

Ατομική Διπλωματική Εργασία

**ΠΕΡΙΛΗΨΗ ΙΑΤΡΙΚΟΥ ΒΙΝΤΕΟ
ΕΝΔΟΣΚΟΠΙΚΗΣ ΕΞΕΤΑΣΗΣ**

Νικολάου Στέλιος

ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΥΠΡΟΥ



ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ

Μάιος 2017

ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΥΠΡΟΥ

ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ

Περίληψη Ιατρικού Βίντεο Ενδοσκοπικής Εξέτασης

Νικολάου Στέλιος

Επιβλέπων Καθηγητής

Πατίχης Κωνσταντίνος

Η Ατομική Διπλωματική Εργασία υποβλήθηκε προς μερική εκπλήρωση των
απαιτήσεων απόκτησης του πτυχίου Πληροφορικής του Τμήματος Πληροφορικής του
Πανεπιστημίου Κύπρου

Μάιος 2017

Ευχαριστίες

Με την παρούσα εργασία ολοκληρώνεται το στάδιο των σπουδών μου ως προπτυχιακός φοιτητής. Στο σημείο αυτό, θα ήθελα να ευχαριστήσω όσους με έχουν βοηθήσει, ο καθένας με το δικό του τρόπο, για την υλοποίηση της εργασίας αυτής.

Αρχικά, θα ήθελα να ευχαριστήσω τον επιβλέποντα καθηγητή μου, Δρ. Κωνσταντίνο Παττίχη. Για την ευκαιρία που μου έδωσε να ασχοληθώ με ένα τόσο ενδιαφέρον θέμα, όπως και για τη εμπιστοσύνη που έδειξε στο πρόσωπό μου καθώς και την πολύ καλή συνεργασία που είχαμε.

Πέραν κάθε αμφιβολίας, ένα τεράστιο ευχαριστώ αξίζει στον Κύριο Αντρέα Παναγίδη, για την καθοδήγησή του και για τη πολύτιμη βοήθεια του στην καταχώρηση των βίντεο για αξιολόγηση στο σύστημα MPEG-DASH.

Ακολούθως, θα ήθελα να εκφράσω τις ευχαριστίες μου στον Κύριο Μάριο Νεοφύτου, για τα βίντεο της ενδοσκοπικής εξέτασης που μου παραχώρησε για τα αποτελέσματα μου.

Επιπρόσθετα, ευχαριστίες αξίζουν και στον ιατρό, Δρ. Βασίλη Τάνο, για την αξιολόγηση των βίντεο περίληψης που έχουν δημιουργηθεί.

Περίληψη

Η διπλωματική μου εργασία καταπιάνεται με την υλοποίηση ενός αλγορίθμου περίληψης βίντεο ενδοσκοπικής χειρουργικής. Στην ενδοσκοπική χειρουργική χρησιμοποιούνται ειδικά εργαλεία, μόνιτορ και κάμερα (ενδοσκόπιο) για την πραγματοποίηση της επέμβασης με ελάχιστα επεμβατικό τρόπο.

Τα βίντεο που έχουν εξαχθεί από την χειρουργική επέμβαση, με τη βοήθεια του αλγορίθμου που ανέπτυξα, μαζί με τη χρήση των φίλτρων Gabor, Tamura, Color and Edge Directivity, Fuzzy Color and Texture Histogram, Auto Color Correlogram, Color Layout, Scalable Color, Edge Histogram και Color Histogram οδήγησαν στην δημιουργία ενός βίντεο περίληψης για τους γιατρούς.

Αυτός ο αλγόριθμος λαμβάνοντας τα κατάλληλα δεδομένα εισόδου εξάγει το βίντεο περίληψης. Αυτό, γίνεται αφού βρεθούν τα στιγμιότυπα με τις μεγαλύτερες διαφορές, ανάλογα με το κάθε φίλτρο. Επιπλέον, έχει δημιουργηθεί και ένας δεύτερος αλγόριθμος ο οποίος αναγνωρίζει με βάση τα χρώματα των εργαλείων τα σημαντικά κομμάτια του βίντεο και τα εξάγει ως περίληψη. Οι δύο αυτοί αλγόριθμοι έχουν υλοποιηθεί στη γλώσσα προγραμματισμού Java με τη βοήθεια του εργαλείου Eclipse.

Για τον έλεγχο του αλγορίθμου χρησιμοποιήθηκαν πέντε βίντεο ενδοσκοπικής εξέτασης. Για κάθε ένα βίντεο έχουμε δοκιμάσει να εξάγουμε περίληψη με υψηλή, μέτρια και χαμηλή συμπίεση και με την εφαρμογή δύο από τα φίλτρα. Τα βίντεο που έχουμε εξάγει έχουν αναρτηθεί στο διαδίκτυο μέσω της χρήσης του εργαλείου MPEG-DASH, με το οποίο θα γίνει και η αξιολόγηση τους.

Η αξιολόγηση του αλγόριθμου που έχει δημιουργηθεί με τα φίλτρα θα αφορά δύο διαφορετικές παραμέτρους. Αυτές είναι η αποτελεσματικότητα στο βαθμό συμπίεσης και η ποιότητα του περιεχομένου του βίντεο.

Περιεχόμενα

Κεφάλαιο 1	Εισαγωγή.....	1
	1.1 Γενική Εισαγωγή	1
	1.2 Ενδοσκοπική Χειρουργική	2
	1.3 Στόχος	3
	1.4 Δομή Εργασίας	3
Κεφάλαιο 2	Αλγόριθμοι.....	5
	2.1 Εισαγωγικά Αλγορίθμων	6
	2.1.1 DCT	6
	2.1.2 Quantization	8
	2.1.3 Zigzag	11
	2.1.4 RGB color model	12
	2.1.5 HSV color space	13
	2.1.6 YIQ color space	14
	2.2 Αλγόριθμος με φίλτρα	15
	2.2.1 Περιγραφή Αλγορίθμου	16
	2.2.2 Φίλτρα	19
	2.2.2.1 Gabor Descriptor	20
	2.2.2.2 Tamura Descriptor	22
	2.2.2.3 CEDD	25
	2.2.2.4 FCTH	25
	2.2.2.5 ACCD	28
	2.2.2.6 CLD	28
	2.2.2.7 SCD	30
	2.2.2.8 EHD	31
	2.2.2.9 CHD	33
	2.2.3 Επιλεγμένα Φίλτρα	34

2.2.4 Επιλογή άλλων παραμέτρων	35
2.3 Αλγόριθμος με χρώματα	35
2.3.1 Περιγραφή Αλγορίθμου	35
Κεφάλαιο 3 Ανάλυση.....	37
3.1 Ανάλυση	37
3.2 Παράμετροι	37
3.3 Αξιολόγηση	39
Κεφάλαιο 4 Ανάλυση Αποτελεσμάτων.....	44
4.1 Περιγραφή Βίντεο	44
4.2 Περιγραφή Συστήματος	45
4.2.1 Λογισμικό Ανάπτυξης - Eclipse	45
4.2.2 Απαιτούμενες τεχνολογίες - Java	46
4.3 Μετρικές Αξιολόγησης	46
4.4 Αποτελέσματα	49
Κεφάλαιο 5 Συμπεράσματα και Μελλοντικές Εργασίες.....	52
5.1 Γενικά Συμπεράσματα	52
5.2 Μελλοντικές Εργασίες	53
Βιβλιογραφία	54

Κεφάλαιο 1

Εισαγωγή

1.1 Γενική Εισαγωγή	1
1.2 Ενδοσκοπική Χειρουργική	2
1.3 Στόχος	3
1.4 Δομή Εργασίας	3

1.1 Γενική Εισαγωγή

Ο τομέας της υγείας αποτελεί έναν από τους βασικότερους για τον άνθρωπο. Είναι γεγονός ότι, με την ανάπτυξη της τεχνολογίας και την εφαρμογή της σε όλους τους βασικούς τομείς της ζωής του ανθρώπου, η ζωή του μπορεί να γίνει καλύτερη.

Για αυτή την εργασία θα μελετηθεί η εφαρμογή της ενδοσκοπικής χειρουργικής. Είναι μια χειρουργική επέμβαση η οποία γίνεται με τη βοήθεια του βίντεο για να βρεθούν τρόποι με τους οποίους θα μπορούν να μειωθούν τα περιττά στοιχεία από τα βίντεο της επέμβασης αυτής.

1.2 Ενδοσκοπική Χειρουργική

Η Ενδοσκοπική Χειρουργική είναι ο ένας από τους τομείς της χειρουργικής. Στο τομέα αυτό χρησιμοποιούνται ειδικά εργαλεία, μόνιτορ και κάμερας (ενδοσκόπιο) για την πραγματοποίηση της επέμβασης με ελάχιστα επεμβατικό τρόπο [1,2].

Αντίθετα, με άλλες χειρουργικές μεθόδους, στην ενδοσκοπική χειρουργική η εισαγωγή των εργαλείων που χρησιμοποιούνται γίνεται συνήθως μέσω φυσιολογικών κοιλοτήτων ή μέσω μικρών τομών στο δέρμα, μήκους λίγων μόνο χιλιοστών [1,2].

Για να γίνει η επέμβαση είναι απαραίτητη η δημιουργία του χώρου που χρειάζεται για την άνετη χρήση των εργαλείων στο σώμα του ασθενή για τη διεξαγωγή της επέμβασης. Αυτό, γίνεται με τη χρήση αερίου ή υγρού διαλύματος. Ακόμη, με την χρήση ειδικών συστημάτων απεικόνισης μεγάλης ανάλυσης, γίνεται μεγέθυνση των ιστών που προβάλλονται έγχρωμα σε οθόνη. Αυτό δίνει τη δυνατότητα στο χειρουργό τόσο για πολύ καλή εξέταση των ιστών όσο και για χειρουργική θεραπεία που έχει βρεθεί άμεσα [1,2].

Τα πλεονεκτήματα του τομέα αυτού της χειρουργικής, σε σχέση με τις συμβατικές μεθόδους είναι:

- καλύτερη ορατότητα, ειδικά σε περιοχές των ιστών που είναι δύσκολα προσβάσιμες
- αρκετά μικρότερο τραύμα, άρα και μικρότερη πιθανότητα για επιπλοκές
- σημαντική μείωση του πόνου μετά την εγχείρηση
- γρηγορότερη ανάρρωση και έξοδος από την κλινική
- γρήγορη επάνοδος στις φυσιολογικές δραστηριότητες και στην εργασία
- καλύτερο αισθητικό αποτέλεσμα λόγω μικρότερων τομών

1.3 Στόχος

Στόχος της παρούσας εργασίας είναι η δημιουργία ενός προγράμματος το οποίο θα δημιουργεί περίληψη ιατρικού βίντεο ενδοσκοπικής χειρουργικής εξέτασης. Το πρόγραμμα που θα δημιουργηθεί δηλαδή θα πρέπει να λαμβάνει ως είσοδο ένα βίντεο ενδοσκοπικής χειρουργικής και να βρίσκει σε αυτό τις σκηνές που πραγματοποιούνται κάποιες λειτουργίες των ιατρικών εργαλείων.

1.3 Δομή Εργασίας

Στο πρώτο κεφάλαιο, γίνεται μια γενική εισαγωγή για το θέμα της διπλωματικής εργασίας. Επιπλέον γίνεται αναφορά στους στόχους της εργασίας αυτής και τη δομή της.

Στο δεύτερο κεφάλαιο, αναλύονται οι δύο αλγόριθμοι οι οποίοι έχουν υλοποιηθεί και γίνεται μια αναφορά στα φίλτρα που έχουν χρησιμοποιηθεί για τον πρώτο αλγόριθμο. Ακόμη, αναλύονται ποια φίλτρα έχουν επιλεγεί ως κατάλληλα και επεξηγούνται οι παράμετροι για την εκτέλεση του προγράμματος.

Στο τρίτο κεφάλαιο, γίνεται η ανάλυση για τις παραμέτρους που έχουν χρησιμοποιηθεί για κάθε εκτέλεση και γίνεται αναφορά στο τρόπο λειτουργίας του εργαλείου MPEG-DASH στο οποίο θα αναρτηθούν τα βίντεο για την αξιολόγηση από το γιατρό.

Στο τέταρτο κεφάλαιο, γίνεται περιγραφή των βίντεο που χρησιμοποιήθηκαν σαν είσοδο και του συστήματος που έχει υλοποιηθεί ο αλγόριθμος. Επιπλέον, σε αυτό το κεφάλαιο γίνεται η αξιολόγηση των βίντεο περίληψης με βάση τη συμπίεση και το περιεχόμενο τους.

Στο πέμπτο κεφάλαιο, αναφέρονται τα συμπεράσματα από την όλη διαδικασία, οι μελλοντικές εργασίες αλλά και βελτιστοποιήσεις που μπορούν να γίνουν στο σύστημα για καλύτερα αποτελέσματα.

Κεφάλαιο 2

Αλγόριθμοι

2.1 Εισαγωγικά Αλγορίθμων	6
2.1.1 DCT	6
2.1.2 Quantization	8
2.1.3 ZigZag	11
2.1.4 RGB color model	12
2.1.5 HSV color space	13
2.1.6 YIQ color space	14
2.2 Αλγόριθμος με Φίλτρα	15
2.2.1 Περιγραφή Αλγορίθμου	16
2.2.2 Φίλτρα	19
2.2.2.1 Gabor Descriptor	20
2.2.2.2 Tamura Descriptor	22
2.2.2.3 CEDD	25
2.2.2.4 FCTHD	25
2.2.2.5 ACCD	28
2.2.2.6 CLD	28
2.2.2.7 SCD	30
2.2.2.8 EHD	31
2.2.2.9 CHD	33
2.2.3 Επιλεγμένα φίλτρα	34
2.2.4 Επιλογή άλλων Παραμέτρων	35
2.3 Αλγόριθμος με Χρώματα	35
2.3.1 Περιγραφή Αλγορίθμου	35

2.1 Εισαγωγικά Αλγορίθμων

2.1.1 DCT

Ο διακριτός μετασχηματισμός συνημιτόνου (DCT) είναι ένας γραμμικός μετασχηματισμός, ο οποίος χαρτογραφεί ένα n -διαστάσεων διάνυσμα σε σύνολο συντελεστών n . Ένας γραμμικός συνδυασμός n γνωστών βασικών σταθμισμένων διανυσμάτων με τους συντελεστές n θα έχει ως αποτέλεσμα το αρχικό διάνυσμα. Τα γνωστά βασικά διανύσματα των μετασχηματισμών από αυτή την κατηγορία είναι ημιτονοειδείς, δηλαδή μπορούν να εκπροσωπούνται από κύματα μορφής ημιτόνου ή, με άλλα λόγια, είναι έντονα εντοπισμένα στο φάσμα συχνοτήτων. Άρα, μιλάμε για μετασχηματισμό στον τομέα συχνοτήτων [3,4,29].

Στη συμπίεση JPEG τα δεδομένα εισόδου είναι δύο διαστάσεων, παρουσιάζονται σε μπλοκ 8×8 . Υπάρχει ανάγκη χρήσης DCT δύο διαστάσεων. Δεδομένου ότι κάθε διάσταση μπορεί να χειριστεί ξεχωριστά, ο δισδιάστατος DCT ακολουθεί απλή μορφή από τη μονοδιάστατη DCT. Μια μονοδιάστατη DCT εκτελείται κατά μήκος των σειρών και έπειτα κατά μήκος των στηλών ή αντιστρόφως [3,4,30].

Ο υπολογισμός για DCT δύο διαστάσεων

$$F(u, v) = C(u)C(v) \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} f(x, y) \cos\left[\frac{\pi(2x+1)u}{2N}\right] \cos\left[\frac{\pi(2y+1)v}{2N}\right]$$

όπου $u=0,1,\dots,N-1$ και $v=0,1,\dots,M-1$

$$C(u), C(v) = \sqrt{\frac{1}{N}} \quad \text{όταν } u, v = 0$$

$$C(u), C(v) = \sqrt{\frac{2}{N}} \quad \text{όταν } u, v \neq 0$$

Όπου N και M οι διαστάσεις της εικόνας.

Η εφαρμογή αυτών των τύπων απαιτεί άμεσα πολλούς υπολογιστικούς πόρους.

