

Ατομική Διπλωματική Εργασία

**ΠΡΟΣΑΡΜΟΣΤΙΚΗ ΜΕΤΑΔΟΣΗ ΙΑΤΡΙΚΟΥ ΒΙΝΤΕΟ
ΣΕ ΠΡΑΓΜΑΤΙΚΟ ΧΡΟΝΟ
ΜΕ ΤΗΝ ΧΡΗΣΗ ΤΟΥ ΠΡΟΤΥΠΟΥ MPEG-DASH**

Νάταλν Ανδρέου

ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΥΠΡΟΥ



ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ

Μάιος 2016

ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΥΠΡΟΥ

ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ

**ΠΡΟΣΑΡΜΟΣΤΙΚΗ ΜΕΤΑΔΟΣΗ ΙΑΤΡΙΚΟΥ ΒΙΝΤΕΟ
ΣΕ ΠΡΑΓΜΑΤΙΚΟ ΧΡΟΝΟ
ΜΕ ΤΗΝ ΧΡΗΣΗ ΤΟΥ ΠΡΟΤΥΠΟΥ MPEG-DASH**

Ανδρέου Νάταλν

Επιβλέπων Καθηγητής
Παττίχης Κωνσταντίντος

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

Μάιος 2016

Ευχαριστίες

Πρώτα απο όλα θα ήθελα να ευχαριστήσω τον επιβλέπων καθηγητή μου κ.Κωνσταντίνο Παττίχη για την ευκαιρία που μου έδωσε να εκπονήσω την παρούσα εργασία και για την εμπιστοσύνη που μου έδειξε. Η καθοδήγηση και υποστήριξη του καθ' όλη την διάρκεια των σπουδών μου στο Πανεπιστήμιο Κύπρου έπαιξαν πρωταρχικό ρόλο στην επιτυχή ολοκλήρωση τους.

Ιδιαίτερα θα ήθελα να ευχαριστήσω τον Δρ. Αντρέα Παναγίδη για τον πολύτιμο χρόνο και την υπομονή που υπέδειξε κατά την διάρκεια εκπόνησης της εργασίας. Η άμεση βοήθεια που πρόσφερε κάθε φορά με τις τόσες πολλές γνώσεις που κουβαλά στον τομέα της Τηλεϊατρικής και της επεξεργασίας βίντεο ήταν ότι πολυτιμότερο. Τον ευχαριστώ επίσης και για την φιλική ατμόσφαιρα που δημιουργούσε σε κάθε συνάντηση μας γιατί μου έδινε κίνητρο να μαθαίνω ακόμα περισσότερα.

Φυσικά, θα ήθελα να ευχαριστήσω τους γονείς μου, την μαμά μου για τις πολύτιμες αναμμένες λαμπάδες και τον μπαμπά μου για τις τόσο σημαντικές συμβουλές. Επίσης ευχαριστώ τις αδερφές μου Βάνα και Άντρια που ήταν δίπλα μου σε όλα δίνοντας μου θάρρος και δύναμη. Θερμές ευχαριστίες επίσης στον Παναγιώτη γιατί η κατανόηση που έδειξε σε όλες τις μέρες άγχους και πίεσης ήταν ο κύριος παράγοντας που με ωθούσε να συνεχίσω σε κάθε σημείο που τα πράγματα ήταν σκούρα. Τέλος, ευχαριστώ όλους τους φίλους μου, ιδιαίτερα όμως την Θεοδώρα και τον Γιώργο για τις ατελείωτες ώρες γέλιου που μου πρόσφεραν εν ώρα διαβάσματος αλλά και χαλάρωσης.

Περίληψη

Στόχος της παρούσας διπλωματικής εργασίας είναι η ανάπτυξη ενός συστήματος ιατρικής εκπαίδευσης και αξιολόγησης ιατρικού βίντεο. Γίνεται προσαρμοστική μετάδοση των βίντεο σε πραγματικό χρόνο, με την χρήση του προτύπου MPEG-DASH. Το σύστημα αυτό μπορεί να χρησιμοποιηθεί για σκοπούς εκπαίδευσης ανθρώπων που ενδιαφέρονται να μάθουν για διάφορα υπερηχογραφήματα.

Στα πλαίσια της εργασίας, διεξάχθηκαν πειράματα για την αξιολόγηση του συστήματος ως προς την ποιότητα των βίντεο που μεταδίδονταν. Τα πειράματα αυτά, έγιναν σε ένα σύνολο από δέκα ασυμπίεστα βίντεο καρωτιδικής αρτηρίας καθώς επίσης και δέκα ασυμπίεστα βίντεο καρδιάς.

Για την αξιολόγηση της ποιότητας του βίντεο, χρησιμοποιήθηκε ο αλγόριθμος αντικειμενικής αξιολόγησης PSNR. Επίσης χρησιμοποιήθηκαν οι συντελεστές συσχέτισης Pearson και Spearman για να βρεθεί ο βαθμός συσχέτισης της αντικειμενικής αξιολόγησης PSNR και της κλινικής αξιολόγησης του παραλήπτη.

Τέλος, από τα αποτελέσματα των πειραμάτων μπορούμε να συμπεράνουμε ότι το σύστημα ιατρικής εκπαίδευσης και αξιολόγησης ιατρικού βίντεο που αναπτύχθηκε, πέτυχε τον στόχο του για αξιόπιστη προσαρμοστική μετάδοση ιατρικού βίντεο σε πραγματικό χρόνο μέσω του προτύπου MPEG-DASH.

ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ

Κεφάλαιο 1	1
Εισαγωγή	1
1.1. Τηλεϊατρική	1
1.2. Στόχος	3
1.3. Δομή εργασίας	4
Κεφάλαιο 2	6
Κωδικοποίηση Και Πρότυπα Κωδικοποίησης Ψηφιακού Βίντεο	6
2.1. Κωδικοποίηση	6
2.1.1. Κωδικοποίηση Εντροπίας	8
2.1.3. Πρόβλεψη Με Εκτίμηση Και Αντιστάθμιση Κίνησης	9
2.2. Τύποι Πλαισίων	10
2.2.1. I-Frames	10
2.2.2. P-Frames	10
2.2.3. B-Frames	10
2.3. Πρότυπα Κωδικοποίησης Ψηφιακού Βίντεο	11
2.3.2. MPEG-1	12
2.3.3. MPEG-2 ή H.262	12
2.3.4. M-JPEG	13
2.3.5. H.263	13
2.3.6. MPEG-4 Part 2	13
2.4. H.264 / AVC MPEG-4 Part 10	14
2.4.1. Αρχιτεκτονική Προτύπου	15
2.4.2. Διαδικασία κωδικοποίησης και αποκωδικοποίησης	16
2.4.3. Profiles και Levels	18
Κεφάλαιο 3	20
Ασύρματες Τεχνολογίες Μετάδοσης Βίντεο και Πρωτόκολλα Επικοινωνίας	20
3.1. Ασύρματα Δίκτυα	20
3.1.1. Πρώτη γενεά 1G	22
3.1.2. Δεύτερη γενεά 2G	22
3.1.3. Δεύτερη και μισό Γενεά 2.5G	23
3.1.4. Τρίτη Γενεά 3G	23
3.1.5. Τρίτη και μισό Γενεά 3.5G	24
3.1.6. Τέταρτη γενεά 4G	25
3.1.7. Πέμπτη Γενεά 5G	25
3.2. Πρωτόκολλα Επικοινωνίας	26
3.2.1. Πρωτόκολλο UDP	26
3.2.2. Πρωτόκολλο TCP	27
3.2.3 Πρωτόκολλο RTP-RTCP	27

Κεφάλαιο 4	29
HTTP adaptive streaming	29
4.1. Adaptive bitrate streaming	29
4.1.1. Adobe HTTP Dynamic Streaming	30
4.1.2. Apple HTTP Adaptive Streaming	30
4.1.3. Microsoft Smooth Streaming	31
4.2. MPEG-DASH	31
4.2.1. Multimedia Presentation Description (MPD)	32
4.2.2. Μορφή των Segments	33
Κεφάλαιο 5	34
5.1. Εφαρμογές ανοικτού κώδικα και λογισμικό	34
5.1.1. FFmpeg	35
5.1.2. ffmpeg	36
5.1.3. x264	36
5.1.4. MP4box	36
5.1.5. Wireshark	37
5.2. Περιγραφή συστήματος ιατρικής εκπαίδευσης και αξιολόγησης ιατρικού βίντεο	37
5.2.1. Κωδικοποιητής	40
5.2.2. Συσκευαστής πολυμέσων	41
5.2.3. Αποστολέας	42
5.2.4. Παραλήπτης	43
5.2.5. Διεπαφή αξιολόγησης ποιότητας βίντεο	45
Κεφάλαιο 6	46
Μεθοδολογία Υλοποίησης	46
6.1. Κωδικοποίηση	46
6.2. Κατακερματισμός και Τμηματοποίηση	49
6.3.1. Αλγόριθμος PSNR	50
6.3.2. Συντελεστής Συσχέτισης Pearson	50
6.3.3. Συντελεστής Συσχέτισης Spearman	51
Κεφάλαιο 7	52
Αποτελέσματα	52
7.1. Αξιολόγηση Κωδικοποιημένων Βίντεο	52
7.2. Συγκεντρωτικά Αποτελέσματα	54
Κεφάλαιο 8	60
Συμπεράσματα και Μελλοντική Εργασία	60
8.1. Συμπεράσματα	60
8.2. Μελλοντική Εργασία	61
Βιβλιογραφία	63

Κεφάλαιο 1

Εισαγωγή

1.1 Τηλεϊατρική	1
1.2 Στόχος	3
1.3 Δομή εργασίας	4

1.1. Τηλεϊατρική

Τηλεϊατρική [1] ονομάζεται η άσκηση της ιατρικής από απόσταση, με την βοήθεια ηλεκτρονικών μέσων. Είναι δηλαδή η εφαρμογή των σύγχρονων τεχνολογιών, των τηλεπικοινωνιών και της πληροφορικής για να παρέχει εξειδικευμένη ιατρική φροντίδα. Ο κύριος λόγος που ώθησε την δημιουργία της τηλεϊατρικής ήταν το γεγονός ότι σε πολλές περιπτώσεις οι ασθενείς βρίσκοντας σε μεγάλη απόσταση από τον χώρο στον οποίο τους παρεχόταν η φροντίδα ή ακόμα βρίσκονταν σε περιοχές οι οποίες δεν είχαν το απαραίτητο ειδικευμένο ιατρικό προσωπικό.

Σε πολλές περιπτώσεις, η θεραπεία και κατεπέκταση η υγεία ενός ασθενή, βασίζεται ιδιαίτερα στην έγκαιρη διάγνωση από έναν ειδικευμένο ιατρό. Έτσι κύριος σκοπός της τηλεϊατρικής είναι να μπορέσει να μεταφέρει ιατρικές πληροφορίες από την τοποθεσία του ασθενή, στην τοποθεσία του ιατρικού προσωπικού. Η μεταφορά αυτή πρέπει όμως να γίνεται άμεσα αφού τις πλείστες φορές ο χρόνος είναι το σημαντικότερο για να σωθεί ή να θεραπευτεί ένας ασθενής.

Η τηλεϊατρική μπορεί να χωριστεί σε τρεις βασικές κατηγορίες ανάλογα με τις υπηρεσίες που παρέχει. Οι κατηγορίες αυτές έχουν ως εξής: αποθήκευση-προώθηση, απομακρυσμένη παρακολούθηση καθώς και διαδραστικές υπηρεσίες.

Στην πρώτη κατηγορία, όπου έχουμε την αποθήκευση και προώθηση [2], εμπεριέχεται ένα ευρύ φάσμα ιατρικών δεδομένων. Το φάσμα αυτό περιλαμβάνει βίντεο υπερηχογραφίας,

εγκεφαλογραφήματα και ηλεκτροκαρδιογραφήματα, βίντεο τραύματος, ιατρικές φωτογραφίες καθώς και καρδιογραφήματα. Τα ιατρικά αυτά δεδομένα αποστέλλονται στον ειδικό ιατρό για να εκτιμήσει την κατάσταση του αρρώστου και να δώσει την κατάλληλη θεραπεία.

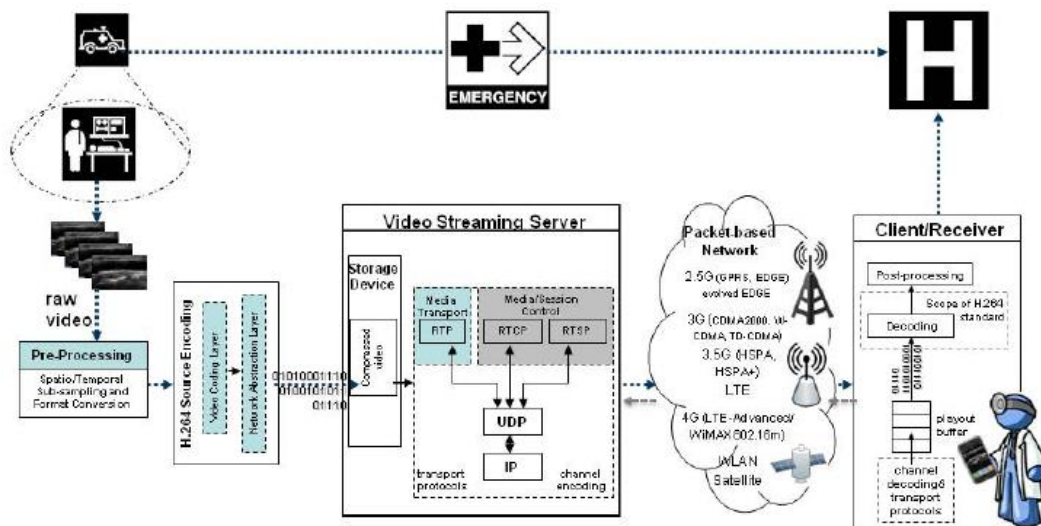
Η απομακρυσμένη παρακολούθηση επιτρέπει την από μακριά παρατήρηση του ασθενούς. Σε νοσοκομεία που βρίσκονται σε απομακρυσμένες περιοχές και το μέγεθος τους δεν είναι ικανό να εργοδοτήσει ιατρούς από κάθε ειδίκευση η απομακρυσμένη παρακολούθηση είναι η λύση του προβλήματος. Αυτό γιατί, σε περίπτωση που χρειαστεί κάποιος ιατρός συγκεκριμένης ειδίκευσης που δεν υπάρχει στο νοσοκομείο, με την βοήθεια της απομακρυσμένης παρακολούθησης ένας γιατρός που βρίσκεται σε κάποια απομακρυσμένη περιοχή θα μπορεί να παρέχει στον ασθενή την απαραίτητη ιατρική φροντίδα. Αυτή η μέθοδος χρησιμοποιείται σε πολλές περιπτώσεις όπου ένας ασθενής έχει κάποιο χρόνιο νόσημα, όπως για παράδειγμα άσθμα, διαβήτη ή κάποιο είδος καρδιοπάθειας.

Ο όρος διαδραστική τηλεϊατρική [3], παρέχει τη δυνατότητα άμεσης και γρήγορης επικοινωνίας μεταξύ του ασθενούς και του θεράποντος ιατρού του.

Με βάση τη χρήση τηλεπικοινωνιακών και πληροφορικών συστημάτων και τη μετατροπή ιατρικής πληροφορίας σε ηλεκτρονική μορφή, διακρίνεται μια πληθώρα εφαρμογών.

Σημαντική εφαρμογή είναι η τηλεπαθολογία μέσω της οποίας γίνεται μεταφορά παθολογικών εξετάσεων για τη γνωμάτευση και θεραπεία από απόσταση. Η πιο συνηθισμένη όμως εφαρμογή, είναι η τηλεακτινολογία [4] η οποία αποτελεί το 50% του όγκου των τηλειατρικών συναλλαγών και υπηρεσιών. Μέσω της τηλεακτινολογίας γίνεται μεταφορά ακτινολογικών εικόνων από ένα σημείο σε άλλο για γνωμάτευση ή παροχή συμβουλών θεραπείας. Μια επιπλέον εφαρμογή είναι η τηλεκαρδιολογία, η οποία αποτελεί την μετάδοση καρδιολογικών εξετάσεων σε περιπτώσεις όπου οι ασθενείς δεν μπορούν να μετακινηθούν. Μια εξειδικευμένη υπηρεσία είναι η τηλεχειρουργική μέσω της οποίας γίνεται μετάδοση εικόνας και ήχου, επιτρέποντας της επικοινωνίας μεταξύ χειρουργών, με χρήση ρομποτικών συσκευών.

Στο σχήμα 1.1 δίνεται μια περιγραφή της αρχιτεκτονικής ενός τηλειατρικού συστήματος για επείγοντα περιστατικά. Η διαδικασία είναι η εξής: Κατά την διάρκεια που το επείγον περιστατικό βρίσκεται μέσα στο ασθενοφόρο, μέσω του εξοπλισμού που υπάρχει εκεί το παραϊατρικό προσωπικό γυρίζει ένα ιατρικό βίντεο του ασθενούς. Στη συνέχεια το βίντεο αυτό μαζί με τα βιοσήματα του ασθενούς, αποστέλλονται μέσω των διαθέσιμων ασύρματων καναλιών, στον ιατρό και τις εγκαταστάσεις του νοσοκομείου στο οποίο θα νοσηλευτεί ο ασθενής. Για να μπορέσει το ιατρικό βίντεο να σταλεί με πιο γρήγορη ταχύτητα, κωδικοποιείται πριν την αποστολή έτσι ώστε να επιτευχθεί η μείωση του μεγέθους του. Ένα σημαντικό όμως κομμάτι της κωδικοποίησης είναι να μην μειωθεί σημαντικά η ποιότητα του συμπίεσμένου βίντεο για να μπορέσει ο ιατρός να κάνει την σωστή διάγνωση.



Σχήμα 1.1: Ασύρματη Μετάδοση Ιατρικού Βίντεο Για Επείγοντα Περιστατικά [5].

1.2. Στόχος

Στόχος της παρούσας διπλωματικής εργασίας είναι η ανάπτυξη ενός συστήματος ιατρικής εκπαίδευσης και αξιολόγησης ιατρικού βίντεο. Τα βίντεο θα στέλνονται από τον αποστολέα με προσαρμοστική μετάδοση σε πραγματικό χρόνο, με την χρήση του προτύπου MPEG-DASH. Το σύστημα αυτό μπορεί να χρησιμοποιηθεί για σκοπούς εκπαίδευσης ανθρώπων που ενδιαφέρονται να μάθουν για διάφορα υπερηχογραφήματα.

Επίσης, γνωρίζοντας την ύπαρξη τεχνολογιών ανοικτού κώδικα, ένας επιπλέον στόχος της εργασίας αυτής είναι οι ανεξάρτητες τεχνολογίες να συνδυαστούν για την δημιουργία του τελικού συστήματος. Ο συνδυασμός τους θα ενσωματώνει επίσης και το πρότυπο προσαρμοστικής μετάδοσης MPEG-DASH. Έτσι θα μπορέσουμε να δούμε, πόσο ικανοποιητικό είναι το αποτέλεσμα του συνδυασμού των διαφόρων τεχνολογιών ανοικτού κώδικα.

Τέλος, ένας επιπλέον στόχος είναι να αξιολογήσουμε τις διάφορες κωδικοποιήσεις των βίντεο σε διαφορετικές αναλύσεις και σε διαφορετικούς ρυθμούς μετάδοσης. Με αυτό τον τρόπο, θα δούμε ποια κωδικοποίηση δίνει το καλύτερο αποτέλεσμα έτσι ώστε τα ιατρικά βίντεο να είναι ιατρικά διαγνώσιμα και να γίνεται αξιόπιστη μετάδοση τους στο διαθέσιμο ασύρματο δίκτυο.

1.3. Δομή εργασίας

Στην συνέχεια της εργασίας αυτής τα κεφάλαια θα παρουσιαστούν με την εξής δομή:

Κεφάλαιο 2^ο : Στο κεφάλαιο αυτό θα αναλυθεί ο όρος της κωδικοποίησης, καθώς επίσης και διάφορες τεχνικές και πρότυπα κωδικοποίησης. Το πρότυπο κωδικοποίησης που χρησιμοποιείται στην εργασία αυτή είναι το H.264 / AVC MPEG-4 Part 10 και για αυτό τον λόγο δίνεται περισσότερη έμφαση στο συγκεκριμένο πρότυπο.

Κεφάλαιο 3^ο: Το κεφάλαιο 3 περιγράφει τις ασύρματες τεχνολογίες μετάδοσης καθώς επίσης και διάφορα πρωτόκολλα επικοινωνίας.

Κεφάλαιο 4^ο: Σε αυτό το κεφάλαιο γίνεται περιγραφή της τρόπων προσαρμοστικής μετάδοσης βίντεο σε πραγματικό χρόνο με HTTP. Έμφαση δίνει στο MPEG-DASH που χρησιμοποιείται στην εργασία αυτή.

Κεφάλαιο 5^ο: Το κεφάλαιο αυτό περιγράφει τις εφαρμογές ανοικτού κώδικα και το λογισμικό που χρησιμοποιήθηκε για την υλοποίηση της εργασίας αυτής, και επίσης δίνεται η περιγραφή του συστήματος.

Κεφάλαιο 6^ο: Στο κεφάλαιο 6 περιέχεται η πλήρης μεθοδολογία που ακολουθήθηκε για την υλοποίηση της εργασίας αυτής.

Κεφάλαιο 7^ο: Σε αυτό το κεφάλαιο παρουσιάζονται τα αποτελέσματα από τα πειράματα που διεξάχθηκαν για τον έλεγχο του συστήματος.

Κεφάλαιο 8^ο: Το τελευταίο αυτό κεφάλαιο, περιέχει τα συμπεράσματα, καθώς επίσης την μελλοντική εργασία που μπορεί να γίνει στην εργασία αυτή.

Κεφάλαιο 2

Κωδικοποίηση Και Πρότυπα Κωδικοποίησης Ψηφιακού Βίντεο

2.1. Κωδικοποίηση	6
2.1.1. Κωδικοποίηση Εντροπίας	8
2.1.2. Κωδικοποίηση Με Πρόβλεψη	9
2.1.3. Πρόβλεψη Με Εκτίμηση Και Αντιστάθμιση Κίνησης	10
2.2. Τύποι Πλαισίων	10
2.2.1. I-Frames	10
2.2.2. P-Frames	10
2.2.3. B-Frames	11
2.3. Πρότυπα Κωδικοποίησης Ψηφιακού Βίντεο	12
2.3.1. H.261	12
2.3.2. MPEG-1	13
2.3.3. MPEG-2 ή H.262	13
2.3.4. M-JPEG	13
2.3.5. H.263	14
2.3.6. MPEG-4 Part 2	15
2.4. H.264 / AVC MPEG-4 Part 10	16
2.4.1. Αρχιτεκτονική Προτύπου	17
2.4.2. Διαδικασία Κωδικοποίησης και Αποκωδικοποίησης	19
2.4.3. Profiles και Levels	19

2.1. Κωδικοποίηση

Τα ψηφιακά βίντεο είναι ουσιαστικά ένα σύνολο από πολλές εικόνες οι οποίες εναλλάσσονται γρήγορα πολλές φορές το δευτερόλεπτα χωρίς εμείς να αντιλαμβανόμαστε την εναλλαγή αυτή, δίνοντας μας την αίσθηση του ότι υπάρχει κίνηση. Για να μπορέσουμε να στείλουμε έναν τέτοιο τεράστιο όγκο δεδομένων μέσω ασύρματου δικτύου χρησιμοποιούμε κάποιες τεχνικές συμπίεσης. Στόχος μας κατά την συμπίεση είναι να

κωδικοποιήσουμε την πληροφορία του βίντεο με λιγότερη ακρίβεια, μειώνοντας τον ρυθμό μετάδοσης (bit-rate) που χρειάζεται για αποθήκευση και μετάδοση χωρίς να αλλοιώσουμε σε μεγάλο βαθμό το τελικό μας αποτέλεσμα.

Σε ότι αφορά συμπίεση ήχου, ο ήχος μπορεί να προσφερθεί για μεγάλα ποσοστά συμπίεσης χωρίς να μπορεί να γίνει αισθητή η υποβάθμιση στην ποιότητα του, κάτι που στην συμπίεση βίντεο δεν συμβαίνει. Στο τομέα του βίντεο η παραμικρή ατέλεια είναι ορατή και έτσι προσπαθούμε οι ατέλειες αυτές να μην είναι ενοχλητικές και να μην αφήνουν πίσω σημαντική πληροφορία. Επομένως, στη συμπίεση ενός βίντεο αναζητείται μια χρυσή τομή ανάμεσα στην ποιότητα και το εύρος ζώνης.

Το πόσο αποδοτική είναι μια τεχνική συμπίεσης το καθορίζουμε με βάση δυο σημαντικούς παράγοντες. Οι παράγοντες αυτοί είναι η ποσότητα πλεονασμού που περιέχεται στα πλαίσια καθώς επίσης και το είδος της τεχνικής συμπίεσης που χρησιμοποιείται για την κωδικοποίηση. Σε πολλά βίντεο υπάρχει πληροφορία που αν δεν συμπεριληφθεί δεν επηρεάζει σε μεγάλο βαθμό την τελική ποιότητα του βίντεο και επίσης υπάρχει πληροφορία που μπορεί να κωδικοποιηθεί με λιγότερη ακρίβεια. Οι πληροφορίες αυτές ονομάζονται πλεονασμός.

Ο πλεονασμός μπορεί να κατηγοριοποιηθεί σε χωρικό πλεονασμό (Spatial Redundancy) και σε χρονικό πλεονασμό (Temporal Redundancy) [6]. Ο χωρικός πλεονασμός προκύπτει από το γεγονός ότι τα γειτονικά pixels είναι πολύ πιθανόν να είναι όμοια μεταξύ τους. Έτσι εκμεταλλεύεται τις ιδιαιτερότητες της ανθρώπινης όρασης. Συγκεκριμένα, το ανθρώπινο μάτι μπορεί να ανεχτεί ένα ποσοστό αλλοίωσης ή παραμόρφωσης μιας εικόνας ή ενός βίντεο, χωρίς αυτό να γίνει αντιληπτό όπως επίσης και το γεγονός ότι το μάτι είναι πολύ πιο ευαίσθητο στην φωτεινότητα μιας εικόνας και όχι στους χρωματισμούς της. Ο χρονικός πλεονασμός έχει να κάνει με το ότι ένα τμήμα της εικόνας μπορεί να σχετίζεται με τμήμα της εικόνας του προηγούμενου πλαισίου. Έτσι μπορούμε απλώς να προσθέσουμε τις διαφορές που έχουν προκύψει χωρίς να κωδικοποιήσουμε τις όμοιες πληροφορίες. Π.χ. σ' ένα δελτίο ειδήσεων μόνο το στόμα του εκφωνητή κινείται και η υπόλοιπη εικόνα μένει ίδια. Η συμπίεση εφαρμόζεται σε ακολουθίες από πλαίσια με κωδικοποίηση των διαφορών ανάμεσά τους και με χωρική κωδικοποίηση στα πλαίσια κλειδιά.

