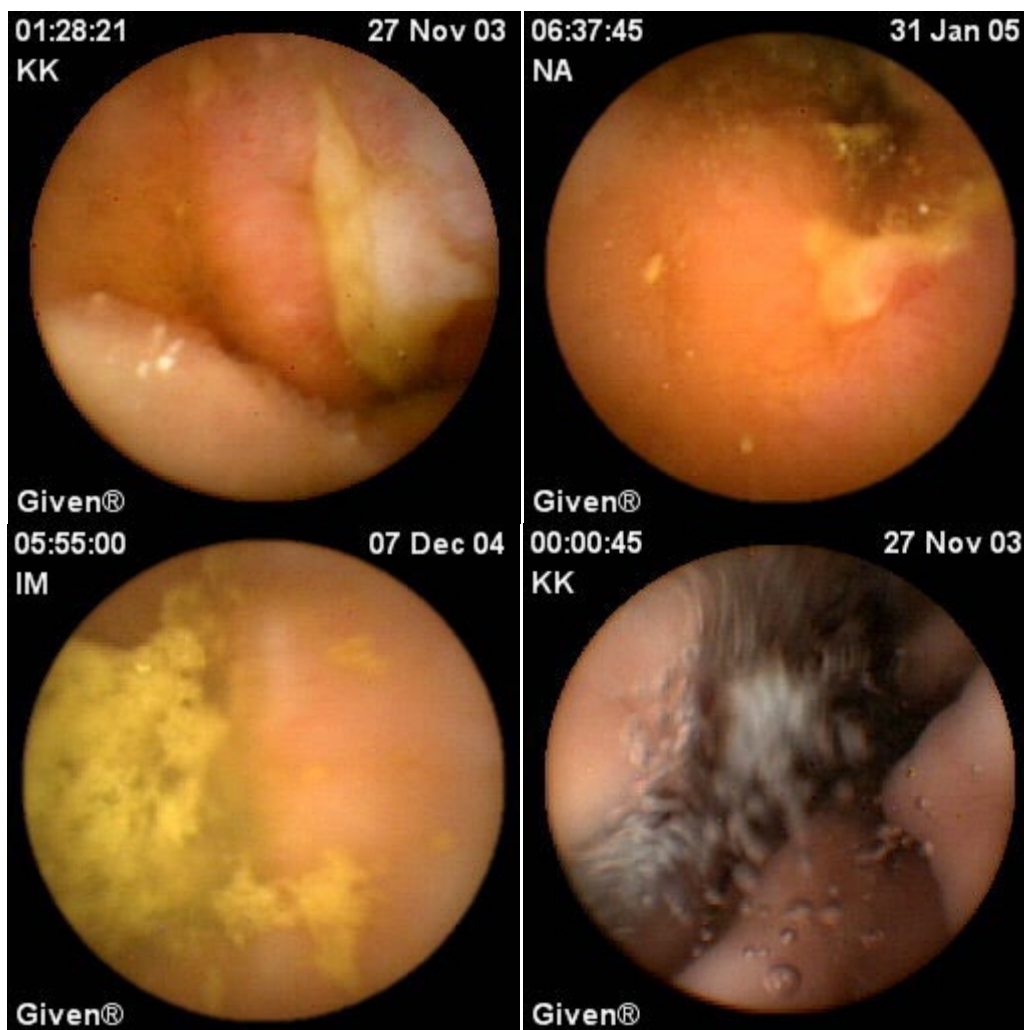
όπου TN είναι το σύνολο των σωστά ταξινομημένων αρνητικών δειγμάτων (υγιών εικόνων στην περίπτωση μας), TP το σύνολο των σωστά ταξινομημένων θετικών δειγμάτων (εικόνων με έλκος), FN είναι το σύνολο των λανθασμένα ταξινομημένων θετικών δειγμάτων ως αρνητικά (εικόνες με έλκος ταξινομημένες ως υγιείς) και FP είναι το σύνολο των λανθασμένα ταξινομημένων αρνητικών δειγμάτων ως θετικά (εικόνες υγιείς ταξινομημένες ως εικόνες με έλκος). Στην περίπτωση της χρήσης ενός ταξινομητή για τη διάγνωση μίας παθολογίας είναι φανερό ότι είναι προτιμότερο να αναγνωριστούν κάποιες υγιείς εικόνες σαν ασθενείς παρά το αντίστροφο. Έτσι, δίνουμε έμφαση στην τιμή της ευαισθησίας σε σχέση με την τιμή της ειδικότητας.

Στην παρούσα εργασία χρησιμοποιήθηκαν οι μηχανές διανυσμάτων υποστήριξης (ΜΔΥ) για την υλοποίηση της διαδικασίας της ταξινόμησης. Βασική παράμετρος του συγκεκριμένου ταξινομητή είναι η συνάρτηση πυρήνα. Επιλέξαμε να μελετήσουμε τη συμπεριφορά από δύο είδη πυρήνων: το γραμμικό (linear) και της ακτινωτής βάσης (radial basis function - rbf). Επίσης, για την περίπτωση της rbf εξετάσαμε το πώς επηρεάζει το αποτέλεσμα η τιμή της παραμέτρου σίγμα (δεν υπάρχει αντίστοιχη παράμετρος για την περίπτωση της linear), η οποία κυμάνθηκε από 1 έως 5 με βήμα 1.

Για την εκπαίδευση και τον έλεγχο του ταξινομητή χρησιμοποιήθηκε η μεθοδολογία 10-fold cross validation. Σύμφωνα με τη μέθοδο αυτή το σύνολο των διαθέσιμων δειγμάτων χωρίζεται τυχαία σε 10 ισομεγέθη υποσύνολα. Από τα υποσύνολα αυτά τα 9 χρησιμοποιούνται για την εκπαίδευση του ταξινομητή ενώ το εναπομείναν λειτουργεί ως υποσύνολο αξιολόγησης για τον έλεγχο της απόδοσης του ταξινομητή. Η όλη διαδικασία επαναλαμβάνεται 10 φορές, με κάθε ένα από τα 10 υποσύνολα να χρησιμοποιείται μόνο μία φορά για τον έλεγχο του ταξινομητή. Στο τέλος κάθε επανάληψης καταγράφεται η τιμή της ακρίβειας, ευαισθησίας και ειδικότητας που προέκυψε. Στη συνέχεια υπολογίζεται ο μέσος όρος των 10 αποτελεσμάτων ώστε να παραχθεί μια ενιαία εκτίμηση (για κάθε δείκτη) της απόδοσης του ταξινομητή. Το πλεονέκτημα του 10-fold cross validation έναντι άλλων αντίστοιχων μεθόδων είναι ότι το σύνολο των δειγμάτων χρησιμοποιείται τόσο για την εκπαίδευση όσο και για τον έλεγχο του ταξινομητή και κάθε δείγμα χρησιμοποιείται μία φορά αποκλειστικά για έλεγχο. Ένα μικρό πρόβλημα που αντιμετωπίζεται στη μεθοδολογία 10-fold cross validation είναι η ανάγκη για τυχαία κατανομή των διαθέσιμων δειγμάτων σε 10 υποσύνολο. Στην πράξη η κατανομή είναι ψευδοτυχαία. Για το λόγο αυτό η παραπάνω διαδικασία επαναλαμβάνεται πενήντα φορές και εξάγουμε το μέσο όρο ώστε τα αποτελέσματα της ταξινόμησης να μη διαφέρουν μέχρι το δεύτερο δεκαδικό ψηφίο.

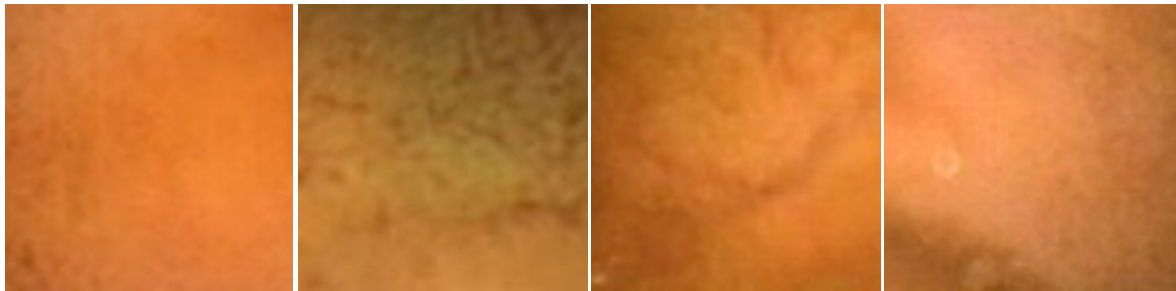
6.6 Δεδομένα

Οι ενδοσκοπικές εικόνες που χρησιμοποιήθηκαν στην παρούσα εργασία για την ανάπτυξη και αξιολόγηση της μεθόδου προέρχονται από τον άτλαντα ενδοσκοπικών εικόνων της εταιρίας Given Imaging. Το σύνολο των δεδομένων αποτελείται από 150 εικόνες που απεικονίζουν έλκος και 150 εικόνες που απεικονίζουν φυσιολογικό ιστό. Οι εικόνες έλκους απεικονίζουν ελκώσεις που προέρχονται είτε από άγνωστη αιτία είτε από διάφορες ασθένειες όπως νόσος του Crohn, ελκώδης κολίτιδα, ελκώσεις από χρήση μη-στεροειδών αντιφλεγμονωδών φαρμάκων. Πρέπει να αναφερθεί ότι οι 150 εικόνες έλκους προέρχονται από 150 διαφορετικά περιστατικά και δεν αποτελούν πολλαπλές απεικονίσεις της ίδιας περιοχής ώστε να επιτευχθεί όσο το δυνατόν μικρότερη ομοιότητα ανάμεσά τους. Οι φυσιολογικές εικόνες εκτός από φυσιολογικό ιστό απεικονίζουν φυσαλίδες από τα γαστρικά υγρά, πτυχώσεις, ίχνη από περιττώματα, κ.α. ώστε να ανταποκρίνονται σε πραγματικές συνθήκες και η διαδικασία αναγνώρισης να γίνει πιο δύσκολη. Γενικά, οι ενδοσκοπικές εικόνες από ασύρματη κάψουλα παρουσιάζουν μεγάλη μεταβλητότητα στα χρώματα, σκοτεινές περιοχές λόγω ανομοιογενούς φωτισμού και διαφορετικής απόστασης της κάμερας από τα τοιχώματα του γαστρεντερικού σωλήνα και λάμπεις από την αντανάκλαση του φωτός των LED της κάψουλας στα γαστρικά υγρά.

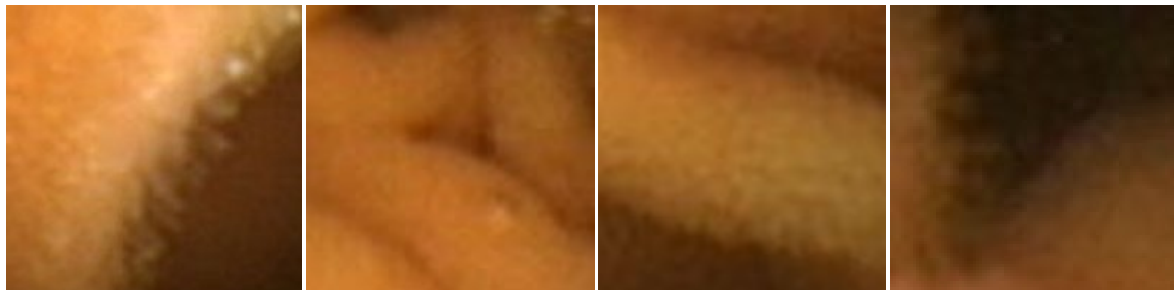


Εικόνα 30. Ενδοσκοπικές εικόνες από ασύρματη ενδοσκοπική κάψουλα

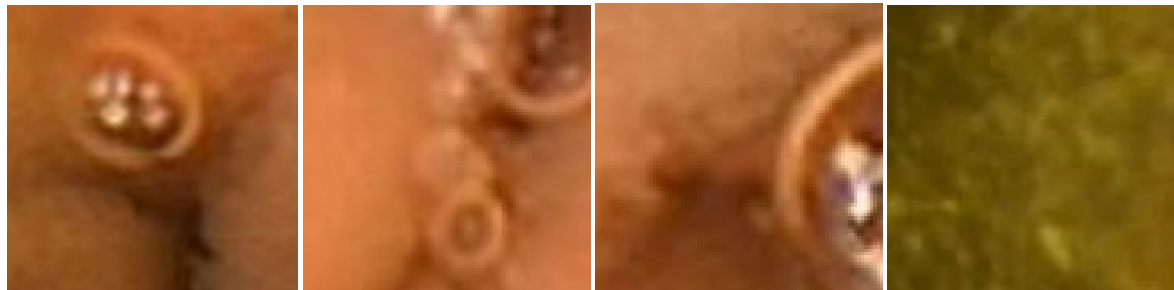
Οι εικόνες που λαμβάνονται από την ενδοσκοπική κάψουλα είναι μεγέθους 320x320 pixel με βάθος χρώματος 8 bit. Στην Εικόνα 30 φαίνονται μερικές ενδοσκοπικές εικόνες όπως καταγράφονται από την κάψουλα και εξετάζονται από τον ιατρό. Ωστόσο, οι εικόνες δε χρησιμοποιήθηκαν στην παραπάνω μορφή, αλλά κατατμήθηκαν χειροκίνητα ώστε να εστιαστεί η περιοχή ενδιαφέροντος και να αφαιρεθεί το μαύρο περίγραμμα που προσδίδει άχρηστη πληροφορία. Στην Εικόνα 31 φαίνονται διάφορες περιπτώσεις εικόνων με



