



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ
ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΣΥΣΤΗΜΑΤΩΝ ΜΕΤΑΔΟΣΗΣ ΠΛΗΡΟΦΟΡΙΑΣ ΚΑΙ
ΤΕΧΝΟΛΟΓΙΑΣ ΥΛΟΚΩΝ

Αξιολόγηση Αλγορίθμων Κωδικοποίησης Πολυμεσικών Αρχείων για
Εφαρμογές στη Βιοϊατρική Τεχνολογία

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Μπεζεριάνος Φιλόθεος

Επιβλέπων: Δ. Κουτσούρης
Καθηγητής Ε.Μ.Π.

Αθήνα, Νοέμβριος 2008

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

Copyright © Μπεζεριάνος Φιλόθεος, 2008
Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης. Η αναφορά του ονόματος του συγγραφέα θα αποτελούσε για αυτόν μια χαρά. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς τον συγγραφέα.

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

Θα ήθελα να ευχαριστήσω θερμά τον καθηγητή μου κ. Δ. Κουτσούρη για τη δυνατότητα που μου έδωσε να εκπονήσω την παρούσα διπλωματική εργασία σε έναν ενδιαφέροντα τομέα της βιοϊατρικής. Επίσης, ευχαριστώ τον υποψήφιο διδάκτορα κ. Κ. Περάκη για την πολύτιμη βοήθεια του και τις πολύτιμες παρατηρήσεις του κατά τη συγγραφή της διπλωματικής εργασίας. Τέλος θα ήθελα να ευχαριστήσω την οικογένειά μου για την στήριξη τους καθ' όλη τη διάρκεια των σπουδών μου.

Μπεζεριάνος Φιλόθεος
Νοέμβριος 2008

Περίληψη

Τα τελευταία χρόνια οι συνεχόμενες εξελίξεις στο χώρο των τηλεπικοινωνιών και της τεχνολογίας έχουν δραματική επίπτωση και στον τομέα της ιατρικής. Από τη διάγνωση μέχρι τη θεραπεία η τεχνολογία παίζει έναν όλο και σημαντικότερο ρόλο. Το ιατρικό βίντεο (medical video) και ο ήχος δεν αποτελούν εξαίρεση. Η εύκολη πρόσβαση σε ιατρικά βίντεο και τρισδιάστατες απεικονίσεις όπως η ηλεκτρονική τομογραφία (computed tomography, CT), η μαγνητική τομογραφία (magnetic resonance imaging, MRI) και το ηχοκαρδιογράφημα (echocardiography) δίνουν στους επαγγελματίες υγείας τη δυνατότητα ευκολότερης ανάλυσης και διάγνωσης της κατάστασης του ασθενούς. Ωστόσο, τα ιατρικά βίντεο και η τρισδιάστατη απεικόνιση έχουν μεγάλες απαιτήσεις σε αποθηκευτική μνήμη και το εύρος ζώνης δικτύου που απαιτείται για την μετάδοσή τους είναι αρκετά μεγάλο. Έτσι είναι απαραίτητη η μείωση των δεδομένων (data) που απαιτούνται για την απεικόνιση του ιατρικού βίντεο.

Ενώ σε κάποιες εφαρμογές μπορούμε να συμβιβαστούμε με χαμηλότερη ποιότητα για να έχουμε μεγαλύτερο ρυθμός συμπίεσης (compression ratio), τα ιατρικά βίντεο απαιτούν και υψηλή ποιότητα και υψηλή συμπίεση. Από τη μία μεριά σε περιπτώσεις διάγνωσης είναι σημαντικό να μην υπάρχουν αισθητές απώλειες λεπτομερειών ή αλλοίωση εικόνας, τα οποία μπορεί να οδηγήσουν σε παρερμηνεία της διάγνωσης και από την άλλη, καθώς το μέγεθος της πληροφορίας του ιατρικού βίντεο είναι μεγάλο και υπάρχουν περιορισμοί τόσο στην αποθήκευση αλλά και στην μεταφορά, είναι απαραίτητο το υψηλό ποσοστό συμπίεσης.

Η συμπίεση των δεδομένων δεν αφορά μόνο τα δεδομένα βίντεο. Πολύ σημαντική είναι και η καλή ποιότητα στη μετάδοση του ήχου πάνω από δίκτυα τηλεϊατρικής. Η καθαρή και χωρίς διακοπές επικοινωνία μεταξύ δύο κόμβων του δικτύου είναι ένας παράγοντας που πρέπει να καλύπτεται για όλες τις περιπτώσεις και τις εφαρμογές. Με χρήση της τεχνολογίας VoIP και του κατάλληλου αλγορίθμου συμπίεσης μπορούμε να έχουμε αξιόπιστη επικοινωνία σε πραγματικό χρόνο.

Σκοπός αυτής της διπλωματικής εργασίας ήταν η ανάλυση της συμπεριφοράς και εξαγωγή συνολικών συμπερασμάτων σχετικά με την απόδοση των δημοφιλέστερων αλγορίθμων συμπίεσης video και ήχου, υπό ορισμένες συνθήκες. Η ανάλυση αυτή περιλαμβάνει τον έλεγχο της συμπεριφοράς των αλγορίθμων για συγκεκριμένες συνθήκες δικτύου και μεταφοράς πακέτων για ένα εύρος εφαρμογών.

Λέξεις κλειδιά

Αλγόριθμοι συμπίεσης δεδομένων, συμπίεση ιατρικού βίντεο, H.264, H.263, H.263+, H.263++, MPEG-4, MPEG-2, Motion JPEG 2000, DivX, WMV, XviD, G.711, G.721, G.722, G.723, G.723.1, G.726, G.728, G.729, G.729A

Abstract

The last few years the evolutions in the field of telecommunications and technology have a great impact on the telemedicine field. From diagnosis to treatment technology has an important role. Medical video and sound are not an exception. Easy access to medical videos and three-dimensional images such as computed tomography, magnetic resonance imaging and echocardiography give the possibility to medical doctors for a better analysis and diagnosis of the patients' condition. However, the requirements of medical video and three-dimensional images transmission are very high, as they need a large amount of storage memory and bandwidth. So, the compression of the data that are needed for the representation of the medical video is necessary.

In some applications we can change the quality with compression ratio, but in telemedicine high quality and high compression are both necessary. During the process of diagnosis is important not to have detail loss or noticeable artifacts, which can lead the medical doctor to false diagnosis. However, as the amount of medical information is huge and there are limitation in memory and bandwidth the high compression ratio is necessary.

Data compression is not only for medical video. The sound and speech quality over telemedicine networks is very important. The clear and continuous communication between the nodes of the network is a parameter that must fulfilled under any circumstances and applications. With VoIP technology and the correct codec we can achieve a reliable real time communication.

The scope of this thesis was the analysis of the behavior and the presentation of some general results regarding the performance of the most famous video and sound compression algorithms, under certain network and transmission settings.

Key words

Data compression algorithms, medical video compression, H.264, H.263, H.263+, H.263++, MPEG-4, MPEG-2, Motion JPEG 2000, DivX, WMV, XviD, G.711, G.721, G.722, G.723, G.723.1, G.726, G.728, G.729, G.729A

Διεθνής ορολογία

Στον Πίνακα 1 παρουσιάζεται μερική διεθνής ορολογία που χρησιμοποιήθηκε μαζί με την απόδοση στα ελληνικά.

adaptive differential pulse code modulation, ADPCM	προσαρμοσμένη διαφορική παλμοκωδική διαμόρφωση
aggregation	συσσωμάτωση
backwards-adaptive codec	codec προσαρμοστικό προς τα πίσω
best-effort services	υπηρεσίες καλύτερης προσπάθειας
bit rate handling	διαχείριση του ρυθμού bit
bitstreams	συνεχείς ροές bits
buffering	προσωρινή αποθήκευση
chrominance components	χρωματικά στοιχεία
compactness	περιεκτικότητα
compression ratio	ρυθμός συμπίεσης
computed tomography, CT	ηλεκτρονική τομογραφία
de-interlace	απόπλεξη
decorrelating	αποσυσχέτιση
demodulation	αποδιαμόρφωση
differentiated Services	διαφοροποιημένες υπηρεσίες
differential pulse code modulation, DPCM	διαφορική παλμοκωδική διαμόρφωση
electrical feedback	ηλεκτρική ανάδραση
encapsulation	ενθυλάκωση
echocardiography	ηχοκαρδιογράφημα
encoding process	διαδικασίας κωδικοποίησης
enhancement layer	επίπεδο ενίσχυσης
forward-adaptive codec	codec προσαρμοστικό προς τα εμπρός
forward error correction, FEC	προοπτική διόρθωση λαθών
frame	πλαίσιο
frame-level rate control	έλεγχος ρυθμού επιπέδου πλαισίου
graveness	κρισημότητα
instrument of zerotrees	εργαλείο μηδενικών δέντρων
interlace	σύμπλεξη
jitter	τρέμουλο ή τρεμούλισμα φωνής
long-term predictor	μηχανισμός μακροπρόθεσμης πρόβλεψης
lossless coding	μη-απωλεστική κωδικοποίηση
lossy coding	απωλεστική κωδικοποίηση
luminance component	στοιχείο φωτεινότητας
macroblock, MB	μακρο-μπλοκ
MB-level rate control	έλεγχος ρυθμού επιπέδου μακρο-μπλοκ
magnetic resonance imaging, MRI	μαγνητική τομογραφία
mean absolute difference ratio, MAD	ποσοστό της απόλυτης διαφοράς
modulated noise reference unit, MNRU	διαμορφωμένη μονάδα θορύβου αναφοράς
motion vectors	διανύσματα κίνησης
nasality	ρινισμός
objective Quality Measures	αντικειμενικές μέθοδοι μέτρησης
partitioning	τεμαχισμός
pulse code modulation, PCM	παλμοκωδική διαμόρφωση
quantization parameter, QP	παράμετρος κβαντοποίησης

quantizing	κβανοποίηση
quantum	κβάντο
rate distortion optimization, RDO	βελτιστοποίηση του ρυθμού παραμόρφωσης
scalability	επεκτασιμότητα
sibilation	σφύριγμα
segmental SNR, SEGSNR	τμηματικός σηματοθορυβικός λόγος
source codecs	codecs πηγής
step size	μέγεθος βήματος
subjective quality measures	υποκειμενικές μέθοδοι μέτρησης
target bit rate	επιθυμητός ρυθμός bit
temporal compression	χρονική συμπίεση
transmission layers	επίπεδα μετάδοσης
voicing	φώνηση
waveform codecs	codecs κυματομορφής