Ανάλογα με τη σχέση που έχει το αρχικό βίντεο με το αποτέλεσμα της αποσυμπίεσης υπάρχουν δύο κατηγορίες μεθόδων συμπίεσης. Οι κατηγορίες αυτές έχουν ως εξής: Τεχνικές

χωρίς απώλειες (lossless compression schemes) και τεχνικές με απώλειες (lossy compression schemes). Κατά την τεχνική χωρίς απώλειες το βίντεο που προκύπτει κατά την αποσυμπίεση είναι ένα πιστό αντίγραφο του αρχικού βίντεο και έτσι δεν υπάρχει μεγάλη διαφορά στο μέγεθος του αρχείου. Κατά την τεχνική με απώλειες πετυχαίνουμε τον αρχικό, σημαντικό στόχο που είναι να μειωθεί σημαντικά το μέγεθος του αρχείου. Στην τεχνική αυτή, η οπτική πλεονάζουσα πληροφορία αφαιρείται χωρίς όμως να γίνεται αντιληπτή η διαφορά στο ανθρώπινο μάτι, όπως έχουμε προαναφέρει, και έτσι μπορούμε να επιτύχουμε μεγαλύτερο βαθμό συμπίεσης ούτως ώστε το μέγεθος του αρχείου να μειωθεί σημαντικά.

Ανάλογα με τη μέθοδο που ακολουθείται για τη συμπίεση βίντεο έχουμε τις ακόλουθες τεχνικές κωδικοποίησης.

2.1.1. Κωδικοποίηση Εντροπίας

Η κωδικοποίηση εντροπίας [7] κατατάσσεται στην κατηγορία τεχνικής χωρίς απώλειες. Η κωδικοποίηση αυτή αναφέρεται σε τεχνικές, οι οποίες δεν λαμβάνουν υπόψη το είδος της πληροφορίας που πρόκειται να συμπιεστεί. Με άλλα λόγια, οι τεχνικές αυτές αντιμετωπίζουν την πληροφορία ως μια απλή ακολουθία από bits. Αυτός είναι ο κύριος λόγος που η κωδικοποίηση εντροπίας μπορεί να εφαρμοσθεί ανεξάρτητα από το είδος της πληροφορίας.

Οι τεχνικές κωδικοποίησης εντροπίας μπορούν να χωριστούν σε δυο βασικές κατηγορίες. Στην κατηγορία περιορισμού των επαναλαμβανόμενων ακολουθιών και στην κατηγορία στατιστικής κωδικοποίησης. Η βασική ιδέα στον περιορισμό των επαναλαμβανόμενων ακολουθιών είναι η ακόλουθη: η πιθανότητα να εμφανιστεί κάποιος επαναλαμβανόμενος χαρακτήρας σε μια ακολουθία από δυαδικά ψηφία (bits) είναι μεγάλη. Επομένως, ο επαναλαμβανόμενος αυτός χαρακτήρας αντικαταστάται από έναν ειδικό χαρακτήρα ο οποίος συνοδεύεται από έναν αριθμό που αντιπροσωπεύει τις φορές που έγινε επανάληψη του χαρακτήρα αυτού. Η στατική κωδικοποίηση βασίζεται στον εντοπισμό των πιο συχνά εμφανιζόμενων ακολουθιών χαρακτήρων και ο τρόπος που λειτουργεί έχει ως εξής: η αναπαράσταση των τιμών των χρωμάτων των εικονοστοιχείων (pixels) αλλάζει ανάλογα με το πόσο συχνά εμφανίζονται. Οι τιμές στα εικονοστοιχεία τα οποία εμφανίζονται σε τακτά χρονικά διαστήματα, αναπαρίστανται με πιο μικρό αριθμό ψηφίων. Οι τιμές στα εικονοστοιχεία τα οποία έχουν πιο μικρή πιθανότητα να εμφανιστούν αναπαρίστανται με περισσότερα ψηφία. Έτσι ο μέσος όρος δυαδικών ψηφίων (bits) που χρησιμοποιούνται για να αναπαραστήσουν τα χρώματα των εικονοστοιχείων, μειώνεται σημαντικά.

2.1.2. Κωδικοποίηση Με Πρόβλεψη

Στην κωδικοποίηση με πρόβλεψη, σε αντίθεση με την κωδικοποίηση που προαναφέραμε, οι μετασχηματισμοί τους οποίους υφίσταται το αρχικό σήμα εξαρτώνται άμεσα από τον τύπο του. Οι μετασχηματισμοί που γίνονται στο σήμα, κάνουν χρήση των ιδιαίτερων σημασιολογικών χαρακτηριστικών που μεταφέρει το σήμα. Η κωδικοποίηση με πρόβλεψη εκμεταλλεύεται το γεγονός ότι στα γειτονικά εικονοστοιχεία (pixels) μέσα στα πλαίσια (frames) ή ενδιάμεσα σε διαδοχικά πλαίσια, υπάρχει πλεονασμός. Ο λόγος που ονομάζεται κωδικοποίηση με πρόβλεψη είναι γιατί, με βάση την πληροφορία που μεταδόθηκε προηγουμένως, γίνεται μια κατά προσέγγιση πρόβλεψη της τιμής του χρώματος κάθε εικονοστοιχείου και ακολούθως γίνεται υπολογισμός της διαφοράς που υπάρχει ανάμεσα στην τιμή που έγινε η πρόβλεψη και στην πραγματική τιμή του εικονοστοιχείου. Η μόνη πληροφορία που κωδικοποιείται είναι η διαφορά που υπολογίστηκε.

Γενικά οι τεχνικές κωδικοποίησης με πρόβλεψη μπορούν να παράγουν μεγαλύτερα ποσοστά συμπίεσης σε σχέση με την κωδικοποίηση εντροπίας. Μειονεκτούν όμως στη σταθερότητα, γιατί το ποσοστό συμπίεσης που επιτυγχάνουν διαφοροποιείται ανάλογα με το αντικείμενο που συμπιέζεται. Τέτοιου είδους τεχνικές μπορεί να είναι τεχνικές με ή χωρίς απώλειες.

2.1.3. Πρόβλεψη Με Εκτίμηση Και Αντιστάθμιση Κίνησης

Ένας τρόπος με τον οποίο μπορούμε να περιγράψουμε πιο αποτελεσματικά εικόνες με κίνηση είναι η τμηματική πρόβλεψη κίνησης. Με τη βοήθεια αυτής της μεθόδου μπορούμε να εκτελέσουμε τη διαπλασιασική κωδικοποίηση και να περιγράψουμε την αλληλουχία εικόνας ως σειρά. Εναλλακτικά, με το να κωδικοποιήσουμε ανεξάρτητα τα δύο πλαίσια μπορούμε να περιγράψουμε τη δεύτερη εικόνα με το να τη χωρίσουμε σε ίσα τμήματα και να φτιάξουμε ένα πίνακα που να περιέχει τμήματα που έχουν μείνει ίδια και διανύσματα που να δείχνουν τη νέα θέση των τμημάτων που έχουν αλλάξει θέση (εκτίμηση κίνησης). Έτσι αν έχουμε ήδη αποστέλλει την πρώτη εικόνα μπορούμε να στείλουμε τη δεύτερη σαν ένα πίνακα διανυσμάτων και ορισμένων σταθερών τμημάτων, που προφανώς έχει πολύ μικρότερο μέγεθος. Το παραπάνω είναι μια καλή προσέγγιση της μεθόδου αλλά στην πραγματικότητα οι πραγματικές εικόνες δεν είναι τόσο όμοιες μεταξύ τους. Θα έχουν κάποια κοινά τμήματα που αλλάζουν θέση από πλαίσιο σε πλαίσιο αλλά θα υπάρχουν και τμήματα που αλλάζουν θέση διατηρώντας το σχήμα τους αλλά μεταβάλλεται το χρώμα τους, καθώς και άλλα που

δεν υπάρχουν σε προηγούμενο πλαίσιο αλλά εμφανίζονται σε κάποιο για πρώτη φορά. Για την κάλυψη αυτών των περιπτώσεων κωδικοποιείται το σφάλμα πρόβλεψης.

2.2. Τύποι Πλαισίων

Κατά την συμπίεση ενός βίντεο, τα frames που αποτελούν το βίντεο αυτό μπορούν να χωριστούν με βάση τις ιδιότητες αλλά και τα χαρακτηριστικά τους, στις εξής τρεις βασικές κατηγορίες που θα αναλύσουμε πιο κάτω [8].

2.2.1. I-Frames

Τα frames αυτά είναι ανεξάρτητα. Αυτό σημαίνει ότι κατά την αποκωδικοποίηση τους δεν χρειάζεται να γίνει αναφορά σε κάποιο προηγούμενο ή επόμενο frame. Είναι ουσιαστικά το πιο βασικό frame αφού σε περίπτωση που υπάρξει σφάλμα κατά την μεταφορά το πλαίσιο (frame) αυτό λειτουργεί σαν σημείο αναφοράς. Επίσης αυτά τα frames εμφανίζονται σε τακτικά χρονικά διαστήματα. Η κωδικοποίηση των πλαισίων αυτών βασίζεται στην λογική του προτύπου JPEG όπου η εικόνα χωρίζεται σε macroblocks και για κάθε block ξεχωριστά, εφαρμόζεται DCT, κβαντοποίηση, zig-zag scanning, run-length-encoding και Huffman Encoding.

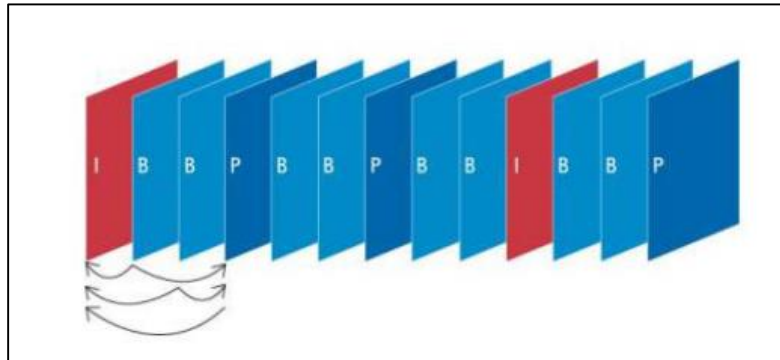
2.2.2. P-Frames

Τα πλαίσια (frames) αυτά ονομάζονται προβλεπόμενα λόγω του ότι υπολογίζονται με πρόβλεψη. Υπολογίζονται δηλαδή με βάση ένα προηγούμενο I-frame ή P-frame. Η δημιουργία τους γίνεται με την βοήθεια αντιστάθμισης κίνησης (motion compensation) που προβλέπει την νέα θέση όποιων macroblocks που έχουν μετακινηθεί. Το μέγεθος τους είναι κατά πολύ μικρότερο από αυτό του I-frame λόγω του ότι έχουν μεγαλύτερο ποσοστό συμπίεσης και δεν έχουν την ίδια ακρίβεια. Επίσης λαμβάνουν μέρος στην δημιουργία και στην διάδοση των σφαλμάτων αφού λειτουργούν σαν σημείο αναφοράς.

2.2.3. B-Frames

Τα frames αυτά ονομάζονται αμφίδρομα λόγω του ότι κάνουν προς τα εμπρός και προς τα πίσω πρόβλεψη των προηγούμενων ή επόμενων I-frame και P-frame. Δεν λαμβάνουν μέρος στην διάδοση σφαλμάτων αφού δεν λειτουργούν σαν σημεία αναφοράς. Τα πλαίσια αυτά

παίρνουν τον μέσο όρο των δύο προηγούμενων πλαισίων και για αυτό τον λόγο μειώνουν σε αρκετά μεγάλο βαθμό τα σφάλματα. Η διαδικασία κωδικοποίησης τους περιλαμβάνει συνδυασμό των αντίστοιχων macroblocks τα οποία παρουσιάζουν μικρές διαφορές με τα αντίστοιχα των πλαισίων αναφοράς.



Σχήμα 2.1: Συνήθης Ακολουθία Πλαισίων [9].

2.3. Πρότυπα Κωδικοποίησης Ψηφιακού Βίντεο

Υπάρχουν πολλά πρότυπα κωδικοποίησης ψηφιακού βίντεο και αυτό που τα κάνει να διαφέρουν είναι η μέθοδος που χρησιμοποιούν για να συμπιέσουν τα δεδομένα. Επομένως, το κάθε πρότυπο κωδικοποίησης διαφέρει από τα άλλα όσο αφορά την ποιότητα του αποτελέσματος, τον χρόνο καθυστέρησης και τον ρυθμό μετάδοσης κατά την αποστολή. Στη συνέχεια θα αναλύσουμε τα βασικά πρότυπα κωδικοποίησης ψηφιακού βίντεο, δίνοντας όμως έμφαση στο πρότυπο H.264 / AVC MPEG-4 Part 10 που χρησιμοποιείται στην υλοποίηση της εργασίας αυτής.

2.3.1. H.261

Το συγκεκριμένο πρότυπο εισάχθηκε το 1990 και στόχος του είναι η εκμετάλλευση χωρικού και χρονικού πλεονασμού έτσι ώστε να επιτυγχάνεται όσο το δυνατόν υψηλότερη συμπίεση. Ο αλγόριθμος συμπίεσης H.261 είναι βασισμένος και μοιάζει με το MPEG. Το πρότυπο αυτό, υποστηρίζει βίντεο χαμηλής ανάλυσης όπως CIF 352x288, QCIF 176x144 και περιορισμένο αριθμό πλαισίων από 7.5 μέχρι 30 πλαίσια το δευτερόλεπτο, λόγω του ότι χρησιμοποιείται για μεταφορά βίντεο με γραμμές όπου το εύρος ζώνης είναι χαμηλό. Χρησιμοποιείται επίσης για να αποκωδικοποιήσει την παραγωγή bit stream για να αναδημιουργήσει τα πλαίσια αναφοράς που χρειάζονται [10].

2.3.2. MPEG-1

Το πρότυπο MPEG1 [11] εισάχθηκε το 1993 και δημιουργήθηκε από την ομάδα MPEG. Το πρότυπο MPEG1 είναι ένα πρότυπο που περιλαμβάνει προδιαγραφές για κωδικοποίηση, συμπίεση, μετάδοση ήχου, βίντεο και δεδομένων σε μια σειρά από συγχρονισμένα και πολυπλεγμένα πακέτα. Η εικόνα έχει ανάλυση NTSC 352x240 με 30 πλαίσια το δευτερόλεπτο ή PAL 352x288 με 25 πλαίσια το δευτερόλεπτο και η ποιότητα της είναι σε επίπεδα VHS βίντεο. Το πρότυπο αυτό βασίζεται στο γεγονός ότι οι συνεχόμενες εικόνες έχουν πολύ μικρή διαφορά μεταξύ τους και έτσι υπάρχει αρκετή πλεονάζουσα πληροφορία. Στόχος του προτύπου αυτού ήταν οι εφαρμογές όπως η αποθήκευση σε CD-ROM ή γραμμές με ταχύτητες T1 από 1,2-1,5 Mbps. Το MPEG-1 σε αντίθεση με το προηγούμενο πρότυπο, χρησιμοποιεί B-frames και έτσι η εκτίμηση κίνησης αρχίζει εδώ να γίνεται πιο περίπλοκη, δημιουργώντας έτσι την ανάγκη χρήσης μιας πιο σύνθετης μονάδας για να υπολογίζεται η αντιστάθμιση της κίνησης. Ο τρόπος υπολογισμού γίνεται μέσω των επόμενων ή προηγούμενων πλαισίων I, P. Τα B πλαίσια εισάχθηκαν για να επιτρέπουν τον περιοδικό συγχρονισμό, να γίνεται ανάκαμψη από σφάλματα και για να μπορεί να γίνεται προσπέλαση σε διάφορα κομμάτια της εικόνας.

2.3.3. MPEG-2 ή H.262

Αυτό το πρότυπο εισάχθηκε το 1995 και είχε ως στόχο την αποτελεσματική κωδικοποίηση κινούμενων εικόνων σε καλύτερη ποιότητα με ρυθμούς κωδικοποίησης μέχρι 10 Mbits/sec. Σε αντίθεση με το MPEG-1, το MPEG-2 σχεδιάστηκε γιατί υπήρχε ανάγκη υποστήριξης μεγαλύτερου εύρους ζώνης, υψηλότερης ποιότητας εικόνας και αρχικά είχε υιοθετηθεί σε συστήματα εκπομπής ψηφιακής τηλεόρασης. Σε σχέση με το πρότυπο MPEG1, το MPEG2 περιέλαβε δύο βασικά επιπλέον στάδια: 1) την αποτελεσματική κωδικοποίηση της πληροφορίας που χάνεται όταν το CCIR601 σήμα υποδειγματοληπτείται κατά την οριζόντια διάσταση στην διαδικασία μετατροπής του σε μορφή MPEG1 και 2) την εκ των υστέρων εκτίμηση των αρτίων γραμμών του αρχικού CCIR 601 σήματος, που παραλήφθηκαν κατά την MPEG1 κωδικοποίηση, από τις περιττές γραμμές με χρησιμοποίηση τεχνικών εκτίμησης και αντιστάθμισης κίνησης. Η βασική ανάλυση της εικόνας είναι NTSC 704x480 ή PAL 704x576. Το MPEG-2 κωδικοποιεί το βίντεο χρησιμοποιώντας πλαίσια I, P, και B, όπως το MPEG-1, με υποστοιχειακή προσέγγιση (μισού εικονοστοιχείου) των διανυσμάτων κίνησης [11].

2.3.4. M-JPEG

Το πρότυπο M-JPEG [12] εισάχθηκε το 1999 και ήταν η πρώτη μέθοδος συμπίεσης που χρησιμοποιήθηκε σε τηλεοπτικά σήματα. Είναι ουσιαστικά μια τεχνική επέκταση της τεχνικής JPEG για τις ακίνητες εικόνες. Πρόκειται για μέθοδο συμπίεσης με απώλειες, για την υλοποίηση της οποίας ακολουθούνται διάφορα βήματα. Κάθε ηλεκτρονικό σήμα μπορεί να αναπαρασταθεί με δύο τρόπους: είτε με βάση τη μεταβολή του στο χρόνο είτε με βάση τη φασματική του σύσταση. Για τις ανάγκες της συμπίεσης επιλέγεται η παράσταση του σήματος ως συνάρτηση του φασματικού του περιεχομένου γιατί στην μορφή αυτή γίνεται πιο εμφανής η πλεονάζουσα πληροφορία και διευκολύνεται η διαδικασία. Πριν γίνει ο μετασχηματισμός κάθε πλαίσιο χωρίζεται σε μπλοκ των 8x8 εικονοστοιχείων. Το μέγεθος αυτό προσφέρει τα περισσότερα πλεονεκτήματα, αφού επιτρέπει γρήγορη διαδικασία αλλά οδηγεί και σε έναν αριθμό εικονοστοιχείων που συνήθως έχουν ικανοποιητική συσχέτιση μεταξύ τους, δηλαδή πολλά γειτονικά εικονοστοιχεία έχουν ίδιες τιμές. Το πρότυπο αυτό σήμερα χρησιμοποιείται για λήψη βίντεο από φωτογραφικές μηχανές ψηφιακής τεχνολογίας και επίσης σε διάφορες κονσόλες παιχνιδιών.

2.3.5. H.263

Είναι ένα πρότυπο που προτάθηκε από την ITU-T και πέρασε από τρεις φάσεις για να εισαχθεί τελικά το 2000. Το πρότυπο αυτό είναι ουσιαστικά μια εξέλιξη του προτύπου H.261 και είχε ως στόχο, όπως όλα τα πρότυπα, να επιτύχει καλύτερη ποιότητα σε χαμηλούς ρυθμούς που κυμαίνονται από 64Kbps μέχρι 128Kbps μέσω του πρωτοκόλλου IP. Το συγκεκριμένο πρότυπο αξιοποιεί πολύ καλύτερα το διαθέσιμο εύρος ζώνης, δίνονται ταυτόχρονα μηχανισμούς οι οποίοι μπορούν να κάνουν ανάκτηση από λάθη. Σε σχέση με το H.261, το πρότυπο αυτό μπορεί συχνά να επιτύχει την ίδια τηλεοπτική ποιότητα με λιγότερο από το μισό αριθμό κομματιών στο κωδικοποιημένο ρεύμα. Μπορεί να παρέχει σχεδόν την ίδια ποιότητα στις μισές τιμές εύρους ζώνης. Αυτός είναι ο λόγος για τον οποίο είναι προτιμότερο τηλεοπτικός κωδικοποιητής-αποκωδικοποιητής από το H.261. Υποστήριζε περισσότερες αναλύσεις σε σχέση με τον προκάτοχο του και μερικές από αυτές είναι Sub-QCIF, QCIF, CIF, 4CIF και 16CIF [13].

2.3.6. MPEG-4 Part 2

Το πρότυπο MPEG-4 ολοκληρώθηκε το 2000 από την ομάδα MPEG και η επίσημη ονομασία

του είναι ISO/IEC 14496. Είναι ένα ανοικτό πρότυπο που αναπτύσσει την συμπίεση ήχου και βίντεο επιτρέποντας την εύκολη μεταφορά τους στο δίκτυο. Η ανάλυση εικόνας που μπορεί να υποστηρίξει είναι 176x144 και η μετάδοση γίνεται σε σχετικά χαμηλούς ρυθμούς που κυμαίνονται ανάμεσα στο 4.8 Kbits/sec και 64 Kbits/sec. Υποστηρίζει πρόσβαση, χειρισμό και ξεχωριστή κωδικοποίηση των αντικειμένων. Είναι το πρότυπο που επιτρέπει κωδικοποίηση βίντεο βασισμένη στο περιεχόμενο. Αντί της frame-based συμπίεσης, το MPEG-4 Part 2 μπορεί να διαχωρίσει το βίντεο σε Video Object Plabes. Η ξεχωριστή κωδικοποίηση αντικειμένων μπορεί να δώσει πολύ μεγαλύτερη συμπίεση. Οι λειτουργίες πρόσβασης και χρήσης αντικειμένων επιτρέπουν τη χρήση γραφικών βίντεος και επιτρέπουν επίσης κωδικοποίηση γραφικών αντικειμένων [14].

2.4. H.264 / AVC MPEG-4 Part 10

Το H.264/MPEG4 Part 10 ή αλλιώς AVC(Advanced Video Coding) είναι ένα πρότυπο συμπίεσης κινούμενης εικόνας και ο τελικός σχεδιασμός αυτού του προτύπου ολοκληρώθηκε το 2003. Το H.264 βασίζεται στην διανυσματική κίνηση των μπλοκ εικοινοστοιχείων και αναπτύχθηκε από την ITU-T(VCEG) μαζί με την επιτροπή ISO/ICE (MPEG). Για να διατηρείται μια συνεκτικότητα στην τεχνική τεκμηρίωση, τα πρότυπα ITU-T H.264 και ISO/IEC MPEG-4 AVC αναπτύσσονται και εξελίσσονται παράλληλα. Το H.264 χρησιμοποιείται σε εφαρμογές όπως οι οπτικοί δίσκοι Blue-ray, δικτυακό video Youtube, καθώς και σε χρήσεις όπως iTunes, DVB επίγειες ή δορυφορικές μεταδόσεις εικόνας, καλωδιακή τηλεόραση ή σε ζωντανές τηλεσυνδιασκέψεις εικόνας .

Σε σχέση με παλαιότερα πρότυπα, το πρότυπο H.264 είχε ως αρχικό σκοπό να παρέχει καλή ποιότητα εικόνας σε σχετικά χαμηλούς ρυθμούς μετάδοσης δεδομένων χωρίς ωστόσο να υπάρξει αύξηση της πολυπλοκότητας του σχεδιασμού, κάτι που θα καθιστούσε μη πρακτικό ή υπερβολικά ακριβό κατά την υλοποίηση του. Ένας ακόμα στόχος που έθεσε το πρότυπο αυτό ήταν να παρέχει ευελιξία ως προς την εφαρμοστικότητα του σε πληθώρα εφαρμογών, κάθε φορά σε διάφορα δίκτυα ή συστήματα, συμπεριλαμβανομένου χαμηλών και υψηλών ρυθμών μετάδοσης ή ανάλυσης εικόνας, εκπομπές, αποθήκευση DVD, RTP/IP δίκτυα καθώς και πολυμεσικά τηλεφωνικά συστήματα τύπου ITU-T.

Το πρότυπο αυτό είναι σχεδιασμένο έτσι ώστε να παρέχει τεχνικές λύσεις συμπεριλαμβάνοντας μεταξύ άλλων την αναμετάδοση μέσω καλωδίων, μέσω δορυφόρου, καλωδιακού modem, DSL και επίγειας, διαδραστική και σειριακή αποθήκευση σε οπτικούς

και μαγνητικούς δίσκους όπως DVD, αποθήκευση και διανομή επαγγελματικών ταινιών και βίντεο, βίντεο κατά παραγγελία(video-on-demand) ή υπηρεσίες πολυμέσων συνεχούς ροής μέσω καλωδιακού μόντεμ, xDSL, τοπικού δικτύου (LAN), ISDN, και ασύρματων δικτύων, συνομιλητικές υπηρεσίες μέσω Ethernet, LAN, xDSL, ISDN, ασύρματων και κινητών δικτύων, και modems, και τέλος υπηρεσιών μηνυμάτων με πολυμέσα (MMS) πάνω σε xDSL, Ethernet, LAN, ISDN, ασύρματα και κινητά δίκτυα. Χρησιμοποιείται ιδιαίτερα για κωδικοποίηση στους δίσκους Blu-ray και για live streaming. Αυτό το ευρύ φάσμα εφαρμογών που καλύπτει το πρότυπο αυτό είναι ο λόγος που είναι ένα από τα πιο διαδεδομένα και επιτυχημένα πρότυπα κωδικοποίησης [15].