(α)



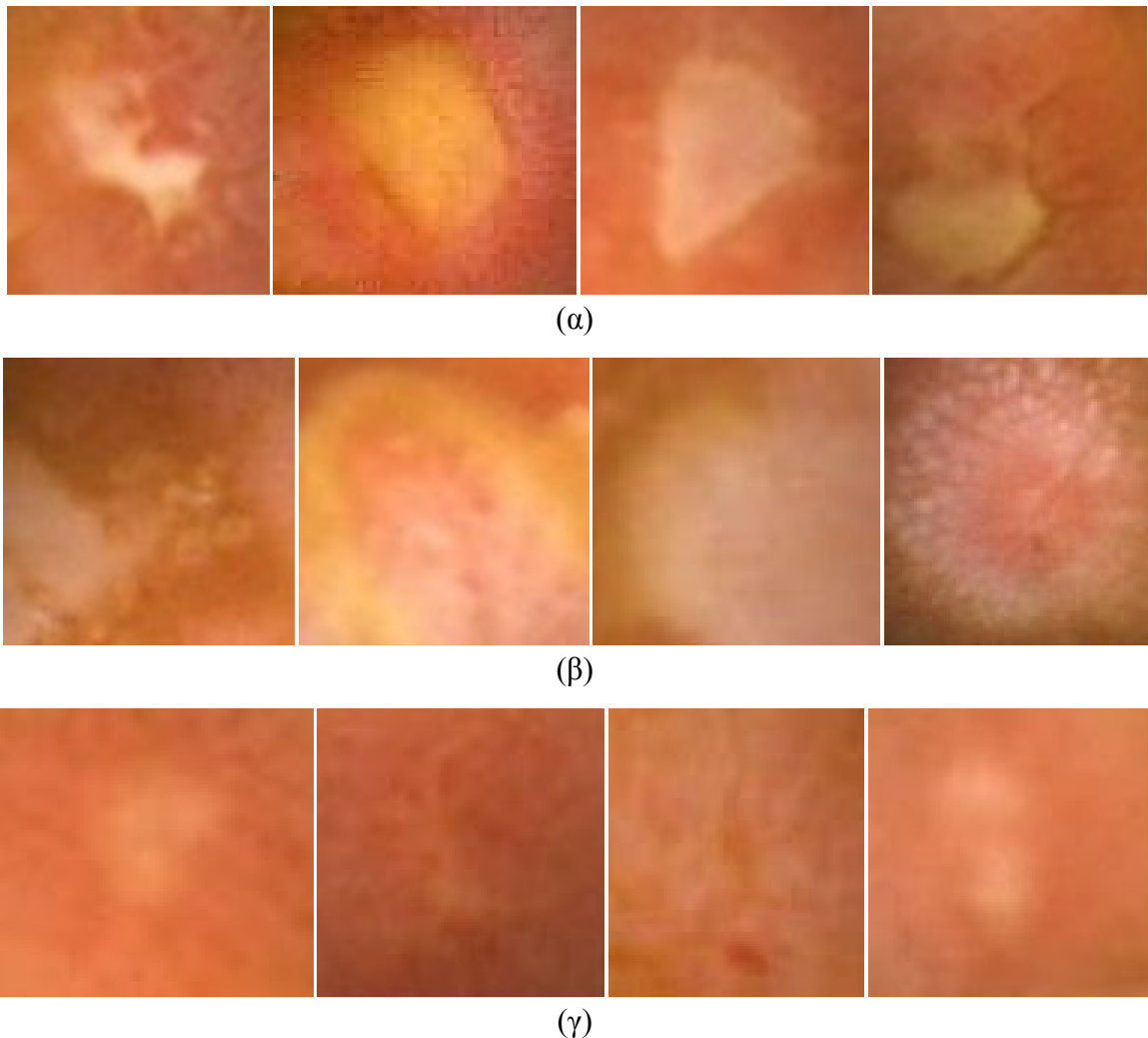
(β)



(γ)

Εικόνα 31. Ενδοσκοπικές εικόνες από ασύρματη ενδοσκοπική κάψουλα που απεικονίζουν φυσιολογικό ιστό διαφορετικής δυσκολίας. (α) χωρίς αλλοίωση, (β) σκοτεινές περιοχές και πτυχώσεις, (γ) φυσαλίδες και γαστρικά υγρά

φυσιολογικό ιστό όπως προέκυψαν μετά την κατάτμηση. Οι εικόνες της σειράς (α) είναι φυσιολογικές εικόνες οι οποίες δεν έχουν αλλοιωθεί σημαντικά από προσμίξεις και μεταβολές στο φωτισμό. Οι εικόνες της σειράς (β) περιέχουν σκοτεινές περιοχές και πτυχώσεις ενώ στη σειρά (γ) παρουσιάζονται εικόνες με σκοτεινές περιοχές, αναδιπλώσεις και λάχνες αντίστοιχα. Σε ιδανικές συνθήκες οι εικόνες θα ήταν όπως στη σειρά (α), αλλά στην πραγματικότητα δε συμβαίνει αυτό. Έτσι, το σύστημα πρέπει να είναι εύρωστο ώστε να αντιμετωπίζει αυτές τις δυσκολίες. Στην Εικόνα 32 παρουσιάζονται εικόνες με έλκος μεταβαλλόμενης δυσκολίας.



Εικόνα 32. Ενδοσκοπικές εικόνες από ασύρματη ενδοσκοπική κάψουλα που απεικονίζουν ελκώδη ιστό διαφορετικής δυσκολίας. (α) εύκολα αναγνωρίσιμο έλκος, (β) μέτριας δυσκολίας αναγνωρίσιμο έλκος, (γ) αναπτυσσόμενο έλκος δύσκολα αναγνωρίσιμο.

6.7 Υλοποίηση της μεθοδολογίας

Η εργασία υλοποιήθηκε στο περιβάλλον προγραμματισμού MATLAB λόγω της εύκολης διαχείρισης εικόνων καθώς και την πληθώρα έτοιμων συναρτήσεων επεξεργασίας σήματος και εικόνων. Μερικές από τις βασικές εντολές - συναρτήσεις του MATLAB που χρησιμοποιήθηκαν είναι οι εξής:

- $im = imread("u1.png");$ Εισαγωγή εικόνας στο workspace
- $imshow(im);$ Απεικόνιση εικόνας σε figure
- $im_new = rgb2ycbcr(im);$ Μετατροπή εικόνας από RGB σε YCbCr
- $im_new2 = rgb2hsv(im);$ Μετατροπή εικόνας από RGB σε HSV
- $C = fdct_wrapping(double(im(:,:,1)),~,4,8);$ Υπολογισμός των συντελεστών (υπο-εικόνων) του μετασχηματισμού Curvelet της πρώτης συνιστώσας της εικόνας im για 4 κλίμακες αποσύνθεσης με 8 γωνίες στη δεύτερη κλίμακα. Σημείωση: Όλες οι εντολές που αφορούν το μετασχηματισμό Curvelet παρέχονται δωρεάν από το www.curvelet.org

- $lbp1 = uLBP(im(:,:,1),P,R)$; Υπολογισμός των ομοιόμορφων ανεξάρτητων περιστροφής προτύπων του πρώτου καναλιού της εικόνας im με P γείτονες και R ακτίνα κύκλου. Σημείωση: Όλες οι εντολές που αφορούν τα LBP παρέχονται δωρεάν από το www.mathworks.com
- $x = hist(lbp1)$; Υπολογισμός του ιστογράμματος των LBP.
- $f1 = mean(x)$; Υπολογισμός της μέσης τιμής του ιστογράμματος
- $f2 = std(x)$; Υπολογισμός της τυπικής απόκλισης του ιστογράμματος
- $f3 = entropy(x)$; Υπολογισμός της εντροπίας του ιστογράμματος
- $f4 = skewness(x)$; Υπολογισμός της λοξότητας του ιστογράμματος
- $f5 = kurtosis(x)$; Υπολογισμός της κύρτωσης του ιστογράμματος
- $SVMStruct = svmtrain(trainset, group, 'Kernel_Function', 'rbf', 'RBF_Sigma', 2)$; Εκπαίδευση ενός SVM ταξινομητή χρησιμοποιώντας το $trainset$ για εκπαίδευση, συνάρτηση πυρήνα την *radial basis function (rbf)* και τιμή 2 για την παράμετρο σίγμα. Το $group$ περιέχει την κλάση στην οποία ανήκουν τα δείγματα του $trainset$.
- $class = svmclassify(SVMStruct, testset)$; Έλεγχος του ταξινομητή που προέκυψε από την προηγούμενη εντολή χρησιμοποιώντας ως δείγματα ελέγχου τα δεδομένα της μεταβλητής $testset$.

7 Αποτελέσματα

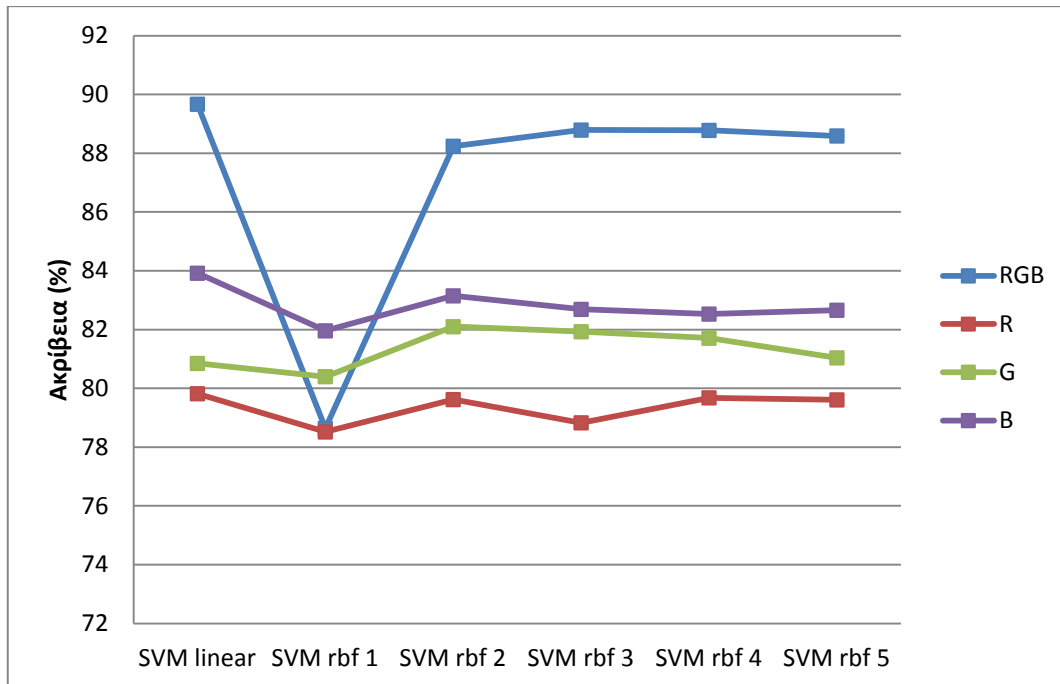
7.1 Χρωματικός χώρος RGB

Αρχικά παρουσιάζουμε τα πειραματικά αποτελέσματα που προέκυψαν όταν η επεξεργασία των εικόνων έγινε στο χρωματικό χώρο RGB. Στον Πίνακα 1 φαίνονται τα ποσοστά της ακρίβειας, ευαισθησίας και ειδικότητας και για τα τέσσερα σενάρια του διανύσματος χαρακτηριστικών και όλα τα είδη ταξινομητή (SVM με συνάρτηση πυρήνα γραμμική, σημειώνεται ως SVM linear, SVM με συνάρτηση πυρήνα ακτινωτής βάσης, σημειώνεται ως SVM rbf, για τιμές της παραμέτρου σίγμα 1 - 5. Στην Εικόνα 33 απεικονίζεται η ακρίβεια για όλους τους ταξινομητές και όλα τα σενάρια. Αντίστοιχα, στις Εικόνες 34 και 35 απεικονίζεται η ευαισθησία και η ειδικότητα. Ο αριθμός που βρίσκεται δίπλα στην ονομασία SVM rbf υποδηλώνει την τιμή που έχει η παράμετρος σίγμα.