Πίνακας 1 - Διεθνής ορολογία και απόδοση στα ελληνικά

Περιεχόμενα

ΠΕΡΙΛΗΨΗ	4
ΔΙΕΘΝΗΣ ΟΡΟΛΟΓΙΑ	6
ΠΡΟΛΟΓΟΣ	13
1. ΕΙΣΑΓΩΓΙΚΑ – ΘΕΩΡΙΑ	15
1.1 ΗΛΕΚΤΡΟΝΙΚΗ ΥΓΕΙΑ	15
1.2 ΔΙΑΔΙΚΑΣΙΑ ΚΩΔΙΚΟΠΟΙΗΣΗΣ – ΠΑΛΜΟΚΩΔΙΚΗ ΔΙΑΜΟΡΦΩΣΗ (PCM)	16
1.2.1 Ανάλυση των λειτουργιών κωδικοποίησης.....	16
1.2.1.1 Δειγματοληψία	16
1.2.1.2 Κβαντοποίηση.....	17
1.2.1.3 Κωδικοποίηση.....	17
1.2.1.4 Αναγέννηση	18
1.2.1.5 Αποκωδικοποίηση	18
1.2.1.6 Φιλτράρισμα.....	19
1.2.1.7 Πολυπλεξία.....	19
1.2.1.8 Συγχρονισμός.....	19
2. ΑΛΓΟΡΙΘΜΟΙ ΚΩΔΙΚΟΠΟΙΗΣΗΣ VIDEO	20
2.1 Το ΠΡΟΤΥΠΟ H.264/AVC	20
2.1.1 Προφίλ του H.264	22
2.2 Τα ΠΡΟΤΥΠΑ H.263, H.263+ ΚΑΙ H.263++	25
2.3 Το ΠΡΟΤΥΠΟ MPEG-4.....	26
2.4 Το ΠΡΟΤΥΠΟ MPEG-2.....	27
2.5 Το ΠΡΟΤΥΠΟ ΜΟΤΙΟΝ JPEG-2000	27
3. ΣΥΓΚΡΙΣΕΙΣ ΑΛΓΟΡΙΘΜΩΝ ΚΩΔΙΚΟΠΟΙΗΣΗΣ VIDEO	29
3.1 ΒΕΛΤΙΣΤΟΠΟΙΗΣΗ ΠΑΡΑΜΟΡΦΩΣΗΣ (ΝΕΑ ΜΕΘΟΔΟΣ ΕΛΕΓΧΟΥ ΡΥΘΜΟΥ BIT)	29
3.1.1 Βελτιστοποίηση παραμόρφωσης.....	29
3.1.2 Νέα μέθοδος για τη βελτιστοποίηση του ρυθμού παραμόρφωσης.....	30
3.1.3 Νέα μέθοδος υπολογισμού της πολυπλοκότητας της κίνησης.....	31
3.1.3 Δέσμευση bits με χρήση αντιληπτών παραμέτρων	32
3.1.4 Νέα μέθοδος ελέγχου του ρυθμού bit.....	33
3.1.5 Πειραματική διαδικασία	33
3.1.6 Αποτελέσματα σύγκρισης μεταξύ των H.264 και MPEG-4	35
3.1.7 Σύγκριση του ελέγχου του ρυθμού bit H014 και της νέας μεθόδου	38
3.1.7.1 Αντικειμενική αξιολόγηση ποιότητας.....	38
3.1.7.2 Υποκειμενική αξιολόγηση ποιότητας	40
3.2 ΣΥΓΚΡΙΣΗ ΤΩΝ CODECS H.264, H.263 ΚΑΙ H.263+.....	41
3.2.1 Πειραματική διαδικασία	41
3.2.2 Αποτελέσματα σύγκρισης.....	41
3.3 ΣΥΓΚΡΙΣΗ ΤΟΥ H.264, ΤΟΥ MPEG-2 ΚΑΙ ΤΟΥ H.263	44
3.3.1 Πειραματική διαδικασία	44
3.3.2 Αποτελέσματα σύγκρισης του H.264 με το MPEG-2.....	44
3.3.3 Αποτελέσματα σύγκρισης του H.264 με το H.263 με βασικό προφίλ.....	47
3.3.4 Αποτελέσματα σύγκρισης του H.264 με το H.263 με προφίλ υψηλής συμπίεσης	50
3.4 ΣΥΓΚΡΙΣΗ ΤΟΥ ΠΡΟΤΥΠΟΥ H.264, ΤΟΥ M-JPEG2000 ΚΑΙ ΤΟΥ PACC	51

3.4.1 PACC	51
3.4.2 Πειραματική διαδικασία	52
3.4.2.1 Ρυθμίσεις του Motion-JPEG2000	53
3.4.2.2 Ρυθμίσεις του H.264/AVC	54
3.4.2.3 Ρυθμίσεις του PACC	54
3.4.3 Αποτελέσματα σύγκρισης	54
3.4.4 Πολυπλοκότητα, Λειτουργικότητα και Εφαρμογές	58
3.4.5 Σύνοψη	60
3.5 ΣΥΝΔΥΑΣΜΟΣ MPEG-2 ΚΑΙ H.264	60
3.5.1 Πειραματική διαδικασία	61
3.5.2 Αποτελέσματα και συμπεράσματα	62
3.5.2.1 Foreman	62
3.5.2.2 Coast Guard	63
3.5.3 Σύνοψη	65
3.6 ΣΥΓΚΡΙΣΗ ΤΟΥ H.264/AVC ΜΕ ΤΑ DIVX, WMV, XVID ΚΑΙ MC-EZBC	65
3.6.1 Πειραματική διαδικασία και ρύθμιση των <i>codexs</i>	65
3.6.2 Μέτρηση της ποιότητας	67
3.6.3 Αποτελέσματα	67
3.6.3.1 H.264/AVC και MC-EZBC	70
3.6.3.2 H.264/AVC, WMV9 και DivX	70
3.6.3.3 H.264/AVC και MPEG-4	72
3.6.4 Σύνοψη	74
4. ΣΥΓΚΡΙΣΕΙΣ ΤΟΥ ΠΡΟΤΥΠΟΥ H.264 ΔΙΑΦΟΡΕΤΙΚΩΝ ΚΑΤΑΣΚΕΥΑΣΤΩΝ ΚΑΙ ΤΟΥ MPEG-4	75
4.1 ΣΥΓΚΡΙΣΗ XVID, MAINCONCEPT H.264, INTEL H.264, x264, AMD H.264 ΚΑΙ ARTEMIS H.264	75
4.1.1 Πειραματική διαδικασία	76
4.1.2 Αποτελέσματα	76
4.1.2.1 Video-διάσκεψη	76
4.1.2.2 Ταινίες (movies)	79
4.1.2.2.1 Ταχύτητα κωδικοποίησης	83
4.1.2.2.2 Ανταλλαγή ταχύτητας / ποιότητας	86
4.1.2.2.3 Διαχείριση ρυθμού bit	91
4.1.2.3 HDTV	95
4.1.3 Συμπεράσματα	97
4.1.3.1 Video-διάσκεψη	97
4.1.3.1.1 Υψηλή ποιότητα	97
4.1.3.1.2 Υψηλή ταχύτητα	97
4.1.3.2 Ταινίες	99
4.1.3.2.1 Υψηλή ποιότητα	99
4.1.3.2.2 Υψηλή ταχύτητα	99
4.1.3.3 HDTV	101
4.1.3.3.1 Υψηλή ποιότητα	101
4.1.3.3.2 Υψηλή ταχύτητα	101
4.1.3.4 Συνολικά συμπεράσματα	103
4.1.4 Σύγκριση του <i>Artemis x264</i> και του <i>x264</i>	104
4.2 ΣΥΓΚΡΙΣΗ DIVX 6.2.5, MAINCONCEPT H.264, INTEL H.264, VSS H.263, x264 ΚΑΙ APPLE H.264	106
4.2.1 Πειραματική διαδικασία	107
4.2.2 Video-διάσκεψη	107
4.2.2.1 Αποτελέσματα	108
4.2.2.2 Συμπεράσματα	112
4.2.3 Ταινίες	112
4.2.3.1 Αποτελέσματα	113