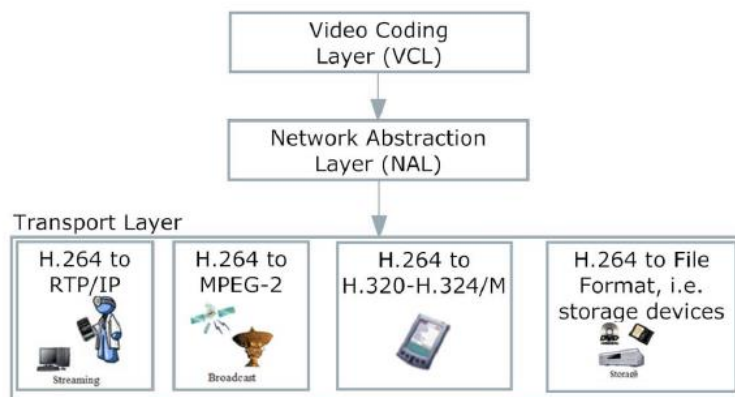
2.4.1. Αρχιτεκτονική Προτύπου

Μια καινοτομία που εισάγει το H.264 είναι ο νοητός διαχωρισμός του κωδικοποιητή σε δύο βασικά στρώματα: το στρώμα κωδικοποίησης εικονοσειρών (VCL-Video Coding Layer) και το στρώμα δικτύου (NAL-Network Abstraction Layer) [16]. Ο νοητός αυτός διαχωρισμός του σε στρώματα του παρέχει την ευελιξία και την προσαρμοστικότητα που είχε ως στόχο. Το στρώμα κωδικοποίησης εικονοσειρών είναι υπεύθυνο για την αναπαράσταση και κωδικοποίηση της εικονοσειράς. Το στρώμα δικτύου αναλαμβάνει να προσαρμόσει την αναπαράσταση του στρώματος VCL και παρέχει πληροφορίες για το header για να μπορεί να γίνει η μεταφορά τους από διάφορα στρώματα μεταφοράς ή η αποθήκευσή τους σε μέσα αποθήκευσης.

Το στρώμα κωδικοποίησης εικονοσειρών ακολουθεί υβριδική προσέγγιση κωδικοποίησης βασισμένη σε μπλοκ στην οποία κάθε εικόνα αναπαριστάτε από ορθογώνιες επιφάνειες που καλούνται macroblocks. Δεν υπάρχει συγκεκριμένο στοιχείο της διαδικασίας κωδικοποίησης που στην οποία οφείλεται η βελτίωση στην ικανότητα συμπίεσης, σε σχέση με τα προηγούμενα πρότυπα συμπίεσης. Αντίθετα, οι μικρές βελτιώσεις κάθε χαρακτηριστικού, όλες μαζί, δίνουν το τελικό αποτέλεσμα που κάνει το πρότυπο αυτό αποδοτικό.

Το στρώμα δικτύου είναι σχεδιασμένο έτσι ώστε να προσφέρει φιλικότητα προς το δίκτυο και να κάνει απλή και αποδοτική την προσαρμογή της πληροφορίας που προέρχεται από το στρώμα κωδικοποίησης σε μια πλειάδα εφαρμογών και συστημάτων. Τα βασικά στοιχεία του στρώματος αυτού παρουσιάζονται ως εξής: RTP/IP για όλα τα είδη ενσύρματων ή ασύρματων υπηρεσιών πραγματικού χρόνου. Διάφοροι τύποι αρχείων όπως για παράδειγμα ISO MP4 και MMS. H.329 για ενσύρματες και ασύρματες διαδραστικές υπηρεσίες. Συστήματα MPEG2 για υπηρεσίες ευρείας μετάδοσης.

Ένα σχηματικό διάγραμμα αυτού του νοητικού διαχωρισμού σε στρώματα μπορούμε να διακρίνουμε στο πιο κάτω σχήμα.



Σχήμα 2.2: Τα στρώματα VCL και NAL του προτύπου H.264/AVC [16].

2.4.2. Διαδικασία κωδικοποίησης και αποκωδικοποίησης

Κατά την κωδικοποίηση [15] το κάθε πλαίσιο του βίντεο επεξεργάζεται σε macroblocks 16x16. Στο αρχικό στάδιο γίνεται πρόβλεψη του macroblock. Αυτό μπορεί να γίνει με δυο τρόπους που έχουν ως εξής: intra prediction όπου η πρόβλεψη του macroblock γίνεται με την χρήση κάποιας προηγούμενης κωδικοποίησης ή ακόμα και από το ίδιο το πλαίσιο. Η δεύτερη πρόβλεψη ονομάζεται inter prediction και γίνεται από πλαίσια που έχουν ήδη κωδικοποιηθεί και έχουν ήδη αποσταλεί

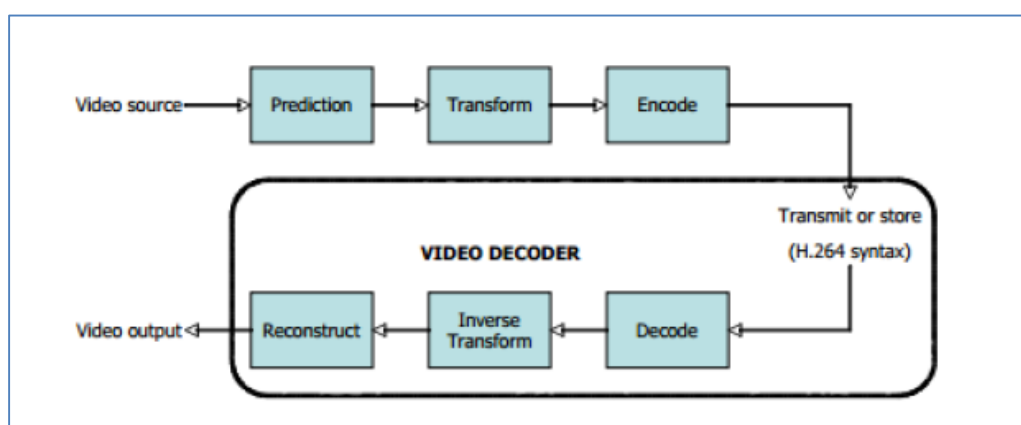
Στην συνέχεια ένα σετ από δείγματα μετασχηματίζεται χρησιμοποιώντας ένα 4x4 ή 8x8 μετασχηματισμό ακέραιων παρόμοιο με τον Διακριτό Μετασχηματισμό Συνημίτονου. Ο μετασχηματισμός αυτός παράγει ένα σύνολο από συντελεστές όπου ο καθένας χρησιμοποιείται σαν βάρος στο σετ των δειγμάτων για την πρότυπη βάση της ακολουθίας. Συνδυάζοντας την πρότυπη βάση με βάρη ξαναδημιουργούνται τα δείγματα στο αρχικό σετ.

Αφού γίνει ο μετασχηματισμός, η έξοδος του μετασχηματισμού κβαντοποιείται, δηλαδή ο κάθε συντελεστής διαιρείται με μια ακέραια τιμή. Κατά την κβαντοποίηση η ακρίβεια του μετασχηματισμού μειώνεται σύμφωνα με την παράμετρο κβαντοποίησης και έτσι μια μεγάλη τιμή της παραμέτρου σημαίνει και μεγαλύτερη συμπίεση, άρα μείωση της ποιότητας της εικόνας.

Κατά την κωδικοποίηση, παράγεται ένας αριθμός από τιμές τις οποίες πρέπει να κωδικοποιήσει σε μορφή του συμπιεσμένου bitstream. Οι τιμές αυτές είναι οι κβαντοποιημένοι μετασχηματισμένοι συντελεστές πληροφορίας και στόχος τους είναι να βοηθήσουν τον αποκωδικοποιητή να αναπαράγει την πρόβλεψη, πληροφορίες που αφορούν την δομή των συμπιεσμένων δεδομένων και των εργαλείων που χρησιμοποιήθηκαν κατά την κωδικοποίηση. Επίσης περιλαμβάνει πληροφορίες που αφορούν ολόκληρη την ακολουθία του βίντεο. Οι πιο πάνω τιμές μετατρέπονται σε δυαδικούς κώδικες χρησιμοποιώντας variable length encoding και/ή arithmetic coding.

Όσο αφορά την αποκωδικοποίηση, ο αποκωδικοποιητής παραλαμβάνει ένα συμπιεσμένο H.264 bitstream και αποκωδικοποιεί το κάθε στοιχείο του και εξάγει πληροφορίες. Στην συνέχεια, για να επανέλθουν οι κβαντοποιημένοι μετασχηματισμένοι συντελεστές στην αρχική τους κλίμακα, ανακλιμακώνονται πολλαπλασιάζοντας τους με μια ακέραια τιμή. Η διαδικασία αυτή είναι η αντίστροφη διαδικασία που ακολουθήθηκε κατά τον μετασχηματισμό και την κβαντοποίηση για να ξαναδημιουργηθεί η ακολουθία των πλαισίων του βίντεο. Για το κάθε macroblock ο αποκωδικοποιητής δημιουργεί μια παρόμοια πρόβλεψη σύμφωνα με αυτή που δημιούργησε ο κωδικοποιητής.

Η διαδικασία κωδικοποίησης και αποκωδικοποίησης που επεξηγήθηκε πιο πάνω, αναπαριστάτε στο πιο κάτω σχήμα.



Σχήμα 2.3: Διαδικασία Κωδικοποίησης και Αποκωδικοποίησης [17].

2.4.3. Profiles και Levels

Το πρότυπο H.264 μπορεί να θεωρηθεί έως μια «οικογένεια προτύπων», τα μέλη των οποίων είναι τα προφίλ. Ένας συγκεκριμένος αποκωδικοποιητής, αποκωδικοποιεί τουλάχιστον ένα, αλλά όχι απαραίτητα όλα τα προφίλ. Το H.264 ορίζει 17 σύνολα χαρακτηριστικών, τα οποία συνήθως συναντώνται έως προφίλ και στοχεύουν κυρίως σε συγκεκριμένες κατηγορίες εφαρμογών. Τα τέσσερα βασικά profiles είναι το Baseline, το Extended, Main και το High.[15]

Το Baseline στοχεύει σε εφαρμογές χαμηλού κόστους που απαιτούν επιπλέον ευρωστία, απώλεια δεδομένων για γίνεται χρήση και αυτού σε κινητές εφαρμογές και τηλεσυνδιασκέψεις. Είχε σχεδιαστεί για εφαρμογές χαμηλής καθυστέρησης, καθώς και για εφαρμογές που τρέχουν σε πλατφόρμες με χαμηλή επεξεργαστική ισχύ και σε περιβάλλον με υψηλή απώλεια πακέτων.

Το Extended προοριζόταν για χρήση σε εφαρμογές δικτυακού video, καθώς το προφίλ αυτό έχει δυνατότητες υψηλής συμπίεσης και ορισμένα επιπλέον χαρακτηριστικά για απώλεια δεδομένων, αναξιοπιστία καναλιού μετάδοσης καθώς και σε περίπτωση εναλλαγής των πηγών στους παροχείς video. Σχεδιάστηκε ουσιαστικά για υπηρεσίες πολυμέσων στο διαδίκτυο.

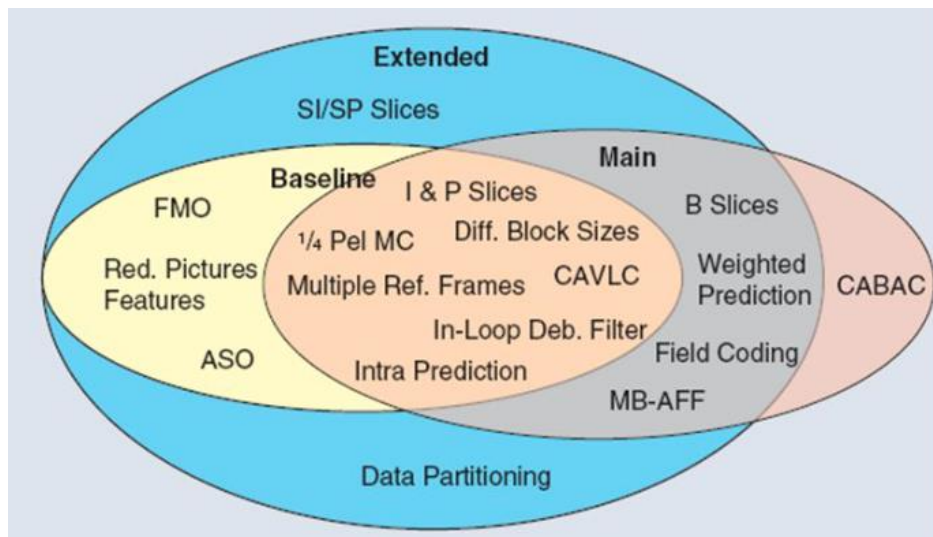
Το Main προφίλ χρησιμοποιείται κυρίως σε εφαρμογές ψηφιακών τηλεοπτικών μεταδόσεων (DVB) με χρήση του φορμά MPEG-4 όπως αυτό ορίζεται στο πρότυπο DVB. Ωστόσο δεν χρησιμοποιείται συνήθως σε μεταδόσεις τηλεόρασης υψηλής ευκρίνειας (HDTV), καθώς αναπτύχθηκε μετέπειτα το High Profile το 2004 για χρήση στις περιπτώσεις εφαρμογών σαν αυτή.

Τέλος το High προφίλ είναι το βασικό προφίλ για εφαρμογές μεταδόσεων και οπτικών δίσκων, συγκεκριμένα για τηλεοπτικά προγράμματα υψηλής ευκρίνειας (HDTV) (π.χ. το προφίλ αυτό υιοθετήθηκε από το φορμά οπτικών δίσκων Blue-ray και το πρότυπο DVB για τηλεοπτικές μεταδόσεις υψηλής ευκρίνειας).

Γενικότερα στα τηλεϊατρικά συστήματα το Baseline θεωρείται το καταλληλότερο προφίλ αφού τις περισσότερες φορές γίνεται χρήση κινητών συσκευών, τα οποία έχουν χαμηλή

επεξεργαστική ισχύ. Ακόμα, σε περιπτώσεις βίντεο streaming όπου υπάρχει περίπτωση να χαθούν αρκετά πακέτα, το Baseline προφίλ είναι και πάλι το ιδανικότερο.

Σαν όρος που έχει χρησιμοποιηθεί κατά την ανάπτυξη του προτύπου H.264, το “επίπεδο” (level) αποτελεί ένα σύνολο περιορισμών που περιγράφουν την ελάχιστη απαίτηση απόδοσης της αποκωδικοποίησης που εκτελεί ένα συγκεκριμένο προφίλ. Για παράδειγμα το επίπεδο υποστήριξης σε ένα προφίλ θα καθορίσει την μέγιστη ανάλυση εικόνας, ρυθμό καρέ (frame rate), και έναν ρυθμό μετάδοσης που δύναται να χρησιμοποιήσει ο αποκωδικοποιητής. Όταν αυτός υπακούει σε ένα συγκεκριμένο επίπεδο των χαρακτηριστικών αυτών θεωρούμε ότι είναι επίσης σε θέση να αποκωδικοποιήσει ροές δεδομένων (bitstreams) του επιπέδου αυτού και όλων των χαμηλότερων επιπέδων.



Σχήμα 2.4: Χαρακτηριστικά των διάφορων profile του προτύπου H.264/AVC [18]

Κεφάλαιο 3

Ασύρματες Τεχνολογίες Μετάδοσης Βίντεο και Πρωτόκολλα Επικοινωνίας

3.1. Ασύρματα Δίκτυα	20
3.1.1. Πρώτη Γενεά 1G	21
3.1.2. Δεύτερη Γενεά 2G	22
3.1.3. Δεύτερη και Μισό Γενεά 2.5G	22
3.1.4. Τρίτη Γενεά 3G	23
3.1.5. Τρίτη και Μισό Γενεά 3.5G	24
3.1.6. Τέταρτη Γενεά 4G	24
3.1.7. Πέμπτη Γενεά 5G	25
3.2. Πρωτόκολλα Επικοινωνίας	25
3.2.1. Πρωτόκολλο UDP	26
3.2.2. Πρωτόκολλο TCP	26
3.2.3. Πρωτόκολλο RTP-RTCP	27

3.1. Ασύρματα Δίκτυα

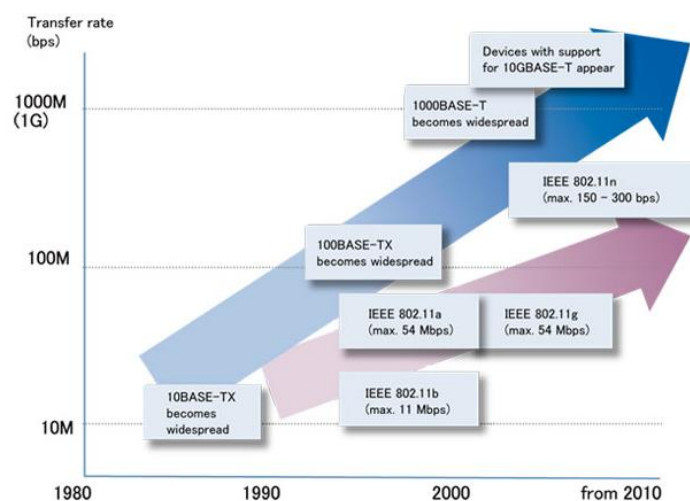
Τα τελευταία χρόνια, ο τρόπος με τον οποίο εργάζονται οι άνθρωποι σε ηλεκτρονικούς υπολογιστές έχει αλλάξει σε μεγάλο βαθμό. Πλέον, η εικόνα ενός ανθρώπου ο οποίος κάθεται στο γραφείο του για να χρησιμοποιήσει τον προσωπικό του υπολογιστή, δεν είναι ένας κοινός τόπος καθώς η ανάγκη για συνεχή χρήση κάποιας μορφής ηλεκτρονικού υπολογιστή, ασχέτως του σημείου που βρίσκεται, είναι διαρκώς αυξανόμενη. Η ανάγκη αυτή για χρήση κάποιου είδους ηλεκτρονικού υπολογιστή, οδήγησε στην δημιουργία αλλά και την ανάπτυξη των ασύρματων δικτύων.

Με τον όρο ασύρματα δίκτυα [19] εννοούμε τα τηλεπικοινωνιακά δίκτυα, στα οποία εντάσσονται συνήθως το τηλεφωνικό δίκτυο ή το δίκτυο υπολογιστών. Τα δίκτυα αυτά χρησιμοποιούν ραδιοκύματα ως φορείς πληροφορίας και τα δεδομένα μεταφέρονται μέσω των ηλεκτρομαγνητικών κυμάτων, με συχνότητα φέροντος η οποία εξαρτάται κάθε φορά από

τον ρυθμό μετάδοσης δεδομένων που απαιτείται να υποστηρίξει το δίκτυο. Η διαφορά των ασύρματων δικτύων με των ενσύρματων δικτύων, είναι το γεγονός ότι στα ασύρματα δίκτυα δεν χρησιμοποιείται ως μέσο μετάδοσης κάποιος τύπος καλωδίου. Στις μέρες μας όλα τα ασύρματα δίκτυα βασίζονται σε ψηφιακή τεχνολογία και, επομένως, κατά μια έννοια, είναι ουσιαστικά δίκτυα υπολογιστών.

Σε γενικές γραμμές, οι ενσύρματες τεχνολογίες έχουν σαφώς καλύτερες αποδόσεις από ότι οι ασύρματες τεχνολογίες [20]. Ο λόγος που συμβαίνει αυτό είναι το γεγονός ότι η ταχύτητα εξαρτάται από το διαθέσιμο φάσμα συχνοτήτων και από το SNR (Signal to Noise Ratio). Η απόδοση των ενσύρματων τεχνολογιών είναι μέχρι και δέκα φορές καλύτερη από την απόδοση των ασύρματων δικτύων. Οι γραμμές από οπτικές ίνες έχουν το μεγαλύτερο φάσμα και τον λιγότερο θόρυβο. Στα ασύρματα περιβάλλοντα ο χρήστης μεταβιβάζεται από μια κυψέλη σε άλλη. Για αυτό τον λόγο, η κινητικότητα είναι η κύρια χρησιμότητα των ασύρματων δικτύων, αλλά αυτή είναι που τα καθιστά πιο αργά από τα ενσύρματα δίκτυα.

Στην πιο κάτω εικόνα φαίνεται η εξέλιξη των ασύρματων τεχνολογιών, σε σχέση με τις ενσύρματες τεχνολογίες.



Σχήμα 3.2: Εξέλιξη ασύρματων και ενσύρματων τεχνολογιών [21].

3.1.1. Πρώτη γενεά 1G

Η πρώτη γενιά (1G) είναι ουσιαστικά η πρώτη γενιά της τεχνολογίας ασύρματης τηλεφωνίας και κινητών τηλεπικοινωνιών. Αυτά είναι τα πρότυπα των αναλογικών τηλεπικοινωνιών, που εισήχθησαν στη δεκαετία του 1980 και συνεχίστηκε μέχρι, που αντικαταστάθηκε από τη δεύτερη γενιά (2G) ψηφιακών τηλεπικοινωνιών. Χρησιμοποιείται τεχνολογία FM(διαμόρφωση συχνότητα), FDD (συχνοδιαιρετική αμφίδρομη επικοινωνία) και FDMA (πολυπλεξία στη συχνότητα).

Η πρώτη γενεά συμπεριλάμβανε διάφορα πρότυπα και το κάθε ένα από τα πρότυπα αυτά κατατασσόταν σε μια διαφορετική χώρα. Για παράδειγμα, κάποια από τα πρότυπα αυτά ήταν το Advanced Mobile Phone Service (AMPS) , Nordic Mobile Telephone (NMT) και Total Access Communication System (TACS). Τα πρότυπα από τα οποία αποτελείτο η πρώτη γενεά, χρησιμοποιούσαν όλα τεχνικές διαμόρφωσης συχνότητας για τις τηλεφωνικές κλήσεις. Για τον λόγο αυτό, όλες οι αποφάσεις που αφορούσαν τον συντονισμό των κλήσεων λαμβάνονταν από κεντρικούς σταθμούς. Το γεγονός ότι χρησιμοποιούταν μόνο για τηλεφωνικές κλήσεις, δεν χρειαζόταν μεγάλη ταχύτητα, έτσι η ταχύτητα του ήταν μέχρι 2.4Kbit/sec. Τέλος, το φάσμα των συχνοτήτων χωριζόταν σε κανάλια και σε κάθε κλήση παραχωρούνταν δύο κανάλια [22].

3.1.2. Δεύτερη γενεά 2G

Η δεύτερη γενεά ασύρματων δικτύων εισήχθηκε το 1991 και είναι ο διάδοχος της πρώτης γενεάς κινητής τηλεφωνίας. Η δημιουργία της δεύτερης γενεάς, προέκυψε λόγω του χαμηλού ρυθμού μετάδοσης δεδομένων και την ανεπαρκή διαβίβαση των δεδομένων καθώς επίσης και λόγω των αναλογικών σημάτων ομιλίας, της πρώτης γενεάς. Έτσι με την γενεά αυτή υπάρχει η δυνατότητα επικοινωνίας δεδομένων υψηλών ταχυτήτων αλλά και η δυνατότητα μετάδοσης φωνής. Για αυτό τον λόγο η γενεά αυτή κυμαίνεται σε πιο γρήγορες ταχύτητες που φτάνουν μέχρι 64Kbits το δευτερόλεπτο.

Η αναλογική τεχνολογία της πρώτης γενιάς αντικαταστάθηκε από την ψηφιακή τεχνολογία σε ασύρματα συστήματα 2G και έτσι αντί των τεχνολογιών που χρησιμοποιούνταν, χρησιμοποιούνται στη δεύτερη γενιά τεχνικές ψηφιακής διαμόρφωσης. Οι τεχνολογίες αυτές, χωρίζονταν σε δύο κατηγορίες ανάλογα με τον τρόπο πολυπλεξίας. Οι δύο αυτές κατηγορίες

έχουν ως εξής: Time Division Multiple Access (TDMA) και Code Division Multiple Access (CDMA). Τα βασικά πρότυπα του ήταν το GSM που λειτουργούσε με TDMA, το cdmaOne που λειτουργούσε με CDMA, D-AMPS που λειτουργούσε με TDMA και το Pacific Digital Cellular που λειτουργούσε με TDMA [23].

3.1.3. Δεύτερη και μισό Γενεά 2.5G

Η δεύτερη και μισό γενεά 2.5G θεωρείται το μεταβατικό στάδιο ανάμεσα στα δίκτυα 2G και 3G. Ουσιαστικά είναι η εξέλιξη της δεύτερης γενεάς η οποία υποστηρίζει μεγαλύτερες ταχύτητες οι οποίες φτάνουν μέχρι και 144 Kbit/sec. Η κύρια διαφορά μεταξύ αυτής της γενεάς και της γενεάς 2G είναι το γεγονός ότι το δίκτυο της 2G ήταν circuit switching ενώ της 2.5G γενεάς ήταν packet switching. Τα πρότυπα που προέκυψαν από την δεύτερη και μισό γενεά, επέτρεπαν στον υπάρχοντα εξοπλισμό της δεύτερης γενεάς να τροποποιηθεί για να μπορεί να υποστηρίζει υπηρεσίες όπως για παράδειγμα πλοήγηση στο διαδίκτυο, αποστολή ηλεκτρονικών μηνυμάτων και πολλά άλλα.

Στην γενεά αυτή εισάχθηκαν τρία βασικά πρότυπα τα οποία έχουν ως εξής: High Speed Circuit-Switched Data (HSCSD), General Packet Radio Service (GPRS) και Enhanced Data Rates for GSM Evolution (EDGE). Το HSCSD είναι μια εξέλιξη του GSM. Ένα τερματικό μπορεί να χρησιμοποιεί δυο, τρεις ή και τέσσερις σχισμές ανά πλαίσιο, προσφέροντας ταχύτητες 28.8, 43.2 και 57.6 kbps. Υποστηρίζει επίσης τις ασύμμετρες συνδέσεις, όταν δηλαδή οι κατωφερείς με τις ανωφερείς συνδέσεις έχουν διαφορά. Το GPRS βασιζόταν στην λειτουργία μεταγωγής πακέτου και έτσι μπορούσε να παρέχει την πρόσβαση στο διαδίκτυο. Χρησιμοποιούσε 8 time slots και είχε θεωρητικές ταχύτητες μετάδοσης δεδομένων 144Kbit/s. Τέλος, το EDGE είναι ένα αντιπροσωπευτικό πρότυπο του 2.7G και επιτυγχάνει ταχύτητες της τάξης 236.8 Kbps για 4 time slots. [23]

3.1.4. Τρίτη Γενεά 3G

Η τρίτη γενεά ασύρματων δικτύων 3G εμφανίστηκε το 1990 όμως εφαρμόστηκε το 2001. Η γενεά αυτή ήταν ουσιαστικά μια καινοτομία αφού με την ανάπτυξη νέας αρχιτεκτονικής, έφερε στο φως ένα σπουδαίο πλήθος εφαρμογών που μέχρι εκείνη τη στιγμή κανένα από τα προηγούμενα δίκτυα δεν είχε καταφέρει να προσφέρει. Αρχικά η γενεά αυτή είχε σχεδιαστεί για να πρόσβαση στο διαδίκτυο με υψηλότερες ταχύτητες και για τα διάφορα είδη εφαρμογών περιήγησης στο διαδίκτυο. Με τον τρόπο αυτό, οι χρήστες μπορούσαν πλέον να πλοηγούνται στο διαδίκτυο με ταχύτητες επιπέδου Megabit, να επικοινωνούν μέσω Voice

over Internet Protocol (VoIP), να κατεβάζουν μουσική και ταινίες και να χρησιμοποιούν μια πληθώρα εφαρμογών. Κάτι ακόμα πιο ελκυστικό που πρόσφερε η γενεά αυτή ήταν η τηλεδιάσκεψη, μέσω της οποίας επιτρεπόταν η επικοινωνία πρόσωπο με πρόσωπο σε πολλαπλούς χρήστες ακόμα και αν βρίσκονταν σε μεγάλη απόσταση.