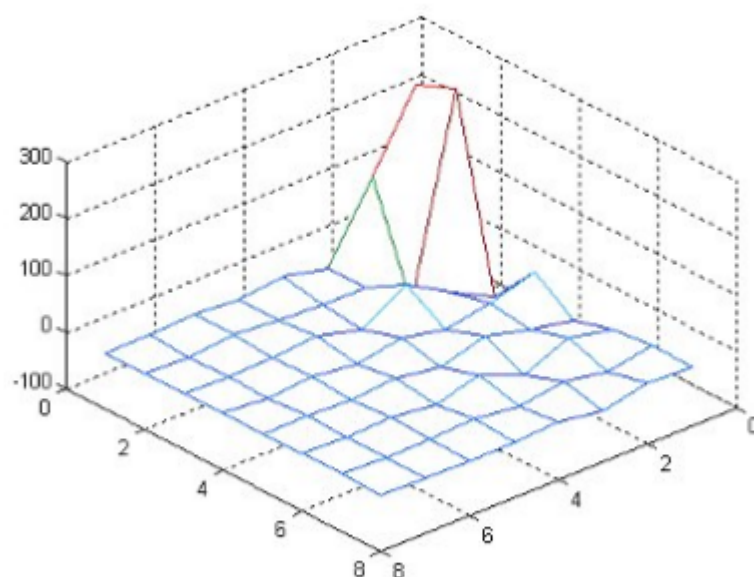
75	76	75	75	69	66	77	71
73	74	73	74	63	64	68	69
69	68	71	72	67	58	48	41
59	55	56	52	47	40	24	9
51	50	45	41	33	22	7	-5
43	37	32	24	15	5	-6	-25
29	21	9	-2	-10	-21	-44	-69
9	-4	-17	-35	-52	-61	-57	-35

Εικόνα 2.1 8x8 μπλοκ πριν το DCT [3]

251	118	-13	6	-2	6	-1	0
279	-68	-8	-7	-1	4	-4	-1
-51	-14	34	-14	5	0	-1	0
27	5	-10	8	-7	4	-5	1
-22	-7	14	-9	4	-2	1	1
-3	15	-18	15	-6	2	-1	2
7	-9	6	-6	4	0	0	2
3	7	-9	3	0	-2	-1	0

Εικόνα 2.2 8x8 μπλοκ μετά το DCT [3]

Όπως φαίνεται στην Εικόνα 2.2, υψηλότερες τιμές συντελεστών μετασχηματισμού συγκεντρώνονται στην πάνω αριστερή γωνία. Στον τομέα των συχνοτήτων φαίνεται ότι η χαμηλή συχνότητα έχει πλεονέκτημα έναντι της υψηλής συχνότητας. Εμφανίζεται στην Εικόνα 2.3 [3,31].



Εικόνα 2.3 Συχνότητες του 8x8 μπλοκ [3]

Όπως μπορούμε να δούμε μόνο λίγα στοιχεία χαμηλής συχνότητας κυριαρχούν πάνω από τους υπόλοιπους συντελεστές. Έτσι επιτρέπει τη μείωση των δεδομένων κατά τα επόμενα στάδια της συμπίεσης JPEG [3,4,31].

2.1.2 Quantization

Η κβαντοποίηση (Quantization) είναι ένα βασικό βήμα στη διαδικασία συμπίεσης, αφού έχουν αφαιρεθεί λιγότερο σημαντικές πληροφορίες. Το πλεονέκτημα της αναπαράστασης στο πεδίο συχνοτήτων είναι ότι, αντίθετα με τον χωρικό τομέα πριν από το DCT, δεν έχει κάθε διάσταση την ίδια σημασία για την οπτική ποιότητα της εικόνας. Η αφαίρεση των στοιχείων υψηλότερων συχνοτήτων θα μειώσει το επίπεδο λεπτομέρειας, αλλά η συνολική δομή παραμένει, δεδομένου ότι κυριαρχείται από τα συστατικά μέρη χαμηλότερης συχνότητας [5,6].

Οι 64 τιμές ενός μπλοκ 8x8 θα διαιρούνται σύμφωνα με τις 64 τιμές ενός 8x8 πίνακα που ονομάζεται πίνακας κβαντισμού. Δεν υπάρχει καμία απώλεια πληροφοριών στη

διαίρεση των ίδιων των συντελεστών, αλλά στη συνέχεια το αποτέλεσμα στρογγυλοποιείται στον επόμενο ακέραιο. Όσο υψηλότερος είναι ο διαιρέτης, τόσο περισσότερες πληροφορίες σχετικά με τον συντελεστή θα τοποθετηθούν μετά το δεκαδικό ψηφίο και έτσι θα χαθούν κατά τη στρογγυλοποίηση [5,29].

Η κβαντοποίηση μπορεί να οριστεί με τον ακόλουθο τύπο:

$$Sq_{vu} = \text{round}\left(\frac{S_{vu}}{Q_{vu}}\right)$$

όπου Sq_{uv} είναι οι συντελεστές μετά την κβαντοποίηση

S_{vu} οι συντελεστές πριν την κβαντοποίηση (μετά το DCT)

Q_{vu} οι τιμές του πίνακα κβαντοποίησης

Ο πίνακας κβαντοποίησης δίνεται ως παράδειγμα στο πρότυπο JPEG (Εικόνα 2.4).

17	18	24	47	99	99	99	99
18	21	26	66	99	99	99	99
24	26	56	99	99	99	99	99
47	66	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99

Εικόνα 2.4 Πίνακας κβαντοποίηση [5]

Πολλοί κωδικοποιητές χρησιμοποιούν απλά παραδείγματα πινάκων κβαντοποίησης, αλλά οι τιμές δεν υποστηρίζονται ότι είναι οι βέλτιστες. Ένας κωδικοποιητής μπορεί να χρησιμοποιήσει οποιοδήποτε άλλο πίνακα κβαντοποίησης, πιθανώς βελτιστοποιημένο αν γίνει πρώτα η ανάλυση της εικόνας. Οι πίνακες κβαντοποίησης αποθηκεύονται στην κεφαλίδα του αρχείου JPEG στον δείκτη "DQT" (Define Quantization Table) για να είναι διαθέσιμος για αποκωδικοποίηση. Οι περισσότεροι κωδικοποιητές JPEG

επιτρέπουν στο χρήστη να επιλέξει ένα επίπεδο συμπίεσης ή μια ρύθμιση ποιότητας. Αυτή η παράμετρος καθορίζει την αναλογία μεταξύ της απόδοσης της συμπίεσης και της σχετικής απώλειας ποιότητας. Με τη ρύθμιση αυτής της παραμέτρου, στην πραγματικότητα καθορίζεται ποιος πίνακας κβαντοποίησης θα χρησιμοποιήσει ο κωδικοποιητής (συνήθως ο πίνακας κβαντοποίησης του παραδείγματος θα πολλαπλασιαστεί με έναν παράγοντα που υπολογίζεται από την επιλεγμένη παράμετρο) [5,6,32].

Ένα παράδειγμα με τη χρήση του πίνακα κβαντοποίησης φαίνεται στην Εικόνα 2.5 για πίνακα 8x8 μετά το DCT από την Εικόνα 2.2 και πίνακα κβαντοποίησης από την Εικόνα 2.4.

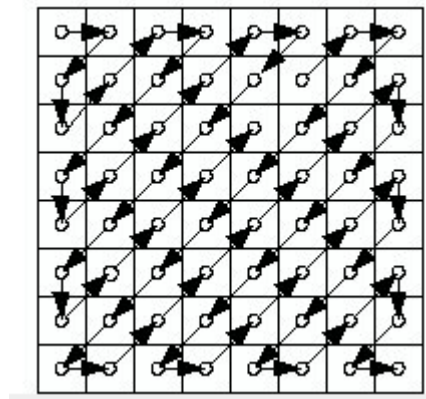
14	6	0	0	0	0	0	0
15	-3	0	0	0	0	0	0
-2	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

Εικόνα 2.5 Αποτέλεσμα μετά τη κβαντοποίηση[5,6]

Σαν αποτέλεσμα, μετά τη κβαντοποίηση οι περισσότεροι συντελεστές είναι ίσοι με μηδέν [5].

2.1.3 Zigzag

Μετά το βήμα κβαντοποίησης, οι περισσότεροι από τους συντελεστές προς την κάτω δεξιά γωνία είναι μηδενικοί (Εικόνα 2.5). Το Zigzag ordering αναδιατάσσει τους συντελεστές σε μια μονοδιάστατη σειρά, με αποτέλεσμα τα περισσότερα από τα μηδενικά να τοποθετηθούν στο τέλος [7,29,33].



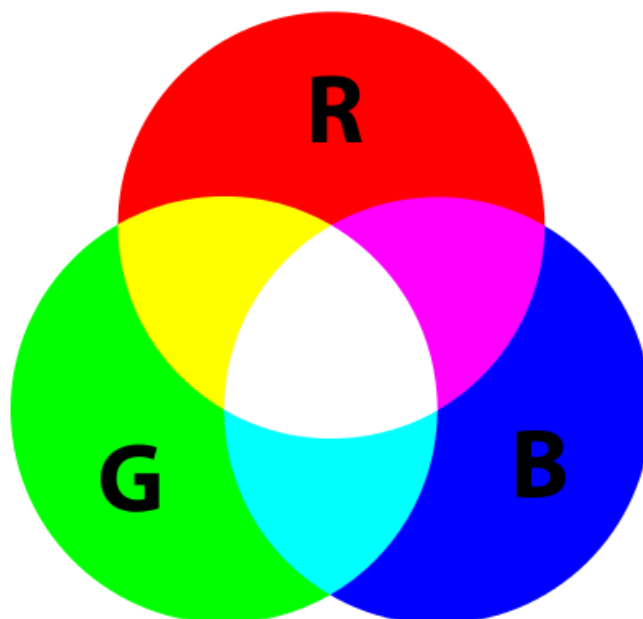
Εικόνα 2.6 Σειρά Zigzag [33]

14	6	15	-2	-3	0	0	0	0	0	0	0
----	---	----	----	----	---	---	---	---	---	---	---

Εικόνα 2.7 Αποτέλεσμα του Zigzag [33]

Αυτή η ροή με τα πολλά διαδοχικά μηδενικά στο τέλος βελτιστοποιείται τώρα για να επιτευχθεί υψηλή συμπίεση στην κωδικοποίηση εντροπίας [7].

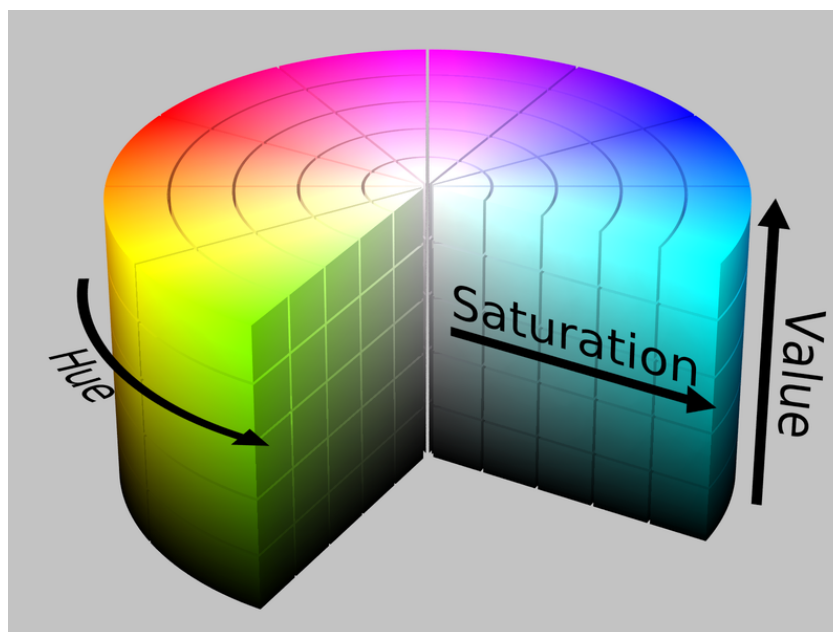
2.1.4 RGB color model



Εικόνα 2.8 Αναπαράσταση χρωμάτων RGB μοντέλου [8]

Το έγχρωμο μοντέλο RGB είναι ένα μοντέλο χρώματος προσθήκης στο οποίο προστίθενται κόκκινο, πράσινο και μπλε φως με διάφορους τρόπους για να αναπαράγουν ένα ευρύ φάσμα χρωμάτων, όπως φαίνεται στην εικόνα. Το όνομα του μοντέλου προέρχεται από τα αρχικά των τριών πρόσθετων βασικών χρωμάτων, κόκκινο, πράσινο και μπλε. Ο κύριος σκοπός του μοντέλου χρωμάτων RGB είναι η ανίχνευση και η απεικόνιση εικόνων σε ηλεκτρονικά συστήματα [8].

2.1.5 HSV color space



Εικόνα 2.9 Αναπαράσταση HSV color space [9]

Ο HSV συμβολίζει την απόχρωση (Hue), τον κορεσμό (Saturation) και την αξία (Value) και συχνά ονομάζεται HSB (B για φωτεινότητα, Brightness). Στον κύλινδρο, η γωνία γύρω από τον κεντρικό κατακόρυφο άξονα αντιστοιχεί σε "απόχρωση", η απόσταση από τον άξονα αντιστοιχεί σε "κορεσμό" και η απόσταση κατά μήκος του άξονα αντιστοιχεί σε "τιμή" [9].

Επειδή ο HSV είναι ένας απλός μετασχηματισμός μοντέλων RGB που εξαρτώνται από τη συσκευή, τα φυσικά χρώματα που καθορίζουν εξαρτώνται από τα χρώματα των κόκκινων, πράσινων και μπλε πρωτευόντων της συσκευής ή του συγκεκριμένου χώρου RGB και από τη διόρθωση γάμμα που χρησιμοποιείται για να αντιπροσωπεύει τις ποσότητες αυτές. Ως αποτέλεσμα, κάθε μοναδική συσκευή RGB έχει μοναδικούς κεντρικούς χώρους χρώματος HSV για να την συνοδεύει (ακριβώς όπως έχει μοναδικό RGB απόλυτο έγχρωμο χώρο για να την συνοδεύει) και οι ίδιες αριθμητικές τιμές HSV (ακριβώς όπως αριθμητικές τιμές RGB) μπορεί να παρουσιάζονται διαφορετικά σε διαφορετικές συσκευές [9].

2.1.6 YIQ color space

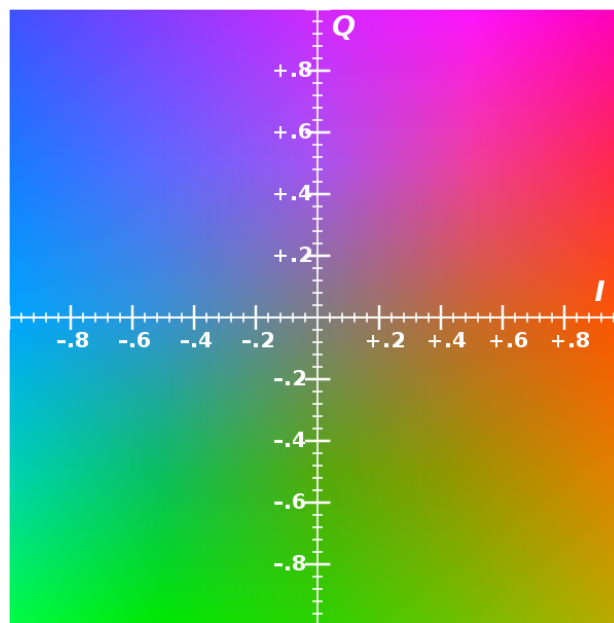
Ο YIQ είναι ο χρωματικός χώρος που χρησιμοποιείται από το σύστημα έγχρωμης τηλεόρασης NTSC. Το I αντιπροσωπεύει τη φάση, ενώ το Q αντιπροσωπεύει τον τετραγωνισμό, δηλαδή αναφέρεται στα συστατικά που χρησιμοποιούνται στη διαμόρφωση πλάτους τετραγωνικής διατομής [11].

Το στοιχείο Y αντιπροσωπεύει τις πληροφορίες luma, το luma αντιπροσωπεύει τη φωτεινότητα μιας εικόνας [10], και είναι το μόνο συστατικό που χρησιμοποιείται από τους ασπρόμαυρους δέκτες τηλεόρασης. Τα I και Q αντιπροσωπεύουν τις πληροφορίες χρώματος. Στο YUV, τα στοιχεία U και V μπορούν να θεωρηθούν ως συντεταγμένες X και Y μέσα στο χώρο χρώματος. Τα I και Q μπορούν να θεωρηθούν ως ένα δεύτερο ζεύγος αξόνων στο ίδιο γράφημα, περιστρεφόμενα κατά 33 °. Επομένως το IQ και το UV αντιπροσωπεύουν διαφορετικά συστήματα συντεταγμένων στο ίδιο επίπεδο [11].