4.2.3.2 Συμπεράσματα	119
4.2.4 HDTV	120
4.2.4.1 Αποτελέσματα	120
4.2.4.2 Συμπεράσματα	123
4.2.5 Συγκεντρωτικά αποτελέσματα και συμπεράσματα	123
4.2.6 Σύγκριση των <i>codecs</i> Apple και Sorenson	126
4.2.6.1 Συμπεράσματα	128
4.3 ΣΥΓΚΡΙΣΗ ΤΩΝ MPEGABLE AVC, MOONLIGHT H.264, MAINCONCEPT H.264, FRAUNHOFER IIS, ATEME MPEG-4 AVC / H.264, VIDEOSOFT H.264 ΚΑΙ DIVX PRO 5.1.1	128
4.3.1 Πειραματική διαδικασία	129
4.3.2 Διαγράμματα PSNR / ρυθμού <i>bit</i>	129
4.3.2.1 Video Bankomatdi	129
4.3.2.2 Video Battle	131
4.3.2.3 Video Bbc3di	133
4.3.2.4 Video Foreman	135
4.3.2.5 Video Susidi	137
4.3.3 Διαγράμματα μεταβολής της φωτεινότητας	139
4.3.3.1 Video Bankomatdi	139
4.3.3.2 Video Battle	140
4.3.3.3 Video Bbc3di	141
4.3.3.4 Video Foreman	142
4.3.3.5 Video Susidi	143
4.3.4 Διαγράμματα διαχείρισης ρυθμού <i>bit</i>	143
4.3.4.1 Video Bankomatdi	144
4.3.4.2 Video Battle	145
4.3.4.3 Video Bbc3di	146
4.3.4.4 Video Foreman	147
4.3.4.5 Video Susidi	148
4.3.5 Σύγκριση ανά πλαίσιο	148
4.3.6 Οπτική σύγκριση των H.264 <i>codecs</i> και των DivX <i>codecs</i>	157
4.3.6.1 Video Bbc3di, frame 280	157
4.3.6.2 Video Foreman, frame 282 και 9	158
4.3.7 Συγκεντρωτικά συμπεράσματα	159
4.3.8 Ένας ανεπίσημος τρόπος σύγκρισης	159
4.4 ΒΙΔΕΟ ΠΟΥ ΧΡΗΣΙΜΟΠΟΙΗΘΗΚΑΝ	162
4.4.1 Videos τηλεδιάσκεψης	162
4.4.1.1 Salesman	162
4.4.1.2 Foreman	163
4.4.1.3 News	164
4.4.1.4 Akiyo	165
4.4.1.5 Carphone	166
4.4.2 Videos ταινιών	167
4.4.2.1 Battle	167
4.4.2.2 Smith	168
4.4.2.3 Iceage	169
4.4.2.4 Lord of the rings	170
4.4.2.5 Rancho	171
4.4.2.6 Futurama	172
4.4.2.7 Matrix	173
4.4.2.8 Bankomatdi	174
4.4.2.9 Bbc3di	175
4.4.2.9 Susidi	176
4.4.3 Video HDTV	177

4.4.3.1 Troy	177
4.4.3.2 Matrix	178
4.4.3.3 Concert	179
5. ΑΛΓΟΡΙΘΜΟΙ ΚΩΔΙΚΟΠΟΙΗΣΗΣ ΉΧΟΥ	180
5.1 ΔΙΚΤΥΑ VOIP	180
5.2 ΠΑΡΑΓΟΝΤΕΣ ΠΟΥ ΕΠΗΡΕΑΖΟΥΝ ΤΗΝ ΠΟΙΟΤΗΤΑ ΤΗΣ ΦΩΝΗΣ ΣΕ VOIP ΔΙΚΤΥΑ	180
5.2.1 Εύρος ζώνης	180
5.2.2 Εξοπλισμός	181
5.2.3 ΑΤΑ / Δρομολογητής	181
5.2.4 Συχνότητα του τηλεφώνου.....	181
5.2.5 Καιρικές συνθήκες	181
5.2.6 Το μέρος όπου βρίσκεται ο εξοπλισμός.....	181
5.2.7 Το codec που χρησιμοποιείται	181
5.3 ΜΕΘΟΔΟΙ ΥΠΟΛΟΓΙΣΜΟΥ ΤΗΣ ΠΟΙΟΤΗΤΑΣ	181
5.3.1 Αντικειμενικές μέθοδοι υπολογισμού της ποιότητας.....	182
5.3.2 Υποκειμενικές μέθοδοι υπολογισμού της ποιότητας	182
5.3.3 Διαμορφωμένη μονάδα θορύβου αναφοράς (MNRU).....	183
5.3.4 Έλεγχοι κατανόησης.....	184
5.4 ΤΟ ΠΡΟΤΥΠΟ G.711	184
5.5 ΤΟ ΠΡΟΤΥΠΟ G.722	185
5.6 ΤΟ ΠΡΟΤΥΠΟ G.723	185
5.7 ΤΟ ΠΡΟΤΥΠΟ G.723.1	185
5.8 ΤΟ ΠΡΟΤΥΠΟ G.726	186
5.9 ΤΟ ΠΡΟΤΥΠΟ G.728	186
5.10 ΤΟ ΠΡΟΤΥΠΟ G.729	186
5.11 ΤΟ ΠΡΟΤΥΠΟ G.729A.....	186
6. ΣΥΓΚΡΙΣΕΙΣ ΑΛΓΟΡΙΘΜΩΝ ΚΩΔΙΚΟΠΟΙΗΣΗΣ ΉΧΟΥ	188
6.1 ΣΥΓΚΡΙΣΗ ΤΩΝ CODECS G.711, G.723.1, G.726, G.729 ΚΑΙ G.729A.....	188
6.1.1 Πειραματική διαδικασία	188
6.1.2 Ρυθμίσεις των codecs.....	189
6.1.3 Αποτελέσματα και συμπεράσματα	190
6.2 ΣΥΓΚΡΙΣΗ ΤΩΝ CODECS G.711, G.721, G.722, G.726, G.729A, MPEG3 ΚΑΙ GSM.....	194
6.2.1 Πειραματική διαδικασία	194
6.2.2 Αποτελέσματα και συμπεράσματα	195
6.3 ΣΥΓΚΡΙΣΗ ΤΟΥ G.711 ΚΑΙ G.728	196
6.3.1 Πειραματική διαδικασία	196
6.3.2 Αποτελέσματα και συμπεράσματα	197
6.3.3 Αναλυτικά συμπεράσματα.....	199
6.4 ΣΥΓΚΡΙΣΗ ΤΩΝ ΠΡΟΤΥΠΩΝ G.723.1, G.728, G.729 ΚΑΙ GSM.....	201
6.4.1 Υποκειμενική αξιολόγηση	201
6.4.2 Σύγκριση της ευαισθησίας στα λάθη.....	203
6.4.3 Αντικειμενική αξιολόγηση.....	204
6.5 ΣΥΓΚΡΙΣΗ ΤΩΝ G.711, G.729 ΚΑΙ AMR	210
6.5.1 Πειραματική διαδικασία	210
6.5.2 Αποτελέσματα και συμπεράσματα	211
7. ΚΥΡΙΟΤΕΡΑ CODECS ΓΙΑ VOIP	215
7.1 ΕΥΡΟΣ ΖΩΝΗΣ / ΟΝΟΜΑΣΤΙΚΟ ΕΥΡΟΣ ΖΩΝΗΣ	218

7.2 ΡΥΘΜΟΣ ΔΕΙΓΜΑΤΟΛΗΨΙΑΣ	218
7.3 ΜΕΓΕΘΟΣ ΩΦΕΛΙΜΟΥ ΦΟΡΤΙΟΥ	218
A. ΠΑΡΑΡΤΗΜΑ.....	219
A1. ΜΕΤΑΔΟΣΗ ΣΕ ΠΡΑΓΜΑΤΙΚΟ ΧΡΟΝΟ	219
A2. ΠΑΛΜΟΚΩΔΙΚΗ ΔΙΑΜΟΡΦΩΣΗ (PCM).....	219
A3. ΔΙΑΦΟΡΙΚΗ ΠΑΛΜΟΚΩΔΙΚΗ ΔΙΑΜΟΡΦΩΣΗ (DPCM).....	219
A4. ΠΡΟΣΑΡΜΟΣΜΕΝΗ ΔΙΑΦΟΡΙΚΗ ΠΑΛΜΟΚΩΔΙΚΗ ΔΙΑΜΟΡΦΩΣΗ (ADPCM)	220
A5. ΥΠΗΡΕΣΙΕΣ ΚΑΛΥΤΕΡΗΣ ΠΡΟΣΠΑΘΕΙΑΣ	220
A6. ΔΙΑΦΟΡΟΠΟΙΗΜΕΝΕΣ ΥΠΗΡΕΣΙΕΣ	220
A7. ΕΠΙΤΑΧΥΜΕΝΗ ΠΡΟΩΘΗΣΗ ΡΗΒ	220
A8. ΠΡΟΟΠΤΙΚΗ ΔΙΟΡΘΩΣΗ ΛΑΘΩΝ	221
A9. CODECS ΚΥΜΑΤΟΜΟΡΦΗΣ	221
A10. CODECS ΠΗΓΗΣ ΚΑΙ VOCODERS	221
A11. ΥΒΡΙΔΙΚΑ CODECS ΗΧΟΥ	221
A12. ΥΠΟΚΕΙΜΕΝΙΚΗ ΜΕΘΟΔΟΣ MOS.....	222
A13. ΠΕΠΛΕΓΜΕΝΟ ΚΑΙ ΜΗ-ΠΕΠΛΕΓΜΕΝΟ VIDEO (INTERLACED/NON-INTERLACED VIDEO)	222
ΕΥΡΕΤΗΡΙΟ ΕΙΚΟΝΩΝ	224
ΕΥΡΕΤΗΡΙΟ ΠΙΝΑΚΩΝ	230
ΒΙΒΛΙΟΓΡΑΦΙΑ.....	233

Πρόλογος

Σκοπός της παρούσας διπλωματικής εργασίας είναι η εξέταση της συμπεριφοράς των δημοφιλέστερων αλγορίθμων συμπίεσης ήχου και video υπό ορισμένες συνθήκες δικτύου και μεταφοράς δεδομένων, έτσι να καταλήξουμε (όπου αυτό είναι δυνατό) σε μία κατάταξη βάση της απόδοσής τους. Κατά την εκπόνηση της διπλωματικής συλλέχθηκαν πειραματικά αποτελέσματα από διάφορες πηγές αναλύθηκαν και εξήχθησαν ενδιαφέροντα συμπεράσματα.