Μέτρο Απόδοσης	Ταξινομητής	Σενάριο Διανύσματος Χαρακτηριστικών			
		RGB	R	G	B
Ακρίβεια	SVM linear	89,67%	79,82%	80,85%	83,92%
	SVM rbf 1	78,64%	78,52%	80,4%	81,96%
	SVM rbf 2	88,24%	79,62%	82,1%	83,15%
	SVM rbf 3	88,79%	78,83%	81,93%	82,69%
	SVM rbf 4	88,78%	79,68%	81,71%	82,53%
	SVM rbf 5	88,59%	79,61%	81,04%	82,66%
Ευαισθησία	SVM linear	89,26%	81,73%	85,42%	87,33%
	SVM rbf 1	64,81%	77,48%	83,04%	83,69%
	SVM rbf 2	85,7%	80,56%	88,14%	84,74%
	SVM rbf 3	89,24%	79,93%	88,66%	86,32%
	SVM rbf 4	90,16%	83,29%	89,33%	86,73%
	SVM rbf 5	90,17%	84,32%	88,68%	87,46%
Ειδικότητα	SVM linear	90,06%	77,9%	76,28%	80,5%
	SVM rbf 1	92,47%	79,57%	77,77%	80,22%
	SVM rbf 2	90,79%	78,68%	76,05%	81,56%
	SVM rbf 3	88,33%	77,73%	75,2%	79,06%
	SVM rbf 4	87,44%	76,06%	74,09%	78,33%
	SVM rbf 5	87,01%	74,9%	73,4%	77,85%

Πίνακας 1. Αποτελέσματα επεξεργασίας των εικόνων στο χρωματικό χώρο RGB.

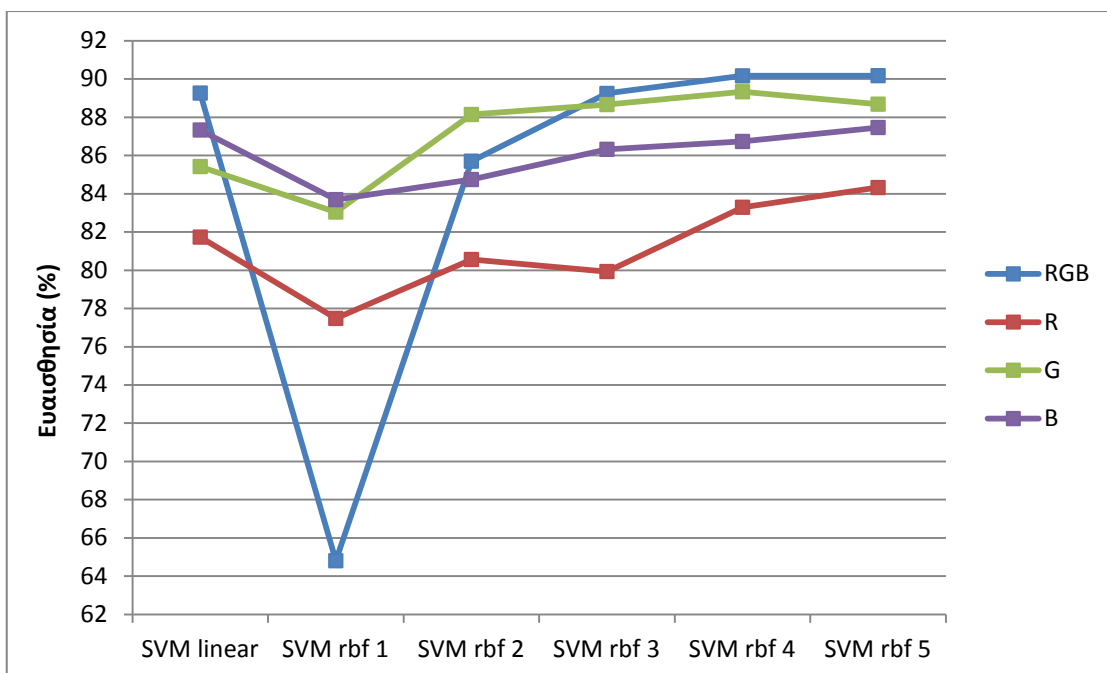
Κοιτώντας την Εικόνα 33 παρατηρούμε ότι τα μεγαλύτερα ποσοστά ακρίβειας της ταξινόμησης επιτυγχάνονται όταν το διάνυσμα χαρακτηριστικών περιλαμβάνει τα χαρακτηριστικά και από τα τρία χρωματικά κανάλια (σενάριο RGB). Εάν εξετάσουμε το κάθε χρωματικό κανάλι ξεχωριστά διαπιστώνουμε ότι το πιο αποτελεσματικό είναι το μπλε, έπειτα το πράσινο και τέλος το κόκκινο. Το μεγαλύτερο ποσοστό ακρίβειας για κάθε σενάριο, όπως προκύπτει από τον Πίνακα 1, είναι: 89.67% για το RGB (ταξινομητής SVM linear), 83.92% για το B (ταξινομητής SVM linear), 82.10% για το G (ταξινομητής SVM rbf 2) και 79.82% για το R (ταξινομητής SVM linear). Από τα παραπάνω αποτελέσματα προκύπτει ότι είναι σημαντικό να χρησιμοποιήσουμε την πληροφορία και από τα τρία χρωματικά κανάλια ταυτόχρονα καθώς επιτυγχάνουμε από 6% έως 10% περίπου μεγαλύτερη



Εικόνα 33. Ποσοστά ακρίβειας για τον χώρο RGB.

ακρίβεια σε σχέση με τη χρήση των πληροφοριών από μεμονωμένα κανάλια. Συγκρίνοντας τα τρία χρωματικά κανάλια μεταξύ τους παρατηρούμε ότι τα αποτελέσματα που επιτυγχάνονται έχουν πολύ μικρές διαφορές. Η επεξεργασία στο μπλε κανάλι επιφέρει, όσον αφορά την ακρίβεια 2%, περίπου υψηλότερη απόδοση σε σχέση με το πράσινο και 4% σε σχέση με το κόκκινο. Επομένως μπορούμε να υποθέσουμε ότι τα χαρακτηριστικά υφής του έλκους κατανέμονται σχεδόν ομοιόμορφα ανάμεσα στα τρία χρώματα.

Όσον αφορά την ευαισθησία (Εικόνα 34), και πάλι ο συνδυασμός των χρωμάτων επιτυγχάνει το υψηλότερο ποσοστό (90.17% με τον ταξινομητή SVM rbf 5) ενώ ακολουθεί

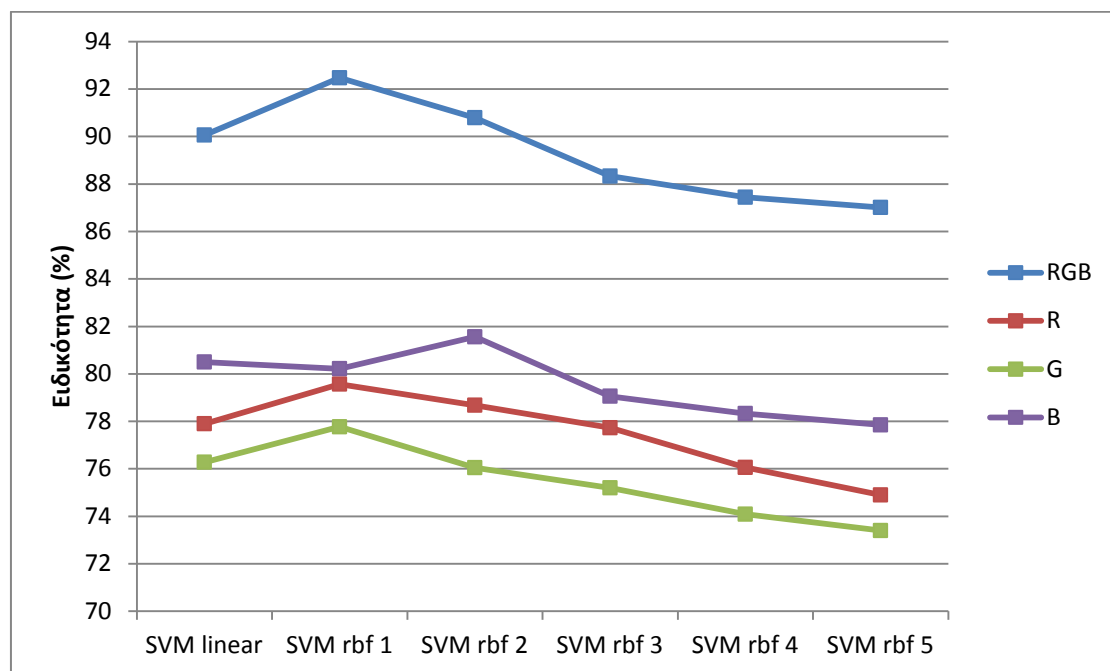


Εικόνα 34. Ποσοστά ευαισθησίας για τον χώρο RGB.

το πράσινο κανάλι (89.33% με τον ταξινομητή SVM rbf 4), το μπλε (87.46% με τον ταξινομητή SVM rbf 5) και το κόκκινο (84.32% με τον ταξινομητή SVM rbf 5). Να σημειώσουμε ότι τώρα η απόδοση του πράσινου καναλιού είναι υψηλότερη από αυτή του μπλε ενώ το κόκκινο είναι και πάλι τελευταίο. Επίσης είναι αξιοσημείωτο ότι η διαφορά του RGB σε σχέση με το πράσινο είναι μόλις 0,84%. Άρα, συμπεραίνουμε ότι εάν κάποιος ενδιαφέρεται αποκλειστικά για το μέτρο της ευαισθησίας μπορεί να χρησιμοποιήσει τα διάλυμα χαρακτηριστικών μόνο του πράσινου καναλιού (66% μικρότερο διάλυμα χαρακτηριστικών) χωρίς να έχει ουσιαστική απώλεια στην απόδοση.

Δε συμβαίνει όμως το ίδιο και με το μέτρο της ειδικότητας. Εδώ, η χρήση των πληροφοριών και από τα τρία κανάλια είναι επιτακτική καθώς επιτυγχάνεται ειδικότητα ύψους 92.47% (με τον ταξινομητή SVM rbf 1) έναντι 81.56%, (με τον ταξινομητή SVM rbf 2) στο μπλε κανάλι (μείωση 11% περίπου), 79.57% (με τον ταξινομητή SVM rbf 1) στο κόκκινο κανάλι (μείωση 13% περίπου) και 77.77% (με τον ταξινομητή SVM rbf 1) στο πράσινο κανάλι (μείωση 15% περίπου).

Όσον αφορά τους ταξινομητές, βγαίνει το συμπέρασμα ότι η συνάρτηση πυρήνα και η τιμή της παραμέτρου σίγμα μπορεί να επηρεάσει από ελάχιστα έως σημαντικά την απόδοση της μεθοδολογίας ανάλογα με την περίπτωση. Για το μέτρο της ακρίβειας ελάχιστα πιο αποτελεσματικός είναι ο SVM linear ταξινομητής καθώς προσφέρει υψηλότερα αποτελέσματα της τάξης του 2%, το μέγιστο. Εξάιρεση αποτελεί ο SVM rbf 1 στο σενάριο RGB ο οποίος προσφέρει 11% χαμηλότερη ακρίβεια. Για το μέτρο της ευαισθησίας ο πιο αποτελεσματικός ταξινομητής είναι ο SVM rbf 5 καθώς πετυχαίνεται έως 7% υψηλότερα ποσοστά. Εξάιρεση αποτελεί και πάλι ο SVM rbf 1 στο σενάριο RGB ο οποίος προσφέρει 26% χαμηλότερη ευαισθησία. Για το μέτρο της ειδικότητας πιο αποτελεσματικός ταξινομητής είναι ο SVM rbf 1 πετυχαίνοντας έως 5% υψηλότερη απόδοση.



Εικόνα 35. Ποσοστά ειδικότητας για τον χώρο RGB.

Σύμφωνα με τα παραπάνω, με μια πιο γενικευμένη ματιά, μπορούμε να ισχυριστούμε ότι τα καλύτερα αποτελέσματα στην αυτοματοποιημένη αναγνώριση περιοχών με έλκος θα την έχουμε εάν χρησιμοποιήσουμε πληροφορίες και από τα τρία χρωματικά κανάλια του RGB χώρου ταυτόχρονα και ταξινομητή SVM linear.