Το πρότυπο αυτό κατηγοριοποιείται στους εξής δυο τύπους: 3GPP και 3GPP2. Το πρότυπο 3GPP περιλαμβάνει το WCDMA, το TD-SCDMA (Time Division Synchronous Code Division Multiple Access) και το EDGE (Enhanced Data for GSM Evolution). Ο ρυθμός μετάδοσης (bitrate) του 3GPP φτάνει θεωρητικά μέχρι και τα 11Mbps. Η τεχνολογία 3GPP2 είναι συμβατή με το πρότυπο 2G CDMA. Το 3G-CDMA2000 1xRTT CDMA υποδηλώνει ένα μοναδικό κανάλι στα 1,25MHz. Ο ρυθμός των δεδομένων είναι θεωρητικά έως και 14.7Mbps. [24]

3.1.5. Τρίτη και μισό Γενεά 3.5G

Η τρίτη και μισό γενεά ασύρματων δικτύων 3.5G εισήχθηκε το 2005 και περιλαμβάνει όλα τα δίκτυα που ενσωματώνουν την τεχνολογία High Speed Downlink Packet Access (HSDPA). Η τεχνολογία HSDPA είναι ουσιαστικά η εξέλιξη των WCDMA δικτύων που συναντήσαμε στην τρίτη γενεά δικτύων, και αυτή η εξέλιξη αύξησε τον ρυθμό ταχύτητας κατεβάσματος με θεωρητική ταχύτητα κατεβάσματος 84Mbps/sec και θεωρητική ταχύτητα ανεβάσματος 21Mbps/sec. Στην τρίτη γενεά ασύρματων δικτύων είχαμε καθυστερήσεις από 180-200ms και στην γενεά αυτή μειώθηκε η καθυστέρηση σε 100ms. Με την εισαγωγή του High Speed Downlink Shared Channel (HS-DSC), αυξήθηκε η χωρητικότητα των δεδομένων που μπορεί να αποθηκευτεί σε κάθε σταθμό βάσης, μέχρι και 5 φορές περισσότερα από ότι ήταν πριν.

Το High Speed Uplink Packet Access (HSUPA) αποτελεί ουσιαστικά μια εξέλιξη του HSDPA. Το HSUPA προσφέρει στον χρήστη γρηγορότερες ταχύτητες όσο αφορά κατέβασμα, μέχρι και 5.76 Mbps και κυκλοφόρησε ως μέρος της 3GPP Release 6 που αναφέραμε πιο πάνω. Το High Speed Uplink Packet Access μειώνει ακόμα, την καθυστέρηση στο uplink καθιστώντας της συνολική καθυστέρηση του συστήματος αποδεκτή για εφαρμογές πραγματικού χρόνου. Τέλος, το HSUPA μπορεί να πετύχει την σπατάλη λιγότερης ενέργειας της συσκευής που το χρησιμοποιεί, αφού κάνει χρήση συστημάτων διαμόρφωσης κατώτερης τάξης [25].

Στην τρίτη και μισό γενεά ασύρματων δικτύων 3.5G ανήκει και το IEEE 802.16e πρωτόκολλο. Το WiMAX σε γενικές γραμμές, λειτουργεί με σχεδόν τον ίδιο τρόπο όπως και

το σύνηθες Wi-fi, με την μόνη διαφορά, ότι εξασφαλίζει εμβέλεια επικοινωνίας 35 και παραπάνω χιλιομέτρων σε αντίθεση με τα 100 μέτρα που εξασφαλίζει το Wi-fi. Το 802.16e δημιουργήθηκε το 2005 και χρησιμοποιεί τύπο συχνότητας scalable orthogonal frequency-division multiple access (SOFDMA). Το πρωτόκολλο IEEE 802.16e-2005 βελτιώνεται με το πρωτόκολλο 802.16-2004 με την διαφορά ότι έχουμε διαβάθμιση του συστήματος FFT(Fast Fourier Transform) στο εύρος ζώνης των καναλιών έτσι ώστε να διατηρείται σταθερό το επίπεδο μεταφοράς δια μέσου καναλιών διαφορετικού εύρους (1.25 -20 MHz). Με το 802.16e έχουμε επίσης την ικανότητα κάλυψης σε περίπτωση που ο χρήστης κινείται [26].

3.1.6. Τέταρτη γενεά 4G

Η τέταρτη γενεά 4G εισάχθηκε για πρώτη φορά το 2008 από την International Telecommunications Union-Radio communications sector (ITU-R). Η τεχνολογία 4G στηρίζεται στις τεχνολογίες WiMAX 802.16m και LTE Advanced. Η γενεά αυτή μειώνει τις καθυστερήσεις σε 10ms από 100ms που συναντήσαμε στην τρίτη και μισό γενεά και έτσι μπορούν πλέον να χρησιμοποιηθούν εφαρμογές όπως κινητή τηλεόραση υψηλής ευκρίνειας, 3D τηλεόραση και cloud computing. Με την γενεά αυτή, λόγω των γρήγορων ταχυτήτων αλλά και της μικρής καθυστέρησης που υπάρχει, θα είναι πλέον δυνατή η μετάδοση βίντεο υψηλής ευκρίνειας.

Το IEEE 802.16m πρότυπο είναι η βασική τεχνολογία της δεύτερης έκδοσης του WiMax και η βασική του διαφορά είναι ότι μπορεί να προσφέρει μέχρι και τέσσερις φορές πιο γρήγορη ταχύτητα από ότι η πρώτη έκδοση WiMax. Η τεχνολογία LTE-Advanced είναι ουσιαστικά η πρόοδος του 3G. Χρησιμοποιεί τις τεχνολογίες MIMO (Multiple-Input Multiple-Output) και OFDM (Orthogonal Frequency Division Multiplexing), τα οποία δίνουν προτεραιότητα στην ελαχιστοποίηση των παρεμβολών μεταξύ των καναλιών [27].

3.1.7. Πέμπτη Γενεά 5G

Οι 5G τεχνολογίες περιλαμβάνουν ούλα τα είδη των προηγμένων χαρακτηριστικών που την κάνουν πιο ισχυρή και με μεγάλη ζήτηση στο εγγύς μέλλον. Τα πρότυπα για το 5G δεν έχουν ακόμη διευκρινιστεί και αυτά είναι που θα καθορίσουν ποιες ασύρματες τεχνολογίες θα μπορούν να λέγονται έτσι και επίσης ποια χαρακτηριστικά θα πρέπει να περιλαμβάνουν αλλά και το πόσο γρήγορα θα είναι. Τα πρότυπα αυτά θα οριστούν το 2018, θα κωδικοποιηθούν το 2019 από την Διεθνή Ένωση Τηλεπικοινωνιών και το 5G θα βγει στην αγορά το 2020.

Αν και ακόμη δεν υπάρχει σαφής προσδιορισμός του 5G γενικά ή των προδιαγραφών του, γνωρίζουμε ότι το κυριότερο χαρακτηριστικό της γενεάς αυτής που θα την κάνει να είναι μια επανάσταση στα ασύρματα δίκτυα, είναι η υψηλές ταχύτητες μετάδοσης. Το δίκτυο αυτό θα προσφέρει την στήριξη μεγάλου αριθμού συνδεδεμένων χρηστών ταυτόχρονα καθώς επίσης και ενεργειακή απόδοση. Οι προδιαγραφές αυτές όμως ίσως να απαιτούν νέα πρωτόκολλα και νέες τεχνολογίες πρόσβασης. Θα παρέχει ακόμα, έως και 1Gbit/s ταχύτητα σύνδεσης σε απόσταση μέχρι και δύο χιλιομέτρων. Όπως βλέπουμε, η γενεά αυτή θα κάνει τα κινητά τηλέφωνα να μην ξεχωρίζουν από τους ηλεκτρονικούς υπολογιστές [28].

3.2. Πρωτόκολλα Επικοινωνίας

Για να μπορεί να υπάρχει η επικοινωνία μεταξύ των διαφόρων υπολογιστικών συστημάτων, είναι απαραίτητη η ύπαρξη ενός συνόλου από φυσικές και λογικές διασυνδέσεις διαφόρων ανεξάρτητων τμημάτων. Μια βασική αρχή που εφαρμόζεται σε κάθε επικοινωνία υπολογιστών, είναι η διάσπαση της προς αποστολή πληροφορίας σε μικρότερα τμήματα πληροφορίας, για παράδειγμα πακέτα, και ακολούθως η τμηματοποιημένη αυτή πληροφορία αποστέλλεται με βοήθεια πρωτοκόλλων επικοινωνίας. Με βάση αυτή την αρχιτεκτονική για την μείωση της πολυπλοκότητας, τα δίκτυα έχουν οργανωθεί σε στρώματα, όπου το κάθε στρώμα χτίζεται πάνω σε κάποιο άλλο και το καθένα παρέχει υπηρεσίες στο ανώτερο του επίπεδο.

Ο αριθμός των επιπέδων, ο ρόλος που επιτελεί το καθένα από αυτά και ο τρόπος που υλοποιεί τις υπηρεσίες του εξαρτάται από το συγκεκριμένο δίκτυο. Σε όλα όμως τα δίκτυα υπάρχει το επίπεδο μεταφοράς, το οποίο αποτελούν τα πρωτόκολλα επικοινωνίας. Το επίπεδο μεταφοράς διαχειρίζεται τη συνομιλία μεταξύ των ομότιμων οντοτήτων της πηγής και του προορισμού. Δηλαδή παίρνει τα δεδομένα από μια εφαρμογή και είναι υπεύθυνο για την ορθή αποστολή τους σε άλλη εφαρμογή μέσω του δικτύου, καθώς και το αντίστροφο.

3.2.1. Πρωτόκολλο UDP

Το User Datagram Protocol (UDP) είναι ένα πρωτόκολλο το οποίο δεν παρέχει καμία επιβεβαίωση (acknowledgment) προς το αποστολέα και δεν εγγυάται αξιόπιστη επικοινωνία. Τα πακέτα UDP που αποστέλλονται από τον αποστολέα, μπορεί να μην φτάσουν καθόλου στον παραλήπτη σε περίπτωση που το δίκτυο είναι υπερφορτωμένο. Υπάρχει επίσης η πιθανότητα τα πακέτα να φτάσουν με λάθος σειρά ή ακόμα και να φτάσουν δυο φορές το καθένα. Το γεγονός αυτό, είναι ο λόγος για τον οποίο το πρωτόκολλο αυτό είναι αρκετά

γρήγορο και αποτελεσματικό. Έτσι, χρησιμοποιείται κυρίως όταν οι χρήστες θέλουν ένα έναν γρήγορο τρόπο να στείλουν δεδομένα και δεν είναι σημαντικό για αυτούς να έχουν αυστηρή επιβεβαίωση για το αν τα πακέτα έφτασαν στον παραλήπτη με την σωστή σειρά και τον σωστό τρόπο. Ο μόνος έλεγχος που μπορεί να προσφέρει το πρωτόκολλο αυτό είναι μέσω του checksum, και δίνει πληροφορία για το αν το πακέτο που έχει παραληφθεί έφθασε ολόκληρο. Μερικές από τις εφαρμογές αυτού του πρωτοκόλλου είναι VoIP, DNS, SNMP, RIP και DHCP.

3.2.2. Πρωτόκολλο TCP

Το Transmission Control Protocol (TCP) είναι ένα αξιόπιστο πρωτόκολλο σύνδεσης το οποίο οργανώνει το εισερχόμενο μήνυμα σε διακριτά μηνύματα και το περνά στο επίπεδο διαδικτύου (Network layer). Το πρωτόκολλο αυτό δημιουργεί σύνδεση ανάμεσα στον αποστολέα και τον παραλήπτη και μετά αποστέλλει τα πακέτα. Όταν λαμβάνει ένα μήνυμα κάνει συναρμολόγηση του μηνύματος αυτού για να παραδοθεί στο επόμενο επίπεδο. Το πρωτόκολλο αυτό σε αντίθεση με το UDP πρωτόκολλο αποστέλλει στον αποστολέα επιβεβαίωση (acknowledgment). Η αξιοπιστία του έγκειται στο γεγονός ότι αν κάποιο πακέτο χαθεί κατά την αποστολή του τότε είναι υπεύθυνο να το ξανά στείλει και επίσης αν κάποιο από τα πακέτα στάλθηκε περισσότερο από μια φορές, τότε είναι υπεύθυνο να απορρίψει τις πλεονάζον αποστολές του. Το πρωτόκολλο αυτό είναι ιδανικό σε περιπτώσεις όπου οι καθυστερήσεις είναι ανεκτές και μερικές εφαρμογές που χρησιμοποιούν TCP είναι FTP, SMTP, HTTP και Telnet.

3.2.3 Πρωτόκολλο RTP-RTCP

Το Real Time Protocol (RTP) παρέχει συνεχείς δικτυακές λειτουργίες μεταφοράς, κατάλληλες για εφαρμογές που μεταδίδουν δεδομένα σε πραγματικό χρόνο όπως για παράδειγμα ο ήχος και το βίντεο. Το πρωτόκολλο αυτό δημιουργήθηκε ειδικά για την αποστολή δεδομένων σε πραγματικό χρόνο λόγω του ότι το πρωτόκολλο UDP δεν παρείχε στον αποστολέα καμία εγγύηση για την σωστή μεταφορά των δεδομένων και το TCP πρωτόκολλο δεν ήταν αρκετά γρήγορο για αποστολή σε πραγματικό χρόνο. Με το πρωτόκολλο αυτό υπάρχει η δυνατότητα προσομοίωσης των δεδομένων σε multicast ή unicast διαδικτυακές υπηρεσίες.

Το RTP δεν καταγράφει την δέσμευση των πόρων και δεν εγγυάται ποιότητα υπηρεσίας όσο αφορά τις πραγματικού χρόνου υπηρεσίες. Έτσι, ο λόγος ύπαρξης του Real Time Control

Protocol (RTCP) είναι για να βελτιώνει την μεταφορά δεδομένων με RTP και να παρέχει ποιότητα υπηρεσίες που δεν παρέχει το RTP. Το πρωτόκολλο αυτό ελέγχου, παρακολουθεί την παράδοση των δεδομένων με έναν κλιμακούμενο τρόπο ανάλογα με το μέγεθος του multicast δικτύου, και προσφέρει κάποιο μικρό ποσοστό ελέγχου καθώς και λειτουργία αναγνώρισης. Οι παραλήπτες και ο αποστολέας σε ταχτά χρονικά διαστήματα αποστέλλουν στατιστικά στοιχεία όπως τον αριθμό των πακέτων που χάθηκαν, αριθμό πακέτων που στάλθηκαν, το jitter και το Round Trip Delay Time (RTT) [29].

Κεφάλαιο 4

HTTP adaptive streaming

4.1. Adaptive bitrate streaming	29
4.1.1. Adobe HTTP Dynamic Streaming	30
4.1.2. Apple HTTP Adaptive Streaming	30
4.1.3. Microsoft Smooth Streaming	30
4.2. MPEG-DASH	31
4.2.1. Multimedia Presentation Description (MPD)	31
4.2.2. Μορφή των Segments	32

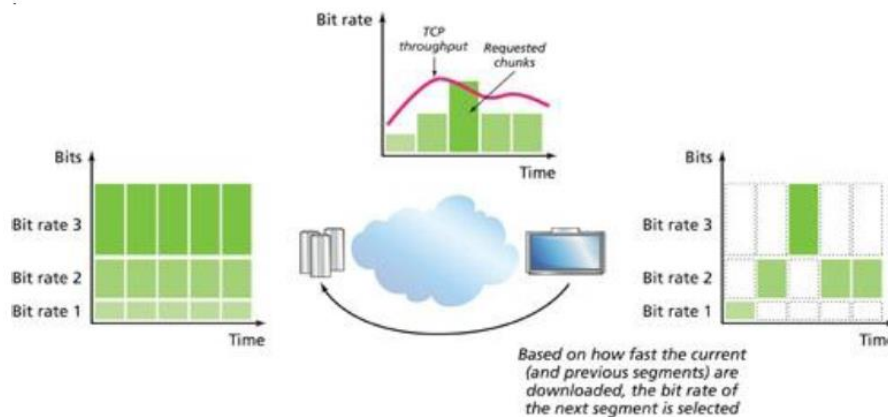
4.1. Adaptive bitrate streaming

Το διαθέσιμο εύρος ζώνης στο διαδίκτυο είναι κάτι εξαιρετικά μεταβλητό. Για να διαχειριστούμε αυτό το χαρακτηριστικό των δικτύων, και να μπορούμε να προσφέρουμε την καλύτερη δυνατή ποιότητα εικόνας σε ένα καλό εύρος ζώνης, είναι απαραίτητη η ύπαρξη μιας adaptive streaming τεχνολογίας. Για να κάνουμε stream επαγγελματικό περιεχόμενο, όπως για παράδειγμα ένα βίντεο με ιατρικά δεδομένο, μέσω του διαδικτύου, μια κοινή προσέγγιση είναι το HTTP Adaptive Streamed (HAS) [30].

Τα adaptive bitrate streaming συστήματα λειτουργούν ανιχνεύοντας το εύρος ζώνης του χρήστη καθώς επίσης και την χωρητικότητα του CPU του σε πραγματικό χρόνο έτσι ώστε να μπορούν να προσαρμόσουν την ποιότητα του βίντεο που θα γίνει stream ανάλογα. Επομένως το βίντεο κωδικοποιείται σε διαφορετικές μορφές και διαφορετικές ποιότητες. Τα HTTP adaptive streaming συστήματα μεταφέρουν τα δεδομένα τους μέσω του πρωτοκόλλου HTTP.

Ένα βίντεο χωρίζεται σε μικρά μη-επικαλυπτόμενα διαστήματα σε μια ακολουθία, μήκους μεταξύ ενός και δέκα δευτερολέπτων, και κάθε χρονικό διάστημα κωδικοποιείται σε διαφορετικά bit rates. Ο χρήστης ενός HTTP adaptive streaming συστήματος ζητά τα κομμάτια αυτά μέσω του HTTP και συνήθως το πρώτο κομμάτι στέλνεται σε ένα χαμηλό bit

rate. Εάν τα κομμάτια παραδίδονται συστηματικά σε χρόνο μικρότερο από το μήκος του διαστήματος, ο Αλγόριθμος Προσδιορισμού του ρυθμού (RDA) στον πελάτη επιλέγει έναν υψηλότερο ρυθμό bit για το επόμενο κομμάτι.



Σχήμα 4.1: HTTP adaptive streaming [30].

Υπάρχουν πολλές υλοποιήσεις του HTTP adaptive Streaming τις οποίες θα δούμε πιο κάτω.

4.1.1. Adobe HTTP Dynamic Streaming

Το HTTP Dynamic Streaming είναι μια διαδικασία αποδοτικής μετάδοσης βίντεο streaming στον χρήστη με δυναμικό switching μέσω διαφορετικών streams που ποικίλουν ως προς την ποιότητα και το μέγεθος κατά την διάρκεια της αναπαραγωγής του βίντεο. Αυτό παρέχει στους χρήστες την καλύτερη δυνατή εμπειρία αναπαραγωγής την οποία μπορεί να υποστηρίξει το εύρος ζώνης τους και η μηχανή που χρησιμοποιούν. Ένας άλλος κύριος στόχος του είναι να κάνει την διαδικασία αναπαραγωγής ομαλή και απρόσκοπτη στους χρήστες, έτσι ώστε όταν είναι αναγκαία η προς τα κάτω ή προς τα πάνω κλιμάκωση της ποιότητας του stream να μην ενοχλείται η κανονική ροή της αναπαραγωγής. Τέλος, προσφέρει υψηλή ποιότητα βίντεο H.264 ή VP6 [31].

4.1.2. Apple HTTP Adaptive Streaming

Το HTTP Live Streaming (HLS) είναι media streaming communication πρωτόκολλο βασισμένο στο HTTP το οποίο υλοποιείται από την Apple Inc. Το HLS υποστηρίζει τόσο την ζωντανή μετάδοση βίντεο όσο και βίντεο σε ζήτηση (on demand). Λειτουργεί, σπάζοντας τα streams ή τα συστατικά του βίντεο σε μικρά MPEG2-TS files με πολλαπλά bit rates και δίνει διάρκεια χρησιμοποιώντας ένα stream ή file τεμαχιστή. Την υλοποίηση ενός τέτοιου

τεμαχιστή παρέχει και η Apple. Ο τεμαχιστής είναι επίσης υπεύθυνος για την παραγωγή ενός συνόλου από index files με M3U8 format το οποίο λειτουργεί σαν playlist αρχείο για τα κομμάτια του βίντεο. Κάθε λίστα σχετίζεται με ένα δεδομένο bitrate και περιέχει τις διευθύνσεις URL στα κομμάτια με το συγκεκριμένο bitrate [32].

4.1.3. Microsoft Smooth Streaming

Το Microsoft Smooth Streaming είναι μια επέκταση του IIS Media Service το οποίο επιτρέπει να γίνεται adaptive streaming στον χρήστη μέσω HTTP. Εισάχθηκε για πρώτη φορά από την Microsoft τον Οκτώβριο του 2008 σαν κομμάτι της αρχιτεκτονικής Silverlight. Δυναμικά, παρακολουθεί το τοπικό εύρος ζώνης και την απόδοση του βίντεο rendering, και με αυτό τον τρόπο το Smooth Streaming είναι ικανό να βελτιστοποιεί το περιεχόμενο της αναπαραγωγής με την αλλαγή της ποιότητας του βίντεο σε πραγματικό χρόνο. Η Microsoft κατάφερε επιτυχώς την παράδοση τόσο της μετάδοσης βίντεο σε πραγματικό χρόνο όσο και την on-demand παράδοση 1080p HD video με το Smooth Streaming σε χρήστες του Silverlight [33].

4.2. MPEG-DASH

Το MPEG-DASH (Dynamic Adaptive Streaming over HTTP) είναι ένα πρότυπο για adaptive streaming μέσω HTTP. Σε αυτό το πρότυπο το περιεχόμενο των multimedia συλλέγεται και αποθηκεύεται σε έναν HTTP server και μεταδίδεται μέσω HTTP. [35] Το περιεχόμενο αυτό αποθηκεύεται σε δύο μέρη τα οποία έχουν ως εξής: το πρώτο μέρος είναι το Media Presentation Description (MPD) και το δεύτερο μέρος είναι τα segments. Το MPD περιγράφει το διαθέσιμο περιεχόμενο, τις διάφορες εναλλακτικές λύσεις όσο αφορά το βίντεο που θα γίνει stream, τις διευθύνσεις URL που βρίσκεται το συγκεκριμένο βίντεο καθώς επίσης και διάφορα άλλα χαρακτηριστικά. Όσο αφορά τα segments, περιέχουν τα πραγματικά multimedia bitstreams σε μορφή τεμαχίων, τα οποία μπορεί να βρίσκονται σε μονά ή πολλαπλά αρχεία.

Για να μπορεί να αναπαραχθεί το περιεχόμενο, το πρώτο και βασικό βήμα είναι ο χρήστης του DASH να αποκτήσει το MPD. Υπάρχουν πολλά είδη μεταφορών μέσω των οποίων το MPD μπορεί να παραδοθεί στον χρήστη, όπως για παράδειγμα μέσω HTTP, thumb drive και broadcast. Στη συνέχεια το MPD αναλύεται και ο χρήστης παίρνει πληροφορίες για την διαθεσιμότητα του media-content, τους τύπους των media, τα resolutions, το μέγιστο και ελάχιστο εύρος ζώνης, τις διάφορες εναλλακτικές κωδικοποιήσεις του περιεχομένου καθώς

επίσης και διάφορα άλλα χαρακτηριστικά του περιεχομένου. Χρησιμοποιώντας αυτές τις πληροφορίες, επιλέγεται η κατάλληλη εναλλακτική κωδικοποίηση και ξεκινά το streaming του περιεχομένου ανακτώντας τα τμήματα μέσω HTTP GET requests.

Ακολουθώντας, ο πελάτης ανακτά τα επόμενα τμήματα και παρακολουθεί ταυτόχρονα τις διακυμάνσεις στο εύρος ζώνης του δικτύου. Ανάλογα με αυτές τις διακυμάνσεις, αποφασίζεται πώς το βίντεο μπορεί να προσαρμοστεί στο διαθέσιμο εύρος ζώνης, ανακτώντας τμήματα διαφορετικών εναλλακτικών λύσεων όπως για παράδειγμα χαμηλότερων ή υψηλότερων bitrates.

4.2.1. Multimedia Presentation Description (MPD)

Το Dynamic HTTP streaming για να γίνει απαιτεί να υπάρχουν στον server διαθέσιμες εναλλακτικές λύσεις, για το περιεχόμενο, για διαφορετικά bitrates. Επίσης το περιεχόμενο μπορεί να αποτελείται από πολλά συστατικά, όπως για παράδειγμα βίντεο ή ήχος, καθένα από τα οποία μπορεί να έχει διαφορετικά χαρακτηριστικά. Στο MPEG-DASH τα χαρακτηριστικά αυτά περιγράφονται μέσω του MPD το οποίο είναι ένα XML έγγραφο.

Το MPD αποτελείται από μια ή περισσότερες περιόδους, όπου κάθε περίοδος είναι ένα χρονικό διάστημα κατά την διάρκεια του χρονικού άξονα. Κάθε περίοδος έχει τον δικό της χρόνο έναρξης, την διάρκεια της και αποτελείται από ένα ή περισσότερα adaptation sets. Το κάθε adaptation παρέχει πληροφορίες για ένα ή περισσότερα συστατικά και τις εναλλακτικές κωδικοποιήσεις για τα συστατικά αυτά. Κάθε adaptation set περιλαμβάνει συνήθως πολλαπλές αναπαραστάσεις.