Το σύστημα YIQ προορίζεται να αξιοποιήσει τα χαρακτηριστικά ανθρώπινης απόκρισης χρώματος. Το μάτι είναι πιο ευαίσθητο στις μεταβολές στην περιοχή πορτοκαλί-μπλε (I) από ό, τι στη περιοχή μοβ-πράσινο (Q) - επομένως απαιτείται μικρότερο εύρος ζώνης για το Q από ότι για το I. Το εκπεμπόμενο NTSC περιορίζει το I έως 1,3 MHz και Q έως 0,4 MHz. Το I και Q είναι συχνότητες που παρεμβάλλονται στο σήμα 4 MHz του Y, το οποίο διατηρεί το εύρος ζώνης του συνολικού σήματος κάτω από 4,2 MHz [11].

Πολύ μεγάλο το κόστος αποκωδικοποίησης των I και Q. Τα I και Q απαιτούν διαφορετικό φίλτρο για να ικανοποιήσει τις διαφορές εύρους ζώνης μεταξύ τους. Αυτές οι διαφορές εύρους ζώνης απαιτούν επίσης ότι το φίλτρο I περιλαμβάνει μια χρονική καθυστέρηση για να ταιριάζει με την μεγαλύτερη καθυστέρηση του φίλτρου Q [11].

Η αναπαράσταση YIQ χρησιμοποιείται σε μετασχηματισμούς επεξεργασίας χρωμάτων εικόνας. Για παράδειγμα, η εφαρμογή εξισορρόπησης ιστογράμματος απευθείας στα κανάλια σε μια εικόνα RGB θα αλλάξει την ισορροπία χρωμάτων της εικόνας. Αντ' αυτού, η ισορροπία ιστόγραμμα εφαρμόζεται στο κανάλι Y της παράστασης YIQ ή YUV της εικόνας, η οποία κανονικοποιεί μόνο τα επίπεδα φωτεινότητας της εικόνας [11].



Εικόνα 2.10 Αναπαράσταση χώρου YIQ [11]

2.2 Αλγόριθμος με Φίλτρα

Για την υλοποίηση του αλγορίθμου που έχει αναπτυχθεί έχουν μελετηθεί κάποια άρθρα και υπάρχων εργασίες για την περίληψη βίντεο. Αυτό έχει γίνει για την ανεύρεση του καλύτερου αλγορίθμου με σκοπό η εργασία αυτή να έχει όσο το δυνατόν καλύτερα αποτελέσματα. Με βάση τα άρθρα αυτά υπάρχουν δύο κύριες επιλογές.

Η πρώτη επιλογή είναι ο διαχωρισμός κάποιον βίντεο από τα οποία θα έχουν ξεχωρίσει τα σημαντικά στιγμιότυπα για περίληψη και αυτά που δεν είναι σημαντικά, δηλαδή τα στιγμιότυπα που δεν περιέχουν κάποια πληροφορία. Με βάση αυτά τα δεδομένα θα

γίνεται εκπαίδευση του προγράμματος με ένα negative και ένα positive dataset. Έτσι θα μπορεί να ελεγχθεί ο αλγόριθμος με βάση τα δεδομένα εκπαίδευσης και να αξιολογηθεί η ποιότητα του [34].

Για να γίνει όμως αυτό θα χρειαστεί μια μεγάλη ποσότητα από βίντεο, να εξαχθούν από αυτά τα στιγμιότυπα του και να κατηγοριοποιηθούν στα δυο datasets. Αυτή είναι μια πολύ χρονοβόρα διαδικασία για να γίνει ενώ παράλληλα για στιγμιότυπα που δεν έχει τη κατάλληλη εκπαίδευση το πρόγραμμα αυτό δεν θα εξάγει με βεβαιότητα τα αποτελέσματα. Για αυτούς τους λόγους δεν έχει επιλεγεί αυτή η μέθοδος [34].

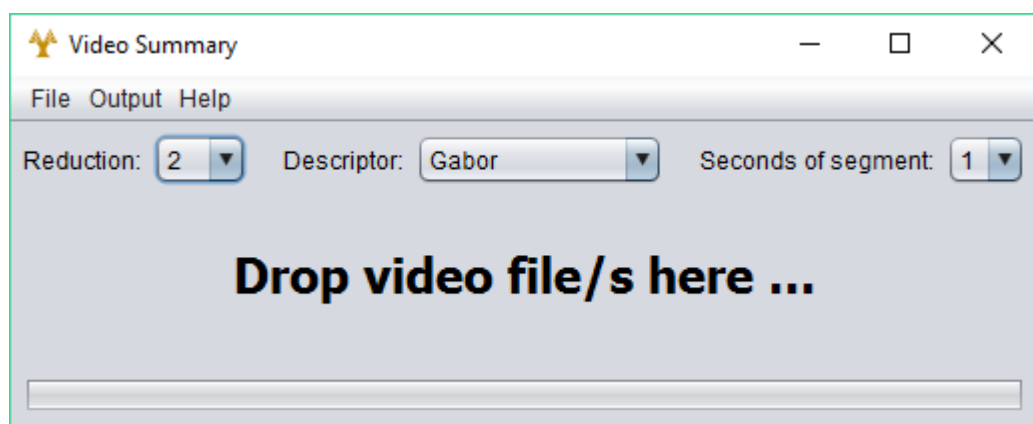
Η δεύτερη μέθοδος είναι με τη χρήση διαφόρων φίλτρων για την εύρεση των σημαντικών αυτών στιγμιότυπων. Αυτή η μέθοδος προτείνει με τη χρήση των φίλτρων να επιλέγονται τα σημαντικά αυτά στιγμιότυπα ανάλογα με τη διαφορά που έχει το ένα με το άλλο, δηλαδή αυτά με τη μεγαλύτερη διαφορά θα μπορούσαν να γίνουν περίληψη [35,36,37].

2.2.1 Περιγραφή Αλγορίθμου

Για την υλοποίηση της εφαρμογής αυτής έχει χρησιμοποιηθεί σαν βάση ένα πρόγραμμα ανοικτού κώδικα, με GPL licence για χρήση [38]. Το licence αυτό μας δείνει τη δυνατότητα να επεξεργαστούμε το κώδικα και να τον εφαρμόσουμε στις ανάγκες της εργασίας αυτής, αρκεί να είναι και αυτή ανοικτού κώδικα [42].

Το πρόγραμμα αυτό παίρνει σαν είσοδο ένα βίντεο και το φίλτρο με το οποίο θα εξάγει ένα αριθμό από στιγμιότυπα, ο οποίος δίνεται από το χρήστη, σαν περίληψη. Σε αυτό το πρόγραμμα έγιναν κάποιες βασικές αλλαγές έτσι ώστε να μπορεί να χρησιμοποιηθεί για το σκοπό της εργασίας αυτής, δηλαδή τη περίληψη ιατρικού βίντεο ενδοσκοπικής εξέτασης [38].

Το διαμορφωμένο αυτό πρόγραμμα παίρνει σαν δεδομένα εισόδου το βίντεο του οποίου θέλουμε να δημιουργηθεί η περίληψη, το φίλτρο με το οποίο θα βρει τα πιο σημαντικά - διαφορετικά στιγμιότυπα από το αρχικό βίντεο, το μέγεθος της μείωσης του βίντεο που θέλει ο χρήστης να πετύχει και ο χρόνος σε δευτερόλεπτα που θέλει να εμφανίζονται μετά από το κάθε σημαντικό στιγμιότυπο. Όλα αυτά τα στοιχεία δίνονται από το χρήστη μέσω ενός διαμορφωμένου μενού με επιλογές οι οποίες διευκρινίζουν στο χρήστη τις παραμέτρους που πρέπει να δώσει. Η επιλογή του βίντεο γίνεται με την χρήση του drag and drop. Όταν ο χρήστης εισάγει το βίντεο τότε αυτόματα ξεκινάει το πρόγραμμα για να ετοιμάσει τη περίληψη. Στο σχήμα 2.11 βλέπουμε το παράθυρο εκτέλεσης του προγράμματος με τις επιλογές αυτές.



Σχήμα 2.11 Μενού εκτέλεσης

Το τελικό μέγεθος του βίντεο που θα έχει η περίληψη υπολογίζεται με βάση τον συνολικό αριθμό σε στιγμιότυπα που έχει το αρχικό βίντεο, τα στιγμιότυπα που εμφανίζονται ανά δευτερόλεπτο του βίντεο, τον αριθμό της μείωσης, που εισάγει ο χρήστης από το μενού που εμφανίζεται όταν εκτελεστεί το πρόγραμμα, και τον αριθμό των δευτερολέπτων που θα εμφανίζονται, το οποίο εισάγει ο χρήστης, μετά από το σημαντικό στιγμιότυπο. Ο αριθμός αυτός σε δευτερόλεπτα δεν είναι απόλυτος γιατί στη περίπτωση που το στιγμιότυπο είναι προς το τέλος του βίντεο θα παρουσιάζονται όσα στιγμιότυπα απομένουν μέχρι το τέλος. Επιπλέον σε περίπτωση που δυο σημαντικά στιγμιότυπα είναι κοντά το ένα στο άλλο τότε θα εμφανίζονται μόνο τα στιγμιότυπα

μέχρι το επόμενο σημαντικό στιγμιότυπο. Έτσι η εξίσωση για των αριθμό των σημαντικών στιγμιότυπων είναι η πιο κάτω:

$$ImportantFrames = \frac{VideoTotalFrames}{VideoReduction * FramesperSecond * SecondstoShow}$$

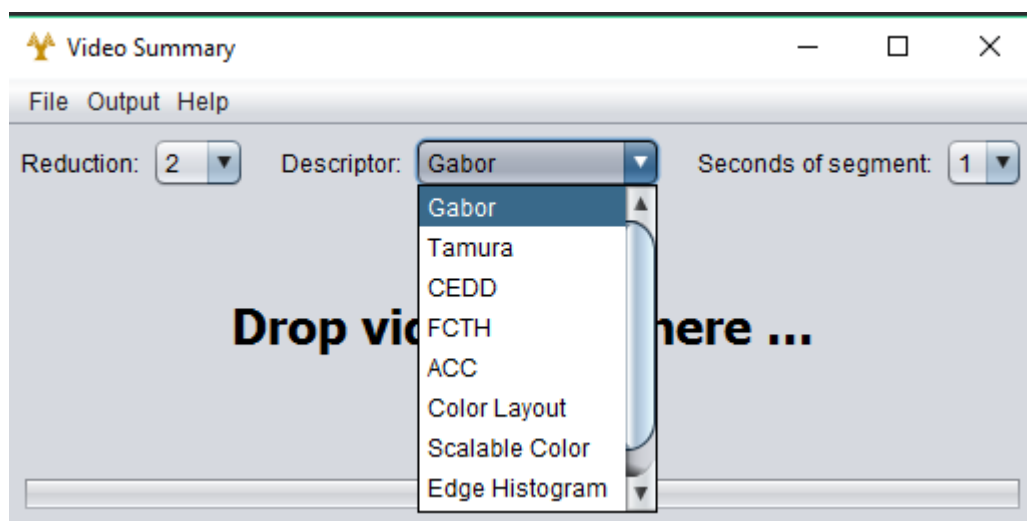
Τα σημαντικά στιγμιότυπα για κάθε βίντεο εξάγονται με βάση το κάθε φίλτρο. Αυτό γίνεται δηλαδή με τον υπολογισμό ενός score για κάθε στιγμιότυπου και για κάθε ζεύγος από στιγμιότυπα υπολογίζεται η διαφορά, απόσταση, των δύο score. Τα στιγμιότυπα με τη μεγαλύτερη συνολική διαφορά αποθηκεύονται ως τα πιθανά για να γίνει η εξαγωγή τους στο βίντεο της περίληψης. Μετά τον υπολογισμό της απόστασης για πρώτη φορά γίνεται ξανά υπολογισμός αν υπάρχουν κάποια άλλα στιγμιότυπα που η συνολική τους απόσταση να είναι μεγαλύτερη σε σχέση με αυτά που έχουν επιλεγεί μέχρι στιγμής. Αυτό επαναλαμβάνεται, όσες φορές χρειαστεί, μέχρι να σταματήσουν να γίνονται αλλαγές στα επιλεγμένα στιγμιότυπα. Όταν τελειώσει αυτή η διαδικασία σημαίνει ότι έχουν επιλεγεί τα στιγμιότυπα με τη μεγαλύτερη συνολική διαφορά, με βάση το επιλεγμένο φίλτρο και αυτά θα πρέπει να εξαχθούν για να δημιουργηθεί η περίληψη.

Αφού έχουν βρεθεί τα στιγμιότυπα που θα αποτελούν την αρχή κάθε κομματιού της περίληψης, το πρόγραμμα τα αποθηκεύει σε μια λίστα. Σε αυτή τη λίστα προσθέτονται και τα στιγμιότυπα για τα επόμενα δευτερόλεπτα από την εμφάνιση του σημαντικού, με βάση τις επιλογές που έχει κάνει ο χρήστης στο αρχικό μενού. Αυτά τα στιγμιότυπα εξάγονται σε εικόνες τύπου mpeg και αποθηκεύονται σε ένα προκαθορισμένο φάκελο από το πρόγραμμα, ο οποίος έχει αδειάσει από πριν. Όταν η εξαγωγή των εικόνων έχει τελειώσει, αυτές διαβάζονται ξανά από το πρόγραμμα και δημιουργείτε ένα βίντεο από αυτές. Αυτό το βίντεο αποτελεί την περίληψη του αρχικού μας βίντεο. Η ανάλυση και τα στιγμιότυπα ανά δευτερόλεπτο του βίντεο περίληψης είναι τα ίδια και με το αρχικό βίντεο.

2.2.2 Φίλτρα

Όπως φαίνεται και στις Εικόνες 2.12 και 2.13 ο χρήστης έχει της δυνατότητα να επιλέξει μέσα από μια λίστα με φίλτρα (επιλογή descriptors). Το κάθε ένα από αυτά τα φίλτρα έχουν διαφορετικές ιδιότητες με αποτέλεσμα να εξάγονται από αυτά διαφορετικά αποτελέσματα ως περίληψη. Τα φίλτρα που έχουν χρησιμοποιηθεί για την εύρεση των διαφορετικών εικόνων είναι χωρισμένα σε δυο κατηγορίες, φίλτρα που λαμβάνουν υπόψη τους για το διαχωρισμό τα χρώματα που υπάρχουν σε κάθε στιγμιότυπο και φίλτρα που λαμβάνουν υπόψη τους για το διαχωρισμό την υφή του κάθε στιγμιότυπο. Επιπλέον υπάρχουν και κάποια που αποτελούν συνδυασμό των δυο αυτών μεθόδων.

Τα φίλτρα αυτά είναι υλοποιημένα στη βιβλιοθήκη lire [39,40]. Σε αυτή τη βιβλιοθήκη έχουν υλοποιηθεί τα φίλτρα Gabor, Tamura, Color and Edge Directivity Descriptor, Fuzzy Color and Texture Histogram, Auto Color Correlogram, Color Layout, Scalable Color, Edge Histogram Color και Histogram και είναι επίσης ανοικτού κώδικα [41].



Εικόνα 2.12 Μενού εκτέλεσης - Επιλογή φίλτρου 1

Εικόνα 2.13 Μενού εκτέλεσης - Επιλογή φίλτρου 2

2.2.2.1 Gabor Descriptor

Το φίλτρο Gabor είναι ένα γραμμικό φίλτρο για την ανίχνευση ακμών. Η αναπαράσταση της συχνότητας και του προσανατολισμού αυτού του φίλτρου είναι παρόμοια με του ανθρώπινου ματιού, και χρησιμοποιείται κυρίως για την αντιπροσώπευση και διάκριση υφής. Στον χωρικό τομέα, ένα φίλτρο 2D Gabor είναι μια συνάρτηση Gaussian πυρήνα διαμορφωμένη από ένα ημιτονοειδές επίπεδο κύμα [12].

Η απόκριση του παλμού ορίζεται από ένα ημιτονοειδές κύμα πολλαπλασιασμένο με μια Gaussian συνάρτηση. Λόγω του πολλαπλασιασμού, ο μετασχηματισμός Fourier της παλμικής απόκρισης του φίλτρου Gabor είναι η συνέλιξη του μετασχηματισμού Fourier της αρμονικής συνάρτησης και του μετασχηματισμού Fourier της Gaussian συνάρτησης. Το φίλτρο έχει ένα πραγματικό (real) και ένα φανταστικό (imaginary) τμήμα που αντιπροσωπεύει ορθογώνιες κατευθύνσεις. Τα δύο τμήματα μπορούν να διαμορφωθούν σε ένα σύνθετο (complex) αριθμό ή να χρησιμοποιηθούν μεμονωμένα [12].