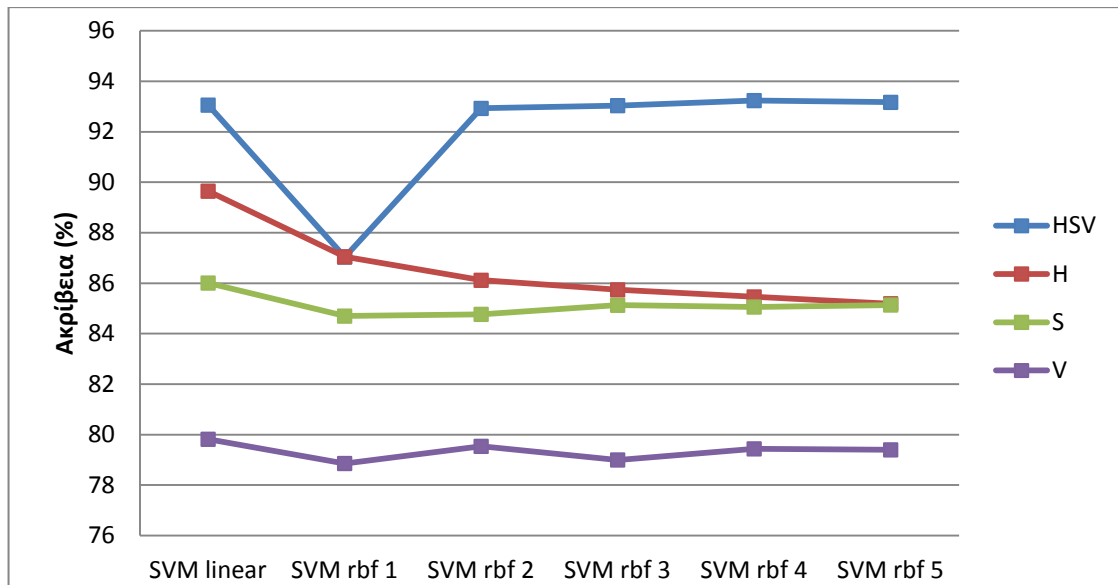
7.2 Χρωματικός χώρος HSV

Στον Πίνακα 2 φαίνονται τα ποσοστά της ακρίβειας, ευαισθησίας και ειδικότητας και για τα τέσσερα σενάρια του διανύσματος χαρακτηριστικών και όλα τα είδη ταξινομητή για την περίπτωση του χρωματικού χώρου HSV. Στην Εικόνα 36 απεικονίζεται η ακρίβεια για όλους τους ταξινομητές και όλα τα σενάρια. Αντίστοιχα, στις Εικόνες 37 και 38 απεικονίζεται η ευαισθησία και η ειδικότητα. Ο αριθμός που βρίσκεται δίπλα στην ονομασία SVM rbf υποδηλώνει την τιμή που έχει η παράμετρος σίγμα.

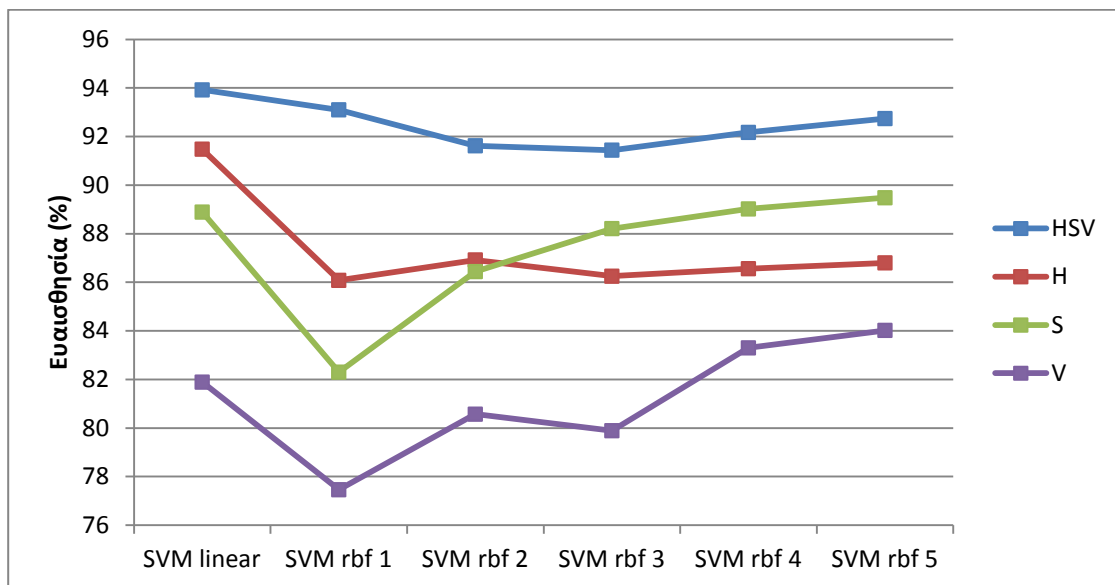
Μέτρο Απόδοσης	Ταξινομητής	Σενάριο Διανύσματος Χαρακτηριστικών			
		HSV	H	S	V
Ακρίβεια	SVM linear	93,06%	89,65%	86,01%	79,82%
	SVM rbf 1	87,04%	87,05%	84,7%	78,86%
	SVM rbf 2	92,93%	86,12%	84,77%	79,54%
	SVM rbf 3	93,04%	85,74%	85,13%	79%
	SVM rbf 4	93,24%	85,46%	85,06%	79,44%
	SVM rbf 5	93,17%	85,19%	85,14%	79,4%
Ευαισθησία	SVM linear	93,92%	91,48%	88,89%	81,89%
	SVM rbf 1	93,1%	86,08%	82,3%	77,46%
	SVM rbf 2	91,62%	86,92%	86,44%	80,57%
	SVM rbf 3	91,44%	86,25%	88,21%	79,89%
	SVM rbf 4	92,17%	86,56%	89,02%	83,3%
	SVM rbf 5	92,74%	86,8%	89,48%	84,02%
Ειδικότητα	SVM linear	92,2%	87,65%	83,13%	77,74%
	SVM rbf 1	80,98%	88,02%	87,1%	80,25%
	SVM rbf 2	94,24%	85,32%	83,1%	78,5%
	SVM rbf 3	94,64%	85,24%	82,05%	78,12%
	SVM rbf 4	94,3%	84,37%	81,1%	75,57%
	SVM rbf 5	93,6%	83,58%	80,8%	74,77%

Πίνακας 2. Αποτελέσματα επεξεργασίας των εικόνων στο χρωματικό χώρο HSV.

Όπως στο RGB έτσι και στο HSV όταν χρησιμοποιούνται όλα τα χρωματικά κανάλια τα ποσοστά επιτυχημένων ταξινομήσεων είναι αρκετά υψηλότερα (από 4% για το κανάλι H έως 7% για το κανάλι V). Πιο συγκεκριμένα, η μέγιστη ακρίβεια φθάνει 93.24% (με τον SVM rbf 4 ταξινομητή) για το σενάριο HSV ενώ για το κάθε χρωματικό κανάλι ξεχωριστά είναι: 89.65% (με τον SVM linear ταξινομητή) για το κανάλι H, 86,01% (με τον SVM linear ταξινομητή) για το κανάλι S και 79,82% (με τον SVM linear ταξινομητή) για το κανάλι V. Συγκρίνοντας τα τρία χρωματικά κανάλια μεταξύ τους παρατηρούμε ότι η συμπεριφορά του H είναι βελτιωμένη κατά 3.5% σε σχέση με το S. Αντίθετα, το κανάλι V (ως μεμονωμένο)



Εικόνα 36. Ποσοστά ακρίβειας για τον χώρο HSV.



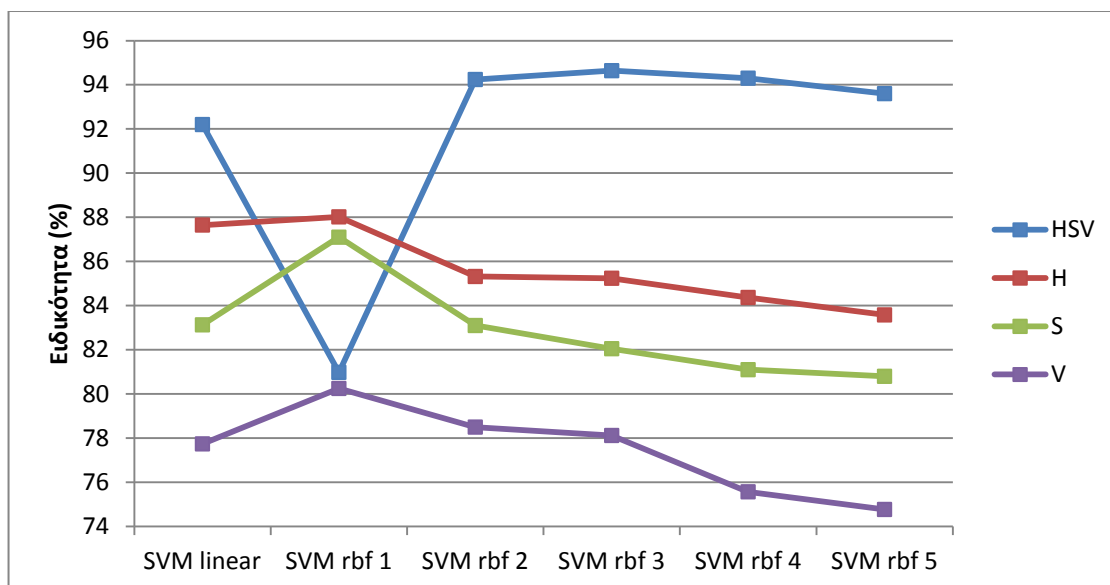
Εικόνα 37. Ποσοστά ευαισθησίας για τον χώρο HSV.

αποδεικνύεται ότι είναι ακατάλληλο για την επεξεργασία ενδοσκοπικών εικόνων με σκοπό την ανίχνευση έλκους καθώς επιτυγχάνει ακρίβεια κατά 10.0% χαμηλότερη από το H. Από τα παραπάνω γίνεται αντιληπτό ότι οι πληροφορίες υφής του έλκους αναπαριστώνται αποτελεσματικότερα στους χώρους που απεικονίζουν την απόχρωση και τον κορεσμό (περιέχουν δηλαδή πληροφορίες χρώματος) παρά τη φωτεινότητα. Αυτό είναι κάτι αναμενόμενο καθώς οι περιοχές με έλκος χαρακτηρίζονται από κιτρινο-πράσινη απόχρωση.

Όσον αφορά την ευαισθησία και πάλι η καλύτερη επίδοση (93.92% με τον ταξινομητή SVM linear) επιτυγχάνεται από τη χρήση πληροφοριών και από τα τρία κανάλια ταυτόχρονα. Ακολουθεί η επίδοση του καναλιού H (91.48 με τον ταξινομητή SVM linear), του S (89.48 με τον ταξινομητή SVM rbf 5), και τέλος του V (84.02 με τον ταξινομητή SVM rbf 5). Και για αυτή την περίπτωση η απόδοση του καναλιού H είναι ελαφρώς υψηλότερη (κατά 2%) από αυτή του καναλιού S ενώ το V παρουσιάζει 7.5% χειρότερα αποτελέσματα.

Αξίζει να αναφερθεί ότι η διαφορά του H από το HSV είναι μόλις 2.5% με αποτέλεσμα να το καθιστά ιδανικό σε περίπτωση που μας ενδιαφέρει αποκλειστικά η σωστή ανίχνευση θετικών δειγμάτων (εικόνων με έλκος) χρησιμοποιώντας μικρό διάνυσμα χαρακτηριστικών (6 αντί 18 χαρακτηριστικά).

Αντίθετα με τα παραπάνω, εάν μας ενδιαφέρει η ικανότητα να ανιχνεύονται σωστά οι φυσιολογικές περιοχές ώστε να μην έχουμε ψεύτικους συναγερμούς για παθολογικές εικόνες (false positive) (Εικόνα 38) τότε πρέπει απαραίτητα να χρησιμοποιήσουμε και τα τρία κανάλια ώστε να επιτύχουμε 94.64% (με τον ταξινομητή SVM rbf 3) απόδοση. Σε αντίθετη περίπτωση, εάν χρησιμοποιήσουμε μόνο το κανάλι H η ειδικότητα είναι 88.02% (με τον ταξινομητή SVM rbf 1), το κανάλι S 87.1% (με τον ταξινομητή SVM rbf 1) και το κανάλι V 80.25% (με τον ταξινομητή SVM rbf 1). Είναι αξιοσημείωτο ότι η διαφορά του καναλιού S από το H είναι κάτω του 1% ενώ για το κανάλι V αγγίζει το 7%.