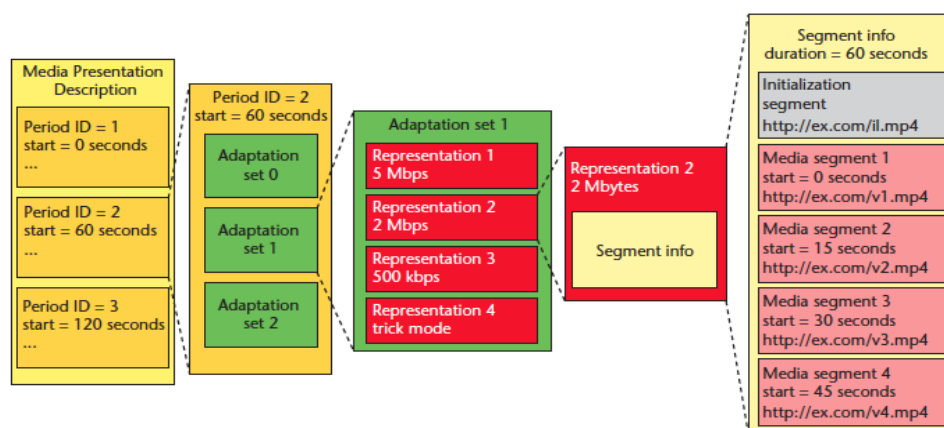
Μια αναπαράσταση είναι ουσιαστικά μια κωδικοποιημένη εναλλακτική του ίδιου συστατικού. Η αναπαράσταση αυτή διαφέρει από τις άλλες ως προς το εύρος ζώνης, το resolution, τον αριθμό των καναλιών ή και άλλα πολλά χαρακτηριστικά. Κάθε αναπαράσταση αποτελείται από ένα ή πολλά segments. Το κάθε segment έχει ένα URL το οποίο είναι μια διευθυνσιοδοτημένη θέση στον server, το οποίο μπορεί να κατέβει μέσω HTTP GET request.

4.2.2. Μορφή των Segments

Το περιεχόμενο είναι προσβάσιμο μέσω μιας συλλογής από segments. Ένα segment ορίζεται ως το σώμα της απόκρισης στο HTTP GET του DASH client. Ένα συστατικό είναι κωδικοποιημένο και μοιράζεται σε πολλά segments. Το πρώτο segment δεν περιλαμβάνει καθόλου πραγματικά δεδομένα για το πολυμέσο. Είναι συνήθως ένα segment αρχικοποίησης που περιλαμβάνει τις απαιτούμενες πληροφορίες για το πώς το ο αποκωδικοποιητής του DASH θα εκκίνηση στην μεριά του client.

Στη συνέχεια το stream διαιρείται σε ένα ή περισσότερα συνεχόμενα segments όπου το καθένα έχει ένα μοναδικό URL, ένα δείκτη, ώρα έναρξης και διάρκεια. Επίσης κάθε segment περιέχει τουλάχιστο ένα σημείο πρόσβασης το οποίο μπορεί να είναι τυχαίας προσπέλασης ή switch-to-point στο stream και η αποκωδικοποίηση μπορεί να αρχίσει χρησιμοποιώντας δεδομένα μόνο από αυτό το σημείο, προς τα εμπρός.

Για να είναι επιτρεπτό το κατέβασμα αυτών των segments σε πολλαπλά μέρη, καθορίζεται μια μέθοδος για σηματοδότηση των sub segments χρησιμοποιώντας ένα index box. Αυτό το index box περιγράφει τα sub segments καθώς και σημεία πρόσβασης του stream στο segment, σηματοδοτώντας της διάρκεια και το byte offset τους. Ο πελάτης DASH μπορεί να χρησιμοποιήσει την πληροφορία αυτή για να ζητήσει τα sub segments χρησιμοποιώντας HTTP GETS.



Σχήμα 4.2: MPD hierarchical data model [34].

Κεφάλαιο 5

Σύστημα Ιατρικής Εκπαίδευσης και Αξιολόγησης Ιατρικού Βίντεο

5.1.	Εφαρμογές ανοικτού κώδικα και λογισμικό	34
5.1.1.	FFmpeg	35
5.1.2.	ffserver	35
5.1.3.	x264	35
5.1.4.	MP4box	36
5.1.5.	Wireshark	36
5.2.	Περιγραφή συστήματος ιατρικής εκπαίδευσης και αξιολόγησης ιατρικού βίντεο	37
5.2.1.	Κωδικοποιητής	39
5.2.2.	Συσκευαστής πολυμέσων	40
5.2.3.	Αποστολέας	41
5.2.4.	Παραλήπτης	42
5.2.5.	Διεπαφή αξιολόγησης ποιότητας βίντεο	44

5.1. Εφαρμογές ανοικτού κώδικα και λογισμικό

Οι εφαρμογές ανοικτού κώδικα [36] είναι ένας όρος που χρησιμοποιείτε στον τομέα της πληροφορικής και των ηλεκτρονικών υπολογιστών πολύ συχνά. Με τον όρο αυτό εννοείται ότι ο πηγαίος κώδικας που εμπεριέχεται στην εφαρμογή αυτή, μπορεί να εξεταστεί ελεύθερα από όποιον το ζητήσει. Τα λογισμικά ανοικτού κώδικα συνοδεύονται πάντοτε από κάποια άδεια χρήσης. Ανάλογα με την κατηγορία της άδειας λογισμικού στην οποία ανήκει το λογισμικό αυτό, οι χρήστες μπορούν είτε να τροποποιήσουν τον κώδικα είτε απλά να τον χρησιμοποιήσουν σε κάποια άλλη εφαρμογή.

Οι άδειες χρήσεις χωρίζονται σε τρεις κατηγορίες ως εξής: Permissive (liberal or minimal): Οι άδειες αυτές διευκρινίζουν μόνο ότι το λογισμικό μπορεί να αναδιανεμηθεί ή να τροποποιηθεί, αλλά δεν επιβάλλει ειδικούς όρους για το πώς θα γίνει η διαδικασία αυτή.

Weak Copyleft: Σε περίπτωση που θέλουμε να κάνουμε αλλαγές σε κάποιο λογισμικό κάτω από την συγκεκριμένη άδεια, το λογισμικό που θα αλλάξουμε πρέπει να έχει την ίδια άδεια διαφορετικά τα παράγωγα της συγκεκριμένης δουλειάς μπορούν να διανεμηθούν κάτω από οποιαδήποτε άδεια. Τέλος η Strong Copyleft: Η άδεια αυτή προϋποθέτει ότι οποιαδήποτε αλλαγή σε κάποιο λογισμικό ή οποιαδήποτε διανομή του λογισμικού που χρησιμοποιεί αυτή την άδεια, τότε πρέπει να έχει την ίδια άδεια.

Η χρήση εφαρμογών ανοικτού κώδικα και λογισμικού προσφέρει πολλά πλεονεκτήματα σε σχέση με τις εφαρμογές κλειστού κώδικα. Το πιο σημαντικό πλεονέκτημα είναι ότι αυξάνεται η αξιοπιστία αφού η εφαρμογή ή το λογισμικό που χρησιμοποιούμε είναι είδη εφαρμοσμένο σε κάποια συστήματα που δουλεύουν, επομένως όλα τα προβλήματα έχουν είδη διορθωθεί. Ακόμα, το σύστημα που χρησιμοποιεί εφαρμογές ανοικτού κώδικα δεν υλοποιείται από το μηδέν, επομένως τόσο ο χρόνος ανάπτυξης όσο και ο χρόνος επικύρωσης μπορεί να μειωθεί. Τέλος, υπάρχει μικρότερη αβεβαιότητα στο κόστος ανάπτυξης, δεδομένου ότι το λογισμικό επαναχρησιμοποιείται και δεν αναπτύχθηκε από το μηδέν.

Οι εφαρμογές ανοικτού κώδικα και το λογισμικό που χρησιμοποιήθηκε για την δημιουργία του συστήματος αυτού, περιγράφονται πιο κάτω.

5.1.1. FFmpeg

Το FFmpeg είναι ένα έργο ελεύθερου λογισμικού που παρέχει βιβλιοθήκες και προγράμματα για χειρισμό των multimedia δεδομένων. Οι πιο βασικές βιβλιοθήκες που περιλαμβάνει είναι η libavcodec, μια βιβλιοθήκη μέσω της οποίας γίνεται κωδικοποίηση ήχου και εικόνας και η οποία χρησιμοποιείται και σε διάφορες άλλες εφαρμογές και η βιβλιοθήκη libavformat η οποία χρησιμοποιείται για πολυπλεξία και αποπολυπλεξία ήχου και βίντεο [36]. Περιέχει επίσης τις βιβλιοθήκες libavutil, libavfilter, libavdevice, libswscale και libswresample. Εκτός από τις βιβλιοθήκες, ο FFmpeg παρέχει και διάφορα εργαλεία όπως ο ffmpeg, ο ffserver, ο ffplay και ο ffprobe. Τα εργαλεία αυτά χρησιμοποιούνται ούτως ώστε να γίνεται επιτυχώς κωδικοποίηση, αποκωδικοποίηση, πολυπλεξία, αποπολυπλεξία, streaming καθώς επίσης και φιλτράρισμα [37]. Οι χρήστες του ffmpeg μπορούν να χρησιμοποιήσουν το εργαλείο αυτό μέσω του command-line για να κωδικοποιήσουν και να αποκωδικοποιήσουν αρχεία ήχου ή βίντεο σε διάφορες μορφές. Για την κωδικοποίηση σε H.264 το ffmpeg χρησιμοποιεί την βιβλιοθήκη x264 η οποία είναι μια εξωτερική βιβλιοθήκη ανοικτού κώδικα. Στόχος του FFmpeg είναι να παρέχει την καλύτερη δυνατή λύση για τους προγραμματιστές εφαρμογών

καθώς επίσης και τους τελικούς χρήστες συνδυάζοντας τις καλύτερες επιλογές από ελεύθερο λογισμικό.

5.1.2. ffmpeg

Ο ffmpeg είναι ένα εργαλείο που όπως έχουμε προαναφέρει, μας το παρέχει ο FFmpeg και είναι ένας εξυπηρετητής πολυμέσων για ζωντανές αναμεταδόσεις σε πολλαπλούς χρήστες. Για να επιτευχθεί αυτό ο ffmpeg μπορεί να ενωθεί και να στέλνει σε κάποιο port το οποίο ορίζουμε εμείς σε ένα configuration file. Σε αυτό το port ο ffmpeg στέλνει τα βίντεο ή τον ήχο που θέλει να κάνει stream. Δεν υπάρχει περιορισμός στα streams αφού μπορεί να κάνει ένα ή περισσότερα. Κάθε stream εισόδου ονομάζεται feed και το ορίζουμε στο configuration file. Για κάθε stream μπορούμε να έχουμε πολλά διαφορετικά stream εξόδου, με διαφορετικό format και τα οποία ονομάζουμε Streams. Ο ffmpeg δουλεύει προωθώντας τα streams κωδικοποιημένα με την βοήθεια του ffmpeg ή μπορεί να στείλει stream τα οποία έχουν κωδικοποιηθεί από πριν [38].

5.1.3. x264

Η x264 [39] είναι μια βιβλιοθήκη δωρεάν λογισμικού η οποία χρησιμοποιείται από πλήθος εφαρμογών και έργων, για κωδικοποίηση video streams σε μορφή H.264/MPEG-4 AVC. Η βιβλιοθήκη αυτή παρέχει στους χρήστες μια command line διεπιφάνεια καθώς επίσης και ένα API. Το API που παρέχει χρησιμοποιείται από πολλές γραφικές διεπιφάνειες χρήστη όπως για παράδειγμα Staxrip και MeGUI. Ο FFmpeg, όπως και πολλές άλλες διεπιφάνειες όμως, χρησιμοποιούν την command line διεπιφάνεια της βιβλιοθήκης αυτής. Η συγκεκριμένη βιβλιοθήκη είναι η καλύτερη επιλογή για κωδικοποίηση σε H.264/AVC αφού παρέχει καλύτερη απόδοση, συμπίεση και ποιότητα από άλλες βιβλιοθήκες της ίδιας κατηγορίας. Τον Απρίλιο του 2010, το έργο x264 ανακοίνωσε την πλήρη Blu-Ray κωδικοποίηση βίντεο και έτσι έγινε το πρώτο δωρεάν Blu-Ray συμβατό λογισμικό για H.264 κωδικοποίηση.

5.1.4. MP4box

Το MP4box είναι ένας συσκευαστής πολυμέσων ο οποίος παρέχει ένα τεράστιο αριθμό λειτουργιών όπως για παράδειγμα μετατροπή, διάσπαση καθώς και πολλά άλλα. Το command-line αυτό εργαλείο, μπορεί να χρησιμοποιηθεί για την δημιουργία περιεχομένου που να είναι συμβατό με τις προδιαγραφές του MPEG-DASH. Με το MP4box υπάρχει η δυνατότητα

παρεμβολής (interleaving), κατακερματισμού (fragmentation), τμηματοποίησης (segmentation) και splitting. Το interleaving είναι όταν ομάδες δειγμάτων από διαφορετικά κομμάτια είναι αποθηκευμένα εναλλάξ στο αρχείο. Ο κατακερματισμός είναι μια προαιρετική διαδικασία που εφαρμόζεται μόνο σε MP4 μορφή αρχείου. Η τμηματοποίηση είναι η διαδικασία δημιουργίας τμημάτων από ένα αρχικό αρχείο που χρησιμοποιείται για την αναπαραγωγή του stream. Τέλος, με το splitting μπορούμε να χωρίσουμε ένα αρχείο και να δημιουργήσουμε ανεξάρτητα αρχεία με δυνατότητα αναπαραγωγής [40].

5.1.5. Wireshark

Το Wireshark [41] είναι δωρεάν λογισμικό και ανοικτού κώδικα έργο, μέσω του οποίου αναλύονται τα πακέτα ενός δικτύου. Κύρια λειτουργία του είναι η καταγραφή των πακέτων που κινούνται μέσα σε ένα δίκτυο, αναλυτικά, παρουσιάζοντας τα με μεγάλη λεπτομέρεια. Το wireshark επιτρέπει στον χρήστη να ορίσει ελεγκτές διεπιφάνειας δικτύου (network interface controllers) οι οποίοι υποστηρίζουν τον ανάμικτο τρόπο λειτουργίας. Έτσι οι χρήστες μπορούν να δουν όλα τα πακέτα που κινούνται στην διεπιφάνεια αυτή, όχι μόνο τα πακέτα που έχουν σαν προορισμό διευθύνσεις της διεπιφάνειας. Επίσης, είναι ορατά τα πακέτα που προορίζονται για όλους τους παραλήπτες (τύπου broadcast) καθώς και τα πακέτα με πολλαπλούς παραλήπτες (τύπου multicast).

Το wireshark μπορεί να χρησιμοποιηθεί με δύο τρόπους. Ο πρώτος τρόπος που είναι και ο πιο γνωστός είναι μέσω του GUI του και ο δεύτερος είναι μέσω command-line. Η command-line έκδοση του Wireshark ονομάζεται T-shark και είναι χρήσιμη όταν οι χρήστες δεν μπορούν να χρησιμοποιήσουν την γραφική διεπαφή ή όταν δεν την χρειάζονται. Όλες οι λειτουργίες που παρέχει το Wireshark όμως είναι διαθέσιμες και στο T-shark.

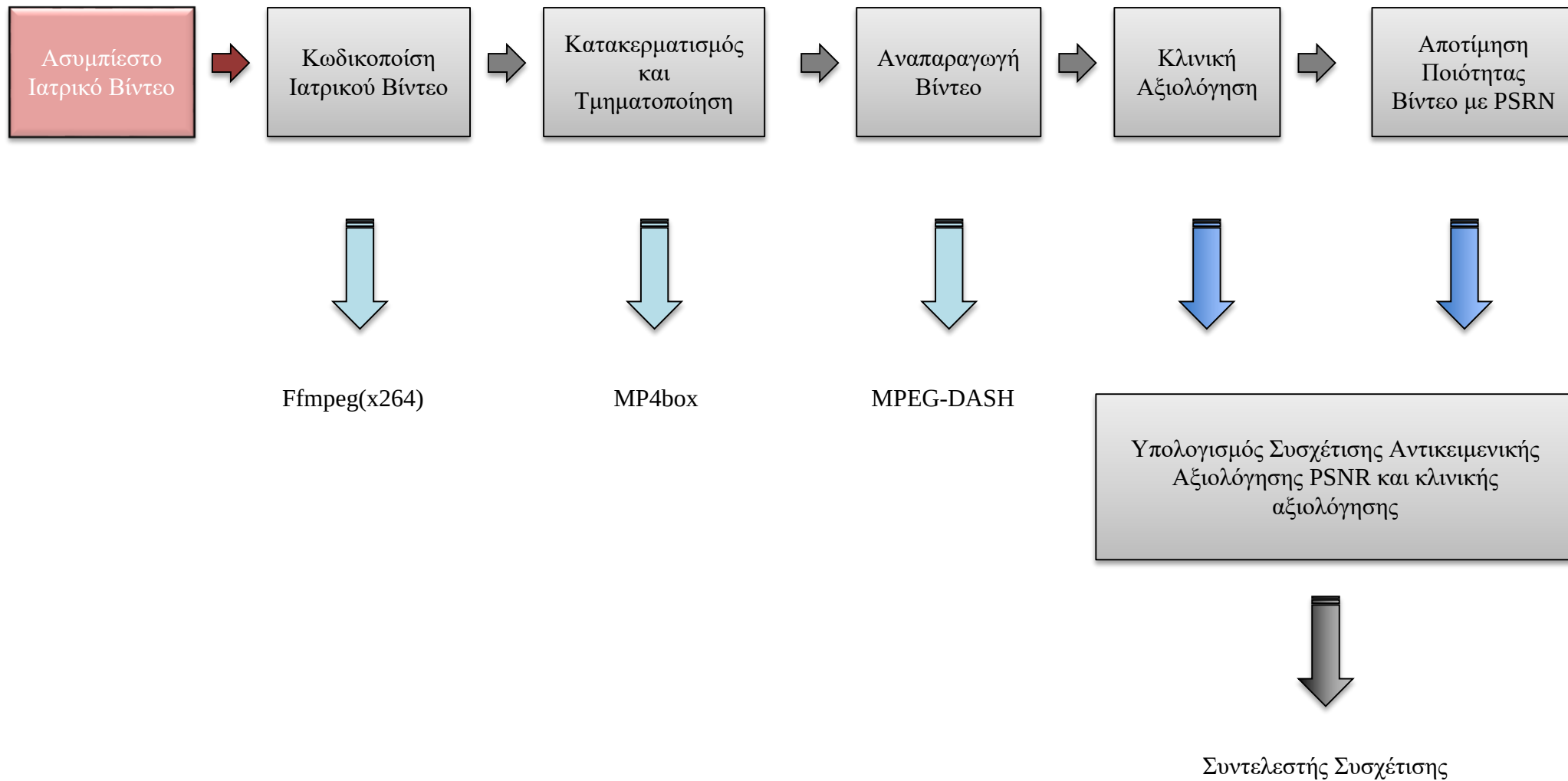
5.2. Περιγραφή συστήματος ιατρικής εκπαίδευσης και αξιολόγησης ιατρικού βίντεο

Για την παρούσα εργασία, η επιλογή εργαλείων ανοικτού κώδικα αποσκοπεί στην ανάπτυξη, αξιολόγηση και διάδοση ενός συστήματος ιατρικής εκπαίδευσης και αξιολόγησης ιατρικού βίντεο χαμηλού κόστους, με ευκολία χρήσης και διείσδυσης στην ερευνητική κοινότητα.

Στο σύστημα που αναπτύχθηκε στα πλαίσια αυτής της εργασίας, η διαδικασία που χρησιμοποιείται έχει ως εξής: Αρχικά, το ασυμπίεστο ιατρικό βίντεο κωδικοποιείται με την χρήση του εργαλείου FFmpeg [36], το οποίο χρησιμοποιεί την βιβλιοθήκη x264[40], με το πρότυπο H.264/AVC. Στη συνέχεια, το κωδικοποιημένο πλέον βίντεο διασπάται σε segments

με την βοήθεια του MP4box[41], το οποίο δημιουργεί ένα περιεχόμενο που είναι συμβατό με τις προδιαγραφές του MPEG-DASH [34]. Ο χρήστης μπορεί να δει τα ιατρικά αυτά βίντεο μέσω του DASH player. Κατά την διάρκεια θέασης του ιατρικού βίντεο χρησιμοποιείται το Wireshark [41] για την παρακολούθηση των πακέτων. Ακολούθως, γίνεται η μέτρηση της ποιότητας του βίντεο που είδε ο χρήστης με αλγόριθμους αξιολόγησης της ποιότητας του βίντεο όπως είναι το PSNR [42], μεταξύ των πλαισίων του βίντεο πριν την κωδικοποίηση και των αντίστοιχων πλαισίων του βίντεο μετά την κωδικοποίηση του. Επίσης ο χρήστης κατά την διάρκεια που βλέπει τα βίντεο, μπορεί να βαθμολογήσει την ποιότητα. Επομένως με την χρήση συγκεκριμένων αλγορίθμων υπολογίζεται ο συντελεστής συσχέτισης μεταξύ του PSNR και της κλινικής αυτής βαθμολογίας. Σκοπός είναι το ιατρικό βίντεο που θα φτάσει στον γιατρό να είναι ικανοποιητικής διαγνωστικής ποιότητας και ο γιατρός προχωρεί στην κλινική αξιολόγηση του.

Στο σχήμα 5.1 που ακολουθεί, παρουσιάζεται η πιο πάνω διαδικασία.

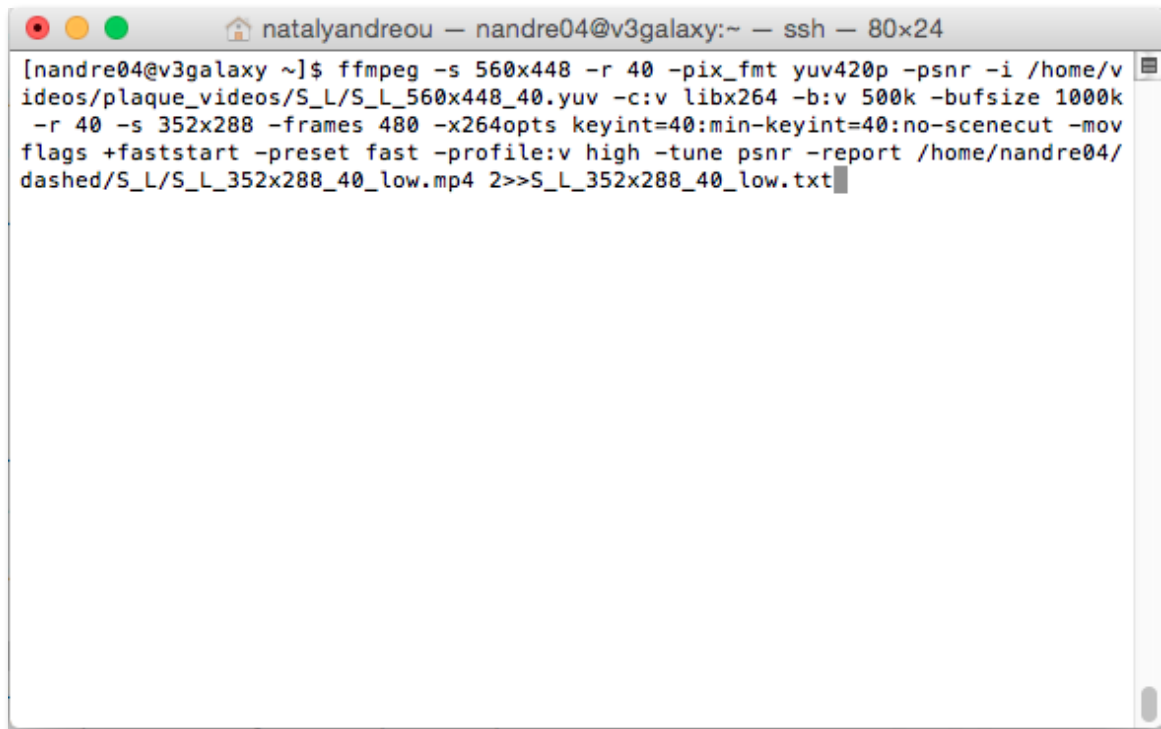


Σχήμα 5.1: Διάγραμμα Συστήματος.

Το σύστημα που δημιουργήθηκε με την χρήση των προαναφερθέντων υπάρχον τεχνολογιών, έχει ως σκοπό την αποστολή ιατρικού βίντεο σε πραγματικό χρόνο, του οποίου η ποιότητα να είναι ικανοποιητικά υψηλή ούτως ώστε να μην χάνεται ιατρική πληροφορία. Το σύστημα αυτό αποτελείται από τα εξής πέντε μέρη: ο κωδικοποιητής, ο συσκευαστής πολυμέσων, ο αποστολέας, ο παραλήπτης και η διεπαφή για αξιολόγηση της ποιότητας του βίντεο.

5.2.1. Κωδικοποιητής

Στο πρώτο βήμα ο κωδικοποιητής καλεί τον ffmpeg μέσω της γραμμής εντολών, δίνοντας τις απαραίτητες παραμέτρους κωδικοποίησης του ασυμπίεστου ιατρικού βίντεο που επιθυμεί. Για να γίνει η κωδικοποίηση του βίντεο, ο ffmpeg καλεί την εξωτερική βιβλιοθήκη x264. Ο κωδικοποιητής όπως φαίνεται στο σχήμα 5.2 δίνει πρώτα τις παραμέτρους για το ασυμπίεστο ιατρικό βίντεο εισόδου που είναι το resolution, το frame rate καθώς επίσης και το format των pixels του. Στην συνέχεια δίνει τις παραμέτρους για το κωδικοποιημένο βίντεο αλλάζοντας το resolution ή/και το bitrate του. Στο παράδειγμα του σχήματος 5.2, το βίντεο αλλάζει resolution από 560x448 σε 352x288 και το bitrate του γίνεται 500k. Επίσης το βίντεο από yuv μετατρέπεται σε mp4 για να μπορεί να προχωρήσει στο επόμενο βήμα. Σε αυτό το σημείο ο κωδικοποιητής, υπολογίζει και το PSNR δημιουργώντας ένα αρχείο εξόδου στο οποίο είναι αποθηκευμένη η τιμή του PSNR για να μπορούν να γίνουν οι απαραίτητοι ελέγχει ποιότητας.