Στο πρώτο κεφάλαιο γίνεται μία αναφορά στους όρους «ηλεκτρονική υγεία» και «τηλεϊατρική» και εξηγούνται οι απαιτήσεις που έχει ένα δίκτυο τηλεϊατρικής. Επίσης παρουσιάζονται και αναλύονται οι διαδικασίες μίας από τις βασικότερες διαμορφώσεις σήματος που χρησιμοποιούνται στην κωδικοποίηση σημάτων, της παλμοκωδικής διαμόρφωση ή όπως είναι πιο γνωστή διαμόρφωση PCM.

Στο δεύτερο κεφάλαιο παρουσιάζονται τα σημαντικότερα χαρακτηριστικά των δημοφιλέστερων προτύπων συμπίεσης video τα οποία ελέχθησαν. Συγκεκριμένα παρουσιάζονται όλα τα νέα χαρακτηριστικά και βελτιώσεις που εισήγαγε το κάθε πρότυπο καθώς επίσης δίνονται και πίνακες με τα προφίλ (ένα σετ χαρακτηριστικών) των προτύπων.

Στο τρίτο κεφάλαιο παρουσιάζονται οι συγκρίσεις των διαφόρων αλγορίθμων συμπίεσης video. Γίνεται αναφορά στην πειραματική διαδικασία κάθε σύγκρισης και στο τέλος παρουσιάζονται τα αποτελέσματα που προέκυψαν. Σε ορισμένα πειράματα προτείνονται και νέοι μέθοδοι συνδυασμού αλγορίθμων ή χαρακτηριστικών αυτών για την βελτίωση της απόδοσής τους.

Στο τέταρτο κεφάλαιο παρουσιάζονται τρία πειράματα που πραγματοποίησε το τμήμα Graphics & Media Lab Video Group του Computer Science Department of Moscow State University, στα οποία συγκρίνεται το πρότυπο H.264 διαφόρων κατασκευαστών. Σαν codec αναφοράς σε όλα τα πειράματα χρησιμοποιείται το πρότυπο MPEG-4. Τα πειράματα πραγματοποιήθηκαν το διάστημα από τον Ιανουάριο 2005 μέχρι τον Δεκέμβριο 2007.

Στο πέμπτο κεφάλαιο παρουσιάζεται η τεχνολογία των δικτύων VoIP και οι παράμετροι οι οποίοι τα επηρεάζουν. Επίσης παρουσιάζονται ορισμένες υποκειμενικές και αντικειμενικές μέθοδοι για την αξιολόγηση των αλγορίθμων συμπίεσης ήχου. Τέλος όπως και στα πρότυπα συμπίεσης video έτσι και εδώ παρουσιάζονται τα σημαντικότερα χαρακτηριστικά των δημοφιλέστερων προτύπων συμπίεσης ήχου.

Στο έκτο κεφάλαιο παρουσιάζονται αναλυτικά οι πειραματικές διαδικασίες και τα αποτελέσματα των συγκρίσεων των αλγορίθμων συμπίεσης ήχου. Για κάθε σύγκριση γίνεται μία αναφορά στη θεωρία που χρησιμοποιείται και εν συνεχεία παρουσιάζονται πίνακες και γραφικές παραστάσεις με τα αποτελέσματα.

Στο κεφάλαιο επτά παρουσιάζεται ένας συνοπτικός πίνακας με τα κυριότερα χαρακτηριστικά των δημοφιλέστερων προτύπων συμπίεσης ήχου. Συγκεκριμένα δίνονται πληροφορίες σχετικά με τη συχνότητα δειγματοληψίας, το εύρος ζώνης, το ονομαστικό εύρος ζώνης, το ωφέλιμο φορτίο, την άδεια χρήσης καθώς επίσης αναφέρονται και κάποια σχόλια (υπέρ και κατά) για κάθε πρότυπο.

Τέλος ακολουθεί το Παράρτημα Α στο οποίο αναλύονται κάποιοι βασικοί όροι, οι οποίοι χρησιμοποιούνται μέσα στην εν λόγω διπλωματική, για λόγους διευκόλυνσης του αναγνώστη.

1. Εισαγωγικά – Θεωρία

1.1 Ηλεκτρονική υγεία

Τα τελευταία χρόνια χρησιμοποιείται όλο και περισσότερο ο όρος ηλεκτρονική υγεία (e-health) αν και δεν υπάρχει ένας σαφής ορισμός για τον όρο αυτό. Σύμφωνα με το [1] ηλεκτρονική υγεία είναι ένας τομέας της ιατρικής πληροφορικής και των τηλεματικών εφαρμογών της, της δημόσιας υγείας και της βιομηχανίας που αναφέρεται σε υπηρεσίες υγείας και πληροφορικής, οι οποίες προσφέρονται ή ενισχύονται μέσω του διαδικτύου και των σχετικών με αυτό τεχνολογιών.

Όταν μιλάμε για ηλεκτρονική υγεία δεν αναφερόμαστε μόνο στην τεχνολογική ανάπτυξη αλλά επεκτείνεται ακόμα παρά πέρα σε ένα τρόπο σκέψης και οργάνωσης με σκοπό την βελτίωση της περίθαλψης με χρήση των τεχνολογιών πληροφορικής και τηλεπικοινωνιών.

Η τηλεϊατρική αποτελεί ένα κομμάτι της ηλεκτρονικής υγείας. Ένας ορισμός της τηλεϊατρικής δίνεται από την Παγκόσμια Ομοσπονδία Υγείας και Τηλεϊατρικής: «Τηλεϊατρική είναι η παροχή ιατρικής περίθαλψης – σε περιπτώσεις που η απόσταση είναι κρίσιμος παράγοντας – από όλους τους επαγγελματίες του χώρου της υγείας, χρησιμοποιώντας τεχνολογίες πληροφοριών και επικοινωνιών για την ανταλλαγή έγκυρης πληροφορίας για τη διάγνωση, αγωγή και πρόληψη ασθενειών, της έρευνα και εκτίμηση, όπως και τη συνεχή εκπαίδευση των επαγγελματιών υγείας των ατόμων και των κοινοτήτων τους.». Με μία ευρύτερη έννοια μπορούμε να πούμε ότι η τηλεϊατρική αναφέρεται στην εφαρμογή σύγχρονων τεχνολογιών τηλεπικοινωνίας και έχει ως σκοπό την παροχή ιατρικής φροντίδας σε ασθενείς απομακρυσμένων περιοχών, την τηλεμετρία και τη διακίνηση της ιατρικής γνώσης μεταξύ των ιατρικών λειτουργιών, μέσα από τη μετάδοση εικόνας και ήχου. Το μεγάλο πλεονέκτημα της τηλεϊατρικής είναι η ανεξαρτησία από τους περιορισμούς του χρόνου και του χώρου. Όλα αυτά έχουν ως σκοπό την ποιοτικότερη και αποτελεσματικότερη φροντίδα των ασθενών από τη στιγμή της διάγνωσης μέχρι και τη στιγμή της θεραπείας.

Κάθε ιατρικό όργανο το οποίο μπορεί να «δώσει» στην έξοδό του την πληροφορία που συλλέγει, μπορεί να χρησιμοποιηθεί ως ένα συστατικό στοιχείο ενός συστήματος τηλεϊατρικής.

Οι υπηρεσίες της τηλεϊατρικής μπορούν να χωριστούν σε δύο είδη:

1. Τις υπηρεσίες πραγματικού χρόνου ή διαδραστικές, όπως η video-διάσκεψη, οι οποίες αναφέρονται στην ταυτόχρονη λειτουργία του κεντρικού και των περιφερειακών κόμβων του συστήματος τηλεϊατρικής
2. Τις εφαρμογές αποθήκευσης και προώθησης (store-and-forward) οι οποίες αποτελούν μια χρονικά μετατοπισμένη επικοινωνία ανάμεσα στους εμπλεκόμενους κόμβους του συστήματος. Έτσι αξιοποιείται καλύτερα το ιατρικό ιστορικό του ασθενούς και συνδυάζοντας δεδομένα εικόνας, βίντεο και ήχου παρέχεται η δυνατότητα αξιολόγησης της κατάστασης του ασθενούς σε μη πραγματικό χρόνο.