Εικόνα 38. Ποσοστά ειδικότητας για τον χώρο HSV.

Όσον αφορά τους ταξινομητές, για την περίπτωση του μέτρου της ακρίβειας στο HSV σενάριο όλοι οι ταξινομητές επιδεικνύουν παρόμοια αποτελέσματα, το οποίο δείχνει ότι το διάνυσμα χαρακτηριστικών είναι αρκετά ισχυρό. Μοναδική εξαίρεση αποτελεί ο SVM rbf 1 ο οποίος αποδίδει 6% χειρότερα. Κάτι αντίστοιχο ισχύει και για τα κανάλια S και V όπου οι διαφορές ανάμεσα στους ταξινομητές κυμαίνονται γύρω στο 1%. Αντίθετα, το κανάλι H επηρεάζεται σημαντικά από τον ταξινομητή (δείχνει προτίμηση στον SVM linear) καθώς η ακρίβεια μειώνεται 4.5% περίπου και τείνει να ταυτιστεί με αυτή του καναλιού S. Στην περίπτωση του μέτρου της ευαισθησίας η επιλογή του ταξινομητή επηρεάζει πολύ λίγο (της τάξης του 2%) το αποτέλεσμα για το σενάριο HSV ενώ η επίδραση είναι πιο σημαντική (από 5.5% έως 7%) για τα σενάρια των μεμονωμένων καναλιών. Είναι χαρακτηριστικό ότι για ορισμένους ταξινομητές (SVM rbf 3 και μετά η απόδοση του καναλιού H γίνεται χειρότερη από αυτή του καναλιού S. Τέλος, για το μέτρο της ειδικότητας ισχύει περίπου ότι αναφέρθηκε και για την περίπτωση του μέτρου της ακρίβειας. Για το σενάριο HSV ο ταξινομητής δεν επηρεάζει ουσιαστικά το αποτέλεσμα (διαφορές της τάξης του 2%) με εξαίρεση τον SVM rbf 1 που παρουσιάζει πτώση 14% περίπου. Για τα σενάρια των μεμονωμένων καναλιών οι διαφορές ποικίλουν από 4.5% έως 6.5%.

Σύμφωνα με τα παραπάνω, με μια πιο γενικευμένη ματιά, μπορούμε να ισχυριστούμε ότι τα καλύτερα αποτελέσματα στην αυτοματοποιημένη αναγνώριση περιοχών με έλκος θα την έχουμε εάν χρησιμοποιήσουμε πληροφορίες και από τα τρία χρωματικά κανάλια του HSV χώρου ταυτόχρονα με τον ταξινομητή να μην παίζει σημαντικό ρόλο.

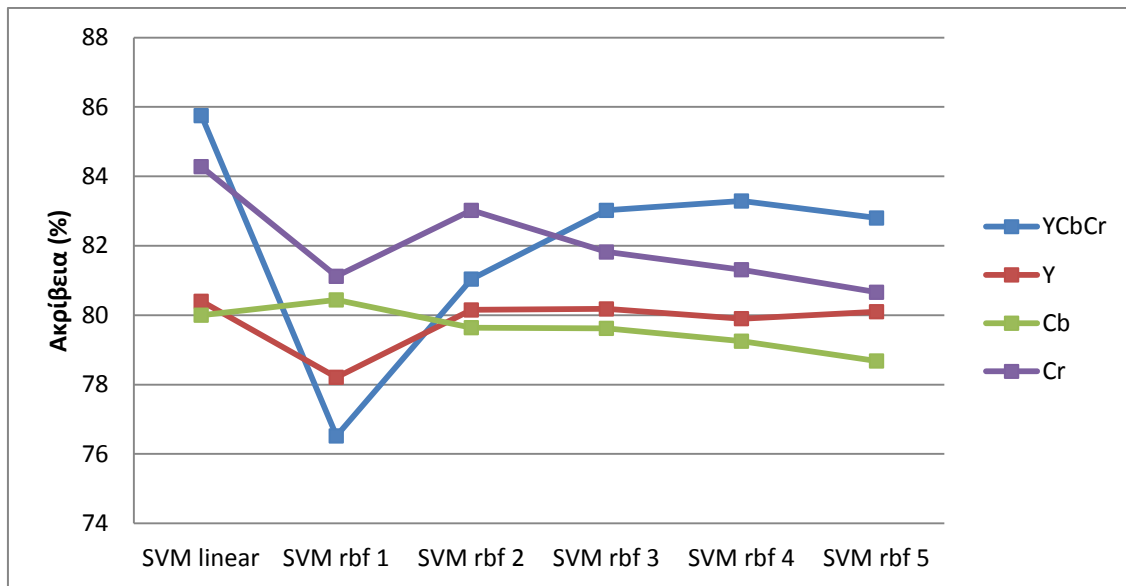
7.3 Χρωματικός χώρος YCbCr

Στον Πίνακα 3 φαίνονται τα ποσοστά της ακρίβειας, ευαισθησίας και ειδικότητας και για τα τέσσερα σενάρια του διανύσματος χαρακτηριστικών και όλα τα είδη ταξινομητή για την περίπτωση του χρωματικού χώρου HSV. Στην Εικόνα 39 απεικονίζεται η ακρίβεια για όλους τους ταξινομητές και όλα τα σενάρια. Αντίστοιχα, στις Εικόνες 40 και 41 απεικονίζεται η ευαισθησία και η ειδικότητα. Ο αριθμός που βρίσκεται δίπλα στην ονομασία SVM rbf υποδηλώνει την τιμή που έχει η παράμετρος σίγμα.

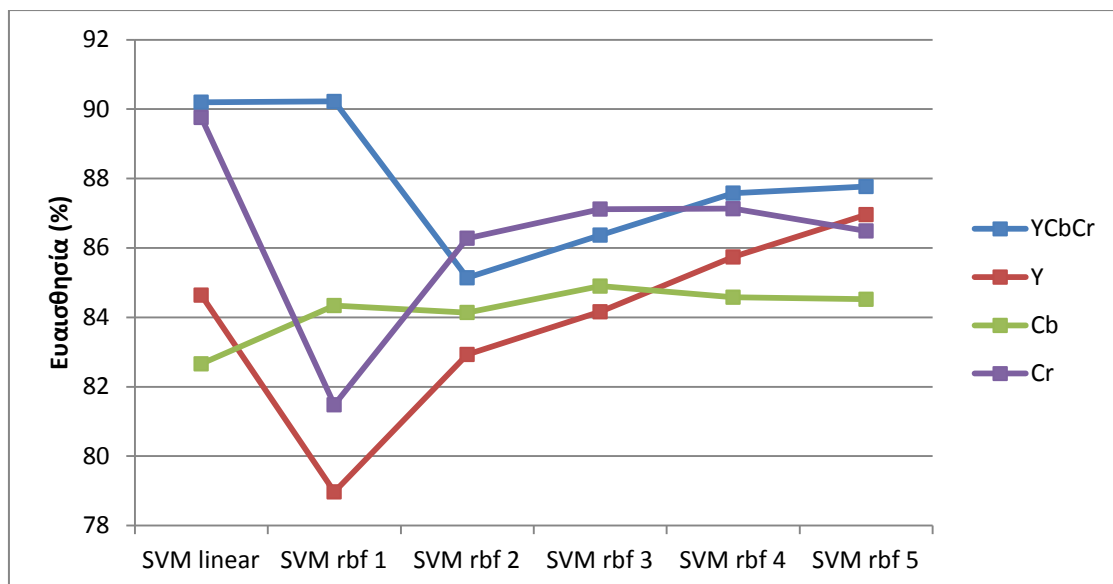
Μέτρο Απόδοσης	Ταξινομητής	Σενάριο Διανύσματος Χαρακτηριστικών			
		YCbCr	Y	Cb	Cr
Ακρίβεια	SVM linear	85,75%	80,4%	80%	84,28%
	SVM rbf 1	76,52%	78,21%	80,44%	81,12%
	SVM rbf 2	81,04%	80,15%	79,64%	83,02%
	SVM rbf 3	83,02%	80,18%	79,62%	81,82%
	SVM rbf 4	83,29%	79,9%	79,25%	81,31%
	SVM rbf 5	82,8%	80,1%	78,68%	80,66%
Ευαισθησία	SVM linear	90,2%	84,64%	82,66%	89,76%
	SVM rbf 1	90,22%	78,97%	84,34%	81,48%
	SVM rbf 2	85,14%	82,93%	84,14%	86,28%
	SVM rbf 3	86,37%	84,16%	84,9%	87,12%
	SVM rbf 4	87,58%	85,74%	84,58%	87,13%
	SVM rbf 5	87,77%	86,96%	84,52%	86,49%
Ειδικότητα	SVM linear	81,3%	76,16%	77,33%	78,8%
	SVM rbf 1	62,81%	77,45%	76,54%	80,77%
	SVM rbf 2	76,93%	77,37%	75,13%	79,76%
	SVM rbf 3	79,66%	76,2%	74,34%	76,52%
	SVM rbf 4	79%	74,05%	73,92%	75,49%
	SVM rbf 5	77,84%	73,25%	72,85%	74,84%

Πίνακας 3. Αποτελέσματα επεξεργασίας των εικόνων στο χρωματικό χώρο YCbCr.

Από την Εικόνα 39 φαίνεται ότι και σε αυτό το χρωματικό χώρο όταν χρησιμοποιούνται οι πληροφορίες και από τα 3 κανάλια ταυτόχρονα πετυχαίνουμε το μεγαλύτερο ποσοστό ακρίβειας (85.75% με τον ταξινομητή SVM linear). Όσον αφορά την επίδοση κάθε χρωματικού καναλιού ξεχωριστά παρατηρούμε ότι το πιο αποδοτικό είναι το Cr με ποσοστό ακρίβειας 84.48% (με ταξινομητή SVM linear) ενώ ακολουθούν με παρόμοια αλλά αρκετά χαμηλότερη επίδοση (μείωση 4% περίπου) το Cb (80.44% με ταξινομητή SVM rbf 1) και το Y (80.4% με ταξινομητή SVM linear). Είναι ξεκάθαρο ότι η επίδοση του καναλιού Cr είναι πολύ κοντά σε αυτή του πλήρους χρωματικού χώρου (υπολείπεται μόνο



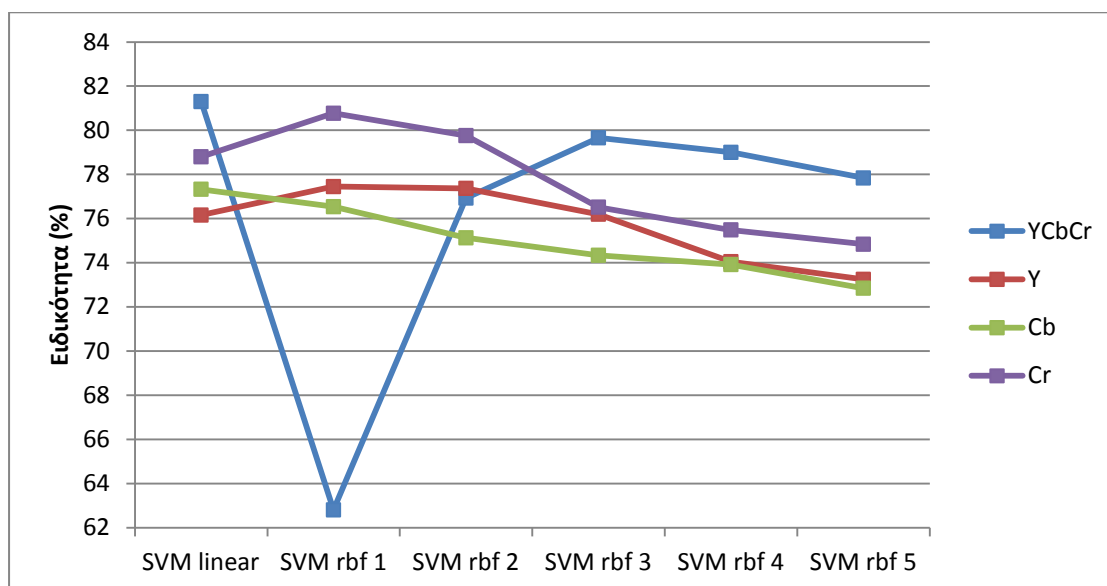
Εικόνα 39. Ποσοστά ακρίβειας για τον χώρο YCbCr.



Εικόνα 40. Ποσοστά ευαισθησίας για τον χώρο YCbCr.