Complex:

$$g(x, y; \lambda, \theta, \psi, \sigma, \gamma) = e^{(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2})} e^{i(2\pi \frac{x'}{\lambda} + \psi)}$$

Real:

$$g(x, y; \lambda, \theta, \psi, \sigma, \gamma) = e^{(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2})} \cos(2\pi \frac{x'}{\lambda} + \psi)$$

Imaginary:

$$g(x, y; \lambda, \theta, \psi, \sigma, \gamma) = e^{(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2})} \sin(2\pi \frac{x'}{\lambda} + \psi)$$

$$\text{Όπου } x' = x \cos \theta + y \sin \theta \quad \text{και} \quad y' = -x \sin \theta + y \cos \theta$$

Στις πιο πάνω εξισώσεις το λ αντιπροσωπεύει το μήκος κύματος του ημιτονοειδούς παράγοντα, το θ αντιπροσωπεύει του κανονικού προς τις παράλληλες λωρίδες μιας συνάρτησης Gabor, το ψ είναι η μετατόπιση φάσης, σ είναι η τυπική απόκλιση του Gaussian φακέλου και γ είναι η αναλογία χωρικών διαστάσεων και καθορίζει την ελλειπτικότητα του υποστηρίγματος της λειτουργίας Gabor [12].

Ένα σύνολο φίλτρων Gabor με διαφορετικές συχνότητες και προσανατολισμούς μπορεί να είναι χρήσιμο για την εξαγωγή χρήσιμων χαρακτηριστικών από μια εικόνα. Στον διακριτό τομέα, δίδονται δισδιάστατα φίλτρα Gabor:

$$G_c[i, j] = B e^{-\frac{(i^2 + j^2)}{2\sigma^2}} \cos(2\pi f(i \cos \theta + j \sin \theta))$$

$$G_s[i, j] = C e^{-\frac{(i^2 + j^2)}{2\sigma^2}} \sin(2\pi f(i \cos \theta + j \sin \theta))$$

Όπου τα B και C είναι παράγοντες ομαλοποίησης που πρέπει να καθοριστούν. Τα φίλτρα 2-D Gabor έχουν πλούσιες εφαρμογές στην επεξεργασία εικόνας, ειδικά στην εξαγωγή χαρακτηριστικών για ανάλυση και τμηματοποίηση υφής. Το f ορίζει τη συχνότητα που αναζητάτε στην υφή. Διαφοροποιώντας το θ , μπορούμε να αναζητήσουμε υφή προσανατολισμένη προς μια συγκεκριμένη κατεύθυνση. Με τη

μεταβολή του σ , αλλάζουμε την υποστήριξη της βάσης ή του μεγέθους της περιοχής εικόνας που αναλύεται [12].

Τα φίλτρα Gabor με διαφορετικές συχνότητες και με προσανατολισμούς σε διαφορετικές κατευθύνσεις έχουν χρησιμοποιηθεί για τον εντοπισμό και την εξαγωγή περιοχών μόνο από κείμενο από σύνθετες εικόνες εγγράφων (τόσο γκρι όσο και χρώματος), καθώς το κείμενο είναι πλούσιο σε συστατικά υψηλής συχνότητας, ενώ οι εικόνες είναι σχετικά ομαλές. Έχει επίσης εφαρμοστεί για την αναγνώριση της έκφρασης του προσώπου. Τα φίλτρα Gabor έχουν επίσης χρησιμοποιηθεί ευρέως σε εφαρμογές ανάλυσης προτύπων. Ο χώρος Gabor είναι πολύ χρήσιμος σε εφαρμογές επεξεργασίας εικόνας, όπως οπτική αναγνώριση χαρακτήρων, αναγνώριση ίριδας και αναγνώριση δακτυλικών αποτυπωμάτων. Οι σχέσεις μεταξύ ενεργοποιήσεων για μια συγκεκριμένη χωρική τοποθεσία είναι πολύ διακριτές μεταξύ αντικειμένων σε μια εικόνα. Επιπλέον, σημαντικές ενεργοποιήσεις μπορούν να εξαχθούν από τον χώρο του Gabor προκειμένου να δημιουργηθεί μια αραιή αντιπροσώπευση αντικειμένων [12].

2.2.2.2 Tamura Descriptor

Το φίλτρο Tamura χρησιμοποιεί ένα σύνολο έξι οπτικών χαρακτηριστικών. Αυτά τα χαρακτηριστικά είναι [13]:

- ακαμψία (*coarseness*)
- αντίθεση (*contrast*)
- κατευθυντικότητα (*directionality*)
- ομοιότητα γραμμής (*linelikeness*)
- κανονικότητα (*regularity*)
- τραχύτητα (*roughness*)

Η ακαμψία σχετίζεται με τις αποστάσεις των αξιοσημείωτων χωρικών μεταβολών των επιπέδων των γκρίζων στο μέγεθος των πρωτόγονων στοιχείων που σχηματίζουν την υφή [13].

Η αντίθεση μετρά τον τρόπο με τον οποίο τα επίπεδα γκρίζου διαφέρουν στην εικόνα και σε ποιο βαθμό η κατανομή τους είναι προκατειλημμένη στο μαύρο ή το λευκό [13].

Ο βαθμός κατευθυντικότητας μετριέται με τη χρήση της κατανομής συχνοτήτων προσανατολισμένων τοπικών ακμών έναντι των κατευθυντικών γωνιών τους. Η δύναμη της ακμής $e(x, y)$ και η γωνία κατεύθυνσης $a(x, y)$ υπολογίζονται χρησιμοποιώντας τον ανιχνευτή άκρου Sobel που προσεγγίζει τα x και y της εικόνας [13,43]:

$$e(x, y) = 0.5(|\Delta_x(x, y)| + |\Delta_y(x, y)|)$$

$$\alpha(x, y) = \tan^{-1}(\Delta_y(x, y)/\Delta_x(x, y))$$

Όπου $\Delta_x(x, y)$ και $\Delta_y(x, y)$ είναι οι οριζόντιες και κάθετες διαφορές στάθμης γκρίζου μεταξύ των γειτονικών εικονοστοιχείων, αντίστοιχα. Οι διαφορές μετρούνται χρησιμοποιώντας τους ακόλουθους χειριστές 3×3 κινούμενων παραθύρων [13]:

$$\begin{array}{ccc} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{array} \quad \begin{array}{ccc} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{array}$$

Ένα ιστόγραμμα $H_{dir}(\alpha)$ των τιμών καθοριζόμενης κατεύθυνσης α κατασκευάζεται μετρώντας αριθμούς των εικονοστοιχείων ακμής με τις αντίστοιχες κατευθυντικές γωνίες και την αντοχή ακμής μεγαλύτερη από ένα προκαθορισμένο κατώφλι. Το ιστόγραμμα είναι σχετικά ομοιόμορφο για εικόνες χωρίς ισχυρό προσανατολισμό και παρουσιάζει κορυφές για εικόνες με μεγάλη κατεύθυνση. Ο βαθμός κατεύθυνσης σχετίζεται με την οξύτητα των κορυφών [13]:

$$F_{dir} = 1 - r n_{peaks} \sum_{p=1}^{n_{peaks}} \sum_{\alpha \in w_p} (\alpha - \alpha_p)^2 H_{dir}(\alpha)$$

Όπου n_p είναι ο αριθμός κορυφών, α_p είναι η θέση της p κορυφής, w_p είναι το εύρος των γωνιών που αποδίδονται στην p κορυφή (δηλαδή, το εύρος μεταξύ κοιλάδων γύρω από την κορυφή), r δηλώνει έναν συντελεστή ομαλοποίησης που σχετίζεται με την

ποσοτικοποίηση των επιπέδων των γωνιών α , και α είναι η ποσοτικοποιημένη κατευθυντική γωνία (κυκλικά σε μορφή 180°) [13].

Τα τρία άλλα χαρακτηριστικά είναι ιδιαίτερα συσχετισμένα με τα παραπάνω τρία χαρακτηριστικά και δεν προσθέτουν πολλά στην αποτελεσματικότητα της περιγραφής της υφής. Το χαρακτηριστικό γραμμικότητας ορίζεται ως μια μέση σύμπτωση των κατευθύνσεων των άκρων (πιο συγκεκριμένα κωδικοποιημένων κατευθυντικών γωνιών) που συνυπάρχουν στα ζεύγη των εικονοστοιχείων που διαχωρίζονται από μια απόσταση d κατά μήκος της κατεύθυνσης ακμής σε κάθε εικονοστοιχείο. Η αντοχή του άκρου αναμένεται να είναι μεγαλύτερη από ένα δεδομένο όριο εξαλείφοντας τις ασήμαντες "αδύναμες" άκρες. Η σύμπτωση μετράται από το συνημίτονο της διαφοράς μεταξύ των γωνιών, έτσι ώστε οι συν-εμφανίσεις στην ίδια κατεύθυνση να μετριοούνται με +1 και εκείνες στις κάθετες κατευθύνσεις κατά -1 [13].

Η συνάρτηση κανονικότητας ορίζεται ως

$$F_{reg} = 1 - r(s_{crs} + s_{con} + s_{dir} + s_{lin})$$

όπου r είναι συντελεστής ομαλοποίησης και κάθε s σημαίνει την τυπική απόκλιση του αντίστοιχου χαρακτηριστικού F σε κάθε υποεικόνα η υφή είναι χωρισμένη. Το χαρακτηριστικό τραχύτητα δίνεται απλά με το άθροισμα των μετρήσεων ακαμψίας και αντίθεσης [13].

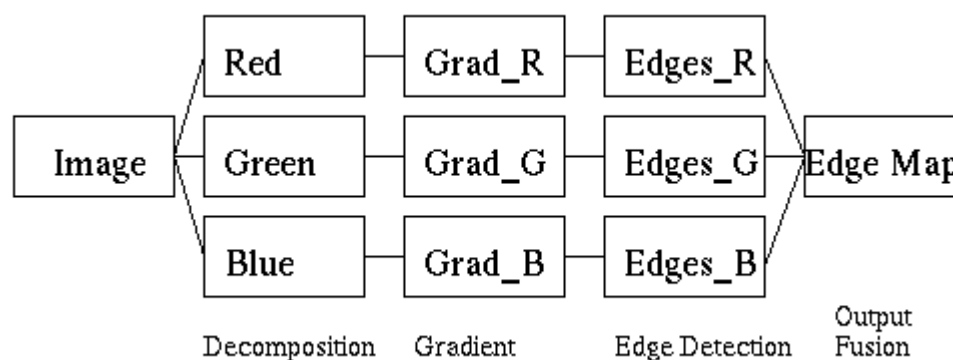
$$F_{rgh} = F_{crs} + F_{con}$$

2.2.2.3 Color and Edge Directivity Descriptor (CEDD)

Το φίλτρο CEDD ενσωματώνει το χρώμα και την υφή σε ένα ιστόγραμμα. Το μέγεθος του περιορίζεται σε 54 bytes ανά εικόνα, κάτι που κάνει αυτό το φίλτρο κατάλληλο για

μεγάλες βάσεις δεδομένων για εικόνες. Ένα από τα πιο σημαντικά χαρακτηριστικά του CEDD είναι η χαμηλή ενέργεια που χρειάζεται για την εξαγωγή του, σε σύγκριση με τους περισσότερους MPEG-7 descriptors [14].

Ο στόχος είναι να πραγματοποιηθεί η ανίχνευση ακμών τρεις φορές, μία φορά για κάθε κόκκινο, πράσινο και μπλε (ή οποιουδήποτε χώρου χρώματος που χρησιμοποιείται) και στη συνέχεια η έξοδος συγχωνεύεται για να σχηματίσει έναν χάρτη ακμής, όπως φαίνεται στην Εικόνα 2.14 [14].



Εικόνα 2.14 Διαδικασία φίλτρου CEDD [14]

2.2.2.4 Fuzzy Color and Texture Histogram Descriptor (FCTH)

Το φίλτρο FCTH ενσωματώνει το χρώμα και την υφή σε ένα ιστόγραμμα. Το αποτέλεσμα αυτό βγαίνει από ένα συνδυασμό 3 ασαφή συστημάτων. Το FCTH περιορίζει το μέγεθος στα 72 bytes ανά εικόνα, καθιστώντας αυτό το φίλτρο κατάλληλο για τη χρήση σε βάσεις με μεγάλες εικόνες. Αυτό το φίλτρο είναι κατάλαλο για την ακριβή ανάκτηση εικόνων ακόμα και σε περιπτώσεις στρέβλωσης, όπως παραμορφώσεις, θορύβου και εξομάλυνσης [15].

Τα χαρακτηριστικά FCTH μπορούν να ληφθούν από το συνδυασμό 3 ασαφή μονάδων. Για να διαιρέσει την εικόνα σε διάφορα μπλοκ, στην πρώτη ασαφή μονάδα με τρία

κανάλια χώρου χρώματος HSV ως είσοδο, το ιστόγραμμα του ασαφούς συστήματος παράγει 10 δοχεία. Στη δεύτερη ασαφής μονάδα για να τροποποιήσει κάθε τόνο χρώματος, το ασαφές σύστημα 24 δοχείων παράγει το τελικό ιστόγραμμα. Στη τρίτη ασαφή μονάδα, η εικόνα χωρίζεται σε μπλοκ με το μετασχηματισμό κυμάτων Haar σε ένα σύνολο στοιχείων υφής και το ασαφές σύστημα στα στοιχεία υφής που έχουν οριστεί ως είσοδος, το ιστόγραμμα 24 κάδων μετατρέπεται σε ιστογράμματα 192 κάδων. Τέλος, πραγματοποιείτε ο υπολογισμός του FCTH [15].

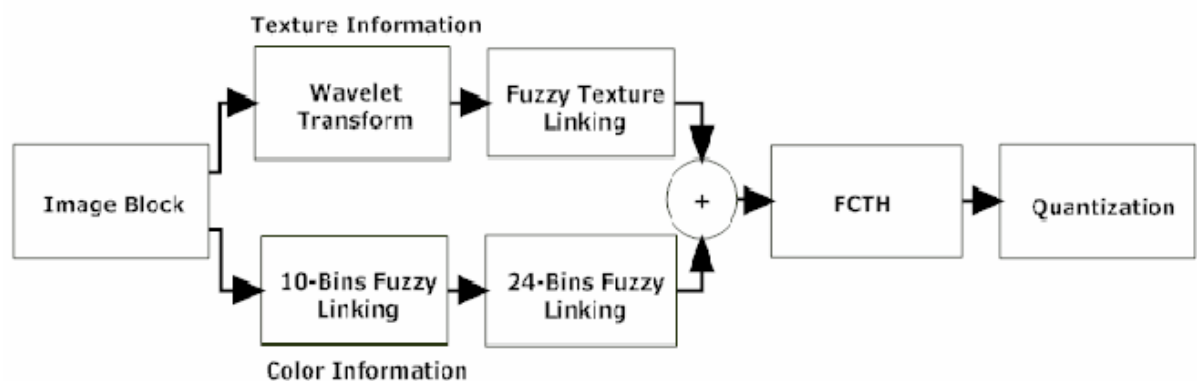
Για το ιστόγραμμα ασαφούς υφής της διαδικασίας επίλυσης, η εικόνα χωρίζεται σε μπλοκ με μετασχηματισμό κυμάτων Haar [44] που έλαβε τρία στοιχεία υφής F_{LH} , F_{HL} και F_{HH} . Τα τρία στοιχεία της υφής ως είσοδος των ασαφών συστημάτων δημιουργούν ένα ιστόγραμμα 8 κάδων, οι 8 κάδοι που αντιστοιχούν στην περιοχή αντίστοιχα είναι [15]:

- Χαμηλής ενέργειας γραμμική περιοχή
- Χαμηλής ενέργειας οριζόντια ενεργοποιημένη περιοχή
- Χαμηλής ενέργειας κάθετη ενεργοποιημένη περιοχή
- Χαμηλή ενέργειας οριζόντια και κάθετη ενεργοποιημένη περιοχή
- Υψηλής ενέργειας γραμμική περιοχή
- Υψηλής ενέργειας οριζόντια ενεργοποιημένη περιοχή
- Υψηλής ενέργειας κάθετη ενεργοποιημένη περιοχή
- Υψηλής ενέργειας οριζόντια και κάθετη ενεργοποιημένη περιοχή

Η διαμόρφωση της FCTH επιλύεται ως εξής:

Το ιστόγραμμα αποτελείται από 8 περιοχές, οι οποίες καθορίζονται από το ασαφές σύστημα που λαμβάνει αποφάσεις σχετικά με την υφή της εικόνας. Κάθε περιοχή αποτελείται από 24 μεμονωμένες περιοχές, όπως αυτές προκύπτουν από το δεύτερο ασαφές σύστημα. Συνολικά, η έξοδος που προκύπτει περιλαμβάνει $8 \times 24 = 192$ κάδους. Με βάση το περιεχόμενο των κάδων, παράγεται το αντίστοιχο τελικό ιστόγραμμα [15].