Οι εφαρμογές πραγματικού χρόνου έχουν μεγάλες απαιτήσεις όσο αναφορά την ταχύτητα και την αξιοπιστία του δικτύου και σε συνδυασμό με τον μεγάλο όγκο δεδομένων, καθιστούν την εκτέλεση του πιο δύσκολη σε σχέση με τις εφαρμογές αποθήκευσης και προώθησης.

Η τεχνική υποδομή που απαιτείται για τις εφαρμογές τηλεϊατρικής περιλαμβάνει τοπικούς σταθμούς εξοπλισμένους με οπτικοακουστικά μέσα και ένα δίκτυο το οποίο είναι σε θέση να παρέχει το απαραίτητο εύρος ζώνης που χρειάζεται η εφαρμογή. Οι τοπικοί σταθμοί πρέπει να είναι εξοπλισμένοι με τον κατάλληλο ιατρικό εξοπλισμό, σταθμούς εργασίας και εργαλεία διαχείρισης του δικτύου και των πόρων του συστήματος. Καθώς ορισμένες εφαρμογές τηλεϊατρικής απαιτούν την υποστήριξη προηγμένων και σύνθετων αλγορίθμων επεξεργασίας και συμπίεσης εικόνας και γραφικών (δυσδιάστατων και τρισδιάστατων) οι σταθμοί εργασίας πρέπει να εκμεταλλεύονται κατά τον καλύτερο δυνατό τρόπο τις δυνατότητες των σύγχρονων υπολογιστικών συστημάτων και του δικτύου.

Η τηλεπικοινωνιακή υποδομή περιλαμβάνει τα μέσα μεταφοράς της πληροφορίας, τα πρωτόκολλα επικοινωνίας και τους μηχανισμούς διαχείρισης του δικτύου. Η ανάπτυξη ενός δικτύου αποκλειστικά για εφαρμογές τηλεϊατρικής εξαρτάται από πολλούς παράγοντες και είναι ένα αρκετά δύσκολη. Μερικοί από τους σημαντικότερους παράγοντες που πρέπει να λαμβάνονται υπόψη στο σχεδιασμό ενός δικτύου τηλεϊατρικής είναι το απαιτούμενο εύρος ζώνης (bandwidth), ο ελάχιστος ρυθμός μετάδοσης και η ασφάλεια. Επίσης πολύ σημαντική είναι η ακεραιότητα της πληροφορίας κατά τη μετάδοσή της, η πιστοποίηση της αυθεντικότητάς της και η διαφύλαξή της από μη εξουσιοδοτημένους χρήστες. Για τη διόρθωση των λαθών κατά τη μετάδοση της πληροφορίας έχουν αναπτυχθεί κατάλληλοι αλγόριθμοι εντοπισμού και διόρθωσης. Τέλος για την ασφάλεια και την ακεραιότητα της ιατρικής πληροφορίας χρησιμοποιούνται προτυποποιημένες τεχνικές κρυπτογράφησης και αυθεντικότητας.

1.2 Διαδικασία κωδικοποίησης – Παλμοκωδική διαμόρφωση (PCM)

Πριν γίνει μετάδοση του αναλογικού σήματος είναι αναγκαία η κωδικοποίησή του και μετατροπή του σε ψηφιακό. Με την αναπαράσταση του αναλογικού σήματος σε ψηφιακό πετυχαίνουμε αντοχή (ruggedness) στο θόρυβο και στην παρεμβολή, αποτελεσματική αναγέννηση του κωδικοποιημένου σήματος και δυνατότητα ομοιόμορφου σχήματος (uniform format) μετάδοσης για διαφορετικά είδη βασικής ζώνης. Τα πλεονεκτήματα αυτά επιτυγχάνονται με το κόστος της αύξησης του απαιτούμενου εύρους ζώνης μετάδοσης και την αύξηση της πολυπλοκότητας του συστήματος.

Στην παλμοκωδική διαμόρφωση (pulse-code modulation, PCM) υπάρχει ένα σύνολο επιτρεπτών τιμών στις οποίες στρογγυλοποιούνται τα δείγματα του σήματος πληροφορίας. Έτσι το σήμα που προκύπτει είναι σε διακριτή μορφή και επιτρέπεται η μετάδοση πληροφορίας μέσω κωδικοποιημένων ηλεκτρικών σημάτων, ξεχωρίζοντας κατ' αυτόν τον τρόπο την PCM από τις άλλες μεθόδους διαμόρφωσης (Διαμόρφωση πλάτους (PAM), Διαμόρφωση διάρκειας παλμών (PDM) και Διαμόρφωση θέσης παλμών (PPM)).

Οι βασικές λειτουργίες ενός πομπού συστήματος PCM είναι η δειγματοληψία (sampling), η κβαντοποίηση (quantizing) και η κωδικοποίηση, ενώ οι βασικές λειτουργίες του δέκτη είναι η αναγέννηση (regeneration) των εξασθετισμένων σημάτων, η αποκωδικοποίηση (decoding) και η αποδιαμόρφωση (demodulation) της ακολουθίας των κβαντισμένων δειγμάτων. Αν χρησιμοποιείται πολυπλεξία με διαίρεση χρόνου, είναι απαραίτητο να συγχρονίσουμε το δέκτη με τον πομπό για την ικανοποιητική λειτουργία του συστήματος.

1.2.1 Ανάλυση των λειτουργιών κωδικοποίησης

1.2.1.1 Δειγματοληψία

Κατά τη διαδικασία της δειγματοληψίας λαμβάνονται δείγματα της εισερχόμενης πληροφορίας (κυματομορφή) με μία ακολουθία στενών ορθογώνιων παλμών, έτσι ώστε να

προσεγγίζουν πολύ καλά τη διαδικασία στιγμιαίας δειγματοληψίας. Για να εξασφαλιστεί η τέλεια ανακατασκευή της πληροφορίας στο δέκτη, ο ρυθμός δειγματοληψίας πρέπει να είναι μεγαλύτερος από το διπλάσιο της υψηλότερης συχνότητας της κυματομορφής της αρχικής πληροφορίας W (θεώρημα δειγματοληψίας, Shannon). Στην πράξη απορρίπτονται συχνότητες μεγαλύτερες από W πριν από τη δειγματοληψία με χρήση ενός βαθυπερατού φίλτρου στην είσοδο του δειγματολήπτη. Έτσι με τη διαδικασία της δειγματοληψίας πετυχαίνουμε τον περιορισμό της συνεχώς μεταβαλλόμενης κυματομορφής της αρχικής πληροφορίας σε ένα πεπερασμένο αριθμό διακριτών τιμών.

1.2.1.2 Κβαντοποίηση

Κβαντοποίηση (quantizing) ονομάζεται η μετατροπή ενός αναλογικού (συνεχούς) σήματος σε μία ψηφιακή (διακριτή) μορφή.

Ένα συνεχές σήμα (όπως η φωνή) έχει συνεχές πεδίο τιμών πλάτους και συνεπώς τα δείγματά του έχουν συνεχές πεδίο τιμών πλάτους. Δηλαδή μέσα στο πεπερασμένο πεδίο τιμών του σήματος βρίσκουμε έναν άπειρο αριθμό σταθμών πλάτους. Στην πραγματικότητα δεν είναι απαραίτητο να μεταδίδουμε τα ακριβή πλάτη των δειγμάτων. Οποιαδήποτε ανθρώπινη αίσθηση (το αυτί ή το μάτι) σαν τελικός δέκτης μπορεί να ανιχνεύσει πεπερασμένες διαφορές έντασης. Αυτό σημαίνει ότι το αρχικό συνεχές σήμα μπορεί να προσεγγιστεί από ένα σήμα το οποίο κατασκευάζεται από διακριτά πλάτη, επιλεγμένα από ένα διαθέσιμο σύνολο με βάση την ελαχιστοποίηση του σφάλματος. Η ύπαρξη ενός πεπερασμένου αριθμού διακριτών σταθμών πλάτους είναι μία βασική συνθήκη της PCM. Εάν καθορίσουμε διακριτές στάθμες πλάτους με αρκετά μικρό βήμα μεταξύ τους, μπορούμε να κάνουμε το τελικό σήμα να μην ξεχωρίζει πρακτικά από το αρχικό συνεχές σήμα.

Η διαφορά μεταξύ των δύο γειτονικών διακριτών τιμών ονομάζεται κβάντο (quantum) ή μέγεθος βήματος (step size). Το σφάλμα κβαντισμού (quantizing error) αποτελείται από τη διαφορά μεταξύ των σημάτων εισόδου και εξόδου του κβαντιστή. Είναι φανερό ότι η μέγιστη στιγμιαία τιμή αυτού του σφάλματος είναι το μισό ενός κβάντου και το συνολικό εύρος της μεταβολής είναι από -μισό έως +μισό βήμα.

1.2.1.3 Κωδικοποίηση

Για να εκμεταλλευτούμε τα πλεονεκτήματα της δειγματοληψίας και της κβαντοποίησης απαιτείται η χρησιμοποίηση μιας διαδικασίας κωδικοποίησης (encoding process) για τη μετατροπή του διακριτού συνόλου των τιμών των δειγμάτων σε μία κατάλληλη μορφή για μετάδοση μέσω ενός διαύλου. Κάθε σχέδιο για αναπαράσταση κάθε διακριτού συνόλου τιμών σαν μια ιδιαίτερη διάταξη διακριτών γεγονότων ονομάζεται κώδικας (code). Ένα από τα διακριτά γεγονότα σε ένα κώδικα ονομάζεται στοιχείο του κώδικα (code element) ή σύμβολο (symbol). Για παράδειγμα, η παρουσία ή η απουσία ενός παλμού είναι ένα σύμβολο. Μια ιδιαίτερη διάταξη συμβόλων, που χρησιμοποιείται σε ένα κώδικα για την παράσταση μιας μόνο τιμής του διακριτού συνόλου ονομάζεται κωδική λέξη (codeword) ή χαρακτήρας (character).