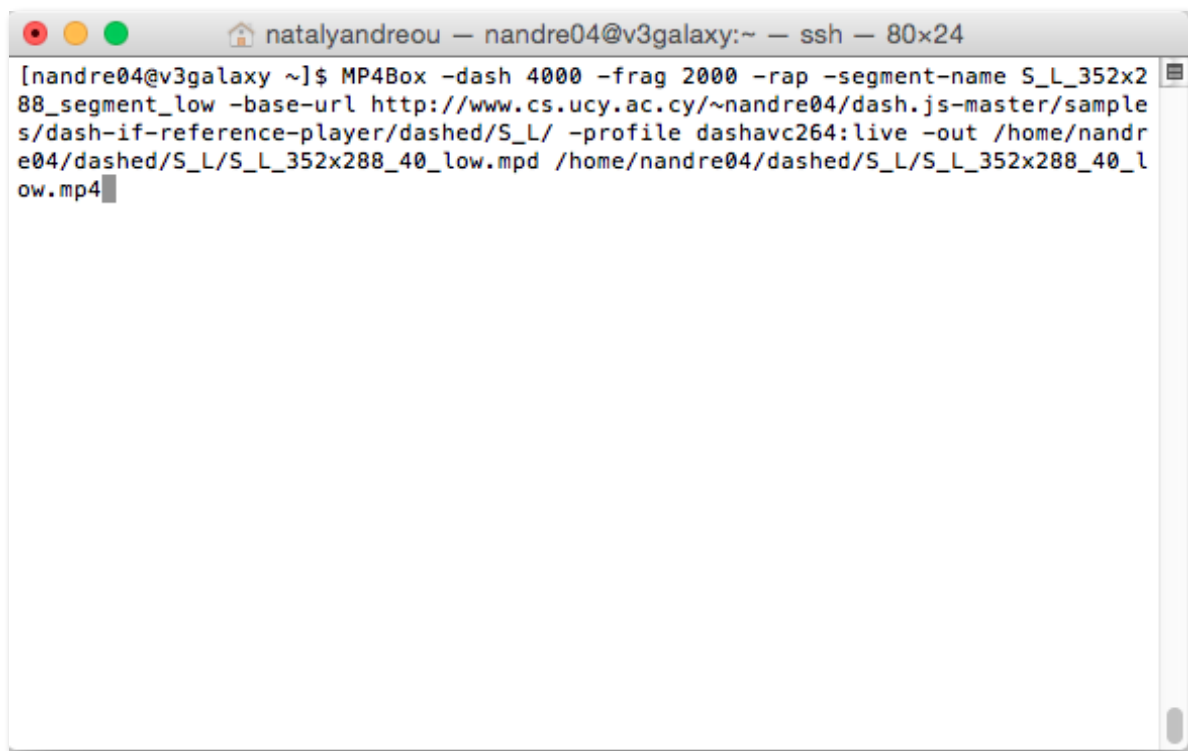


```
natalyandreou — nandre04@v3galaxy:~ — ssh — 80x24
[nandre04@v3galaxy ~]$ ffmpeg -s 560x448 -r 40 -pix_fmt yuv420p -psnr -i /home/v
ideos/plaque_videos/S_L/S_L_560x448_40.yuv -c:v libx264 -b:v 500k -bufsize 1000k
-r 40 -s 352x288 -frames 480 -x264opts keyint=40:min-keyint=40:no-scenecut -mov
flags +faststart -preset fast -profile:v high -tune psnr -report /home/nandre04/
dashed/S_L/S_L_352x288_40_low.mp4 2>>S_L_352x288_40_low.txt
```

Σχήμα 5.2: Γραμμή Εντολών Κωδικοποιητή.

5.2.2. Συσκευαστής πολυμέσων

Στη συνέχεια ο συσκευαστής πολυμέσων ο οποίος αναπαρίσταται στο σχήμα 5.3, καλεί το MP4box και πάλι μέσω της γραμμής εντολών. Με την χρήση του MP4box γίνεται κατακερματισμός και τμηματοποίηση, του κωδικοποιημένου πλέον βίντεο, δίνοντας τις απαραίτητες παραμέτρους για να δημιουργηθεί ο σωστός αριθμός segments με σωστή διάρκεια. Στο παράδειγμα που φαίνεται στο σχήμα 5.3, ο συσκευαστής πολυμέσων καθορίζει ότι τα segments θα είναι μεγέθους τεσσάρων δευτερολέπτων και ότι τα subsegments θα είναι μεγέθους δύο δευτερολέπτων το καθένα. Με την βοήθεια του `-segment-name` ο συσκευαστής πολυμέσων δίνει το όνομα που επιθυμεί στα segments που θα δημιουργηθούν. Έτσι στο σημείο αυτό, δημιουργείται ένα περιεχόμενο που είναι συμβατό με τις προδιαγραφές του MPEG-DASH το οποίο είναι το `mpd` αρχείο.

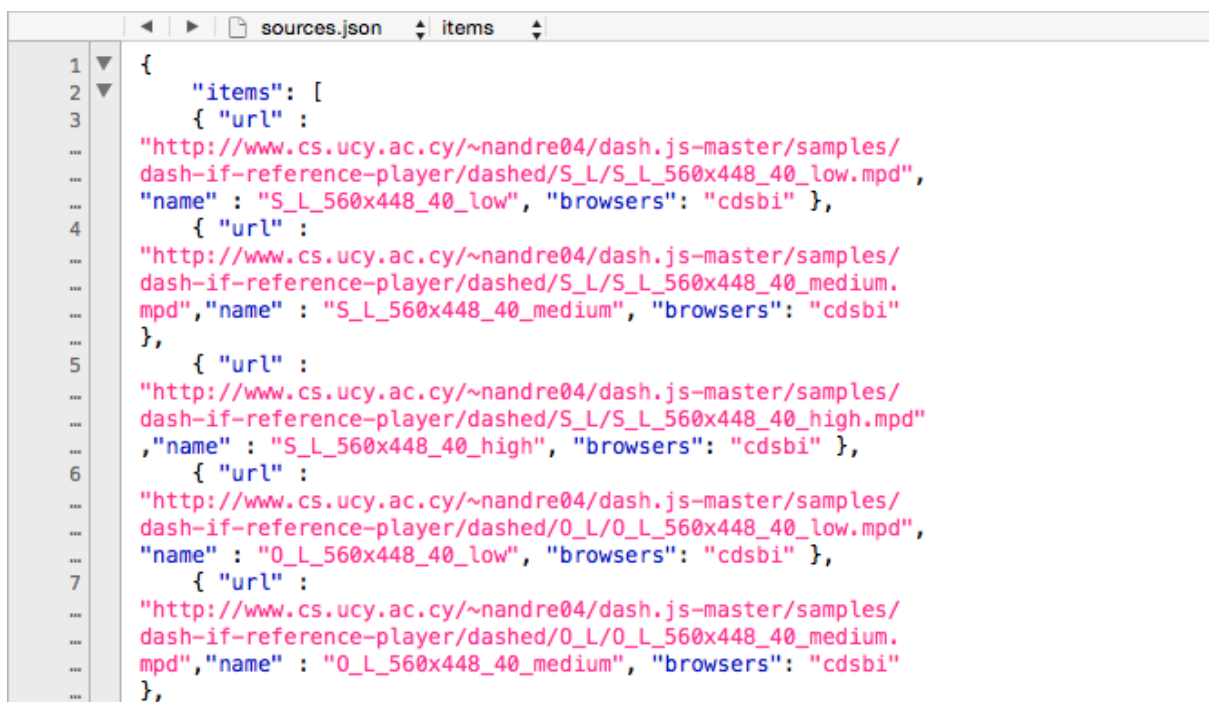


```
natalyandreou — nandre04@v3galaxy:~ — ssh — 80x24
[nandre04@v3galaxy ~]$ MP4Box -dash 4000 -frag 2000 -rap -segment-name S_L_352x288_segment_low -base-url http://www.cs.ucy.ac.cy/~nandre04/dash.js-master/sample s/dash-if-reference-player/dashed/S_L/ -profile dashavc264:live -out /home/nandre04/dashed/S_L/S_L_352x288_40_low.mpd /home/nandre04/dashed/S_L/S_L_352x288_40_low.mpd
```

Σχήμα 5.3: Γραμμή Εντολών Συσκευαστή Πολυμέσων.

5.2.3. Αποστολέας

Το MPEG-DASH πρότυπο adaptive streaming, όπως έχουμε προαναφέρει, συλλέγει και αποθηκεύει το περιεχόμενο των multimedia σε έναν HTTP server. Επομένως η δουλειά του αποστολέα που παρουσιάζεται στο σχήμα 5.4, είναι να αποθηκεύσει σε έναν HTTP server το ιατρικό βίντεο που έχει κωδικοποιηθεί, κατακερματιστεί και τμηματοποιηθεί. Αυτό γίνεται μέσω ενός αρχείου JSON μορφής όπου ο αποστολέας πρέπει να προσθέσει το url του βίντεο καθώς και ένα όνομα που επιθυμεί να του δώσει. Το url δείχνει το μονοπάτι στο οποίο βρίσκεται το βίντεο που θα μπορεί να δει ο παραλήπτης και το name είναι το όνομα το οποίο θα αντιπροσωπεύει του συγκεκριμένο url στην λίστα με τα βίντεο που θα μπορεί να δει ο παραλήπτης.

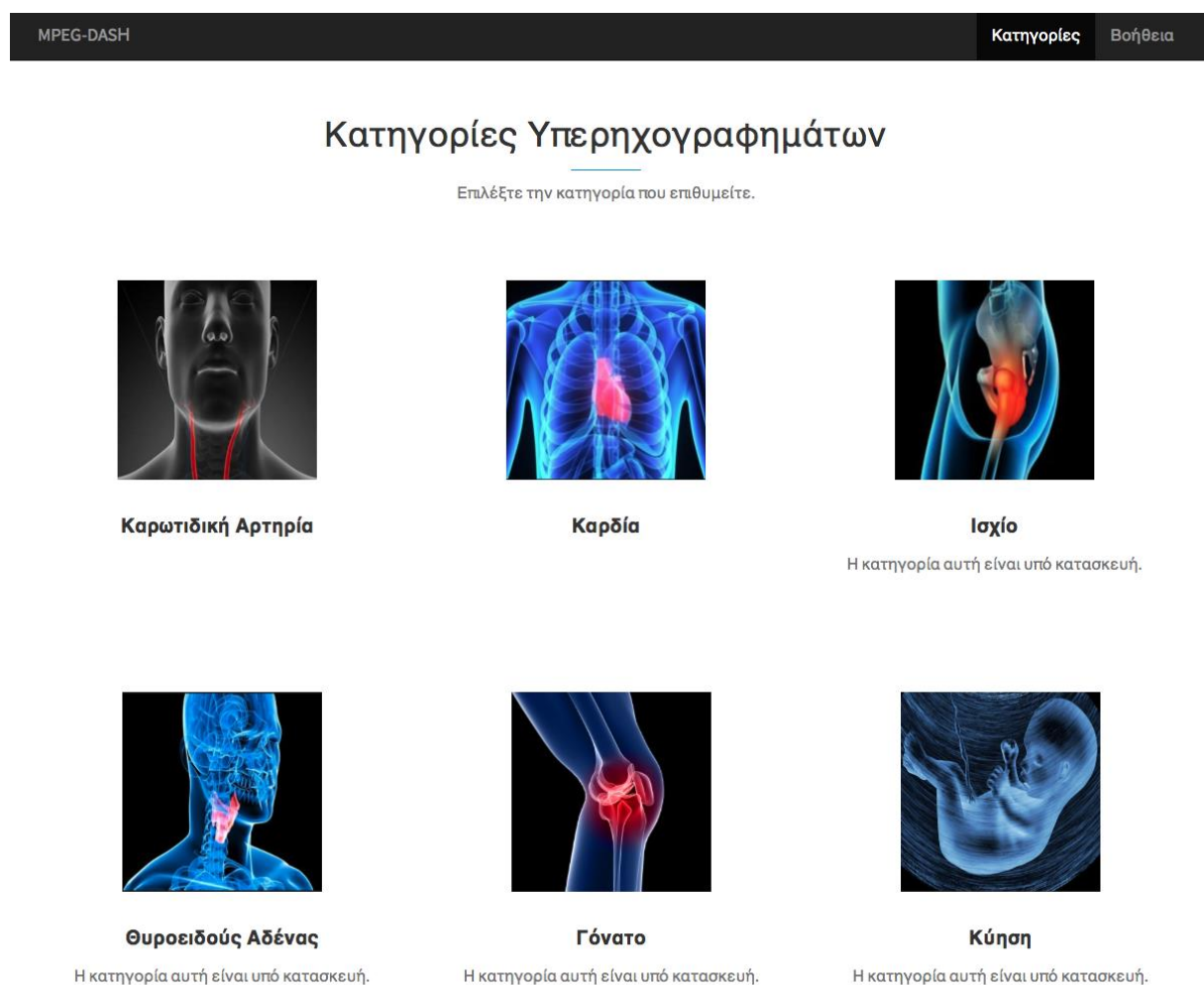


```
1 {  
2   "items": [  
3     { "url" :  
4       "http://www.cs.ucy.ac.cy/~nandre04/dash.js-master/samples/  
5       dash-if-reference-player/dashed/S_L/S_L_560x448_40_low.mpd",  
6       "name" : "S_L_560x448_40_low", "browsers": "cdsbi" },  
7     { "url" :  
8       "http://www.cs.ucy.ac.cy/~nandre04/dash.js-master/samples/  
9       dash-if-reference-player/dashed/S_L/S_L_560x448_40_medium.  
10      mpd", "name" : "S_L_560x448_40_medium", "browsers": "cdsbi"  
11    },  
12    { "url" :  
13      "http://www.cs.ucy.ac.cy/~nandre04/dash.js-master/samples/  
14      dash-if-reference-player/dashed/S_L/S_L_560x448_40_high.mpd"  
15      , "name" : "S_L_560x448_40_high", "browsers": "cdsbi" },  
16    { "url" :  
17      "http://www.cs.ucy.ac.cy/~nandre04/dash.js-master/samples/  
18      dash-if-reference-player/dashed/O_L/O_L_560x448_40_low.mpd",  
19      "name" : "O_L_560x448_40_low", "browsers": "cdsbi" },  
20    { "url" :  
21      "http://www.cs.ucy.ac.cy/~nandre04/dash.js-master/samples/  
22      dash-if-reference-player/dashed/O_L/O_L_560x448_40_medium.  
23      mpd", "name" : "O_L_560x448_40_medium", "browsers": "cdsbi"  
24    },  
25  ],  
26}
```

Σχήμα 5.4: Αποστολέας.

5.2.4. Παραλήπτης

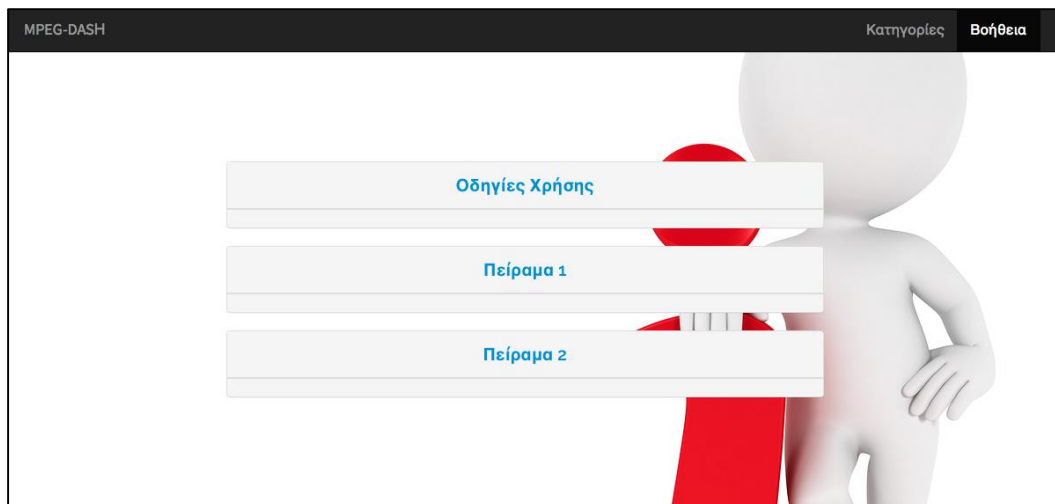
Ο παραλήπτης είναι υπεύθυνος για την παραλαβή του βίντεο. Αρχικά μέσω της διεπιφάνειας που φαίνεται στο σχήμα 5.5 ο παραλήπτης καλείται να επιλέξει ανάμεσα σε διάφορες κατηγορίες υπερηχογραφήμάτων την κατηγορία που επιθυμεί. Επιλέγοντας της κατηγορία, τα βίντεο που θα το εμφανίζονται για παραλαβή είναι μόνο όσα ανήκουν στην κατηγορία που επέλεξε. Στη συνέχεια μέσω της διεπιφάνειας που φαίνεται στο σχήμα 5.6, ο χρήστης μπορεί να επιλέξει από μια λίστα με ιατρικά βίντεο, το βίντεο που επιθυμεί να παρακολουθήσει και το βλέπει μέσω του DASH.



Σχήμα 5.5: Διεπιφάνεια Κατηγοριών.

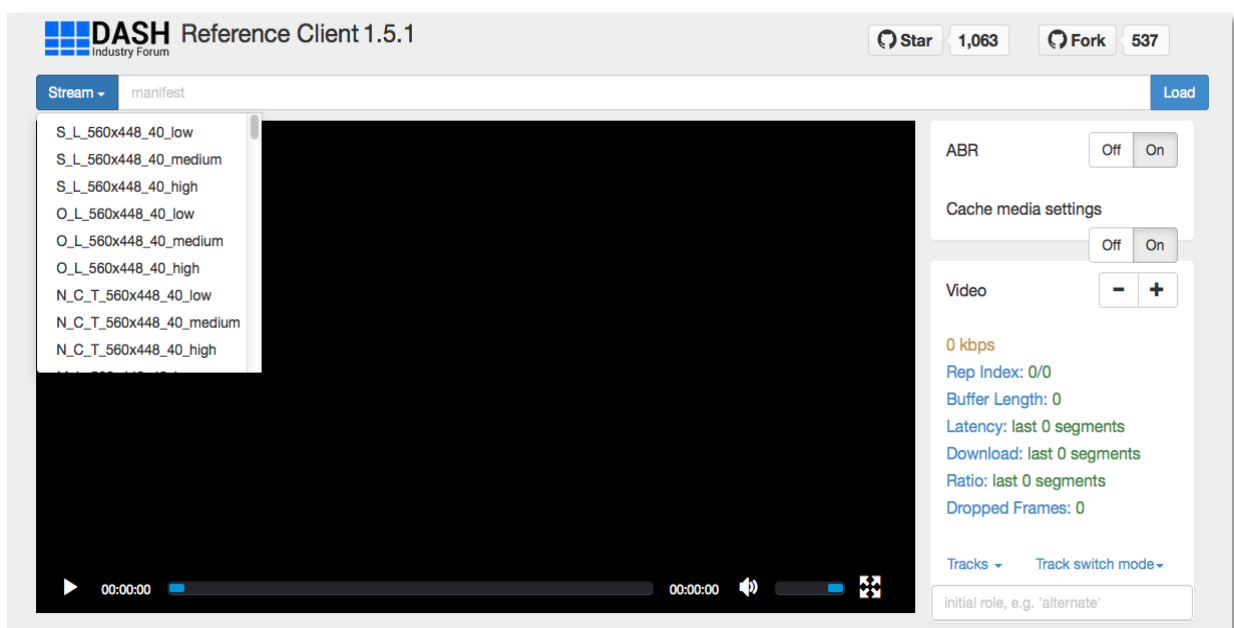
Στο σχήμα 5.6. που βλέπουμε, φαίνεται η διεπιφάνεια στην οποία ο χρήστης μπορεί να βρεί σχετικές πληροφορίες για την σελίδα και να βοηθηθεί. Για ια παράδειγμα μπορεί να βρεί πληροφορίες για τνο τρόπο με τον οποίο γίνεται χρήση της σελίδας και επίσης να ενημερωθεί

για τα σχετικά πειράματα που γίνονται μέσω της σελίδας.



Σχήμα 5.6: Διεπιφάνεια βοήθειας.

Το Dash player που βλέπουμε στο σχήμα 5.7, είναι μία πρωτοβουλία που πήρε το DASH Industry Forum θέλοντας να δημιουργήσει ένα ποιοτικό πλαίσιο για να μπορεί να γίνεται η αναπαραγωγή βίντεο και ήχου που το περιεχόμενο είναι συμβατό με το MPEG-DASH. Η δημιουργία αυτού του player έγινε με την χρήση Javascript βιβλιοθηκών οι οποίες τρέχουν στην μεριά του client. Τα βίντεο που εμφανίζονται στην λίστα και μπορούν να αναπαραχθούν με το συγκεκριμένο player, έχουν περάσει από το στάδιο κωδικοποίησης, κατάτμησης και κατακερματισμού καθώς και αποστολής όπως είδαμε στα πάνω βήματα.

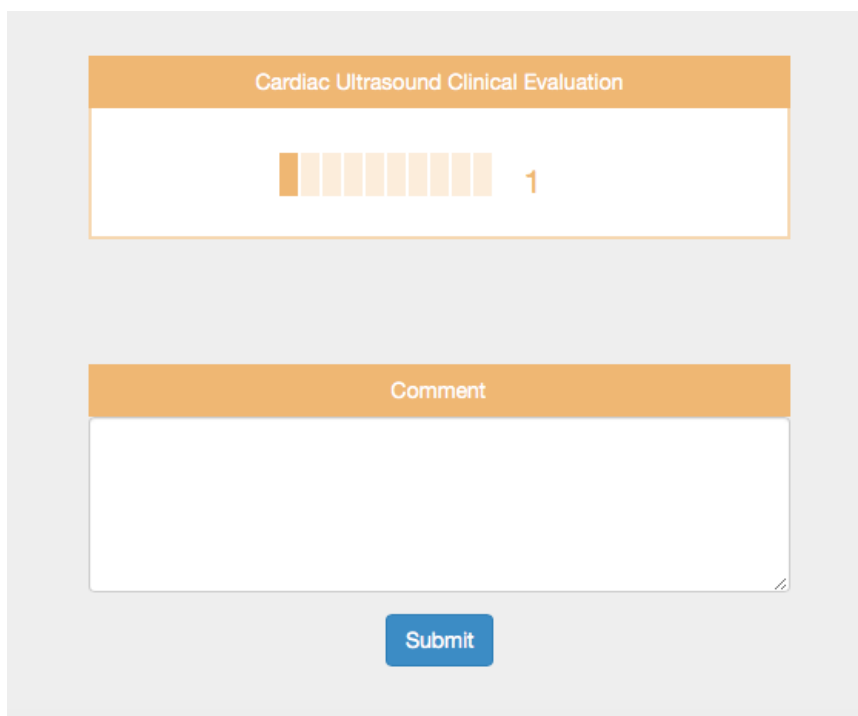


Σχήμα 5.7: Διεπιφάνεια παραλήπτη.

5.2.5. Διεπαφή αξιολόγησης ποιότητας βίντεο

Τέλος, στο κάτω μέρος της διεπιφάνειας του παραλήπτη, αφού ο χρήστης έχει παρακολουθήσει το βίντεο, μπορεί να αξιολογήσει την ποιότητα του με τον τρόπο που φαίνεται στο σχήμα 5.7. Του δίνεται η δυνατότητα να βαθμολογήσει την ποιότητα του βίντεο με τιμές από 1 μέχρι 10 και μπορεί επίσης αν επιθυμεί να προσθέσει κάποιο σχόλιο στο πλαίσιο σχολίων.

Η μπάρα αξιολόγησης που φαίνονται στο σχήμα 5.7. είναι κάποια select fields τα οποία μετατράπηκαν με την βοήθεια CSS κώδικα σε γραφική μπάρα αξιολόγησης. Όταν ένας χρήστης δίνει την βαθμολογία του στην πιο κάτω μπάρα, αποθηκεύεται η βαθμολογία μαζί με το όνομα του χρήστη καθώς επίσης και την ημερομηνία της βαθμολογίας, σε ένα text αρχείο. Αυτό γίνεται με την βοήθεια PHP κώδικα και το αρχείο αποθηκεύεται στην μεριά του server έτσι ώστε να υπάρχουν για να μπορεί να γίνει έρευνα για την ποιότητα του βίντεο.



The image shows a web form titled "Cardiac Ultrasound Clinical Evaluation". It contains a rating scale with 10 vertical bars, the first of which is filled orange, with the number "1" to its right. Below the rating scale is a section titled "Comment" with a large text input area and a blue "Submit" button at the bottom.

Σχήμα 5.7: Διεπιφάνεια Αξιολόγησης Ποιότητας.

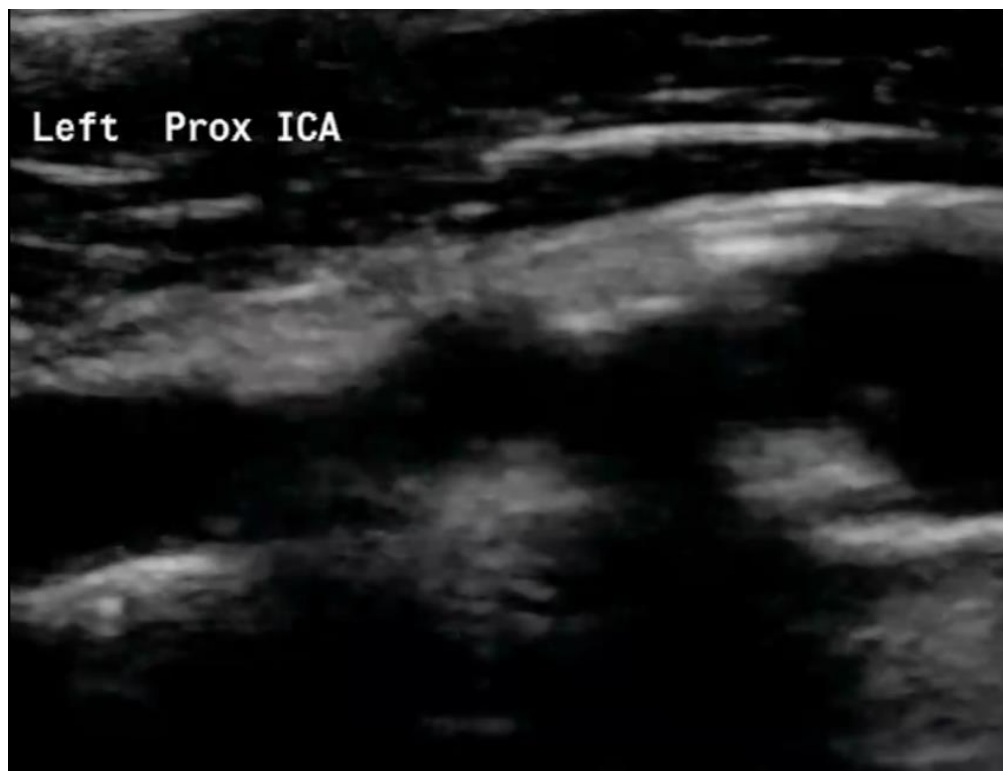
Κεφάλαιο 6

Μεθοδολογία Υλοποίησης

6.1.	Κωδικοποίηση	46
6.2.	Κατακερματισμός και Τμηματοποίηση	48
6.3.	Αξιολόγηση Αποτελεσμάτων	49
6.3.1.	Αλγόριθμος PSNR	50
6.3.2.	Συντελεστής Συσχέτισης Pearson	50
6.3.3.	Συντελεστής Συσχέτισης Spearman	51

6.1. Κωδικοποίηση

Πάνω στο σύστημα που αναπτύχθηκε, διεξάχθηκαν κάποια πειράματα στα οποία έλαβαν μέρος δέκα ιατρικά βίντεο υπερήχων της καρωτιδικής αρτηρίας όπως φαίνεται στο σχήμα 6.1. καθώς επίσης και δέκα ιατρικά βίντεο υπερήχων καρδιάς όπως φαίνεται στο σχήμα 6.2.