Για να διαμορφώσουμε το ιστόγραμμα, πρώτα αποστέλλουμε την εικόνα σε 1600 μπλοκ. Αυτός ο αριθμός μπλοκ επιλέχθηκε ως συμβιβασμός μεταξύ της λεπτομέρειας εικόνας και της υπολογιστικής ζήτησης. Κάθε μπλοκ περνά διαδοχικά από όλα τα ασαφή συστήματα. Αν ορίζουμε τον κάδο που προκύπτει από το ασαφές σύστημα ανίχνευσης υφής ως N και ως M ο κάδος που προκύπτει από το ασαφές σύστημα που σχηματίζει ιστόγραμμα χρωματικής σύνδεσης με ασαφή 24 κάδους, τότε κάθε μπλοκ τοποθετείται στη θέση του δοχείου: $N \times 24 + M$. Η Εικόνα 2.15 απεικονίζει ολόκληρη τη διαδικασία, η οποία περιγράφεται στη συνέχεια [15].



Εικόνα 2.15 Διαδικασία φίλτρου FCTH [15]

Κάθε μπλοκ μεταφέρεται στο χώρο χρώματος YIQ και μετασχηματίζεται με το μετασχηματισμό Haar Wavelet. Οι τιμές F_{LH} , F_{HL} και F_{HH} υπολογίζονται και μέσω του τρίτου ασαφούς συστήματος αυτό το μπλοκ κατατάσσεται σε έναν από τους 8 κάδους εξόδου. Υποθέστε, για παράδειγμα, ότι η ταξινόμηση αποδίδει αυτό το μπλοκ στον δεύτερο κάδο που ορίζει την οριζόντια δράση χαμηλής ενέργειας. Στη συνέχεια, το ίδιο μπλοκ μεταφέρεται στο χώρο χρώματος HSV και υπολογίζονται οι μέσες τιμές μπλοκ H , S και V . Αυτές οι τιμές αποτελούν εισροές στο ασαφές σύστημα που διαμορφώνει το ιστόγραμμα ασαφούς 10-δοχείων. Ας υποθέσουμε ξανά ότι η ταξινόμηση κατατάσσει αυτό το μπλοκ στον τέταρτο κάδο που ορίζει ότι το χρώμα είναι κόκκινο. Στη συνέχεια, το δεύτερο ασαφές σύστημα, χρησιμοποιώντας τις μέσες τιμές των S και V καθώς και την τιμή του κάδου (ή κάδους) που προκύπτει από το σύστημα ασαφούς σύνδεσης 10 κόμβων, υπολογίζει την απόχρωση του χρώματος και διαμορφώνει τα ασαφή ιστογράμματα 24 κάδων. Ας υποθέσουμε ξανά ότι το σύστημα ταξινομεί αυτό το μπλοκ

στον τέταρτο κάδο που ορίζει ότι το χρώμα είναι το σκούρο κόκκινο. Ο συνδυασμός των 3 ασαφών συστημάτων θα ταξινομήσει τελικά το μπλοκ στον 27ο κάδο ($1 \times 24 + 3$). Η διαδικασία επαναλαμβάνεται για όλα τα μπλοκ της εικόνας. Κατά την ολοκλήρωση της διαδικασίας, το ιστόγραμμα κοινωνικοποιείται στο διάστημα $\{0-1\}$ και κβαντισμένο σε 3 bits / κάδο [15].

2.2.2.5 Auto Color Correlogram Descriptor (ACCD)

Το φίλτρο Auto Color Correlogram χρησιμοποιείται για την εύρεση και σύγκριση εικόνων. Αυτό το φίλτρο συγκρίνει τη χωρική συσχέτιση των χρωμάτων, και είναι αποτελεσματικό και χωρίς μεγάλο κόστος για αυτόματη ανάκτηση εικόνων. Το φίλτρο αυτό βρίσκει μεγάλες αλλαγές στην εμφάνιση των σχημάτων με την αλλαγή της θέσης που βλέπουμε το αντικείμενο, από μεγέθυνση της κάμερας κ.α. Πειραματικά έχει αποδειχτεί ότι αυτό το φίλτρο είναι καλύτερο από το ιστόγραμμα χρωμάτων και του ιστογράμματος εύρεσης- ανάκτησης εικόνων [16].

2.2.2.6 Color Layout Descriptor (CLD)

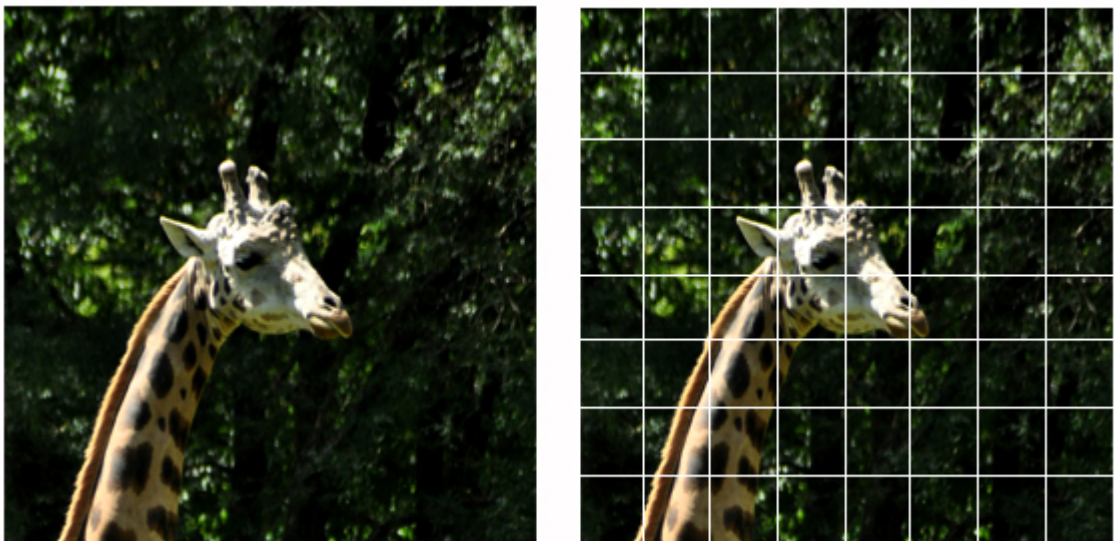
Το φίλτρο Color Layout υπολογίζει τη χωρική κατανομή του χρώματος σε μια εικόνα. Η διαδικασία εξαγωγής χαρακτηριστικών αποτελείται από δύο μέρη, με βάση πλέγμα αντιπροσωπευτικής επιλογής χρώματος και διακριτό μετασχηματισμό συνημίτονου με κβαντισμό [17].

Το χρώμα είναι η πιο βασική ποιότητα του οπτικού περιεχομένου, επομένως είναι δυνατή η χρήση χρωμάτων για την περιγραφή και την αναπαραγωγή μιας εικόνας. Το πρότυπο MPEG-7 έχει δοκιμάσει την πιο αποτελεσματική διαδικασία για να περιγράψει το χρώμα και έχει επιλέξει εκείνες που έχουν δώσει πιο ικανοποιητικά αποτελέσματα. Αυτό το πρότυπο προτείνει διαφορετικές μεθόδους για την απόκτηση αυτών των περιγραφών και ένα εργαλείο που ορίζεται για να περιγράψει το χρώμα είναι το CLD, το οποίο επιτρέπει την περιγραφή της σχέσης χρώματος μεταξύ ακολουθιών ή ομάδων εικόνων [17].

Το CLD καταγράφει τη χωρική διάταξη των αντιπροσωπευτικών χρωμάτων σε ένα πλέγμα που τοποθετείται πάνω σε μια περιοχή ή εικόνα. Η αναπαράσταση βασίζεται σε συντελεστές του διακριτού μετασχηματισμού συνημιτόνου (DCT). Αυτός είναι ένας πολύ συμπαγής περιγραφέας που είναι ιδιαίτερα αποδοτικός στις εφαρμογές γρήγορης περιήγησης και αναζήτησης. Μπορεί να εφαρμοστεί σε ακίνητες εικόνες καθώς και σε τμήματα βίντεο [17].

Η διαδικασία εξαγωγής αυτού του περιγραφέα χρώματος αποτελείται από τέσσερα στάδια [17]:

- Διαίρεση εικόνας (Εικόνα 2.16)
- Αντιπροσωπευτική επιλογή χρώματος (Εικόνα 2.17)
- Μετασχηματισμός DCT
- Σάρωση ζιγκ-ζαγκ



Εικόνα 2.16 Διαίρεση εικόνας φίλτρου CLD [17]



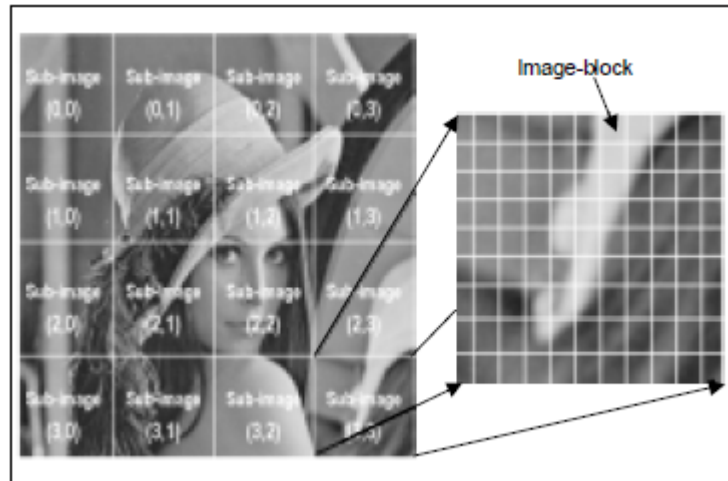
Εικόνα 2.17 Αντιπροσωπευτική επιλογή χρώματος φίλτρου CLD [17]

2.2.2.7 Scalable Color Descriptor (SCD)

Το φίλτρο scalable color είναι ένα χρωματικό ιστόγραμμα που εξάγεται στο HSV χώρο χρώματος, και κωδικοποιήτε για αποδοτικότητα αποθήκευσης. Η εξαγωγή του περιγραφικού στοιχείου ξεκινά με τον υπολογισμό του χρώματος Ιστόγραμμα με 256 κάδους στο χώρο χρώματος HSV με απόχρωση (H) συστατικό κβαντισμένο σε 16 κάδους, και κορεσμό (S) και τιμή (V) κβαντισμένο σε 4 κάδους έκαστο. Αυτή η αρχική έκδοση διαβιβάζεται μέσω μιας σειράς μετασχηματισμών 1-D Haar, ξεκινώντας από τον άξονα H, ακολουθούμενη από S, V. Το αποτέλεσμα είναι ένα σύνολο από 16 συντελεστές χαμηλής διέλευσης και πάνω από 240 συντελεστές υψηλής διέλευσης. Λόγω της πλεονασμού του αρχικού ιστόγραμμου, οι συντελεστές υψηλής διέλευσης τείνουν να έχουν χαμηλές (θετικές και αρνητικές) τιμές. Τυπική εφαρμογή του περιγραφικού πεδίου περιλαμβάνει αναζήτηση ομοιότητας σε βάσεις δεδομένων πολυμέσων και περιήγηση μεγάλων βάσεων δεδομένων. Το κλιμακωτό χρώμα αποτελεί επίσης τη βάση της ομάδας των πλαισίων / εικόνων περιγραφής χρωμάτων [18].

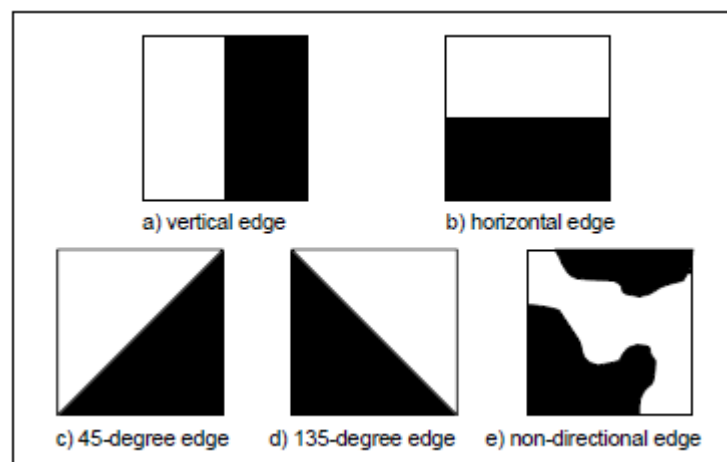
2.2.2.8 Edge Histogram Descriptor (EHD)

Το φίλτρο Edge Histogram αντιπροσωπεύει βασικά τη διανομή 5 τύπων ακμών σε κάθε τοπική περιοχή που ονομάζεται υπο-εικόνα [19].



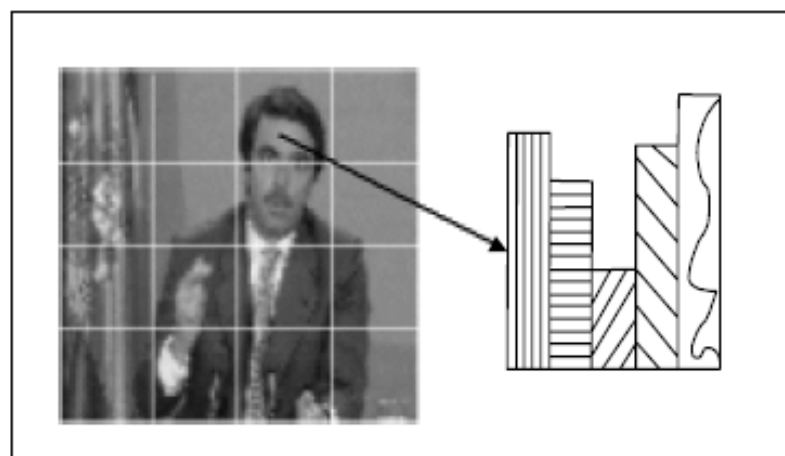
Εικόνα 2.18 Διαχωρισμός του χώρου για το φίλτρο EHD [19]

Όπως φαίνεται στην Εικόνα 2.18, η δευτερεύουσα εικόνα ορίζεται από το διαχωρισμό του χώρου εικόνας σε 4x4 μη επικαλυπτόμενα μπλοκ. Έτσι, το διαμέρισμα εικόνας αποδίδει πάντα 16 υπο-εικόνες ίσου μεγέθους ανεξάρτητα από το μέγεθος της αρχικής εικόνας. Για να χαρακτηρίσουμε την υπο-εικόνα, τότε παράγουμε ένα ιστόγραμμα κατανομής ακμών για κάθε υπο-εικόνα. Οι ακμές στις υπο-εικόνες κατηγοριοποιούνται σε 5 τύπους: κάθετη, οριζόντια, διαγώνια 45 μοιρών, διαγώνια 135 μοιρών και μη κατευθυντικές άκρες (Εικόνα 2.19) [19].



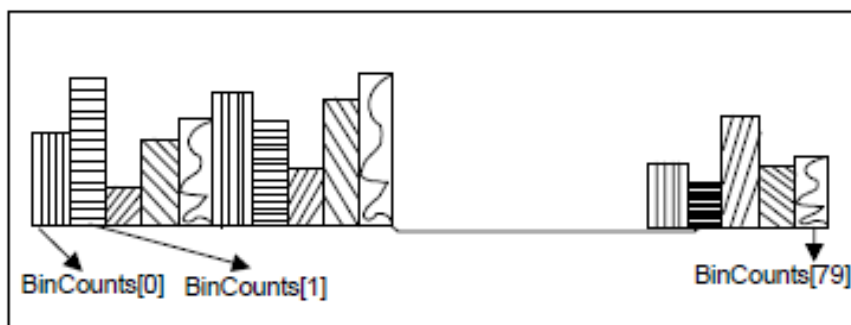
Εικόνα 2.19 Κατηγορίες ακμών [19]

Έτσι, το ιστόγραμμα για κάθε δευτερεύουσα εικόνα αντιπροσωπεύει τη σχετική συχνότητα εμφάνισης των 5 τύπων ακμών στην αντίστοιχη δευτερεύουσα εικόνα. Ως αποτέλεσμα, όπως φαίνεται στην Εικόνα 2.20, κάθε τοπικό ιστόγραμμα περιέχει 5 κάδους [19].



Εικόνα 2.20 Τοπικό ιστόγραμμα ακμών [19]

Κάθε κάδος αντιστοιχεί σε έναν από τους 5 τύπους άκρων. Δεδομένου ότι υπάρχουν 16 υπο-εικόνες στην εικόνα, απαιτούνται συνολικά $5 \times 16 = 80$ δοχεία ιστογράμμων (Εικόνα 2.21) [19].