Σε ένα δυαδικό κώδικα (binary code) κάθε σύμβολο μπορεί να πάρει μια από δύο διακριτές τιμές, οι οποίες μπορούν να αντιστοιχούν στην παρουσία ή απουσία ενός παλμού. Τα δύο σύμβολα ενός δυαδικού κώδικα συνήθως συμβολίζονται με 0 και 1. Σε ένα τριαδικό κώδικα (ternary code) αντίστοιχα, κάθε σύμβολο μπορεί να πάρει μία από τρεις διακριτές τιμές και ούτω καθ' εξής. Το μέγιστο πλεονέκτημα του δυαδικού κώδικα είναι ότι επειδή ένα δυαδικό σύμβολο αντέχει μια σχετικά υψηλή στάθμη θορύβου, είναι εύκολο να αναγεννηθεί παρ' όλη την παρουσία θορύβου.

Αν σε ένα δυαδικό κώδικα κάθε λέξη αποτελείται από n bits, μπορούμε να παραστήσουμε ένα σύνολο 2^n διακριτών λέξεων. Για παράδειγμα ένα δείγμα κβαντισμένο σε 128 στάθμες μπορεί να παρασταθεί με μία λέξη των 7 bits. Υπάρχουν πάρα πολύ τρόποι για τη δημιουργία μιας αντιστοίχησης ένα προς ένα μεταξύ κβαντισμένων σταθμών και κωδικών λέξεων. Ένας βολικός τρόπος είναι να εκφράσουμε τον αριθμό τάξης της κβαντισμένης στάθμης σαν ένα δυαδικό αριθμό.

1.2.1.4 Αναγέννηση

Ένα σημαντικό χαρακτηριστικό των συστημάτων PCM είναι η ικανότητα να ελέγχουν τις επιδράσεις της παραμόρφωσης και του θορύβου που παράγονται κατά τη μετάδοση μιας κυματομορφής PCM. Η δυνατότητα επιτυγχάνεται με την ανακατασκευή της κυματομορφής PCM μέσω μιας αλυσίδας αναγεννητικών επαναληπτών (regenerative repeaters) που είναι τοποθετημένοι σε αρκετά μικρή απόσταση κατά μήκος της διαδρομής μετάδοσης. Οι τρεις βασικές λειτουργίες που εκτελούνται από έναν αναγεννητικό επαναλήπτη είναι η ισοστάθμιση (equalization), ο χρονισμός (timing) και η λήψη απόφασης (decision making).

Στην ισοστάθμιση μορφοποιούνται οι λαμβανόμενοι παλμοί, έτσι ώστε να αντισταθμιστούν οι επιδράσεις τις παραμόρφωσης πλάτους και της φάσης που παράγονται κατά τη μετάδοση. Κατά το χρονισμό παρέχεται μια περιοδική ακολουθία παλμών που παράγεται από τους λαμβανόμενους παλμούς, για τη δειγματοληψία των ισοσταθμισμένων παλμών σε χρονικές στιγμές όπου ο λόγος σήματος προς θόρυβο είναι μέγιστος. Τέλος η λήψη απόφασης ενεργοποιείται όταν στο χρόνο δειγματοληψίας που καθορίζεται από το κύκλωμα χρονισμού, το πλάτος του ισοσταθμισμένου παλμού με το θόρυβο ξεπερνά μια προκαθορισμένη στάθμη τάσης.

Για παράδειγμα σε ένα σύστημα PCM με σηματοδοσία on-off ο επαναλήπτης παίρνει μια απόφαση σε κάθε διάρκεια bit κατά πόσο υπάρχει ή όχι παλμός. Εάν η απόφαση είναι «ναι» μεταδίδεται ένας καινούργιος καθαρός παλμός προς τον επόμενο επαναλήπτη. Αν από την άλλη πλευρά η απόφαση είναι «όχι» μεταδίδεται μια καθαρά μηδενική στάθμη. Κατ' αυτόν τον τρόπο, με την προϋπόθεση ότι η παρενόχληση δεν είναι τόσο μεγάλη ώστε να προκαλέσει σφάλμα στη διαδικασία λήψης απόφασης, απαλείφεται πλήρως η συσσώρευση παραμόρφωσης και θορύβου σε ένα διάστημα μεταξύ επαναληπτών.

Στην ιδανική περίπτωση εκτός από την καθυστέρηση, το αναδημιουργούμενο σήμα είναι ακριβώς το ίδιο με το αρχικό. Στη πράξη όμως το αναδημιουργούμενο σήμα διαφέρει από το αρχικό σήμα για δύο κυρίως λόγους:

1. την παρουσία θορύβου κατά τη μετάδοση και των λανθασμένων αποφάσεων του επαναλήπτη, εισάγοντας έτσι λανθασμένα bits (bit error) στο αναδημιουργούμενο σήμα
2. εάν η απόσταση μεταξύ των λαμβανόμενων παλμών αποκλίνει από την αρχική τιμή της, εισάγεται ένα τρέμουλο (jitter) στη θέση του αναδημιουργούμενου παλμού προκαλώντας κατ' αυτόν τον τρόπο παραμόρφωση

1.2.1.5 Αποκωδικοποίηση

Όταν το σήμα φτάσει στο δέκτη η πρώτη λειτουργία είναι η αναγέννηση του λαμβανόμενου σήματος. Οι καθαροί παλμοί που παράγονται ομαδοποιούνται ξανά σε κωδικές λέξεις και αποκωδικοποιούνται (δηλαδή αντιστοιχούνται) σε ένα κβαντισμένο σήμα PAM. Η διαδικασία αποκωδικοποίησης (decoding) περιλαμβάνει τη δημιουργία ενός παλμού, το πλάτος του οποίου είναι το γραμμικό άθροισμα όλων των παλμών στη κωδική λέξη, με τον κάθε παλμό να έχει βάρος την τιμή της θέσης του ($2^0, 2^1, 2^2, 2^3, \dots$) στον κώδικα.

1.2.1.6. Φιλτράρισμα

Το φιλτράρισμα είναι η τελική λειτουργία στον δέκτη. Η έξοδος του αποκωδικοποιητή περνάει από ένα βαθυπερατό φίλτρο, του οποίου η συχνότητα αποκοπής είναι ίση με το εύρος ζώνης πληροφορίας W . Υποθέτοντας ότι δεν υπάρχουν σφάλματα κατά τη μετάδοση, το σήμα που προκύπτει δεν περιλαμβάνει θόρυβο, με εξαίρεση την αρχική παραμόρφωση που εισάγεται από τη διαδικασία κβαντοποίησης.

1.2.1.7 Πολυπλεξία

Εφαρμογές που χρησιμοποιούν PCM, είναι φυσικό να πολυπλέκουν διάφορες πηγές με διαίρεση χρόνου, ενώ κάθε πηγή κρατάει την ανεξαρτησία της καθ' όλη τη διαδρομή από τον πομπό μέχρι τον δέκτη. Όσο ο αριθμός των ανεξάρτητων πηγών πληροφορίας αυξάνει το διάστημα χρόνου που διατίθεται σε κάθε μια πηγή μειώνεται καθώς πρέπει να εξυπηρετηθούν όλες οι πηγές μέσα σε ένα χρονικό διάστημα ίσο με το αντίστροφο του ρυθμού δειγματοληψίας. Αυτό σημαίνει ότι μειώνεται η επιτρεπόμενη διάρκεια της κωδικής λέξης που παριστάνει ένα απλό δείγμα.

Όσο η διάρκεια μειώνεται η παραγωγή και η μετάδοση των παλμών τείνει να γίνει δυσκολότερη. Επιπλέον, εάν οι παλμοί γίνουν πολύ μικροί, τα σφάλματα στο μέσο μετάδοσης αρχίζουν να παρεμβάλουν στην ορθή λειτουργία του συστήματος. Συνεπώς, είναι απαραίτητος ο περιορισμός των ανεξάρτητων πηγών πληροφορίας που πολυπλέκονται με διαίρεση χρόνου.

1.2.1.8 Συγχρονισμός

Για την ικανοποιητική λειτουργία ενός συστήματος PCM με πολυπλεξία διαίρεσης χρόνου είναι απαραίτητο οι λειτουργίες χρονισμού στον δέκτη, εκτός από τον χρόνο που χάνεται στη μετάδοση και στην αναγέννηση, να ακολουθούν τις αντίστοιχες λειτουργίες στον πομπό. Αυτό ισοδυναμεί με ότι το τοπικό ρολόι στον δέκτη πρέπει να κρατάει τον ίδιο χρόνο με ένα μακρινό σταθερό ρολόι στον πομπό, εκτός του ότι το τοπικό ρολόι καθυστερεί κατά μία ποσότητα που αντιστοιχεί στο χρόνο που απαιτείται για τη μετάδοση του σήματος πληροφορίας από τον πομπό στο δέκτη.