Σχήμα 6.1. Υπερηχογράφημα καρωτιδικής αρτηρίας ασθενούς.



Σχήμα 6.2. Υπερηχογράφημα καρδιάς ασθενούς.

Κατά την διάρκεια μιας εξέτασης ασθενούς για καρδιά, γίνονται ουσιαστικά ταυτόχρονα τέσσερις απεικονίσεις. Οι τέσσερις αυτές απεικονίσεις έχουν ως εξής: Η πρώτη είναι η απεικόνιση δυο διαστάσεων (2D). Η ηχογραφική δέσμη, που παράγεται από τον πιεζοηλεκτρικό κρύσταλλο, δεν έχει σταθερή διεύθυνση αλλά κάνει συνεχή σάρωση ενός τομέα 45-90 μοιρών. Έτσι απεικονίζονται τα διάφορα στοιχεία, που υπάρχουν στην περιοχή αυτού του τομέα, σε πραγματική κίνηση και σε επίπεδο 2 διαστάσεων. Το δεύτερο είναι η απεικόνιση Mode και είναι η κίνηση σε συνάρτηση με τον χρόνο. Είναι κατάλληλη για μετρήσεις σε διάφορες φάσεις όπως είναι η συστολή και διαστολή της καρδιάς, επειδή αποτυπώνει καλά τις μετακινήσεις των διαφόρων δομικών στοιχείων της καρδιάς στη διάρκεια του χρόνου. Με αυτό τον τρόπο μπορεί να αναδείξει περιοχές υποκινησίας, να γίνει εκτίμηση της συσπαστικότητας και να υπολογισθεί το κλάσμα εξωθήσεως της αριστερής κοιλίας. Το τρίτο είναι η Doppler απεικόνιση μέσω της οποίας μελετιούνται οι ροές αίματος ενδοκαρδιακά. Τέλος, η απεικόνιση έγχρωμου Doppler που είναι ένα παλμικό Doppler πολλών και ταυτόχρονων δειγμάτων ροής με την μέτρηση όμως των μέσων ταχυτήτων ροής στα δείγματα και ακολούθως την παράθεσή τους με την μορφή χρώματος.

Τα δέκα βίντεο της καρωτιδικής αρτηρίας είχαν αρχικά ανάλυση εικόνας 560x448 και frame rate 40 ενώ τα δέκα βίντεο της καρδιάς είχαν αρχικά ανάλυση εικόνας 800x600 και frame rate 30. Το κάθε ένα από τα ασυμπίεστα ιατρικά βίντεο, κωδικοποιήθηκε σε H264/MPEG-4 AVC με την βοήθεια του ffmpeg. Τα βίντεο καρωτιδικής αρτηρίας κωδικοποιήθηκαν σε τέσσερις διαφορετικές αναλύσεις εικόνας ενώ τα βίντεο καρδιάς σε με μία μόνο ανάλυση εικόνας, και κάθε ανάλυση εικόνας των βίντεο αυτών κωδικοποιήθηκε με τέσσερα διαφορετικά bitrates. Επομένως υπάρχουν δεκαέξι κωδικοποιήσεις για το κάθε ασυμπίεστο ιατρικό βίντεο καρωτιδικής αρτηρίας και τέσσερις κωδικοποιήσεις για το κάθε ασυμπίεστο βίντεο καρδιάς.

Στο σχήμα 6.3. υπάρχουν οι παράμετροι κωδικοποίησης που χρησιμοποιήθηκαν σε κάθε ένα από τα δέκα βίντεο της καρωτιδικής αρτηρίας και στο σχήμα 6.4. υπάρχουν οι παράμετροι κωδικοποίησης που χρησιμοποιήθηκαν στα τέσσερα βίντεο καρδιάς.

Video Resolution	352x288, 560x448, 640x480, 704x576
Video Frame Rate	40
Bitrate (Kbps)	256, 500, 1000, 2000
Buffer Size	1000
Intra Update	40
Profile	High
Video	10
Σύνολο Κωδικοποιήσεων	160

Σχήμα 6.3: Πίνακας παραμέτρων Κωδικοποίησης Υπερηχογραφημάτων Καρωτιδικής Αρτηρίας.

Video Resolution	800x600
Video Frame Rate	30
Bitrate (kbps)	256, 500, 1000, 2000
Buffer Size	1000
Intra Update	30
Profile	high
Video	10
Σύνολο Κωδικοποιήσεων	40

Σχήμα 6.4: Πίνακας παραμέτρων Κωδικοποίησης Υπερηχογραφήματος Καρδιάς.

6.2. Κατακερματισμός και Τμηματοποίηση

Το κάθε ένα από τα βίντεο που έχουν κωδικοποιηθεί, σε αυτό το σημείο περνά από τον συσκευαστή πολυμέσων για να κατακερματιστεί και να τμηματοποιηθεί ανάλογα με τις παραμέτρους που φαίνονται στον πίνακα 6.3. Ο συσκευαστή πολυμέσων που χρησιμοποιεί αυτές τις παραμέτρους είναι ο MP4box. Επομένως, τα κωδικοποιημένα βίντεο μετά από αυτό το βήμα, έχουν segments μεγέθους τεσσάρων δευτερολέπτων, subsegments μεγέθους 2 δευτερολέπτων και είναι συμβατά με το πρότυπο MPEG-DASH. Στον πίνακα που φαίνεται πιο κάτω σαν profile χρησιμοποιείται το dashavc264:live λόγω του ότι το βίντεο που βλέπει ο παραλήπτης είναι σε πραγματικό χρόνο.

Dash Duration	4000ms
Fragmentation	2000ms
Profile	dashavc264:live

Πίνακας 6.3: Παράμετροι Κατακερματισμού και Τμηματοποίησης.

6.3. Αξιολόγηση Αποτελεσμάτων

Η αξιολόγηση του συστήματος έγινε με έναν από τους πιο διαδεδομένους αλγόριθμους αντικειμενικής αξιολόγησης της ποιότητας του βίντεο. Ο αλγόριθμος που χρησιμοποιήθηκε ονομάζεται PSNR, και είναι ένας τρόπος να μπορέσουμε να μετρήσουμε την ποιότητα των ιατρικών βίντεο που κωδικοποιήθηκαν.

Ο παραλήπτης όπως έχουμε προαναφέρει, μπορεί να βαθμολογήσει την ποιότητα του ιατρικού βίντεο που έχει παραλάβει και βαθμολογία αυτή αποθηκεύεται στο σύστημα. Για να μπορέσουμε να υπολογίσουμε πιο αντιπροσωπευτικά την ποιότητα του βίντεο θα πρέπει να γίνει μια σύγκριση ανάμεσα στο PSNR των κωδικοποιημένων βίντεο και της βαθμολογίας που έδωσε ο παραλήπτης. Έτσι στην αξιολόγηση των αποτελεσμάτων προστέθηκαν και δύο άλλοι αλγόριθμοι όπου, ο πρώτος ονομάζεται Συντελεστής Συσχέτισης Pearson και ο δεύτερος ονομάζεται Spearman.

6.3.1. Αλγόριθμος PSNR

Με την βοήθεια του PSNR [42] υπολογίζουμε το λόγο του σήματος προς τον θόρυβο (Peak Signal-to-Noise Ratio). Αν και ο αλγόριθμος αυτός είναι τόσο διαδεδομένος, κυρίως λόγω της ταχύτητας υλοποίησης και εκτέλεσης, συχνά δεν αποτυπώνει την πραγματική του ποιότητα. Για να μπορέσουμε όμως να καταλάβουμε ακριβώς πως δουλεύει ο αλγόριθμος αυτός θα κάνουμε μια αναφορά στο μαθηματικό υπόβαθρο του.

Η κύρια χρήση του PSNR είναι η μέτρηση της ποιότητας σε περίπτωση συμπίεσης ενός βίντεο ή μιας εικόνας. Ορίζεται πιο εύκολα μέσω μιας ποσότητας του μέσου τετραγωνικού σφάλματος (MSE) το οποίο φαίνεται στην Σχήμα 6.1. Όσο αφορά του MSE, έστω I το αρχικό ασυμπίεστο βίντεο και K είναι το βίντεο που προβλήθηκε στον παραλήπτη. Επίσης έστω ότι $m \times n$ είναι η ανάλυση του βίντεο.

Αρχικά υπολογίζουμε τον MSE και στην συνέχεια με την βοήθεια του MSE υπολογίζουμε το τελικό PSNR όπως φαίνεται και πάλι στο Σχήμα 6.1. Το MAX_I ορίζεται ως η μέγιστη τιμή που μπορεί να πάρει ένα εικονοστοιχείο. Για παράδειγμα αν τα εικονοστοιχεία αναπαρίστανται με 8bits ανά δείγμα, τότε η τιμή του ορίζεται στο 255. Η μονάδα μέτρησης του PSNR είναι τα dB (decibel).

$$MSE = \frac{1}{m \cdot n} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i, j) - K(i, j)]^2$$
$$PSNR = 10 \cdot \log_{10} \left(\frac{MAX_I^2}{MSE} \right)$$

Σχήμα 6.1: Υπολογισμός PSNR.

6.3.2. Συντελεστής Συσχέτισης Pearson

Ο συντελεστής συσχέτισης Pearson είναι μια μέθοδος για να γίνει μέτρησης της γραμμικής συσχέτισης μεταξύ δυο μεταβλητών X και Y . Ο Αυτός ο συντελεστής συσχέτισης μπορεί να δώσει τιμές από -1 μέχρι 1 συμπεριλαμβανομένων. Σε περίπτωση που δώσει την τιμή 1 τότε σημαίνει ότι οι μεταβλητές έχουν απολύτως θετική συσχέτιση. Αν επιστρέψει την τιμή 0, τότε δεν υπάρχει συσχέτιση και αν επιτρέψει την τιμή -1 σημαίνει ότι οι δυο μεταβλητές

έχουν απολύτως αρνητική συσχέτιση. Τις πλείστες φορές το γράμμα r αντιπροσωπεύει τον συντελεστή συσχέτισης Pearson και ο τρόπος υπολογισμού του φαίνεται στο σχήμα 6.2.

$$r = r_{xy} = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n \sum x_i^2 - (\sum x_i)^2} \sqrt{n \sum y_i^2 - (\sum y_i)^2}}.$$

Σχήμα 6.2: Υπολογισμός Συντελεστή Συσχέτισης Pearson.

6.3.3. Συντελεστής Συσχέτισης Spearman

Ο συντελεστής συσχέτισης Spearman είναι ένας τρόπος να αξιολογηθεί το πόσο καλά μπορεί να περιγραφεί η σχέση μεταξύ δύο μεταβλητών χρησιμοποιώντας μια μονότονη συνάρτηση. Τις περισσότερες φορές συμβολίζεται με το γράμμα ρ , και είναι ένα μη-παραμετρικό μέτρο της στατιστικής εξάρτησης μεταξύ δύο μεταβλητών.

Οι n πρώτες βαθμολογίες X_i, Y_i μετατρέπονται σε κατατάξεις x_i, y_i , και το ρ υπολογίζεται με τον τρόπο που βλέπουμε στο σχήμα 6.2. Το πρόσημο της συσχέτισης Spearman δείχνει την κατεύθυνση της σχέσης μεταξύ της X που είναι η ανεξάρτητη μεταβλητή και της Y που είναι η εξαρτημένη μεταβλητή. Ο συντελεστής συσχέτισης Spearman είναι θετικός στις

περιπτώσεις όπου η Y τείνει να αυξάνεται όταν η X αυξάνεται. Αρνητικός συντελεστής συσχέτισης Spearman προκύπτει όταν η Y τείνει να μειώνεται όταν η X αυξάνεται. Μια μηδενική συσχέτιση Spearman δείχνει ότι δεν υπάρχει τάση για την Y είτε να αυξηθεί ή να μειωθεί, όταν η X αυξάνεται.

$$\rho = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2 \sum_i (y_i - \bar{y})^2}}.$$

Σχήμα 6.3: Υπολογισμός Συντελεστή Συσχέτισης Spearman.

Κεφάλαιο 7

Αποτελέσματα

7.1.	Αξιολόγηση Κωδικοποιημένων Βίντεο	52
7.2.	Συγκεντρωτικά Αποτελέσματα	54

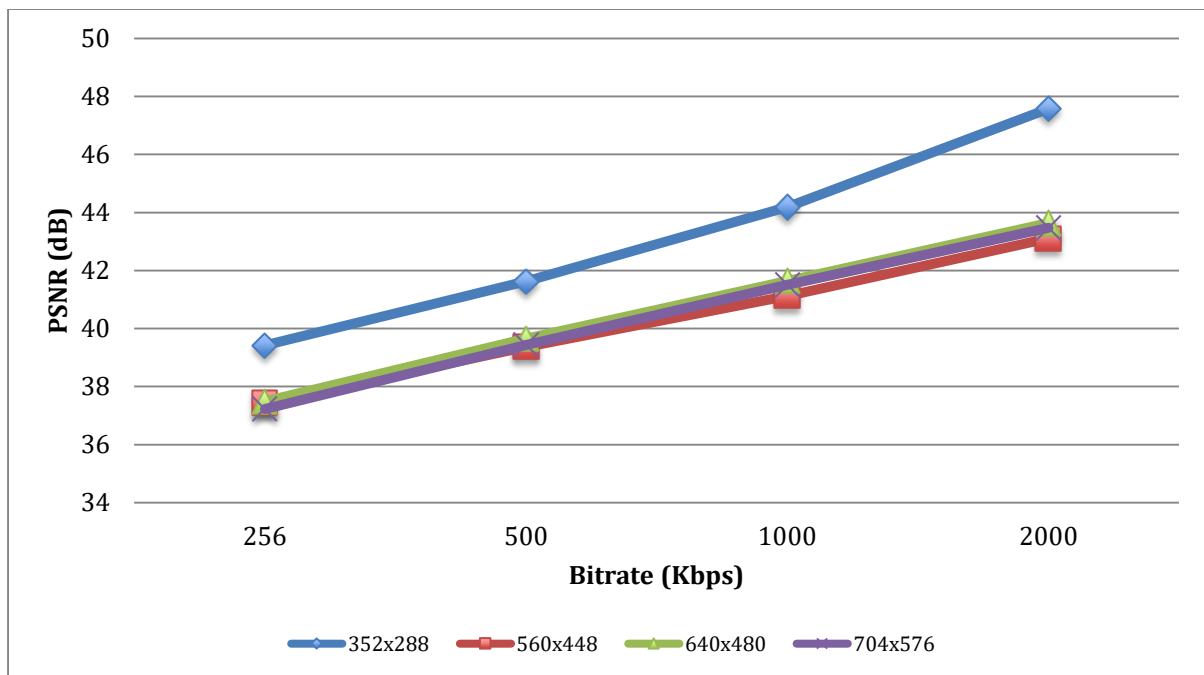
7.1. Αξιολόγηση Κωδικοποιημένων Βίντεο

Στο σχήματα 7.1. που παρουσιάζεται πιο κάτω, φαίνονται οι τιμές του PSNR για ένα από τα δέκα βίντεο καρωτιδικής αρτηρίας που χρησιμοποιήθηκαν για την παρούσα εργασία, μετά την συμπίεση. Η γραφική αυτή παράσταση είναι αντιπροσωπευτική για όλα τα δέκα βίντεο της καρωτιδικής αρτηρίας. Ο λόγος που είναι αντιπροσωπευτική, είναι ότι όλα τα σχόλια που θα γίνουν για αυτό το σχήμα, ισχύουν και για όλα τα υπόλοιπα βίντεο καρωτιδικής αρτηρίας.

Αρχικά, να αναφέρουμε ότι ένα ιατρικό βίντεο καρωτιδικής αρτηρίας, για να θεωρείται διαγνωστικά επαρκή, δεν πρέπει η τιμή του PSNR του να είναι πιο κάτω από 35dB [43]. Αν η τιμή είναι πιο κάτω από αυτό το κατώφλι, σημαίνει ότι ο παραλήπτης δεν θα μπορεί να πάρει από το κωδικοποιημένο βίντεο, την ιατρική πληροφορία που χρειάζεται.

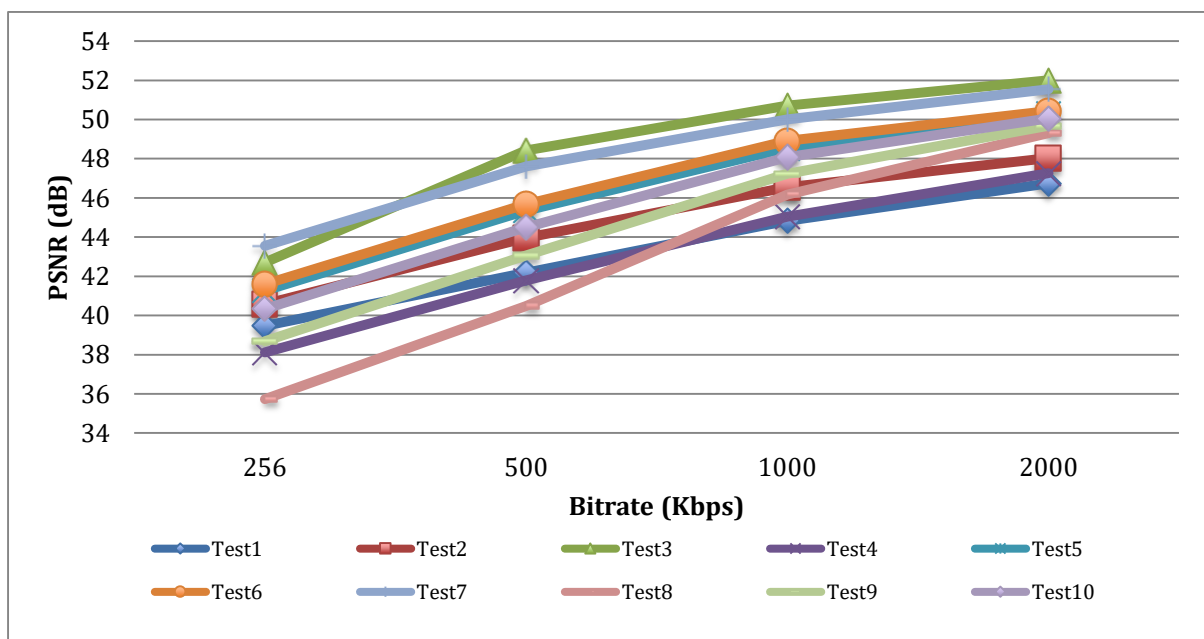
Από το σχήμα που βλέπουμε, παρατηρούμε ότι για όλα τα bitrate σε όλα τα resolutions η τιμή του PSNR είναι πάντοτε πιο μεγάλη από την τιμή που θέσαμε πιο πάνω σαν ελάχιστη τιμή. Αυτό σημαίνει ότι η συμπίεση των βίντεο ήταν επιτυχής σε όλες τις κωδικοποιήσεις, αφού όλα θεωρούνται διαγνωστικά επαρκή.

Ακόμα, παρατηρούμε ότι η τιμή του PSNR αυξάνεται ανάλογα με το πόση είναι η τιμή του bitrate. Δηλαδή όσο αυξάνεται το bitrate, τόσο πιο καλή είναι και η ποιότητα του βίντεο. Μέγιστη τιμή του PSNR έχουν όλα τα βίντεο όταν το bitrate τους είναι 2000k/bps. Συγκρίνοντας επίσης τα τέσσερα σχήματα μεταξύ τους, βλέπουμε ότι όταν η ανάλυση του βίντεο είναι 352x288 η τιμές του PSNR είναι πιο μεγάλες από ότι στις υπόλοιπες περιπτώσεις, και αντίστοιχα πιο μικρές τιμές PSNR έχουν στην ανάλυση 704x576.



Σχήμα 7.1. Γραφική Παράσταση PSNR-Bitrate για διαφορετικές αναλύσεις εικόνες για βίντεο καρωτιδικής αρτηρίας.

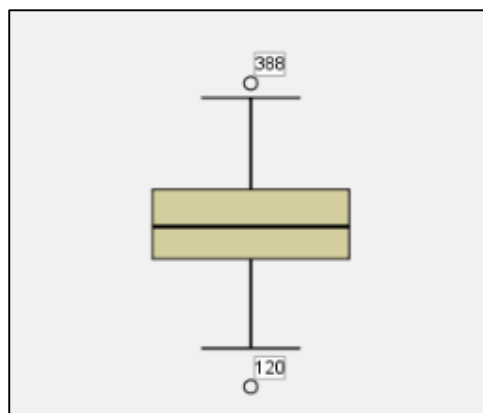
Στο σχήμα 7.2. που ακολουθεί, παρουσιάζονται οι τιμές του PSNR για τα δέκα βίντεο καρδιάς που χρησιμοποιήθηκαν μετά την κωδικοποίηση. Από το σχήμα μπορούμε να συμπεράνουμε ότι μέγιστη PSNR τα βίντεο καρδιάς παίρνουν στο μεγαλύτερο bitrate που έχουν κωδικοποιηθεί, δηλαδή στα 2000Kbits/sec. Αντιθέτως ελάχιστη τιμή παίρνουν στο πιο χαμηλό bitrate, δηλαδή στα 256Kbits/sec.



Σχήμα 7.2. Γραφική Παράσταση PSNR-Bitrate για βίντεο καρδιάς.

7.2. Συγκεντρωτικά Αποτελέσματα

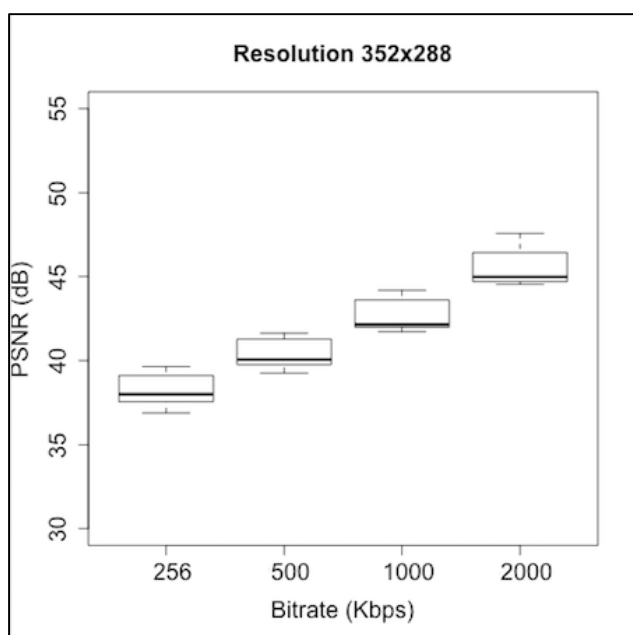
Ένα θηκόγραμμα (box plot) [44] είναι ένα γράφημα μέσω του οποίου μπορούμε να συνοψίσουμε βασικά περιγραφικά μέτρα όπως είναι η διάμεσος, τα τεταρτημόρια, το ενδοτεταρτημοριακό εύρος, καθώς επίσης και τις ακραίες τιμές. Ένα box plot είναι το ορθογώνιο παραλληλόγραμμο, που φαίνεται στο σχήμα 7.2. και αποτελείται από τα εξής δύο κομμάτια: Το πρώτο κομμάτι είναι μια οριζόντια γραμμή στο εσωτερικό του που αντιπροσωπεύει την διάμεσο. Το δεύτερο κομμάτι είναι οι οριζόντιες γραμμές που ονομάζονται φράκτες. Αν κάποια τιμή βρίσκεται εξωτερικά των φρακτών ονομάζεται ακραίο σημείο (outliers) και συμβολίζεται με μια κενή κουκκίδα όπως φαίνεται στο σχήμα 7.2.



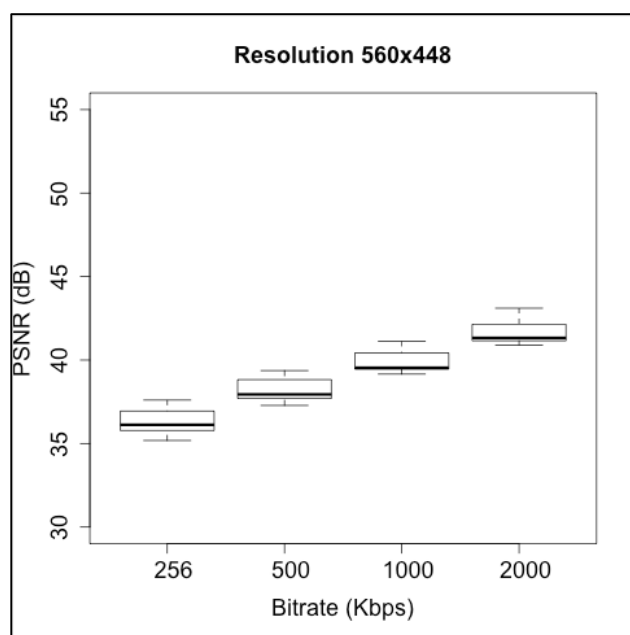
Σχήμα 7.2. Τυπικό θηκόγραμμα [44].

Στα σχήματα 7.4.-7.7 που ακολουθούν, παρουσιάζονται τα boxplots για τα συγκεντρωτικά αποτελέσματα των τιμών του PSNR που έχουν τα βίντεο καρωτιδικής αρτηρίας μετά την κωδικοποίηση. Κάθε box plot αντιπροσωπεύει μια συγκεκριμένη ανάλυση εικόνας και παρουσιάζονται οι τιμές για όλα τα bitrates και για τα δέκα βίντεο που χρησιμοποιήθηκαν.