κατά 1.3%). Έτσι μπορούμε να συμπεράνουμε ότι το μεγαλύτερο μέρος της πληροφορίας της υψής του έλκους αποτυπώνεται αποτελεσματικότερα στο κανάλι που εκφράζει το κατά πόσο κόκκινη-φούξια ή πράσινο-μπλε είναι η περιοχή ενδιαφέροντος. Αυτό είναι αναμενόμενο καθώς όπως αναφέρθηκε και πιο πάνω ο αλλοιωμένος από έλκος ιστός χαρακτηρίζεται από κιτρινοπράσινη απόχρωση ενώ ο φυσιολογικός ιστός χαρακτηρίζεται από κόκκινο-ροζ απόχρωση. Επομένως, εάν επιθυμούμε να έχουμε μικρό σε μέγεθος διάνυσμα χαρακτηριστικών μπορούμε χωρίς ενδοιασμούς να χρησιμοποιήσουμε το κανάλι Cr χωρίς να έχουμε ουσιαστική επίπτωση στην αποτελεσματικότητα της μεθόδου.

Όσον αφορά το μέτρο της ευαισθησίας και πάλι ο συνδυασμός YCbCr πετυχαίνει οριακά το υψηλότερο ποσοστό (90.22% με ταξινομητή SVM rbf 1) ενώ το κανάλι Cr υπολείπεται μόλις κατά 0.46% πετυχαίνοντας ποσοστό 89.76% με ταξινομητή SVM linear. Στη συνέχεια, υψηλότερο ποσοστό ευαισθησίας προσφέρει το κανάλι της φωτεινότητας (Y)



Εικόνα 41. Ποσοστά ειδικότητας για τον χώρο YCbCr.

με ποσοστό 86.96% με ταξινομητή SVM rbf 5 ενώ το λιγότερο αποτελεσματικό είναι το κανάλι Cb με ποσοστό που δεν ξεπερνά το 85% (84.9%) με ταξινομητή SVM rbf 3. Τα αποτελέσματα αυτά ενισχύουν το συμπέρασμα ότι οι περισσότερες απαραίτητες πληροφορίες βρίσκονται στο κανάλι Cr.

Παρόμοια είναι η κατάσταση και για το μέτρο της ειδικότητας. Η καλύτερη συμπεριφορά επιδεικνύεται από το YCbCr με ποσοστό 81.3% με ταξινομητή SVM linear ενώ ακολουθεί με μικρή διαφορά με ποσοστό 80.77% και ταξινομητή SVM rbf 1 το κανάλι Cr. Τα κανάλια Y και Cb καταφέρνουν παρόμοια επίδοση με ποσοστά 77.45% (με ταξινομητή SVM rbf 1) και 77.33% (με ταξινομητή SVM linear), αντίστοιχα.

Συγκρίνοντας τους ταξινομητές μεταξύ τους παρατηρούμε ότι ο πιο αποτελεσματικός για τον συγκεκριμένο χρωματικό χώρο είναι ο SVM linear. Η επιλογή της συνάρτησης πυρήνα και της παραμέτρου σίγμα παίζουν ουσιαστικό ρόλο στη συνολική απόδοση του συστήματος. Για την περίπτωση του μέτρου της ακρίβειας στα σενάρια YCbCr και Cr η επιρροή του ταξινομητή είναι ξεκάθαρη (Εικόνα 39) καθώς προκαλεί διαφορές ύψους 2.5% - 9% και 1.5% - 4%, αντίστοιχα. Δεν ισχύει το ίδιο για τα κανάλια Y και Cb όπου η διαφορά δεν ξεπερνούν το 1.5%. Στην περίπτωση του μέτρου της ευαισθησίας (Εικόνα 40) η επίδραση του ταξινομητή είναι σημαντική για όλα τα κανάλια, ίσως με εξαίρεση το Cb. Πιο συγκεκριμένα, στο σενάριο YCbCr εμφανίζονται διαφορές έως 5%, στο σενάριο Y έως 9%, στο σενάριο Cb έως 2.2% και στο σενάριο Cr έως 8%. Τέλος, στην περίπτωση του μέτρου της ειδικότητας (Εικόνα 41) οι διαφορές για το σενάριο YCbCr είναι από 1.7% έως 4.4% με εξαίρεση τον ταξινομητή SVM rbf 1 ο οποίος υστερεί κατά 18.5%, για το σενάριο Y είναι από 0.1% έως 4.2%, για το σενάριο Cb είναι από 0.8% έως 4.5% και για το σενάριο Cr είναι από 1% έως 5.9%.

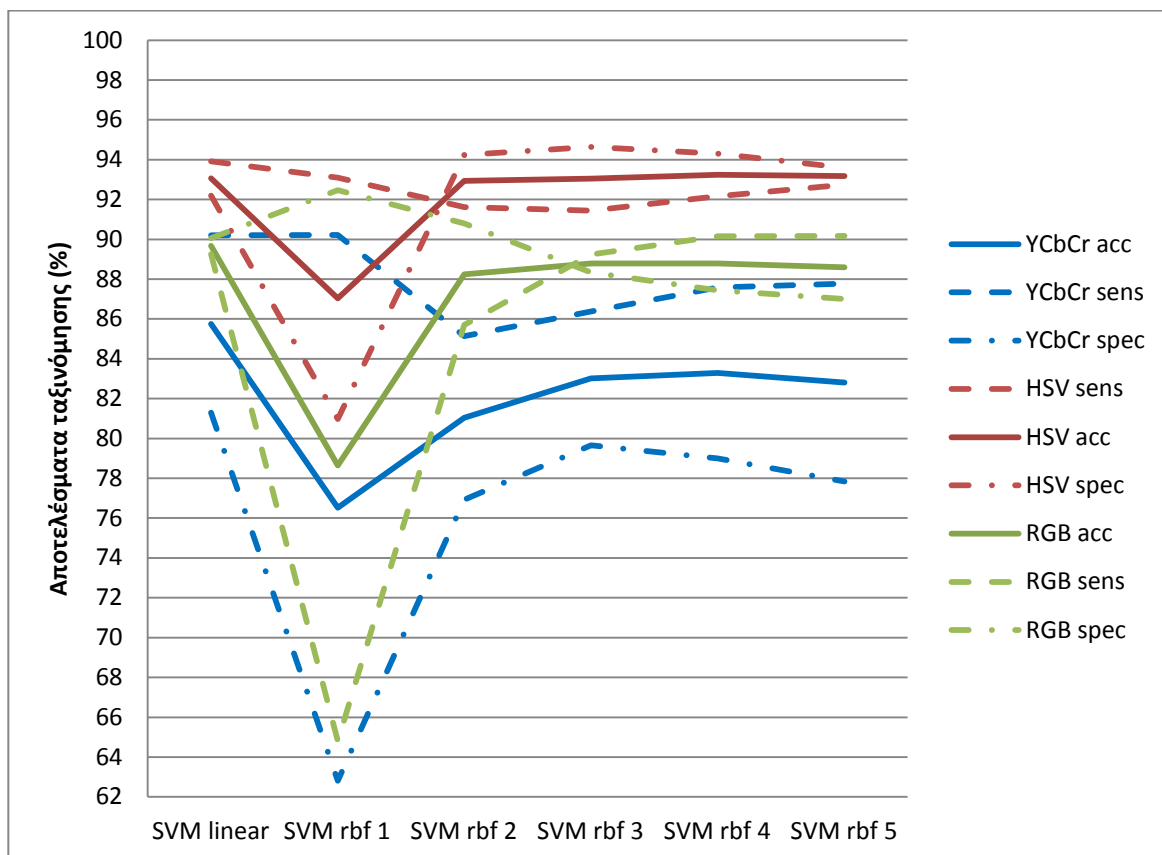
Ολοκληρώνοντας, σύμφωνα με τα παραπάνω, αξίζει να αναφέρουμε ότι στο χρωματικό χώρο YCbCr τη βέλτιστη απόδοση την πετυχαίνουμε με τον συνδυασμό των τριών καναλιών αλλά με το μεγαλύτερο ποσοστό των σημαντικών πληροφοριών να βρίσκονται στο κανάλι Cr καθιστώντας το ικανό για την αποκλειστική επεξεργασία των

εικόνων. Οι μεγάλες διαφορές στα ποσοστά επιτυχίας που παρατηρούνται από τους διάφορους ταξινομητές σημαίνει ότι το διάνυσμα χαρακτηριστικών δεν είναι τόσο ισχυρό και η απόδοση του εξαρτάται ξεκάθαρα από τις δυνατότητες του ταξινομητή.

7.4 Σύγκριση των χρωματικών χώρων

Ολοκληρώνοντας την παρουσίαση των αποτελεσμάτων παραθέτουμε τα ποσοστά ακρίβειας, ευαισθησίας και ειδικότητας για όλους τους χρωματικούς χώρους όταν χρησιμοποιούνται τα χαρακτηριστικά και από τα τρία χρωματικά κανάλια (Εικόνα 42). Είναι ξεκάθαρο ότι η μεθοδολογική προσέγγιση με χρήση του μετασχηματισμού Curvelet και εξαγωγή χαρακτηριστικών υψής με τα ομοιόμορφα ανεξάρτητα περιστροφής τοπικά δυαδικά πρότυπα που εφαρμόστηκε στην παρούσα εργασία έχει τα υψηλότερα αποτελέσματα όταν η επεξεργασία των εικόνων γίνεται στο χρωματικό χώρο HSV (ακρίβεια: 93.24%, ευαισθησία: 93.92%, ειδικότητα: 94.64%).

Ο HSV επιτυγχάνει όχι μόνο υψηλότερα ποσοστά στην ακρίβεια της ταξινόμησης αλλά και στην ειδικότητα και ευαισθησία. Όσον αφορά την ακρίβεια τα αποτελέσματα είναι από 3.5% έως 14.5% υψηλότερα σε σχέση με το χώρο RGB και από 7.5% έως 16% υψηλότερα σε σχέση με το χώρο YCbCr, για όλους τους ταξινομητές. Ομοίως, ανεξαρτήτως ταξινομητή, η ευαισθησία είναι υψηλότερη από τις αντίστοιχες των RGB και YCbCr τουλάχιστον κατά 3.7%. Για το μέτρο της ειδικότητας ο χώρος HSV προσφέρει υψηλότερα



Εικόνα 42. Ποσοστά ακρίβειας, ευαισθησίας, ειδικότητας για τους χρωματικούς χώρους RGB, HSV, YCbCr.

αποτελέσματα κατά 13% - 32% από τον χώρο YCbCr για όλους τους ταξινομητές. Σε σύγκριση όμως με το χώρο RGB, ο HSV επιτυγχάνει υψηλότερα αποτελέσματα κατά 4% - 7.5% για τους ταξινομητές SVM rbf 2 έως 5 ενώ για τους ταξινομητές SVM linear και SVM rbf 1 ο RGB επιτυγχάνει υψηλότερα ποσοστά κατά 0.3% - 12%.

Όσον αφορά τους ταξινομητές, τα πιο καλά αποτελέσματα συγκεντρωτικά μπορούμε να ισχυριστούμε ότι τα προσφέρει ο SVM linear εάν μας ενδιαφέρει η ευαισθησία να είναι υψηλότερη από την ειδικότητα ή ο SVM rbf 5 εάν απαιτείται μια ισορροπημένη συμπεριφορά στη σωστή αναγνώριση παθολογικών και φυσιολογικών εικόνων. Άξιος προσοχής είναι ο ταξινομητής SVM rbf 1 ο οποίος για την πληθώρα των περιπτώσεων παρουσιάζει ανικανότητα στη σωστή κατηγοριοποίηση των εικόνων.

Τέλος, κάτι περίεργο που επίσης αξίζει να σημειωθεί είναι ότι στους χώρους HSV και RGB για κάποιους ταξινομητές τα ποσοστά της ευαισθησίας είναι υψηλότερα από τα ποσοστά της ειδικότητας ενώ για κάποιους άλλους ισχύει το αντίθετο. Στην περίπτωση του YCbCr η ευαισθησία έχει σταθερά υψηλότερες τιμές από την ειδικότητα, κάτι που είναι απόλυτα θεμιτό και προτιμότερο από το αντίστροφο.