Εικόνα 2.21 Συνολικός αριθμός κάδων ακμών [19]

Σημειώστε ότι καθένας από τους κάδους των 80 ιστογραμμάτων έχει τη δική του σημασιολογία ως προς την τοποθεσία και τον τύπο ακμής. Για παράδειγμα, ο κάδος για το άκρο οριζόντιου τύπου στην υπο-εικόνα που βρίσκεται στο (0,0) στην Εικόνα 2.18 μεταφέρει τις πληροφορίες του σχετικού πληθυσμού των οριζόντιων άκρων στην πάνω αριστερή τοπική περιοχή της εικόνας [19].

2.2.2.9 Color Histogram Descriptor (CHD)

Το φίλτρο Color Histogram περιγράφει την αναλογία διαφορετικών χρωμάτων στην αναλογία ολόκληρης της εικόνας και δεν ενδιαφέρεται για το ποια είναι η χωρική θέση κάθε χρώματος, δηλαδή δεν περιγράφει το αντικείμενο ή αντικείμενα στην εικόνα. Το ιστόγραμμα χρώματος είναι ιδιαίτερα κατάλληλο για τις εικόνες που είναι δύσκολο να περιγράψουν την αυτόματη κατάτμηση. Το χρωματικό ιστόγραμμα βασίζεται στον τύπο του χρωματικού χώρου και στο σύστημα συντεταγμένων. Ο συνηθέστερος χρωματικός χώρος είναι ένας χρωματικός χώρος RGB, επειδή εκφράζεται η πλειοψηφία της ψηφιακής εικόνας σε αυτόν τον χρωματικό χώρο. Είναι απαραίτητο να υπολογιστεί το χρωματικό ιστόγραμμα. Ο χρωματικός χώρος διαιρείται σε αρκετά μικρά διαστήματα

χρώματος μεταξύ κάθε κελιού σε ένα κάδο του ιστογράμματος. Αυτή η διαδικασία ονομάζεται χρωματική κβαντισμοποίηση. Στη συνέχεια, υπολογίζοντας τον αριθμό των εικονοστοιχείων μεταξύ των χρωμάτων κάθε κελιού εμπίπτει στο χρωματικό ιστόγραμμα [20].

2.2.3 Επιλεγμένα Φίλτρα

Με βάση τις πληροφορίες που έχουν αναφερθεί πιο πάνω επιλέχθηκαν τα δυο φίλτρα που κρίθηκαν πιο κατάλληλα και με βάση τα οποία θα γίνουν οι περιλήψεις. Τα φίλτρα που επιλέχθηκαν είναι το Gabor και το Tamura, τα οποία αποτελούν και τα δυο φίλτρα υφής. Παρόλο που είναι και τα δύο φίλτρα υφής, λαμβάνουν υπόψη τους διαφορετικές παραμέτρους για την αξιολόγηση του σκορ για κάθε εικόνα, με αποτέλεσμα στις δυο αυτές περιπτώσεις να παίρνουμε διαφορετικά αποτελέσματα σε κάθε εκτέλεση.

Ο λόγος που έχει επιλεγεί το φίλτρο Gabor είναι η αξιολόγηση των εικόνων με βάση την συχνότητα που υπάρχει σε αυτές. Έτσι, για στιγμιότυπο που υπάρχει εργαλείο η συχνότητα θα είναι χαμηλότερη σε σχέση με τα στιγμιότυπα που δεν υπάρχει εργαλείο. Από την άλλη όμως, η επιλογή θα πρέπει να γίνει με προσοχή, γιατί στις περιπτώσεις που το εργαλείο είναι αρκετά κοντά στη κάμερα η συχνότητα θα είναι πιο μικρή. Για αυτό το στιγμιότυπο θα είναι στα σημαντικά, αλλά το γεγονός αυτό είναι λάθος γιατί το πρόγραμμα θα πρέπει να βρίσκει τα στιγμιότυπα που εκτελείται κάποια λειτουργία από το εργαλείο.

Ο λόγος που έχει επιλεγεί το φίλτρο Tamura είναι γιατί το φίλτρο αυτό χρησιμοποιεί τα διάφορα στοιχεία της υφής του στιγμιότυπου που έχουμε αναφέρει με αποτέλεσμα να είναι πιο εύκολη η αναγνώριση περιοχών με εργαλεία σε αυτές.

Τέλος, έχουν επιλεγεί αυτά τα δυο φίλτρα υφής γιατί έχει υλοποιηθεί και ένας δεύτερος αλγόριθμος, ο οποίος έχει προσαρμοστεί για να βρίσκει τα ιατρικά εργαλεία που υπάρχουν στο βίντεο εισόδου με βάση τα χρώματα των ιατρικών εργαλείων αυτών που υπάρχουν σε κάποιο συγκεκριμένο αρχείο.

2.2.4 Επιλογή άλλων Παραμέτρων

Για την επιλογή της συμπίεσης - μείωσης που θα έχει το βίντεο έχουμε επιλέξει 3 διαφορετικές τιμές. Θα έχουμε χαμηλή συμπίεση, δηλαδή η τιμή θα είναι 2 και το βίντεο περίληψης θα είναι το $\frac{1}{2}$ από το αρχικό, μεσαία συμπίεση, δηλαδή η τιμή θα είναι 4 και το βίντεο περίληψης θα είναι το $\frac{1}{4}$ από το αρχικό και μεγάλη συμπίεση, δηλαδή η τιμή θα είναι 8 και το βίντεο περίληψης θα είναι το $\frac{1}{8}$ από το αρχικό.

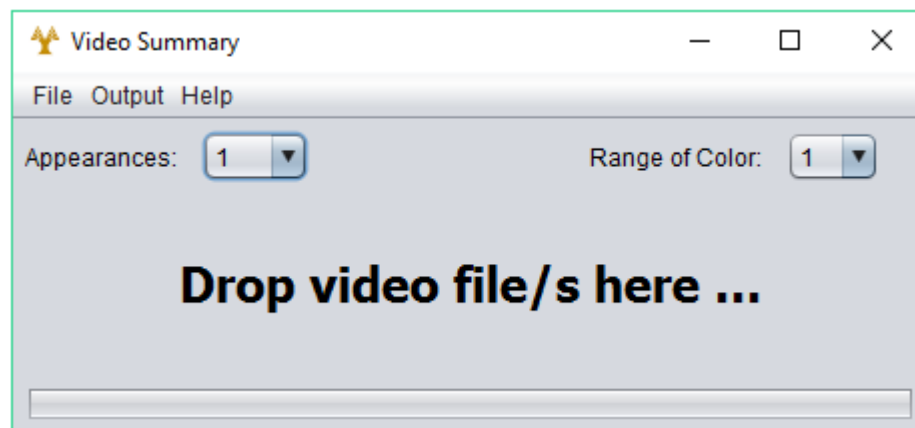
Όσο αφορά τη παράμετρο για το χρόνο που θα έχουμε για όλα τα βίντεο περίληψης τιμή για 3 δευτερόλεπτα Δεν έχει βρεθεί κάποιο άρθρο το οποίο να προτείνει τον κατάλληλο χρόνο που χρειάζεται για να παρουσιαστεί ένα κομμάτι από βίντεο. Θα δημιουργηθούν περιλήψεις για 5 διαφορετικά βίντεο με συνδυασμούς των παραμέτρων που έχουν αναφερθεί και πιο πάνω.

2.3 Αλγόριθμος με χρώματα

Η αρχική ιδέα για την δημιουργία αυτού του νέου αλγόριθμου για περίληψη ήταν να γίνεται η επιλογή με βάση τα χρώματα των εργαλείων. Αυτά έχουν αρκετά διαφορετικό χρώμα από ότι άλλο υπάρχει στο ανθρώπινο σώμα. Έτσι, για να επιτευχθεί η αναγνώριση των χρωμάτων χρησιμοποιήθηκε το RGB μοντέλο χρωμάτων .

2.3.1 Περιγραφή Αλγορίθμου

Το πρόγραμμα αυτό παίρνει ως δεδομένα εισόδου από το χρήστη το βίντεο του οποίου θέλουμε να γίνει η περίληψη, τον αριθμό των φορών που πρέπει να βρεθεί το χρώμα ενός εργαλείου σε ένα στιγμιότυπο και ένα αριθμό για την απόσταση (range) που μπορεί να αποκλίνει το χρώμα που έχει βρεθεί στο στιγμιότυπο σε σχέση με αυτό του εργαλείου. Το μενού αυτό είναι όπως φαίνεται στην εικόνα πιο κάτω.



Εικόνα 2.22 Μενού αλγορίθμου με χρώματα

Τα χρώματα των εργαλείων διαβάζονται από ένα συγκεκριμένο αρχείο στο οποίο ο χρήστης μπορεί να προσθέσει ή και να αφαιρέσει τα εργαλεία που θέλει. Στο αρχείο αυτό υπάρχουν σε κάθε γραμμή τα χαρακτηριστικά του χρώματος του κάθε εργαλείου, σε μορφή RGB. Η πρώτη στήλη είναι για το κόκκινο χρώμα, η δεύτερη για πράσινο και η τρίτη για το μπλε.

Αφού, έχουν διαβαστεί από το αρχείο τα εργαλεία που μπορεί να βρεθούν στο βίντεο και οι παράμετροι για τον αριθμό εμφανίσεων και την απόσταση από το αρχικό μενού, τότε το πρόγραμμα διαβάζει το βίντεο και σε κάθε εικόνα ψάχνει pixel - pixel και με βάση το χρώμα του αυξάνει ανάλογα τον αριθμό εμφανίσεων για το κάθε ένα από τα πιθανά εργαλεία που μπορεί να βρεθεί. Πιο συγκεκριμένα, υπάρχει ένας πίνακας στον οποίο υπάρχει το άθροισμα των φορών που έχει βρεθεί κάποιο χρώμα σε αυτή. Αν κάποιο από τα χρώματα φτάσει στον αριθμό που έχει εισάγει στο αρχικό μενού ο χρήστης σαν αριθμό εμφανίσεων τότε το πρόγραμμα εξάγει την εικόνα με μορφή jpeg σε ένα συγκεκριμένο φάκελο. Όταν διαβαστεί όλο το βίντεο και έχουν αποθηκευτεί όλες οι εικόνες στις οποίες έχει βρεθεί το χρώμα από τα εργαλεία, τότε δημιουργείται με βάση αυτές τις εικόνες, το βίντεο περίληψης.

Κεφάλαιο 3

Ανάλυση

3.1 Ανάλυση	37
3.2 Παράμετροι	37
3.3 Αξιολόγηση	39

3.1 Ανάλυση

Με βάση την εφαρμογή των φίλτρων θα εξαχθούν κάποια βίντεο. Από αυτά, θα διαπιστωθεί αν τα αποτελέσματα με την αλλαγή διάφορων παραμέτρων είναι αυτά που πρέπει να εξάγει το πρόγραμμα, δηλαδή είναι η περίληψη με τα σημαντικά σημεία του βίντεο.

3.2 Παράμετροι

Για την εξαγωγή των βίντεο περίληψης χρησιμοποιήθηκαν πέντε διαφορετικά βίντεο με διαφορετικές παραμέτρους κάθε φορά, ώστε να δημιουργηθεί μια διαφορετική περίληψη σε κάθε εκτέλεση. Οι παράμετροι που αλλάζονται είναι το βίντεο εισόδου, το φίλτρο επεξεργασίας και το μέγεθος της μείωσης του βίντεο, δηλαδή το μέγεθος της συμπίεσης. Η επιλογή για το χρόνο σε δευτερόλεπτα που θα έχει κάθε κομμάτι, που θεωρείται σημαντικό από το κάθε φίλτρο, θα είναι πάντα 3 δευτερόλεπτα.

Όνομα Βίντεο	Μέγεθος Συμπίεσης	Φίλτρο
1	2	Gabor
1	4	Gabor
1	8	Gabor
1	2	Tamura
1	4	Tamura
1	8	Tamura
2	4	Gabor
2	8	Gabor
2	4	Tamura
2	8	Tamura
Clip Mioma	2	Gabor
Clip Mioma	4	Gabor
Clip Mioma	8	Gabor
Clip Mioma	2	Tamura
Clip Mioma	4	Tamura
Clip Mioma	8	Tamura
Gs_RC_23-02-2006_clip_0004	2	Gabor
Gs_RC_23-02-2006_clip_0004	4	Gabor
Gs_RC_23-02-2006_clip_0004	8	Gabor

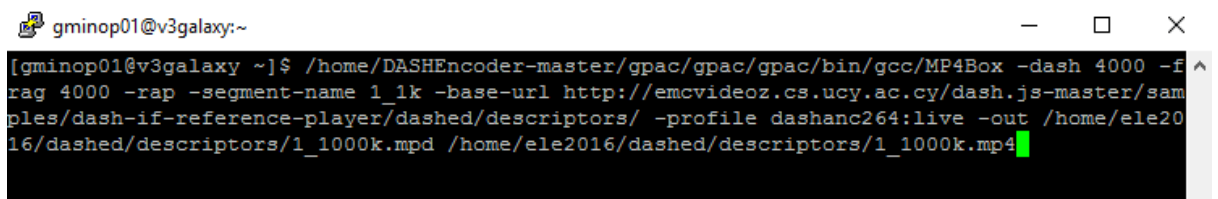
Gs_RC_23-02-2006_clip_0004	2	Tamura
Gs_RC_23-02-2006_clip_0004	4	Tamura
Gs_RC_23-02-2006_clip_0004	8	Tamura
De La Cruz MIOMA	8	Gabor
De La Cruz MIOMA	8	Tamura

Πίνακας 3.1 Παράμετροι εκτέλεσης

Ο Πίνακας 3.1 παρουσιάζει της παραμέτρους που έχουν χρησιμοποιηθεί. Στο πίνακα φαίνεται το όνομα του βίντεο, το μέγεθος και το φίλτρο που έχουμε χρησιμοποιήσει κάθε φορά για τη συμπίεση.

3.3 Αξιολόγηση

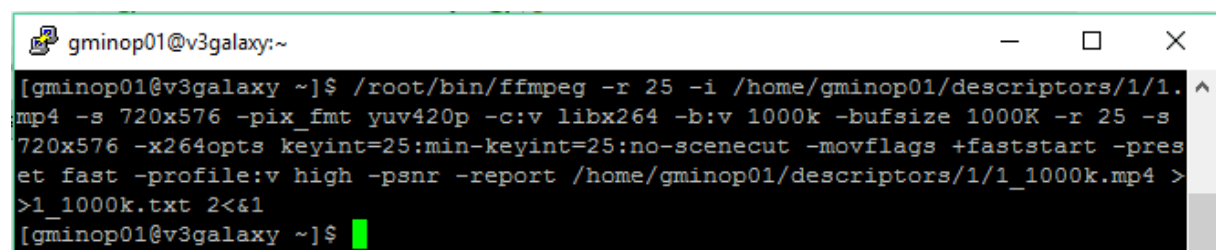
Η αξιολόγηση των βίντεο του πίνακα θα γίνει με τη βοήθεια του εργαλείου MPEG - DASH το οποίο μπορεί να μεταδώσει ένα βίντεο με τη χρήση του πρωτοκόλλου HTTP [25]. Αυτό γίνεται με την αποθήκευση των βίντεο σε έναν HTTP server και μεταδίδονται μέσω HTTP. Αυτό, επιτυγχάνεται με την αποθήκευση του περιεχόμενου αυτού σε δύο μέρη. Το πρώτο μέρος είναι το Media Presentation Description (MPD) το οποίο περιγράφει το διαθέσιμο περιεχόμενο, τις εναλλακτικές λύσεις για το βίντεο που θα προβληθεί, τις διευθύνσεις που βρίσκεται το βίντεο όπως και διάφορα άλλα χαρακτηριστικά [26]. Το δεύτερο μέρος αφορά τα segments του βίντεο τα οποία αποτελούν τα multimedia bitstreams σε μορφή κομματιών και μπορεί να είναι μόνα τους ή σε πολλά αρχεία [26]. Αυτό, γίνεται με την εκτέλεση της εντολής που βλέπουμε στην Εικόνα 3.1 για κάθε ένα από τα βίντεο [21].



```
gminop01@v3galaxy:~$ /home/DASHEncoder-master/gpac/gpac/gpac/bin/gcc/MP4Box -dash 4000 -frag 4000 -rap -segment-name 1_1k -base-url http://emcvideoz.cs.ucy.ac.cy/dash.js-master/samples/dash-if-reference-player/dashed/descriptors/ -profile dashanc264:live -out /home/ele2016/dashed/descriptors/1_1000k.mpd /home/ele2016/dashed/descriptors/1_1000k.mpd
```