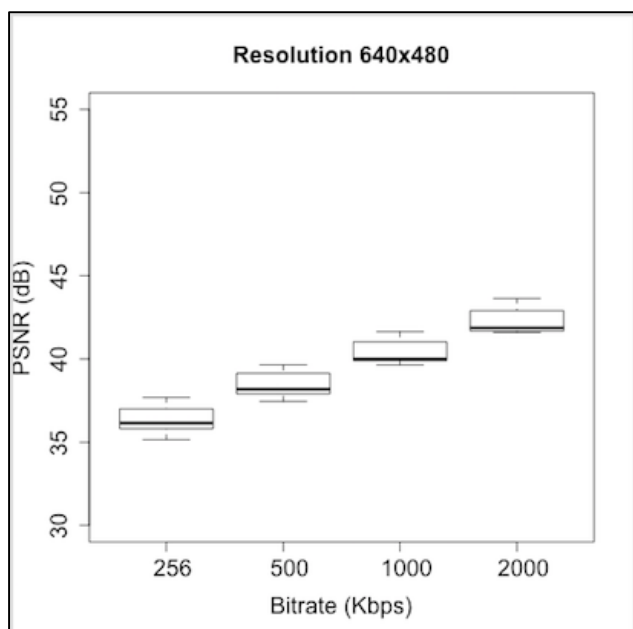
Από τα box plots που βλέπουμε, παρατηρούμε ότι καμία τιμή, σε κανένα θηκόγραμμα δεν είναι ακραίο σημείο άρα καμία τιμή δεν αγνοείται. Επίσης, σε όλες τις αναλύσεις εικόνας παρατηρούμε ότι όσο πιο μεγάλο είναι το bitrate τόσο πιο μεγάλο είναι και το PSNR, επομένως η μέγιστη τιμή που παρουσιάζεται είναι για bitrate 2000Kbits/sec. Τέλος, βλέπουμε ότι καμία τιμή δεν είναι κάτω από το 35dB που είναι το κατώφλι.



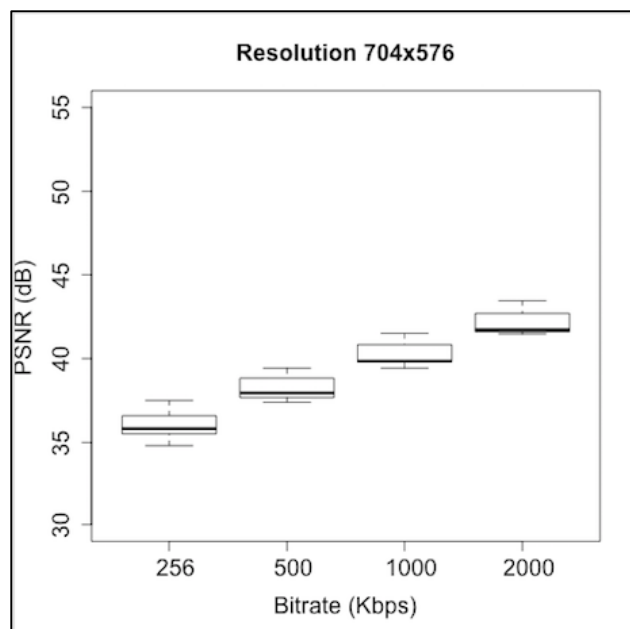
Σχήμα 7.4. Boxplot για PSNR βίντεο καρωτιδικής αρτηρίας με resolution 352x288.



Σχήμα 7.5. Boxplot για PSNR βίντεο καρωτιδικής αρτηρίας με resolution 560x448.

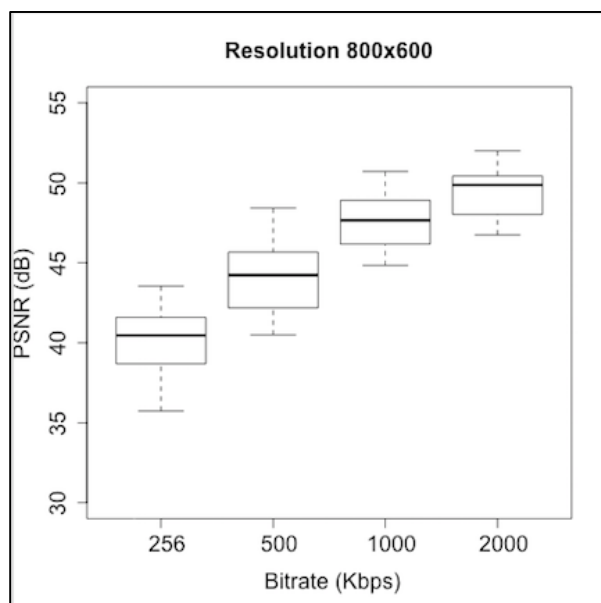


Σχήμα 7.6. Boxplot για PSNR βίντεο καρωτιδικής αρτηρίας με resolution 640x480.



Σχήμα 7.7. Boxplot για PSNR βίντεο καρωτιδικής αρτηρίας με resolution 704x576.

Στο σχήμα 7.8 που φαίνεται πιο κάτω, βλέπουμε το boxplot για τα συγκεντρωτικά αποτελέσματα των τιμών του PSNR που έχουν τα βίντεο καρδιάς μετά την κωδικοποίηση. Όπως και στα βίντεο καρωτιδικής αρτηρίας, έτσι και στα βίντεο καρδιάς δεν έχουμε τιμές που να είναι που να είναι ακραία σημεία, και έτσι καμία τιμή δεν αγνοείται. Παρατηρούμε ότι οι τιμές του PSNR για bitrate 2000Kbits/sec είναι πιο συγκεντρωμένες από ότι στα υπόλοιπα bitrates, αφού οι τιμές καταλαμβάνουν πιο μικρό εύρος τιμών και επίσης στο bitrate αυτό, τα βίντεο παίρνουν την πιο ψηλή τιμή PSNR.



Σχήμα 7.8. Boxplot για PSNR βίντεο καρδιάς με resolution 800x600.

Στο σχήμα 7.9. που βλέπουμε πιο κάτω, φαίνεται ο συγκεντρωτικός πίνακας αποτελεσμάτων για όλες τις τιμές PSNR για τα βίντεο καρωτιδικής αρτηρίας. Στην πρώτη στήλη βλέπουμε το όνομα του βίντεο και το bitrate με το οποίο είναι κωδικοποιημένο. Στις επόμενες τέσσερις στήλες βλέπουμε την ανάλυση στην οποία έχει κωδικοποιηθεί. Όπως βλέπουμε στον πίνακα χρησιμοποιήθηκαν δέκα ασυμπίεστα βίντεο καρωτιδικής αρτηρίας και για κάθε βίντεο υπάρχουν δεκαέξι διαφορετικές κωδικοποιήσεις.

Ακόμα, στο σχήμα 7.10. παρουσιάζεται ο συγκεντρωτικός πίνακας αποτελεσμάτων για όλες τις τιμές του PSNR για τα βίντεο καρδιάς. Στην πρώτη στήλη παρουσιάζεται το όνομα κάθε βίντεο, και στις επόμενες τέσσερις στήλες παρουσιάζεται το bitrate στο οποίο έχουν κωδικοποιηθεί. Όσο αφορά τα βίντεο καρδιάς, χρησιμοποιήθηκαν δέκα βίντεο καρδιάς και για κάθε βίντεο υπάρχουν τέσσερις διαφορετικές κωδικοποιήσεις.

Video Name bitrate	Resolution 352x288	Resolution 560x448	Resolution 640x480	Resolution 704x576
Video_1 256k	37.556	35.768	35.811	35.509
Video_1 500k	39.763	37.640	37.907	37.671
Video_1 1000k	41.994	39.279	39.824	39.648
Video_1 2000k	44.694	40.898	41.609	41.474
Video_2 256k	37.331	35.777	35.776	35.458
Video_2 500k	39.598	37.703	37.875	37.662
Video_2 1000k	42.030	39.546	39.969	39.797
Video_2 2000k	45.069	41.444	41.977	41.825
Video_3 256k	39.119	36.941	37.001	36.588
Video_3 500k	41.279	38.827	39.142	38.836
Video_3 1000k	43.608	40.424	41.028	40.836
Video_3 2000k	46.632	42.144	42.900	42.704
Video_4 256k	38.335	36.569	36.593	36.282
Video_4 500k	40.632	38.500	38.791	38.562
Video_4 1000k	42.958	40.159	40.718	40.573
Video_4 2000k	45.798	41.883	42.598	42.482
Video_5 256k	37.838	35.971	35.970	35.648
Video_5 500k	39.891	37.865	38.080	37.877
Video_5 1000k	41.887	39.456	39.910	39.809
Video_5 2000k	44.535	41.154	41.683	41.599
Video_6 256k	38.123	36.224	36.229	35.898

Video_6 500k	40.204	38.018	38.269	38.014
Video_6 1000k	42.267	39.529	40.031	39.885
Video_6 2000k	44.891	41.207	41.766	41.636
Video_7 256k	39.420	37.442	37.492	37.230
Video_7 500k	41.633	39.369	39.649	39.438
Video_7 1000k	44.190	41.124	41.640	41.504
Video_7 2000k	47.576	43.106	43.639	43.467
Video_8 256k	38.876	36.020	36.073	35.741
Video_8 500k	39.928	37.859	38.107	37.877
Video_8 1000k	42.041	39.469	39.963	39.816
Video_8 2000k	44.711	41.171	41.737	41.613
Video_9 256k	39.655	37.602	37.685	37.505
Video_9 500k	41.631	39.245	39.577	39.410
Video_9 1000k	43.704	40.660	41.211	41.105
Video_9 2000k	46.434	42.335	42.963	42.822
Video_10 256k	36.887	35.183	35.158	34.798
Video_10 500k	39.263	37.282	37.447	37.386
Video_10 100k	41.722	39.158	39.634	39.439
Video_10 2000k	44.706	40.938	41.640	41.495

Σχήμα 7.9. Πίνακας συγκεντρωτικών αποτελεσμάτων για το PSNR των βίντεο καρωτιδικής αρτηρίας μετά την κωδικοποίηση.

Video Name	Bitrate 256k	Bitrate 500k	Bitrate 1000k	Bitrate 2000k
Test1	39.473	42.183	44.831	46.749
Test2	40.585	43.961	46.539	48.027
Test3	42.711	48.419	50.717	52.003
Test4	38.123	41.799	45.047	47.258
Test5	41.266	45.353	48.548	50.280
Test6	41.592	45.673	48.907	50.431
Test7	43.554	47.604	50.017	51.550
Test8	35.732	40.485	46.176	49.300
Test9	38.688	43.064	47.229	49.679
Test10	40.340	44.509	48.086	50.057

Σχήμα 7.10. Πίνακας συγκεντρωτικών αποτελεσμάτων για το PSNR των βίντεο καρδιάς μετά την κωδικοποίηση.

Στο σχήμα 7.11. που βλέπουμε πιο κάτω φαίνονται τα αποτελέσματα των συντελεστών συσχέτισης Pearson και Spearman της κλινικής και αντικειμενικής αξιολόγησης για τα βίντεο καρωτιδικής αρτηρίας. Υπολογίζεται ο συντελεστής συσχέτισης ανάμεσα στην αντικειμενική αξιολόγηση PSNR και στην κλινική αξιολόγηση του παραλήπτη. Οι συντελεστές συσχέτισης όπως έχουμε προαναφέρει παίρνουν τιμές απο -1 μέχρι 1 συμπεριλαμβανομένων. Η τιμή 1 σημαίνει ότι οι δυο μεταβλητές έχουν απολύτως θετική συσχέτιση.

Correlation	Resolution 352x288	Resolution 560x448	Resolution 640x480	Resolution 704x576
Pearson	0.415	0.431	0.477	0.418
Spearman	0.466	0.459	0.472	0.475

Σχήμα 7.11. Πίνακας συντελεστών συσχέτισης.

Κεφάλαιο 8

Συμπεράσματα και Μελλοντική Εργασία

8.1 Συμπεράσματα	60
8.2 Μελλοντική Εργασία	61

8.1. Συμπεράσματα

Στόχος της παρούσας διπλωματικής εργασίας ήταν η ανάπτυξη ενός συστήματος ιατρικής εκπαίδευσης και αξιολόγησης ιατρικού βίντεο όπου τα βίντεο θα στέλνονταν από τον αποστολέα με προσαρμοστική μετάδοση σε πραγματικό χρόνο, με την χρήση του προτύπου MPEG-DASH.

Επίσης, γνωρίζοντας την ύπαρξη τεχνολογιών ανοικτού κώδικα, ένας επιπλέον στόχος της εργασίας αυτής ήταν οι ανεξάρτητες τεχνολογίες να συνδυαστούν για την δημιουργία του τελικού συστήματος. Ο συνδυασμός τους θα ενσωμάτωνε επίσης και το πρότυπο προσαρμοστικής μετάδοσης MPEG-DASH.

Με τις τεχνολογίες ανοικτού κώδικα που χρησιμοποιήθηκαν, το κύριο πλεονέκτημα είναι ότι το λογισμικό ανοικτού κώδικα είναι δωρεάν και δεν αποτελεί κάποια οικονομική επιβάρυνση για αυτόν που το υλοποιεί. Εκτός αυτού, λόγο του ότι ο κώδικας είναι ελεύθερα διαθέσιμος και υπάρχει η δυνατότητα αλλαγής του, μπορεί εύκολα να προσαρμοστεί στις ανάγκες οποιασδήποτε εφαρμογής.

Το τελικό σύστημα ιατρικής εκπαίδευσης και μετάδοσης ιατρικού βίντεο σε πραγματικό χρόνο, έχει πετύχει τον στόχο του στο να συνδυάσει σωστά τις τεχνολογίες ανοικτού κώδικα και να ενσωματώσει το πρότυπο προσαρμοστικής μετάδοσης MPEG-DASH. Ένα από τα πλεονεκτήματα του συστήματος αυτού είναι ότι, είναι εύκολα προσβάσιμο σε οποιονδήποτε θελήσει να το χρησιμοποιήσει. Αυτό γιατί, δεν χρειάζεται να γίνει εγκατάσταση του συστήματος, χρειάζεται απλά ο ενδιαφερόμενος να επισκεφθεί την σελίδα του συστήματος. Ακόμα ένα πλεονέκτημα, είναι ότι μέσω του συστήματος αυτού παρέχεται η καλύτερη

δυνατή ποιότητα του βίντεο στο διαθέσιμο εύρος ζώνης του δικτύου, αφού το σύστημα ανιχνεύει το διαθέσιμο εύρος ζώνης και διαλέγει την κατάλληλη κωδικοποίηση.

Όπως σε όλα τα συστήματα που ασχολούνται με την αποστολή ιατρικού βίντεο, έτσι και σε αυτό, ένας από τους πιο σημαντικούς στόχος ήταν να γίνεται αξιόπιστη μετάδοση του ιατρικού βίντεο. Η μετάδοση του ιατρικού βίντεο θα έπρεπε να προσαρμόζεται στις συνθήκες του διαθέσιμου ασύρματου δικτύου έτσι ώστε το βίντεο να είναι ιατρικά διαγνώσιμο.

Στο σύστημα έγιναν πειραματικές αξιολογήσεις για να γίνει η μέτρηση της ποιότητας του ιατρικού συμπίεσμένου βίντεο. Τα πειράματα αυτά έγιναν πάνω σε δέκα ασυμπίεστα βίντεο καρωτιδικής αρτηρίας και σε δέκα ασυμπίεστα βίντεο καρδιάς. Η αντικειμενική αξιολόγηση της ποιότητας του βίντεο έγινε με τον αλγόριθμό PSNR και έγινε επίσης κλινική αξιολόγηση της ποιότητας του βίντεο από τους παραλήπτες. Τα πειράματα έδειξαν ότι η ποιότητα των βίντεο είναι τέτοια ώστε το βίντεο να είναι ιατρικά διαγνώσιμο.

8.2. Μελλοντική Εργασία

Στην παρούσα εργασία όπως αναφέραμε στο υποκεφάλαιο 5.2.4. ο παραλήπτης αρχικά καλείται να επιλέξει από ένα σύνολο κατηγοριών υπερηχογραφημάτων. Σαν μελλοντική εργασία θα μπορούσε όλες οι κατηγορίες, όπου την δεδομένη στιγμή είναι υπό κατασκευή, να ολοκληρωθούν έτσι ώστε η εργασία να περιλαμβάνει ιατρικά βίντεο διαφόρων κατηγοριών. Έτσι στο μέλλον η εργασία αυτή θα μπορεί να χρησιμοποιείται από φοιτητές ιατρικής ως ένα εκπαιδευτικό υλικό. Οι φοιτητές θα μπορούν να βλέπουν διάφορα υπερηχογραφήματα στα πλαίσια κάποιων μαθημάτων για να παίρνουν την απαραίτητη γνώση.

Επίσης, θα μπορούσε στην σελίδα του παραλήπτη να υπάρχει χώρος στον οποίο να υπάρχει μια περιγραφή του συγκεκριμένου υπερηχογραφήματος και πάλι για σκοπούς εκπαίδευσης ή για σκοπούς ενημέρωσης του παραλήπτη. Ακόμα, στην συγκεκριμένη σελίδα, προτείνεται να εισαχθούν επιπλέον μπάρες αξιολόγησης για να μπορεί ο παραλήπτης να αξιολογεί το βίντεο και για διάφορα άλλα είδη. Επιπλέον προτεινόμενη μελλοντική εργασία για την διεπιφάνεια αυτή, είναι να προστεθεί και ένα πλαίσιο στο οποίο ο παραλήπτης θα μπορεί να εισάγει τα δικά του σχόλια έτσι ώστε οι επόμενοι επισκέπτες να μπορούν να τα βλέπουν.

Τέλος, λόγο του ότι στην συγκεκριμένη εργασία ο αριθμός των ιατρικών βίντεο που χρησιμοποιήθηκαν δεν ήταν αρκετά μεγάλος για να μπορούν τα αποτελέσματα να είναι πιο

αντιπροσωπευτικά. Επομένως σαν προτεινόμενη μελλοντική εργασία, είναι να γίνουν περισσότερα πειράματα με μεγαλύτερο αριθμό βίντεο και περισσότερες κωδικοποιήσεις όσο αφορά την ανάλυση των βίντεο καθώς επίσης και τα bitrates.

Βιβλιογραφία

- [1] “Τηλειατρική.” Wipidea: The Free Encyclopedia. Wikimedia Foundation, Inc. 6th January 2016. Web. 10th January 2016. Available: https://el.wikipedia.org/wiki/Τηλεϊατρική#cite_note-3
- [2] American Telemedicine Association. Telemedicine/Telehealth Terminology, American Telemedicine Association website, Washington, D.C. Retrieved August 21, 2011.
- [3] Sachpazidis, Ilias Image and Medical Data Communication Protocols for Telemedicine and Teleradiology (dissertation), Department of Computer Science, Technical University of Darmstadt, Germany, 10 July 2008.
- [4] Kontaxakis, George; Visvikis, Dimitris; Ohl, Roland; Sachpazidis, Ilias; Suarez, Juan Pablo; Selby, Boris Peter; et al.: "Integrated Telemedicine Applications and Services for Oncological Positron Emission Tomography", Oncology Reports, Vol.15: 1091–1100, 2006
- [5] A. Panayides, Diagnostically Resilient Encoding, Wireless Transmition, and Quality Assessment of Medical Video, Cyprus: PhD dissertation, University of Cyprus, 2011.
- [6] Νικόλας Τσαπατσούλης, ΒΕΣ 04: Συμπίεση και Μετάδοση Πολυμέσων, Συμπίεση Βίντεο: Αρχές και Πρότυπα Συμπίεσης. Ελλάδα 2006. Available: <http://users.iit.demokritos.gr/~ntsap/courses/bes04/lectures/mm10.pdf>
- [7] Νικόλας Τσαπατσούλης, ΔΤΨΣ 150: Ψηφιακή Επεξεργασία Εικόνας, Ελλάδα: Πανεπιστήμιο Πειραιώς, 2005.
- [8] Νικόλας Τσαπατσούλης, ΕΠΛ 422: Συστήματα Πολυμέσων, Cyprus: Univeristy of Cyprus, 2005. Available: <https://www.cs.ucy.ac.cy/~nicolast/courses/cs422/lectures/mm21.pdf>
- [9] G. S. Jens-Rainer Ohm, "The Moving Picture Experts Group website," July 2005. [Online]. Available: <http://mpeg.chiariglione.org/standards/mpeg-4/advanced-video-coding>.
- [10] Θεοχαρίδου Φωτεινή, Πτυχιακή Εργασία: Πρότυπο τηλεδιάσκεψης H.323, Άρτα

2006. Available:
http://apothetirio.teiep.gr/xmlui/bitstream/handle/123456789/135/tlp_000408.pdf?sequence=1.
- [11] Stefanos Kollias, National Technical University of Athens, GREECE, for contributions to intelligent systems for multimedia content analysis and human-machine interaction, 2015. Available: <http://www.image.ntua.gr>
 - [12] W. B. Pennebaker and J. L. Mitchell, JPEG still image data compression standard, 1993.
 - [13] J. D. Glibson, T. Berger, T. Lookabaugh, D. Lindbergh and R. L. Baker, Digital Compression for Multimedia: Principles and Standards.
 - [14] Κ. Παττίχης Κωνσταντίνος, ΕΠΑ 445, Ψηφιακή Επεξεργασία Εικόνας, “Ψηφιακό Βίντεο”.
 - [15] Nick Kounis, Digital Quanta Of Noise Shared With The World, 8th December, 2015. Available: <https://quantaofnoise.wordpress.com/2015/12/08/κεφάλαιο-2-ανάλυση-προτύπων-h-264/>
 - [16] Μπάλλου Αλέξανδρου, Διπλωματική Εργασία: Υλοποίηση συστήματος P2P multimedia Streaming. University of Macedonia, 2014
 - [17] L. Richardo, “An Overview of H.264 Advanced,” Vcodex White Paper.
 - [18] http://images.slideplayer.com/8/2409537/slides/slide_49.jpg
 - [19] “Ασύρματο Δίκτυο.” Wipidea: The Free Encyclopedia. Wikimedia Foundation, Inc. 5th December 2015. Web. 13th January 2016. Available: https://el.wikipedia.org/wiki/Ασύρματο_δίκτυο

- [20] Ειρήνη Ελευθερίου, Διατριβή Μεταπτυχιακού: Ολοκληρωμένο σύστημα μετάδοσης ιατρικού βίντεο σε πραγματικό χρόνο χρησιμοποιώντας τεχνολογίες ανοικτού κώδικα, Πανεπιστήμιο Κύπρου, Ιούνιος 2012.
- [21] https://product.tdk.com/en/techlibrary/archives/vol03_alt/img/con02_02.jpg
- [22] "What is 1G or First generation of wireless telecommunication technology?," [Online]. Available: <http://www.clear doubts.com/technology/what-is-1g-or-first-generation-of-wireless-telecommunication-technology/>.
- [23] Κόλλια Αναστασία, Προπτυχιακή εργασία στο μάθημα «Ευρυζωνικές Τεχνολογίες» με θέμα «Πέμπτης γενιάς κινητά δίκτυα επικοινωνιών ή Πέμπτη γενιά ασύρματων δικτύων (5G)», Πάτρα, Ιούνιος 2014.
- [24] Λύκου Άννα Αικατερίνη, Πτυχιακή Εργασία: «Οι γενεές των δικτύων κινητής τηλεφωνίας και παρουσίαση των συστημάτων 4^{ης} γενεάς LTE και LTE-Advanced», ΑΤΕΙ Θεσσαλίας, Λάρισα, 2014.
- [25] M. Jaloun and Z. Guennoun, "Wireless Mobile Evolution to 4G Network," *Wireless Sensor Network*, vol. 2, pp. 309-317, 2010.
- [26] "WiMAX." Wipedia: The Free Encyclopedia. Wikimedia Foundation, Inc. 30th June 2014. Web. 15th April 2016. Available: <https://el.wikipedia.org/wiki/WiMAX>.
- [27] <http://www.newsbomb.gr/bombplus/tehnologia/story/464673/tehnologia-4g-osa-prepei-na-gnorizete>
- [28] 4G Americas, "Recommendations on 5G Requirements and Solutions", October 2014.
- [29] Αθανάσης Καρούλης, «Διαδυκτιακά πρωτόκολλα ροής (Streaming Protocols) RTP και RTSP», Ελλάδα Ιανουάριος 2001.
- [30] 4G Americas, «Supporting WirelessVideo Growth and Trends », April 2013.
- [31] Hassoun, David. Adobe Developer Connection. Adobe Systems.. Blog. [Accessed 24 June 2010].

- [32] Jordan, Larry (10 June 2013). "The Basics of HTTP Live Streaming". Larry's Blog. Larry Jordan & Associates. Retrieved 18 June 2013.
- [33] "Smooth Streaming". IIS.net. Archived from the original on 15 June 2010.
- [34] Anthony Vetro, Mitsubishi Electric Research Labs "The MPEG-DASH Standard for Multimedia Streaming Over The Internet", October-December 2011.
- [35] "Λογισμικό Ανοικτού Κώδικα" Wipidea: The Free Encyclopedia. Wikimedia Foundation, Inc. 19th January 2016. Web. 3th April 2016. Available: https://el.wikipedia.org/wiki/Λογισμικό_ανοικτού_κώδικα.
- [36] "FFmpeg." Wipidea: The Free Encyclopedia. Wikimedia Foundation, Inc. 28th March 2016. Web. 30th March 2016. Available: <https://en.wikipedia.org/wiki/FFmpeg>.
- [37] "FFmpeg. A complete, cross-platform solution to record, convert and stream audio and video" Available: <http://ffmpeg.org>.
- [38] "ffserver Documentation" Available: <https://ffmpeg.org/ffserver.html>.
- [39] "x264" Wipidea: The Free Encyclopedia. Wikimedia Foundation, Inc. 21th March 2016. Web. 31th March 2016. Available: <https://en.wikipedia.org/wiki/X264>.
- [40] "MP4Box: fragmentation, segmentation, splitting and interleaving", Φεβρουάριος 2011, Available: <https://gpac.wp.mines-telecom.fr/2011/02/02/mp4box-fragmentation-segmentation-splitting-and-interleaving/>.
- [41] "Wireshark" Wipidea: The Free Encyclopedia. Wikimedia Foundation, Inc. 11th March 2016. Web. 31th March 2016. Available: <https://en.wikipedia.org/wiki/Wireshark>.
- [42] Q. Huynh-Thu and M. Ghanbari, "Scope of validity of PSNR in image/video quality assessment," The Institution of Engineering and Technology, vol.44, no. 13, pp.80 0-801, 2008.
- [43] A. Panayides, M. S. Pattichis, C. P. Pattichis, M. Pantziaris and A. Pitsillides,

“Atherosclerotic Plaque Ultrasound Video Encoding, Wireless Transmission, and Quality Assessment Using H.264,” *IEEE Transactions on Information Technology in Biomedicine*, vol. 15, no. 3, pp. 387-397, May 2011.

- [44] Δημήτρης Σωτηρόπουλος, Δημοκρήτσιο Πανεπιστήμιο Θράκης, Τμήμα Διοίκησης Επιχειρήσεων “Μέτρα διασποράς”, 30 Νοεμβρίου 2012.