Μια διαδικασία για τον συγχρονισμό των ρολογιών πομπού και δέκτη είναι να προσαρμοσθεί ένα στοιχείο του κώδικα ή παλμός στο τέλος ενός πλαισίου (frame) (το οποίο είναι η διαδοχή κωδικών λέξεων που προέρχονται από κάθε μία από τις ανεξάρτητες πηγές πληροφορίας) και να μεταδίδει αυτόν τον παλμό σε κάθε πλαίσιο. Ο δέκτης περιλαμβάνει ένα κύκλωμα το οποίο ψάχνει για το σχηματισμό εναλλασσόμενων 1 και 0 με ρυθμό το μισό του ρυθμού πλαισίων και έτσι επιτυγχάνει συγχρονισμό μεταξύ του πομπού και του δέκτη.

Όταν η μετάδοση διακόπτεται είναι αρκετά απίθανο ότι τα ρολόγια πομπού και δέκτη θα συνεχίσουν να δείχνουν τον ίδιο χρόνο για αρκετό χρονικό διάστημα. Συνεπώς, κατά την εκτέλεση μιας διαδικασίας συγχρονισμού πρέπει να ορίσουμε μια βοηθητική διαδικασία για την ανίχνευση του παλμού συγχρονισμού. Η διαδικασία συνίσταται στην παρατήρηση στοιχείων του κώδικα ένα προς ένα, μέχρι να ανιχνευτεί ο παλμός συγχρονισμού. Δηλαδή, μετά την παρατήρηση κάποιου συγκεκριμένου στοιχείου του κώδικα για αρκετό χρόνο ώστε να βεβαιωθούμε ότι δεν είναι παλμός συγχρονισμού, το ρολόι του δέκτη οπισθοχωρεί κατά ένα στοιχείο κώδικα και παρατηρείται το επόμενο στοιχείο κώδικα. Αυτή η διαδικασία αναζήτησης επαναλαμβάνεται μέχρις ότου ανιχνευτεί ο παλμός συγχρονισμού. Προφανώς, ο χρόνος που απαιτείται για τον συγχρονισμό εξαρτάται από την στιγμή κατά την οποία επανακαθίσταται η σωστή μετάδοση.

2. Αλγόριθμοι κωδικοποίησης video

Σε αυτό το κεφάλαιο παρουσιάζονται ορισμένα βασικά χαρακτηριστικά των δημοφιλέστερων προτύπων κωδικοποίησης video. Κάθε πρότυπο έχει δικές του διαδικασίες κωδικοποίησης και συμπίεσης οι οποίες αποδίδουν διαφορετικά κάτω από τις ίδιες συνθήκες. Η κατάταξη των αλγορίθμων δεν είναι πάντα εφικτή καθώς σημαντικό ρόλο παίζουν οι συνθήκες του δικτύου, το υλικό που χρησιμοποιείται (hardware) και το υλικό που πρόκειται να κωδικοποιηθεί.

2.1 Το πρότυπο H.264/AVC

Το H.264/AVC είναι ένα πρότυπο για την κωδικοποίηση video γνωστό και ως MPEG-4 Part 10 ή MPEG-4 AVC (Advanced Video Technology). Πρόκειται για ένα πρότυπο κωδικοποίησης το οποίο στηρίζεται στην κίνηση και στα μπλοκ (block-oriented motion-compensation-based codec standard) και αναπτύχθηκε από την ομάδα ITU-T Video Coding Experts Group (VCEG) σε συνεργασία με την ISO/IEC Moving Picture Experts Group (MPEG). Η συνεργασία αυτή είναι γνωστότερη με την ονομασία Joint Video Team (JVT).

Ο σκοπός του H.264/AVC ήταν να δημιουργηθεί ένα πρότυπο το οποίο θα παρέχει καλύτερη ποιότητα σε χαμηλότερους ρυθμούς bit (bit rates) σε σχέση με τα προγενέστερα πρότυπα, χωρίς όμως την αύξηση της πολυπλοκότητας σχεδιασμού τόσο ώστε η εφαρμογή του να γίνει ασύμφορη. Ένας άλλος στόχος ήταν η δυνατότητα χρησιμοποίησης του προτύπου σε διαφόρων ειδών εφαρμογές, δικτύων και συστημάτων, ανεξαρτήτως ρυθμού bit και ποιότητας.

Το H.264/AVC έχει χαρακτηριστικά τα οποία του επιτρέπουν να συμπιέζει video πολύ πιο αποτελεσματικά σε σχέση με τα προηγούμενα από αυτό codecs και του παρέχουν μεγάλη ευελιξία εφαρμογής σε μια μεγάλη ποικιλία δικτύων και εφαρμογών. Μερικά χαρακτηριστικά του είναι:

- εσωτερική πρόβλεψη εικόνας (inter-picture prediction)
- χρησιμοποιούνται προηγούμενες κωδικοποιημένες εικόνες σαν αναφορές με έναν ευέλικτο τρόπο, επιτρέποντας μέχρι και 16 πλαίσια αναφοράς (ή 32 πεδία αναφοράς αν πρόκειται για συμπλεγμένες εικόνες (interlaced pictures))
- δύναται διαφορετικά μεγέθη μπλοκ (block-size) τα οποία μπορεί να είναι μεγάλα μέχρι 16x16 έως πολύ μικρά 4x4. Τα επιτρεπόμενα μεγέθη είναι: 16x16, 16x8, 8x16, 8x8, 8x4, 4x8 και 4x4 τα οποία μπορούν να συνδυαστούν μέσα σε ένα μακρο-μπλοκ (macroblock, MB)
- δίνεται η δυνατότητα χρήσης διανυσμάτων κίνησης (motion vectors) ανά μακρο-μπλοκ, τα οποία για 8x8 ή μεγαλύτερο τεμαχισμό μπορούν να «δείχνουν» σε διαφορετικές εικόνες αναφοράς
- υπάρχει η δυνατότητα χρήσης οποιουδήποτε τύπου μακρο-μπλοκ σε πλαίσια τύπου-B (B-frames), ακόμα και μακρο-μπλοκ τύπου-I (I-macroblocks), και έτσι έχουμε αποδοτικότερη συμπίεση όταν χρησιμοποιούμε πλαίσια τύπου-B
- δυνατότητα φιλτραρίσματος διαίρεσης των φωτεινών δειγμάτων μισού pixel (derivation of half-pel luma sample prediction), δημιουργώντας καλύτερη αναπαράσταση της κίνησης

- δυνατότητα ακρίβειας διαιρεμένου pixel (quarter-pixel precision) για την απόδοση της κίνησης, το οποίο μας επιτρέπει ακριβή αναπαράσταση των κινούμενων περιοχών
- πρόβλεψη με βαρύτητες (weighted prediction) η οποία επιτρέπει στον κωδικοποιητή την αλλαγή των παραμέτρων επεκτασιμότητας (scaling) και έναρξης (offset), η οποία οδηγεί σε σημαντική βελτίωση της απόδοσης υπό ορισμένες συνθήκες
- χωρική πρόβλεψη (spatial prediction) από τις άκρες των γειτονικών blocks.
- μη-απωλεστική κωδικοποίηση ενός μακρο-μπλοκ (lossless macroblock coding)
- μη-απωλεστική αναπαράσταση PCM (Pulse Code Modulation) του μακρο-μπλοκ, αφού τα δεδομένα απεικονίζονται απευθείας, επιτρέποντας πολύ καλή απεικόνιση συγκεκριμένων περιοχών και την εφαρμογή αυστηρών ορίων ως προς το πλήθος των δεδομένων που χρησιμοποιούνται για κάθε μακρο-μπλοκ
- ένα πιο ευέλικτο είδος κωδικοποίησης το οποίο δεν αυξάνει το απαιτούμενο εύρος ζώνης (flexible interlaced-scan video coding), όπου μπορεί να χρησιμοποιηθεί ένα ζευγάρι μακρο-μπλοκ για εικόνες κωδικοποιημένες ως πλαίσια (macroblock-adaptive frame-field, MBAFF coding) ή μπορεί να χρησιμοποιηθεί ένα μίγμα από κωδικοποιημένες εικόνες ως MBAFF πλαίσια με τις εικόνες κωδικοποιημένες σαν ξεχωριστά πεδία (fields) του πεπλεγμένου video (interlaced video)¹ (picture-adaptive frame-field, PAFF coding)
- νέοι μέθοδοι μετασχηματισμού όπως ο 4x4 ή 8x8 ακέραιος μετασχηματισμός ακριβείας (exact-match integer 4x4 or 8x8 spatial block transform), όπου μπορεί να υπάρξει και εναλλαγή μεταξύ των δύο αυτών μετασχηματισμών για καλύτερα αποτελέσματα. Υπάρχει και ο μετασχηματισμός Hadamard ο οποίος επιτρέπει ακόμα καλύτερη συμπίεση στις λείες (smooth) επιφάνειες
- νέος σχεδιασμός κβάντοποίησης όπου χρησιμοποιείται λογαριθμικός έλεγχος βήματος (logarithmic step size control) για ευκολότερο έλεγχο του ρυθμού μετάδοσης
- επαναληπτικό φίλτρο διαχωρισμού μπλοκ (loop deblocking filter) που βοηθάει στην εξάλειψη των αλλοιώσεων και έτσι έχουμε καλύτερη απόδοση της εικόνας.
- νέοι σχεδιασμοί κωδικοποίησης όπως η δυαδική κωδικοποίηση γενικευμένου πλαισίου (Context-adaptive binary arithmetic coding, CABAC) όπου τα στοιχεία κωδικοποιούνται βάσει πιθανοτήτων σε σχέση με τα άλλα στοιχεία σε ένα δοσμένο πλαίσιο και η μεταβλητού μήκους κωδικοποίηση γενικευμένου πλαισίου (Context-adaptive variable-length coding, CAVLC) όπου είναι χαμηλότερης πολυπλοκότητας από την CABAC αλλά και χαμηλότερης συμπίεσης. Ωστόσο η CAVLC αποδίδει καλύτερα από τις προηγούμενες τυπικές κωδικοποιήσεις
- ένα στρώμα δικτύου NAL (network abstraction layer, NAL) που επιτρέπει τη χρήση της ίδιας σύνταξης video σε διαφορετικά δίκτυα, παρέχοντας έτσι μεγαλύτερη ευελιξία σε σχέση με τα προηγούμενα codecs
- ευέλικτη οργάνωση των μακρο-μπλοκ (flexible macroblock ordering, FMO), τεχνικές οι οποίες επιτρέπουν την παρουσίαση των μακρο-μπλοκ μέσα στην εικόνα