Εικόνα 3.1 Εντολή διαχωρισμού βίντεο σε κομμάτια

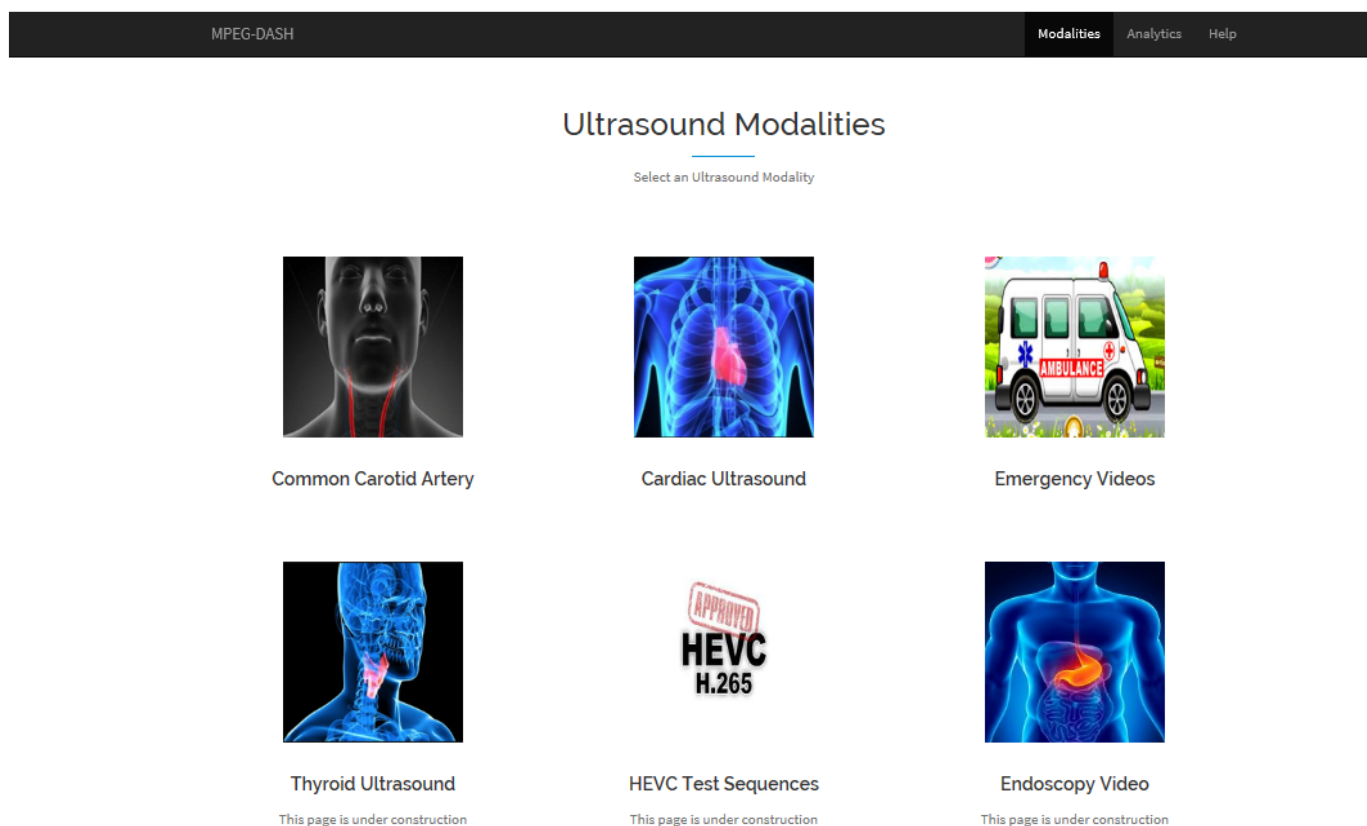
Για να γίνει όμως αυτό, τα βίντεο θα πρέπει να είναι σε κατάλληλη κωδικοποίηση, δηλαδή motion jpeg video. Για να γίνει αυτό, εκτελέσθηκε η εντολή που φαίνεται στην Εικόνα 3.1 μέσω της γραμμής εντολών του συστήματος.



```
gminop01@v3galaxy:~$ /root/bin/ffmpeg -r 25 -i /home/gminop01/descriptors/1/1_1000k.mpd -s 720x576 -pix_fmt yuv420p -c:v libx264 -b:v 1000k -bufsize 1000K -r 25 -s 720x576 -x264opts keyint=25:min-keyint=25:no-scenecut -movflags +faststart -preset fast -profile:v high -psnr -report /home/gminop01/descriptors/1/1_1000k.mpd > >1_1000k.txt 2<&1
```

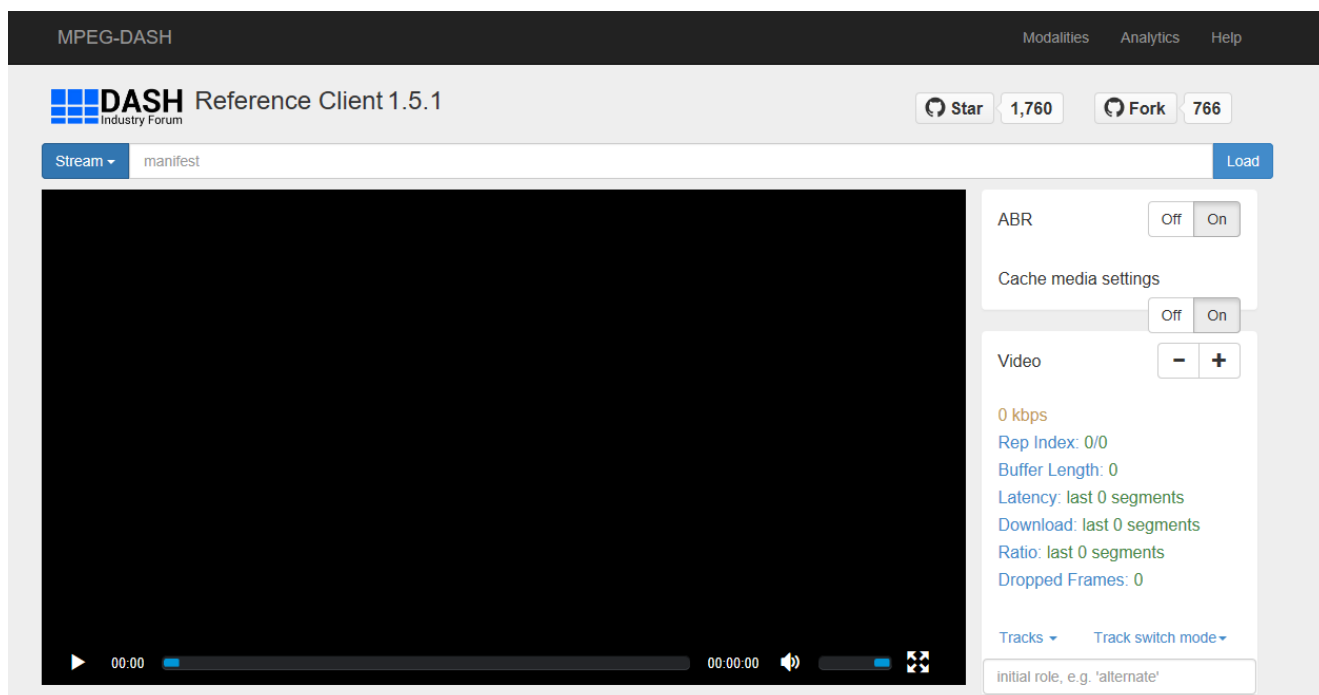
Εικόνα 3.2 Εντολή κωδικοποίησης

Για να αξιολογηθούν τα βίντεο περίληψης ο γιατρός θα πρέπει να εισέλθει στην ιστοσελίδα του πανεπιστημίου που είναι το σύστημα MPEG-DASH. Η αρχική σελίδα είναι όπως φαίνεται στην εικόνα πιο κάτω. Αφού ο χρήστης εισέλθει στην αρχική σελίδα επιλέγει το εικονίδιο κάτω δεξιά για τα βίντεο της ενδοσκοπίας (Endoscopy Video), Εικόνα 3.3.



Εικόνα 3.3 Αρχική σελίδα MPEG-DASH [27]

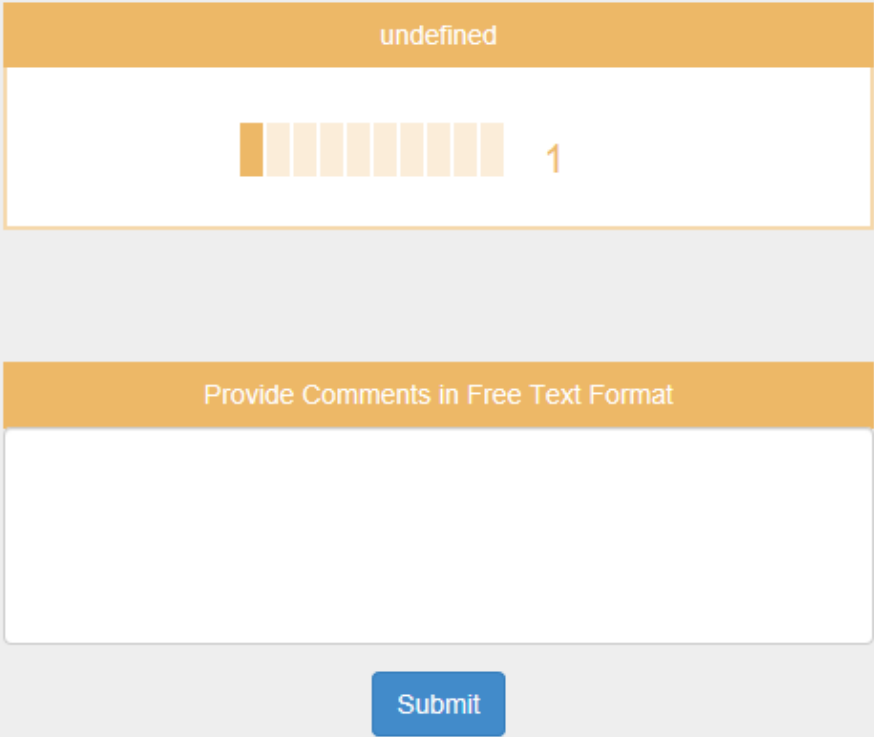
Αφού επιλέξει από το μενού για τα βίντεο ενδοσκοπικής χειρουργικής τότε στο χρήστη εμφανίζεται ένα νέο μενού, από το οποίο επιλέγει το βίντεο που θέλει όπως φαίνεται στην Εικόνα 3.4.



Εικόνα 3.4 Σελίδα αναπαραγωγής βίντεο [28]

Όπως φαίνεται και από την Εικόνα 3.4, ο χρήστης επιλέγει από το κουμπί Stream ποιο από τα βίντεο θέλει να παρακολουθήσει. Στη συνέχεια πατάει το κουμπί Load για να φορτώσει το βίντεο στο παράθυρο που υπάρχει. Μετά, ο χρήστης μπορεί να δει το βίντεο που έχει επιλέξει.

Λίγο πιο κάτω από το video player υπάρχει η επιλογή για τη βαθμολόγηση του βίντεο. Έτσι, θα βαθμολογηθούν τα βίντεο περίληψης από το γιατρό. Η βαθμολογία είναι από το 1 μέχρι το 10 και ο γιατρός θα βαθμολογήσει τα βίντεο με βάση το πόσες σημαντικές πληροφορίες υπάρχουν στο αρχικό βίντεο και στο βίντεο περίληψης που έχει δημιουργηθεί. Επιπλέον, υπάρχει η δυνατότητα να γράψει κάποιο σχόλιο για τη βαθμολογία, για παράδειγμα εάν είναι καλό το βίντεο αυτό ή τι θα μπορούσε να έχει επιπλέον από το αρχικό βίντεο. Μόλις εισάγει όλα τα σχόλια και τη βαθμολογία, πρέπει να πατήσει το κουμπί Submit για να καταχωρηθεί στο σύστημα με τις βαθμολογίες.



undefined

1

Provide Comments in Free Text Format

Submit

Εικόνα 3.5 Φόρμα αξιολόγησης βίντεο [28]

Κεφάλαιο 4

Ανάλυση Αποτελεσμάτων

4.1 Περιγραφή βίντεο	44
4.2 Περιγραφή Συστήματος	45
4.2.1 Λογισμικό Ανάπτυξης - Eclipse	45
4.2.2 Απαιτούμενες τεχνολογίες - Java	46
4.3 Μετρικές Αξιολόγησης	46
4.4 Αποτελέσματα	49

4.1 Περιγραφή βίντεο

Για τη δημιουργία των βίντεο περίληψης έχουν χρησιμοποιηθεί πέντε διαφορετικά βίντεο ενδοσκοπικής εξέτασης. Τα βίντεο αυτά έχουν τα χαρακτηριστικά που φαίνονται στο πίνακα πιο κάτω.

Όνομα Βίντεο	Διάρκεια σε δευτερόλεπτα	Ανάλυση	Frames per Second
1.wmv	40	720 x 576	25
2.wmv	324	720 x 576	25
Clip Mioma.m4v	92	960 x 540	25
De La Cruz MIOMA.m4v	455	960 x 540	25
Gs_RC_23-02-2006_clip_0004.mpeg	58	720 x 576	25

Πίνακας 4.1 Χαρακτηριστικά βίντεο

4.2 Περιγραφή συστήματος

Για την υλοποίηση των στόχων της διπλωματικής εργασίας έχουν χρησιμοποιηθεί κάποιες τεχνολογίες και εφαρμογές. Η γλώσσα προγραμματισμού που χρησιμοποιήθηκε για την ανάπτυξη του αλγορίθμου είναι η Java και το εργαλείο για αυτή είναι το Eclipse.

4.2.1 Λογισμικό Ανάπτυξης - Eclipse EE

Το περιβάλλον Eclipse αποτελεί ένα από τα κορυφαία προγράμματα υλοποίησης κώδικα για τις γλώσσες Java και C++. Το Eclipse αυτό αποτελεί το περιβάλλον που χρησιμοποιεί η πλειοψηφία των προγραμματιστών Java. Ίσως το πιο σημαντικό χαρακτηριστικό του είναι η δυνατότητα που προσφέρει στο χρήστη του για την ολοκλήρωση όλων των σταδίων του κύκλου υλοποίησης ενός προγράμματος με τη χρήση ενός μόνο περιβάλλον. Δηλαδή ο προγραμματιστής χρησιμοποιώντας αυτό το περιβάλλον μπορεί να συντάξει το κώδικα του, να το μεταγλωττίσει, να τον εκτελέσει και να ελέγξει για τα σφάλματα (debugging) και έτσι γίνεται πιο εύκολη η υλοποίηση.

Επίσης δίνει συμβουλές στο προγραμματιστή κατά τη σύνταξη, επιλογές για να διορθώσει τα λάθη του, version control, refactoring κ.α. Διατίθεται για το λειτουργικό σύστημα Windows όσο και για το Linux/Unix [22].

4.2.2 Απαιτούμενες Τεχνολογίες - Java

Η Java είναι μια αντικειμενοστραφής γλώσσα προγραμματισμού, δηλαδή υπάρχει η δυνατότητα για το σχεδιασμό ενός προγράμματος με τις πληροφορίες συγκεντρωμένες σε αντικείμενα. Αυτό, αποτελεί ίσως το πιο σημαντικό χαρακτηριστικό της γλώσσας αυτής, που την κάνει πιο χρήσιμη σε κάποιες περιπτώσεις από άλλες γλώσσες. Επιπλέον, η Java έχει και το πλεονέκτημα του garbage collector ο οποίος είναι υπεύθυνος για τη διαχείριση της μνήμης του προγράμματος, δηλαδή αποδεσμεύει από τη μνήμη τα δεδομένα που δεν χρειάζονται πλέον. Στη γλώσσα αυτή, αφού είναι πλέον πολύ διαδεδομένη, υπάρχουν πάρα πολλές βιβλιοθήκες ανοικτού κώδικα τις οποίες μπορούμε να χρησιμοποιήσουμε για τα προγράμματα που θα υλοποιηθούν [23,24].

4.3 Μετρικές Αξιολόγησης

Στο Πίνακα 4.2 παρουσιάζονται τα ποσοστά μείωσης που είχαν τα βίντεο ανάλογα με τις παραμέτρους που έχουν χρησιμοποιηθεί. Φαίνεται δηλαδή, αν η μείωση είναι η επιθυμητή, αφού σε κάθε περίπτωση, με βάση τη διάρκεια του βίντεο και το βαθμό συμπίεσης, υπολογίζεται η τελική διάρκεια του βίντεο. Με βάση αυτή τη μετρική, παρατηρείται ότι σε μεγάλο βαθμό τα βίντεο έχουν επιτύχει την επιθυμητή συμπίεση, χωρίς να υπάρχουν μεγάλες αποκλίσεις για κάποιο συγκεκριμένο φίλτρο ή κάποιο συγκεκριμένο βαθμό συμπίεσης.