8 Αναφορές

- [1] Βιοϊατρική τεχνολογία, en.wikipedia.org.
- [2] World Health Organisation, www.who.int
- [3] Wayne, J.D. (1992). Small bowel endoscopy. *Endoscopy*, 24:68-72.
- [4] Iddan, G., Meron, G., Glukhovsky, A., & Swain, P. (2000). Wireless capsule endoscopy. *Nature*, 405-417.
- [5] Καλαϊτζής, Ν. & Αυγερινός, Α. (2006). *Εντεροσκόπηση με Ασύρματη Κάψουλα*. Αθήνα: Εκδόσεις Βήτα.
- [6] Τζάθας, Χ. & Χαραλαμπίδης, Σ. (2012). Ενδοσκόπηση με ασύρματη κάψουλα. *Επιστημονικά Χρονικά*, 17(3): 136-138.
- [7] Βιβλίο Βιολογίας Α' Τάξης Γενικού Λυκείου Γενικής Κατεύθυνσης.
- [8] <http://www.protonpharma.gr/>
- [9] Jung, Y.S., Kim, Y.H., Lee, D.H., & Kim, J.H. (2008). Active Blood Detection in a High Resolution Capsule Endoscopy Using Color Spectrum Transformation. *2008 IEEE International Conference on BioMedical Engineering and Informatics*.
- [10] Baopu, Li. & Meng, M.Q.H. (2009). Computer-Aided Detection of Bleeding Regions for Capsule Endoscopy Images. *Biomedical Engineering IEEE*, 56(4):1032-1039.
- [11] Sainju, S., Bui, F.M., & Wahid, K. (2013). Bleeding detection in wireless capsule endoscopy based on color features from histogram probability. *Electrical and Computer Engineering IEEE*, 1-4.
- [12] Cauvin, J.M., Guillou, C.L., & Solaiman, B. (2004). An intelligent Web atlas in digestive endoscopy. *MEDINFO 2004*, Amsterdam.
- [13] Cauvin, J.M., Guillou, C.L., Solaiman, B., Robaszkiewicz, M., Le Beux, P., & Roux, C. (2003). Computer-Assisted Diagnosis System in Digestive Endoscopy. *IEEE Transactions on information technology in biomedicine*, 7(4).
- [14] Karkanis, S.A., Iakovidis, D.K., Maroulis, D.E., Karras, D.A., & Tzivras, M. (2003). Computer-Aided Tumor Detection in Endoscopic Video Using Color Wavelet Features. *IEEE Transactions on information technology in biomedicine*, 7(3).
- [15] Coimbra, M. T., & Cunha, J. S. (2006). MPEG-7 visual descriptors—contributions for automated feature extraction in capsule endoscopy. *Circuits and Systems for Video Technology, IEEE Transactions on*, 16(5):628-637.
- [16] Xiaoying Liu et. al. (2012). A new approach to detecting ulcer and bleeding in Wireless capsule endoscopy images. *Biomedical and Health Informatics IEEE*, 737-740.
- [17] Karargyris, A. & Bourbakis, N. (2009). Identification of ulcers in Wireless Capsule Endoscopy videos. *Biomedical Imaging: From Nano to Macro, IEEE*, 554-557.
- [18] Vieira, P., Ramos, J., Barbosa, D. et. al. (2012). Segmentation of small bowel tumor tissue in Capsule Endoscopy Images by using the MAP algorithm. *34th Annual International Conference IEEE EMBS*, 4010-4013.
- [19] Vilarino, F., Spyridonos, P., Pujol, O., Vitria, J., & Radeva, P. (2006). Automatic Detection of Intestinal Juices in Wireless Capsule Video Endoscopy. *18th IEEE International Conference on Pattern Recognition (ICPR'06)*.

- [20] Tsevas, S., Iakovidis, D.K., Maroulis, D., & Pavlakis, E. (2008). Automatic Frame Reduction of Wireless Capsule Endoscopy Video. *8th IEEE International Conference on BioInformatics and BioEngineering*.
- [21] Iakovidis, D.K., Tsevas, S., Maroulis, D., & Polydorou, A. (2008). Unsupervised summarisation of capsule endoscopy video. *4th IEEE International Conference on Intelligent Systems*, 1:3-15-3-20.
- [22] Dhandra, B.V., Hegadi, R., Hangarge, M., & Malelath, V.S. (2006). Analysis of abnormality in endoscopic images using combined HSI color space and watershed segmentation. *18th Int. Conf. IEEE ICPR*, 4:695-698.
- [23] Li, B. & Meng, M.Q.H. (2009). Texture analysis for ulcer detection in capsule endoscopy images. *Image and Vision Computing*, 27:1336–1342.
- [24] Παπαμάρκος, Ν. (2010). *Ψηφιακή Επεξεργασία και Ανάλυση Εικόνας*.
- [25] Smith, A. R. (1978). Color gamut transform pairs. *Computer Graphics*, 12(3):12-19.
- [26] Candes, J. & Donoho, D.L. (2000). *Curvelets – A Surprisingly Effective Nonadaptive Representation For Objects with Edges*. Stanford University Campus, Department of Statistics.
- [27] Candes, J. & Donoho, D.L. (2002). New tight frames of curvelets and optimal representations of objects with piecewise- C^2 singularities. *Comm. Pure Appl. Math.*, 57:219-266.
- [28] Candes, J., Demanet, L., Donoho, D., & Ying, L. (2006). Fast Discrete Curvelet Transforms. *Multiscale Modeling & Simulation*, 5(3):861-899.
- [29] Materka, A. & Strzelecki, M. (1998). *Texture Analysis Methods – A Review*, Technical University of Lodz, Institute of Electronics, COST B11 report, Brussels.
- [30] Ojala, T., Pietikainen, M., & Harwood, D. (1996). A Comparative Study of Texture Measures with Classification based on Feature Distributions. *Pattern Recognition*, 29:51-59.
- [31] Ojala, T., Pietikainen, M., & Maenpaa, T. (2002). Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 24(7):971-987.
- [32] Boser, B. E., Guyon, I. M., & Vapnik, V. N. (1992). A training algorithm for optimal margin classifiers. *Proceedings of the fifth annual workshop on Computational learning theory*, 144-152, ACM.
- [33] Cortes, C. & Vapnik, V.P. (1995). Support-vector networks. *Machine Learning*, 20(3):273-297.

9 Παράρτημα

9.1 Κώδικας MATLAB

```

addpath('dataset')
clear
clc
close all

P=16;
R=4;

input1=[];
input2=[];
h = waitbar(0,'Please wait');
for i=1:150
    waitbar(i/150,h,['calculating ' num2str(i) ' of 150'])
    imname = ['n' int2str(i) '.png'];
    im = imread(imname);
    feat = extract_features(im, P, R, 1,'yes');
    input1 = [input1;feat];

    imname = ['u' int2str(i) '.png'];
    im = imread(imname);
    feat = extract_features(im, P, R, 1,'yes');
    input2 = [input2;feat];
end
total_input = [input1;input2];
total_target = [ones(150,1); 2*ones(150,1)];
matname = ['P' int2str(P) '_R' int2str(R) '_YCbCr_uLBP.mat' ];
% Save Input to .mat file for later use
save(['matfiles/' matname], 'total_input_YCbCr_uLBP')
close(h)

clear
clc
close all

P=16;
R=4;

input1=[];
input2=[];
h = waitbar(0,'Please wait');
for i=1:150
    waitbar(i/150,h,['calculating ' num2str(i) ' of 150'])
    imname = ['n' int2str(i) '.png'];
    im = imread(imname);
    feat = extract_features(im, P, R, 2,'yes');
    input1 = [input1;feat];

    imname = ['u' int2str(i) '.png'];
    im = imread(imname);
    feat = extract_features(im, P, R, 2,'yes');
    input2 = [input2;feat];
end
total_input = [input1;input2];
total_target = [ones(150,1); 2*ones(150,1)];
matname = ['P' int2str(P) '_R' int2str(R) '_HSV_uLBP.mat' ];

```

```

% Save Input to .mat file for later use
save(['matfiles/' matname], 'total_input_HSV_uLBP ')
close(h)
clear
clc
close all

P=16;
R=4;
input1 = [];
input2=[];
h = waitbar(0, 'Please wait');
for i=1:150
    waitbar(i/150,h,['calculating ' num2str(i) ' of 150'])
    imname = ['n' int2str(i) '.png'];
    im = imread(imname);
    feat = extract_features(im, P, R, 3, 'yes');
    input1 = [input1;feat];

    imname = ['u' int2str(i) '.png'];
    im = imread(imname);
    feat = extract_features(im, P, R, 3, 'yes');
    input2 = [input2;feat];
end

total_input = [input1;input2];
total_target = [ones(150,1); 2*ones(150,1)];
matname = ['P' int2str(P) '_R' int2str(R) '_RGB_uLBP.mat' ];
% Save Input to .mat file for later use
save(['matfiles/' matname], 'total_input_RGB_uLBP ')
close(h)

mode=6;

normal = total_input_RGB_uLBP(1:150,:);
ulcer = total_input_RGB_uLBP(151:end,:);

total_input1 = [ulcer;normal];

GROUP = cell(300,1);
for i=1:300
    if i<=150
        GROUP{i} = '1';
    else
        GROUP{i} = '0';
    end
end
h = waitbar(0, 'Please wait');
i=1;
waitbar(i/24,h,['calculating ' num2str(i) ' of 24 at RGB'])
[results_RGB(i,1,1),results_RGB(i,2,1),results_RGB(i,3,1),results_RGB(i,4,1)
,results_RGB(i,5,1),results_RGB(i,6,1)] =
myclassifyWCEsvm(total_input1,GROUP, 'linear',1);
for i=2:6
    waitbar(i/24,h,['calculating ' num2str(i) ' of 24 at RGB'])

[results_RGB(i,1,1),results_RGB(i,2,1),results_RGB(i,3,1),results_RGB(i,4,1)
,results_RGB(i,5,1),results_RGB(i,6,1)] =
myclassifyWCEsvm(total_input1,GROUP, 'rbf',i-1);

```

```

end
for j=2:4
    i=1;
    waitbar((i+(j-1)*6)/24,h,['calculating ' num2str(i) ' of 24 at RGB'])

    [results_RGB(i,1,j),results_RGB(i,2,j),results_RGB(i,3,j),results_RGB(i,4,j)
    ],results_RGB(i,5,j),results_RGB(i,6,j)] =
    myclassifyWCESvm(total_input1(:,(j-2)*mode+1:(j-1)*mode),GROUP,'linear',1);
    for i=2:6
        waitbar((i+(j-1)*6)/24,h,['calculating ' num2str(i) ' of 24 at
    RGB'])

    [results_RGB(i,1,j),results_RGB(i,2,j),results_RGB(i,3,j),results_RGB(i,4,j)
    ],results_RGB(i,5,j),results_RGB(i,6,j)] =
    myclassifyWCESvm(total_input1(:,(j-2)*mode+1:(j-1)*mode),GROUP,'rbf',i-1);
    end
end

normal = total_input_HSV_uLBP(1:150,:);
ulcer = total_input_HSV_uLBP(151:end,:);

total_input1 = [ulcer;normal];

GROUP = cell(300,1);
for i=1:300
    if i<=150
        GROUP{i} = '1';
    else
        GROUP{i} = '0';
    end
end

i=1;
waitbar(i/24,h,['calculating ' num2str(i) ' of 24 at HSV'])
[results_HSV(i,1,1),results_HSV(i,2,1),results_HSV(i,3,1),results_HSV(i,4,1)
],results_HSV(i,5,1),results_HSV(i,6,1)] =
myclassifyWCESvm(total_input1,GROUP,'linear',1);
for i=2:6
    waitbar(i/24,h,['calculating ' num2str(i) ' of 24 at HSV'])

[results_HSV(i,1,1),results_HSV(i,2,1),results_HSV(i,3,1),results_HSV(i,4,1)
],results_HSV(i,5,1),results_HSV(i,6,1)] =
myclassifyWCESvm(total_input1,GROUP,'rbf',i-1);
end
for j=2:4
    i=1;
    waitbar((i+(j-1)*6)/24,h,['calculating ' num2str(i) ' of 24 at HSV'])

    [results_HSV(i,1,j),results_HSV(i,2,j),results_HSV(i,3,j),results_HSV(i,4,j)
    ],results_HSV(i,5,j),results_HSV(i,6,j)] =
    myclassifyWCESvm(total_input1(:,(j-2)*mode+1:(j-1)*mode),GROUP,'linear',1);
    for i=2:6
        waitbar((i+(j-1)*6)/24,h,['calculating ' num2str(i) ' of 24 at
    HSV'])

    [results_HSV(i,1,j),results_HSV(i,2,j),results_HSV(i,3,j),results_HSV(i,4,j)
    ],results_HSV(i,5,j),results_HSV(i,6,j)] =
    myclassifyWCESvm(total_input1(:,(j-2)*mode+1:(j-1)*mode),GROUP,'rbf',i-1);
    end
end