Όνομα Βίντεο	Αρχική Διάρκεια Βίντεο	Μέθοδος	Συμπίεση	Επεξεργασμέ νη Διάρκεια
1	40 sec	Gabor	Χαμηλή Συμπίεση	12 sec
1	40 sec	Gabor	Μέτρια Συμπίεση	9 sec
1	40 sec	Gabor	Υψηλή Συμπίεση	3 sec
1	40 sec	Tamura	Χαμηλή Συμπίεση	17 sec
1	40 sec	Tamura	Μέτρια Συμπίεση	8 sec
1	40 sec	Tamura	Υψηλή Συμπίεση	3 sec
2	324 sec	Gabor	Μέτρια Συμπίεση	68 sec
2	324 sec	Gabor	Υψηλή Συμπίεση	33 sec
2	324 sec	Tamura	Μέτρια Συμπίεση	80 sec
2	324 sec	Tamura	Υψηλή Συμπίεση	36 sec
Clip Mioma	92 sec	Gabor	Χαμηλή Συμπίεση	37 sec

Clip Mioma	92 sec	Gabor	Μέτρια Συμπίεση	30 sec
Clip Mioma	92 sec	Gabor	Υψηλή Συμπίεση	15 sec
Clip Mioma	92 sec	Tamura	Χαμηλή Συμπίεση	37 sec
Clip Mioma	92 sec	Tamura	Μέτρια Συμπίεση	19 sec
Clip Mioma	92 sec	Tamura	Υψηλή Συμπίεση	9 sec
Gs_RC_23-02 -2006_clip_00 04	58 sec	Gabor	Χαμηλή Συμπίεση	21 sec
Gs_RC_23-02 -2006_clip_00 04	58 sec	Gabor	Μέτρια Συμπίεση	11 sec
Gs_RC_23-02 -2006_clip_00 04	58 sec	Gabor	Υψηλή Συμπίεση	6 sec
Gs_RC_23-02 -2006_clip_00 04	58 sec	Tamura	Χαμηλή Συμπίεση	25 sec
Gs_RC_23-02 -2006_clip_00 04	58 sec	Tamura	Μέτρια Συμπίεση	12 sec

Gs_RC_23-02 -2006_clip_00 04	58 sec	Tamura	Υψηλή Συμπίεση	6 sec
De La Cruz MIOMA	455 sec	Gabor	Υψηλή Συμπίεση	50 sec
De La Cruz MIOMA	455 sec	Tamura	Υψηλή Συμπίεση	52 sec

Πίνακας 4.2 Χαρακτηριστικά συμπιεσμένων βίντεο

4.4 Αποτελέσματα

Στο πιο κάτω πίνακα φαίνονται οι βαθμολογίες για την αξιολόγηση ποιότητας με βάση τη βαθμολογία που έχουμε πάρει μετά από την προβολή τους από τον γιατρό για την εφαρμογή της περίληψης των βίντεο με βάση τα φίλτρα.

Όνομα Video	Μέθοδος	Συμπίεση	Ποιότητα
1	Gabor	Χαμηλή Συμπίεση	
1	Gabor	Μέτρια Συμπίεση	
1	Gabor	Υψηλή Συμπίεση	
1	Tamura	Χαμηλή Συμπίεση	
1	Tamura	Μέτρια Συμπίεση	
1	Tamura	Υψηλή Συμπίεση	
2	Gabor	Μέτρια Συμπίεση	

2	Gabor	Υψηλή Συμπύεση	
2	Tamura	Μέτρια Συμπύεση	
2	Tamura	Υψηλή Συμπύεση	
Clip Mioma	Gabor	Χαμηλή Συμπύεση	
Clip Mioma	Gabor	Μέτρια Συμπύεση	
Clip Mioma	Gabor	Υψηλή Συμπύεση	
Clip Mioma	Tamura	Χαμηλή Συμπύεση	
Clip Mioma	Tamura	Μέτρια Συμπύεση	
Clip Mioma	Tamura	Υψηλή Συμπύεση	
Gs_RC_23-02-200 6_clip_0004	Gabor	Χαμηλή Συμπύεση	
Gs_RC_23-02-200 6_clip_0004	Gabor	Μέτρια Συμπύεση	
Gs_RC_23-02-200 6_clip_0004	Gabor	Υψηλή Συμπύεση	
Gs_RC_23-02-200 6_clip_0004	Tamura	Χαμηλή Συμπύεση	
Gs_RC_23-02-200 6_clip_0004	Tamura	Μέτρια Συμπύεση	
Gs_RC_23-02-200 6_clip_0004	Tamura	Υψηλή Συμπύεση	
De La Cruz MIOMA	Gabor	Υψηλή Συμπύεση	

De La Cruz MIOMA	Tamura	Υψηλή Συμπίεση	
---------------------	--------	----------------	--

Πίνακας 4.3 Βαθμολογία ποιότητας συμπιεσμένων βίντεο από γιατρό

Κεφάλαιο 5

Συμπεράσματα και Μελλοντικές Εργασίες

5.1 Γενικά Συμπεράσματα	52
5.2 Μελλοντικές εργασίες	53

5.1 Συμπεράσματα

Από τα αποτελέσματα για τον αλγόριθμο των φίλτρων, υπάρχει μια πολύ καλή και πετυχημένη συμπίεση στο βαθμό που πρέπει, με βάση τις παραμέτρους που εισήχθησαν στο σύστημα.

Ο αλγόριθμος χρωμάτων δεν έχει τη δυνατότητα ελεγχόμενου χρόνου του βίντεο περίληψης που παράγει. Αυτό συμβαίνει, γιατί δεν υπάρχει κάποια παράμετρος που να καθορίζει τον αριθμό των στιγμιότυπων που θα εξαχθούν στο τελικό βίντεο περίληψης.

Από την άλλη, τα βίντεο δεν έχουν αξιολογηθεί από το γιατρό, με αποτέλεσμα να μην έχουμε μια εικόνα για την ορθότητα του περιεχομένου τους. Με βάση όμως τον σκοπό που έχει υλοποιηθεί αυτός έχει αρκετά καλά αποτελέσματα διότι με συγκεκριμένες παραμέτρους σε κάθε βίντεο, το βίντεο περίληψης που παράγεται περιλαμβάνει ένα μεγάλο ποσοστό από αυτά τα στιγμιότυπα του βίντεο.

5.2 Μελλοντικές Εργασίες

Οι μελλοντικές εργασίες που θα μπορούσαν να γίνουν στην εργασία αυτή για τον αλγόριθμο περίληψης ιατρικού βίντεο ενδοσκοπικής χειρουργικής με βάση τα φίλτρα είναι:

- Χρησιμοποίηση και άλλων φίλτρων υφής και χρώματος, εκτός από αυτά που έχουν χρησιμοποιηθεί, για την εξαγωγή καλύτερων αποτελεσμάτων
- Χρησιμοποίηση φίλτρων διαφορετικών ειδών από αυτά που έχουν χρησιμοποιηθεί. Μια καλή δοκιμή θα ήταν φίλτρα κίνησης.
- Υλοποίηση του αλγόριθμου με νήματα για γρηγορότερη επεξεργασία και εμφάνιση αποτελεσμάτων
- Τροποποίηση αλγορίθμου για καλύτερη διαχείριση της μνήμης, ώστε να μπορεί να επεξεργαστεί μεγαλύτερα βίντεο.

Οι μελλοντικές εργασίες που θα μπορούσαν να γίνουν στην εργασία αυτή για τον αλγόριθμο περίληψης ιατρικού βίντεο ενδοσκοπικής χειρουργικής με βάση τα χρώματα είναι:

- Μετατροπή αλγορίθμου σε αλγόριθμο αναγνώρισης αντικειμένων (π.χ. Αλγόριθμο Viola–Jones object detection)
- Τροποποίηση του αλγορίθμου για να ελέγχει και το μέγεθος της συμπίεσης

Βιβλιογραφία

- [1] “Endoscopy”, Wikipedia. [Online]. Available:
<https://en.wikipedia.org/wiki/Endoscopy>
- [2] “Βίντεο λαπαροσκοπικών - αρθροσκοπικών χειρουργικών επεμβάσεων”, Γιώργος Τσιλεδάκης. [Online]. Available:
<http://chaniacclinic.com/article-about-media>
- [3] “JPEG Discrete Cosine Transform”, Prime Core Solution. [Online]. Available:
<http://www.pcs-ip.eu/index.php/main/edu/5>
- [4] “JPEG Discrete Cosine Transform”, Wikipedia. [Online]. Available:
https://en.wikipedia.org/wiki/JPEG#Discrete_cosine_transform
- [5] “JPEG Quantization”, Prime Core Solution. [Online]. Available:
<http://www.pcs-ip.eu/index.php/main/edu/6>
- [6] “JPEG Quantization”, Wikipedia. [Online]. Available:
<https://en.wikipedia.org/wiki/JPEG#Quantization>
- [7] “JPEG Zigzag ordering”, Prime Core Solution. [Online]. Available:
<http://www.pcs-ip.eu/index.php/main/edu/7>
- [8] “RGB color model”, Wikipedia. [Online]. Available:
https://en.wikipedia.org/wiki/RGB_color_model
- [9] “HSL and HSV”, Wikipedia. [Online]. Available:
https://en.wikipedia.org/wiki/HSL_and_HSV
- [10] “Luma (video)”, Wikipedia. [Online]. Available:

[https://en.wikipedia.org/wiki/Luma_\(video\)](https://en.wikipedia.org/wiki/Luma_(video))

- [11] “YIQ”, Wikipedia. [Online]. Available:
<https://en.wikipedia.org/wiki/YIQ>
- [12] “Gabor filter”, Wikipedia. [Online]. Available:
https://en.wikipedia.org/wiki/Gabor_filter
- [13] “CBIR: Texture Features - Tamura Texture features”. [Online]. Available:
<https://www.cs.auckland.ac.nz/courses/compsci708s1c/lectures/Glect-html/topic4c708FSC.htm#tamura>
- [14] “A Short History of Color Edge Detection”, Mark A. Ruzon. [Online]. Available:
<http://ai.stanford.edu/~ruzon/compass/color.html>
- [15] Savvas A. Chatzichristofis and Yiannis S. Boutalis, FCTH: Fuzzy Color and Texture Histogram - A Low Level Feature for Accurate Image Retrieval. In: WIAMIS '08 Proceedings of the 2008 Ninth International Workshop on Image Analysis for Multimedia Interactive Services. Pages 191-196. May 07-09, 2008. [Online]. Available: <http://chatzichristofis.info/files/papers/C2.pdf>
- [16] “Image Indexing Using Color Correlograms”,. [Online]. Available:
<http://www.cs.cornell.edu/~rdz/Papers/Huang-CVPR97.pdf>
- [17] "Color Layout Descriptor." Wikipedia. [Online]. Available:
https://en.wikipedia.org/wiki/Color_layout_descriptor
- [18] “Color descriptors from compressed images”, Vinay Modi. School of Informatics, The University of Edinburgh, UK. [Online]. Available:

http://www.vinaymodi.com/uploads/7/5/7/0/7570030/color_descriptors_from_compressed_images_-_vinay_modi.pdf

- [19] Chee Sun Won , Dong Kwon Park , and Soo-Jun Park , "Efficient Use of MPEG-7 Edge Histogram Descriptor," ETRI Journal, vol. 24, no. 1, Feb. 2002, pp. 23-30. [Online]. Available:
<http://dx.doi.org/10.4218/etrij.02.0102.0103>

- [20] “Color histogram”, Wikipedia. [Online]. Available:
https://en.wikipedia.org/wiki/Color_histogram

- [22] "Eclipse (software)." Wikipedia. Wikimedia Foundation, 14 May 2017. Web. 15 May 2017. [Online]. Available:
https://en.wikipedia.org/wiki/Eclipse_%28software%29

- [23] "What Is Java? - Definition from WhatIs.com." SearchMicroservices. Web. 15 May 2017. [Online]. Available:
<http://searchmicroservices.techtarget.com/definition/Java>

- [24] “Εισαγωγή στη Γλώσσα Προγραμματισμού Java”, Εργαστήριο Τεχνολογίας Πολυμέσων ΕΜΠ. [Online]. Available:
http://aetos.it.teithe.gr/~sfetsos/java_book_EMP.pdf

- [25] “Dynamic Adaptive Streaming over HTTP”, Wikipedia. [Online]. Available:
https://en.wikipedia.org/wiki/Dynamic_Adaptive_Streaming_over_HTTP

- [26] “Media presentation description and segment formats: MPEG Dynamic Adaptive Streaming over HTTP (DASH)”. [Online]. Available:
<http://mpeg.chiariglione.org/standards/mpeg-dash/media-presentation-description-and-segment-formats>

- [27] Κυρίως ιστοσελίδα MPEG-DASH. [Online]. Available:
<http://emcvideoz.cs.ucy.ac.cy/>
- [28] Ιστοσελίδα MPEG-DASH - Ενδοσκοπικά βίντεο. [Online]. Available:
<http://emcvideoz.cs.ucy.ac.cy/dash.js-master/samples/dash-if-reference-player/dash.html#endoscopy>
- [29] “Image Compression and the Discrete Cosine Transform”, Ken Cabeen, Peter Gent. [Online]. Available:
<https://www.math.cuhk.edu.hk/~lmlui/dct.pdf>
- [30] “The Discrete Cosine Transform (DCT)”, Dave Marshall. [Online]. Available:
<https://users.cs.cf.ac.uk/Dave.Marshall/Multimedia/node231.html>
- [31] “The Discrete Cosine Transform (DCT)”. [Online]. Available:
<https://cs.stanford.edu/people/eroberts/courses/soco/projects/data-compression/lossy/jpeg/dct.htm>
- [32] “Coefficient Quantization”. [Online]. Available:
<https://cs.stanford.edu/people/eroberts/courses/soco/projects/data-compression/lossy/jpeg/coeff.htm>
- [33] “JPEG Zigzag”. [Online]. Available:
<http://www.pcs-ip.eu/index.php/main/edu/7>
- [34] “Algorithm for Video Summarization of Bronchoscopy Procedure”, BioMedical Engineering OnLine. [Online]. Available:
<https://biomedical-engineering-online.biomedcentral.com/articles/10.1186/1475-925X-10-110>
- [35] “Video summarization using color features and efficient adaptive threshold

technique”, Stevica Cvetkovic, Marko Jelenkovic, Sasa V. Nikolic. [Online]. Available: <http://pe.org.pl/articles/2013/2a/53.pdf>

- [36] “A New Method for Static Video Summarization Using Local Descriptors and Video Temporal Segmentation”, Edward J. Y. Cayllahua Cahuina, Guillermo Camara Chavez. [Online]. Available: http://www.ucsp.edu.pe/sibgrapi2013/e proceedings/technical/114951_2.pdf
- [37] “Video summarization using fuzzy descriptors and a temporal segmentation”, Mickael Guironnet, Dennis Pellerin and Patricia Ladret. [Online]. Available: http://www.gipsa-lab.grenoble-inp.fr/~denis.pellerin/06_publicis/04_acivs_guironnet_rev.pdf
- [38] “Video summary generation as open source”, SemanticMetadata.net. [Online]. Available: <http://www.semanticmetadata.net/2008/12/18/video-summary-generation-as-open-source/>
- [39] “Lire library”, grepcode. [Online]. Available: <http://grepcode.com/file/repo1.maven.org/maven2/de.lmu.ifi.dbs.jfeaturelib/JFeatureLib/1.0.1/net/semanticmetadata/lire/imageanalysis/LireFeature.java>
- [40] “LIRE: Lucene Image Retrieval – An Extensible Java CBIR Library”, Mathias Lux, Savvas A. Chatzichristofis. In proceedings of the 16th ACM International Conference on Multimedia, pp. 1085-1088, Vancouver, Canada, 2008. [Online]. Available: <http://dl.acm.org/citation.cfm?id=1459577>
- [41] “LIRE”, Github. [Online]. Available: <https://github.com/dermotte/LIRE>

- [42] “GNU GENERAL PUBLIC LICENSE”, GNU. [Online]. Available:
<https://www.gnu.org/licenses/gpl-3.0.en.html>

- [43] “Sobel operator”, Wikipedia. [Online]. Available:
https://en.wikipedia.org/wiki/Sobel_operator

- [44] “Haar wavelet”, Wikipedia. [Online]. Available:
https://en.wikipedia.org/wiki/Haar_wavelet