```

```

end

normal = total_input_YCbCr_uLBP(1:150,:);
ulcer = total_input_YCbCr_uLBP(151:end,:);

total_input1 = [ulcer;normal];

GROUP = cell(300,1);
for i=1:300
    if i<=150
        GROUP{i} = '1';
    else
        GROUP{i} = '0';
    end
end
i=1;
waitbar(i/24,h,['calculating ' num2str(i) ' of 24 at YCbCr'])
[results_YCbCr(i,1,1),results_YCbCr(i,2,1),results_YCbCr(i,3,1),results_YCbCr(i,4,1),results_YCbCr(i,5,1),results_YCbCr(i,6,1)] =
myclassifyWCEsvm(total_input1,GROUP,'linear',1);
for i=2:6
    waitbar(i/24,h,['calculating ' num2str(i) ' of 24 at YCbCr'])

[results_YCbCr(i,1,1),results_YCbCr(i,2,1),results_YCbCr(i,3,1),results_YCbCr(i,4,1),results_YCbCr(i,5,1),results_YCbCr(i,6,1)] =
myclassifyWCEsvm(total_input1,GROUP,'rbf',i-1);
end
for j=2:4
    i=1;
    waitbar((i+(j-1)*6)/24,h,['calculating ' num2str(i) ' of 24 at YCbCr'])

[results_YCbCr(i,1,j),results_YCbCr(i,2,j),results_YCbCr(i,3,j),results_YCbCr(i,4,j),results_YCbCr(i,5,j),results_YCbCr(i,6,j)] =
myclassifyWCEsvm(total_input1(:,(j-2)*mode+1:(j-1)*mode),GROUP,'linear',1);
    for i=2:6
        waitbar((i+(j-1)*6)/24,h,['calculating ' num2str(i) ' of 24 at YCbCr'])

[results_YCbCr(i,1,j),results_YCbCr(i,2,j),results_YCbCr(i,3,j),results_YCbCr(i,4,j),results_YCbCr(i,5,j),results_YCbCr(i,6,j)] =
myclassifyWCEsvm(total_input1(:,(j-2)*mode+1:(j-1)*mode),GROUP,'rbf',i-1);
    end
end
end

```

```

function feat = extract_features(im , P, R, mode, curvelet)
% Extract features for classification using Uniform Rotation Invariant
% Local Binary Patterns on Curvelet Coefficients.

% Change Color Space (default is RGB)
if (mode==1)
    im = rgb2ycbcr(im); % YCbCr
elseif (mode==2)
    im = rgb2hsv(im);    % HSV
elseif (mode==3)
    im=im;                % RGB
end

% Calculate 6 statistical measurements of LBP histogram for 2 DCT scales ,
% for all 3 channels. Dimensionality of feature vector is 6 x 2 x 3 = 36.

feat=[];

if strcmp(curvelet,'yes') % first calculate curvelet scales and then uLBP
    for chan=1:3
        % Calculate forward DCT transform for 4 scales.
        C = fdct_wrapping(double(im(:,:,chan)),1,2,4,8);

        % LBP and histogram of first scale

        LBP1 = double(uLBP(C{1}{1},P,R));
        x1 = hist(LBP1(:),numel(unique(LBP1)));

        % The statistical measurements used are standard deviation, skew,
kurtosis,
        % entropy, energy, and mean.
        temp1 =
[mean(x1),std(x1),mean(abs(x1).^2),entropy(x1),skewness(x1),kurtosis(x1)];

        % LBP and histogram of second scale
        LBP2 = double(uLBP(C{2}{1},P,R));
        x2 = hist(LBP2(:),numel(unique(LBP2)));

        temp2 =
[mean(x2),std(x2),mean(abs(x2).^2),entropy(x2),skewness(x2),kurtosis(x2)];

        feat = [feat,temp1,temp2];
    end
else % no curvelet transform takes place
    for chan=1:3
        LBP1 = double(uLBP(im(:,:,chan),P,R));
        x1 = hist(LBP1(:),numel(unique(LBP1)));
        temp1 =
[mean(x1),std(x1),mean(abs(x1).^2),entropy(x1),skewness(x1),kurtosis(x1)];
        temp2=[];
        feat = [feat,temp1,temp2];
    end
end
end

```

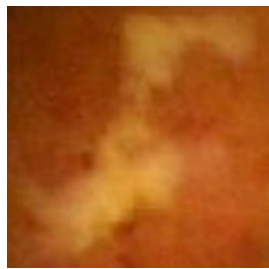
```

% Καθε grammi toy pinaka DATA antistoixei sta features enos deigmatos. Dld
% ean o DATA exei diastaseis MxN tote M einai to plithos tw n deigmatwn kai
% N einai to plithos tw n features.
% O pinakas GROUP einai ena cell array diastasewn Mx1 kai periexei thn
% kathgoria sthn opoia anhkei ka8e deigma toy pinaka DATA. Sth sygkekrimenh
% synarthsh 8ewroyme oti ta deigmata anhkei eite sthn kathgoria '1' (elkos)
% eite sthn kathgoria '0' (normal)
% Sto method bazoume: 'linear' h 'rbf'
function
[mean_ACCURACY,std_ACCYRACY,mean_SENSITIVITY,std_SENSITIVITY,mean_SPECIFICI
TY,std_SPECIFICITY] = myclassifyWCESvm(DATA,GROUP,method,sigma)
index=50;
ACCURACY(1:index) = 0;
SENSITIVITY(1:index) = 0;
SPECIFICITY(1:index) = 0;
test_size(1:10)=0;
numofnormalintestset(1:10)=0;
err(1:10)=0;
FP_temp(1:10)=0;
normal_samples =0;
for i=1:length(GROUP)
    if GROUP{i}=='0'
        normal_samples = normal_samples+1;
    end
end
%normal_samples;
% index times 10 fold cross validation
for kk=1:index
    FP_temp(1:10) = 0; % false negative
    indices = crossvalind('Kfold',DATA(:,1),10);
    for i = 1:10
        test = (indices == i); % test dataset (binary matrix)
        train = ~test; % train dataset (binary matrix)
        test_size(i) = sum(test);
        len = length(test);
        numofnormalintestset(i) = sum(test(len-normal_samples+1:len)); %
SOS!!! proupo8esi oti oi normal eikones briskontai stis teleytaies grammes
toy pinaka DATA
        if (strcmp(method,'rbf')==1)
            SVMStruct =
svmtrain(DATA(train,:),GROUP(train,:), 'Kernel_Function',method,'RBF_Sigma',
sigma,'autoscale','true');
        else
            SVMStruct =
svmtrain(DATA(train,:),GROUP(train,:), 'Kernel_Function',method,'autoscale',
'true');
        end
        class = svmclassify(SVMStruct,DATA(test,:));
        err(i) = sum(~strcmp(class,GROUP(test,:))); % upologismos tw n
esfalmenwn katataksewn
        % upologismos tw n false negative katataksewn
        for j = 1:length(class)
            AAA = GROUP(test);
            if (strcmp(class(j),'1') && strcmp(AAA(j),'0'))
                FP_temp(i) = FP_temp(i)+1;
            end
        end
    end
    FP = sum(FP_temp);
    FN = sum(err) - FP;
    ACCURACY(kk) = (1 - sum(err)/sum(test_size))*100;
end

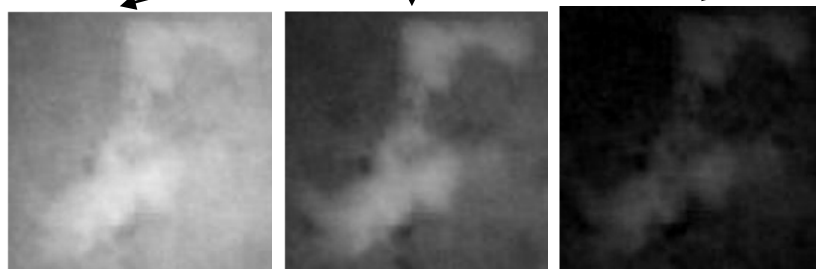
```

```
    SENSITIVITY(kk) = (1 - FN/(sum(test_size)-  
sum(numofnormalintestset)))*100;  
    SPECIFICITY(kk) = (1 - FP/sum(numofnormalintestset))*100;  
end  
  
mean_ACCURACY = mean(ACCURACY);  
std_ACCYRACY = std(ACCURACY);  
mean_SENSITIVITY = mean(SENSITIVITY);  
std_SENSITIVITY = std(SENSITIVITY);  
mean_SPECIFICITY = mean(SPECIFICITY);  
std_SPECIFICITY = std(SPECIFICITY);
```

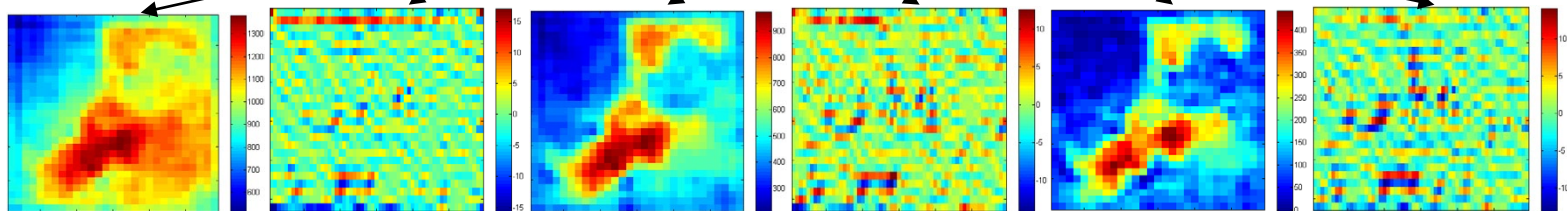
9.2 Διάγραμμα ροής με εικόνα



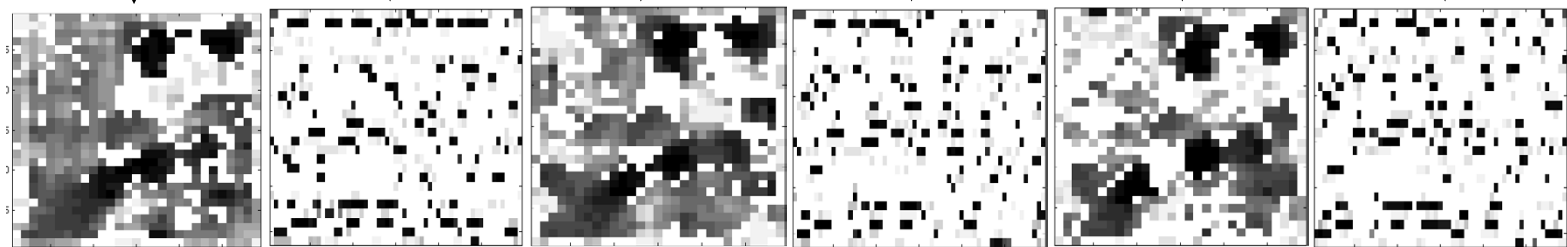
Εξαγωγή χρωματικών καναλιών



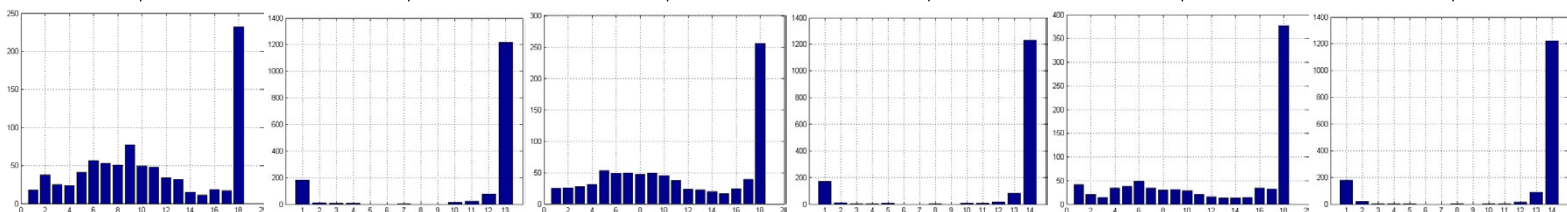
Μετασχηματισμός Curvelet



Ομοιόμορφα ανεξάρτητα περιστροφής ΤΔΠ



Ιστόγραμμα



Υπολογισμός 6 στατιστικών μεγεθών

46,72222	49,49724	4496,833	0	3,100902	12,29113
118	334,67	117312,3	0,890492	3,055733	10,5863

46,72222	53,61497	4897,833	0	3,553058	14,48082
109,5714	326,1964	110809,7	0,749595	3,214497	11,58209

46,72222	83,09159	8703,611	0	3,776843	15,55661
109,5714	324,1433	109569,9	0,749595	3,196856	11,50225

Κατασκευή διανύσματος χαρακτηριστικών ανάλογα με το ποιο/ποια κανάλια θέλουμε να χρησιμοποιήσουμε