¹ Βλ. Παράρτημα A13

- διαίρεση των δεδομένων (data partitioning, DP) όπου επιτρέπεται ο διαχωρισμός των λιγότερο σημαντικών δεδομένων σε ξεχωριστά πακέτα και έτσι βελτιώνεται ο ρυθμός διόρθωσης λαθών
- δυνατότητα του κωδικοποιητή να στέλνει επιπλέον αντιπροσώπευση μιας περιοχής της εικόνας (συνήθως χαμηλότερης πιστότητας) η οποία μπορεί να χρησιμοποιηθεί σε περίπτωση που η απεικόνιση είναι ελλιπής (redundant slices, RS)
- αρίθμηση πλαισίων (frame numbering) η οποία επιτρέπει τη δημιουργία «υπο-ακολουθιών» επιτρέποντας και χρονικό διαχωρισμό πολλών εικόνων και έτσι εκμεταλλευόμαστε καλύτερα το δίκτυο
- δυνατότητα του κωδικοποιητή να καθοδηγήσει τον αποκωδικοποιητή σε ένα επερχόμενο κομμάτι video (switching slices, SP and SI slices)
- δυνατότητα κωδικοποίησης ξεχωριστών χρωματικών επιπέδων σαν διαφορετικές εικόνες με διαφορετικές δομές, μακρο-μπλοκ, διανύσματα κίνησης κτλ. επιτρέποντας στους κωδικοποιητές να σχεδιάζονται με απλή παράλληλη αρχιτεκτονική
- αρίθμηση σειράς των εικόνων (picture order count) όπου δίνεται η δυνατότητα να κρατείται η σειρά των εικόνων και οι τιμές των δειγμάτων των κωδικοποιημένων εικόνων ανεξάρτητα από την χρονική πληροφορία, επιτρέποντας έτσι να μεταφέρεται/αλλάζεται/ελέγχεται ανεξάρτητα η χρονική πληροφορία χωρίς να επηρεάζεται το περιεχόμενο της κωδικοποιημένης εικόνας

2.1.1 Προφίλ του H.264

Το H.264/AVC περιλαμβάνει τα παρακάτω σύνολα χαρακτηριστικών (profiles) σχεδιασμένα για συγκεκριμένες εφαρμογές:

- Baseline profile (BP): το πρωταρχικό προφίλ για εφαρμογές χαμηλού ρυθμού bit με περιορισμένες υπολογιστικές δυνατότητες. Συνήθως το προφίλ αυτό χρησιμοποιείται σε εφαρμογές video-διάσκεψης (videoconferencing) και κινητές εφαρμογές (mobile application)
- Main profile (MP): το πιο συνηθισμένο προφίλ για εφαρμογές μετάδοσης και αποθήκευσης, το οποίο άρχισε να «χάνεται» όταν αναπτύχθηκε το High profile (HiP)
- Extended profile (XP): το πιο συνηθισμένο προφίλ για συνεχή ροή video (video streaming), αφού περιλαμβάνει χαρακτηριστικά που του επιτρέπουν υψηλή συμπίεση, διόρθωση λαθών κ.α.
- High Profile (HiP): το πρωταρχικό προφίλ για εφαρμογές μετάδοσης και αποθήκευσης, ειδικά σε εφαρμογές υψηλής ανάλυσης (high definition applications) (για παράδειγμα αυτό το προφίλ χρησιμοποιείται στο HD DVD και στο Blu-ray)
- High 10 Profile (Hi10P): πηγαίνοντας ένα βήμα παραπέρα το HiP σε αυτό το προφίλ έχουμε 10 bits παραπάνω ανά δείγμα για μεγαλύτερη ακρίβεια της αποκωδικοποιημένης εικόνας
- High 4:2:2 Profile (Hi422P): αναπτύχθηκε πάνω στο Hi10P προσθέτοντας υποστήριξη για 4:2:2 χρωματική δειγματοληψία (4:2:2 chroma sampling) για εφαρμογές που χρησιμοποιούν πεπλεγμένο video

- High 4:4:4 Predictive Profile (Hi444PP): αναπτύχθηκε πάνω στο Hi442P υποστηρίζοντας 4:4:4 χρωματική δειγματοληψία, 14 bits ανά δείγμα και επιπλέον αποδοτικότερη κωδικοποίηση των περιοχών της εικόνας και την κωδικοποίηση της εικόνας σαν τρία χρωματικά επίπεδα

	Baseline	Extended	Main	High	High 10	High 4:2:2	High 4:4:4 Predictive
I και P κομμάτια	Ναι	Ναι	Ναι	Ναι	Ναι	Ναι	Ναι
B κομμάτια	Όχι	Ναι	Ναι	Ναι	Ναι	Ναι	Ναι
SI και SP κομμάτια	Όχι	Ναι	Όχι	Όχι	Όχι	Όχι	Όχι
Πολλαπλά πλαίσια αναφοράς	Ναι	Ναι	Ναι	Ναι	Ναι	Ναι	Ναι
Φίλτρο διαχωρισμού μπλοκ, εσωτερικού βρόχου	Ναι	Ναι	Ναι	Ναι	Ναι	Ναι	Ναι
Κωδικοποίηση CAVLC	Ναι	Ναι	Ναι	Ναι	Ναι	Ναι	Ναι
Κωδικοποίηση CABAC	Όχι	Όχι	Ναι	Ναι	Ναι	Ναι	Ναι
Ευέλικτη σειρά μακρο-μπλοκ (FMO)	Ναι	Ναι	Όχι	Όχι	Όχι	Όχι	Όχι
Διατητική σειρά κομματιών (ASO)	Ναι	Ναι	Όχι	Όχι	Όχι	Όχι	Όχι
Πλεονάζουσα κομμάτια	Ναι	Ναι	Όχι	Όχι	Όχι	Όχι	Όχι
Διαμερισμός δεδομένων	Όχι	Ναι	Όχι	Όχι	Όχι	Όχι	Όχι
Interlaced κωδικοποίηση (PicAFF, MBAFF)	Όχι	Ναι	Ναι	Ναι	Ναι	Ναι	Ναι
4:2:0 χρωματικό format	Ναι	Ναι	Ναι	Ναι	Ναι	Ναι	Ναι
Μονοχρωματικό format (4:0:0)	Όχι	Όχι	Όχι	Ναι	Ναι	Ναι	Ναι
4:2:2 χρωματικό format	Όχι	Όχι	Όχι	Όχι	Όχι	Ναι	Ναι
4:4:4 χρωματικό format	Όχι	Όχι	Όχι	Όχι	Όχι	Όχι	Ναι
Βάθος δείγματος 8 bit	Ναι	Ναι	Ναι	Ναι	Ναι	Ναι	Ναι
Βάθος δείγματος 9 και 10 bit	Όχι	Όχι	Όχι	Όχι	Ναι	Ναι	Ναι
Βάθος δείγματος 11 και 14 bit	Όχι	Όχι	Όχι	Όχι	Όχι	Όχι	Ναι
Προσαρμογή 8x8 αντί 4x4 μετασχηματισμού	Όχι	Όχι	Όχι	Ναι	Ναι	Ναι	Ναι
Επεκτασιμότητα κβαντοποίησης	Όχι	Όχι	Όχι	Ναι	Ναι	Ναι	Ναι
Διαχωρισμένοι έλεγχοι Cb και QP	Όχι	Όχι	Όχι	Ναι	Ναι	Ναι	Ναι
Διαχωρισμένο χρωματική κωδικοποίηση	Όχι	Όχι	Όχι	Όχι	Όχι	Όχι	Ναι
Προληπτική μη-απωλεστική κωδικοποίηση	Όχι	Όχι	Όχι	Όχι	Όχι	Όχι	Ναι

Πίνακας 2 - Προφίλ του προτύπου H.264

2.2 Τα πρότυπα H.263, H.263+ και H.263++

Το πρότυπο H.263 αναπτύχθηκε από την ITU-T Video Coding Experts Group (VCEG) κυρίως για εφαρμογές μικρού ρυθμού bit όπως είναι η video-διάσκεψη. Βρήκε όμως πολλές εφαρμογές στο internet και κυρίως στα Flash Videos, όπως αυτά που χρησιμοποιούνται στο YouTube, Google Video, My Space κτλ.

Ορισμένα χαρακτηριστικά του H.263 είναι:

- Αντίστροφος μετασχηματισμός προσδιορισμού ακρίβειας (inverse transform accuracy specification)
- Υποθετική αναφορά αποκωδικοποιητή (hypothetical reference decoder)
- Εξέταση πολλαπλών σημείων (considerations for multipoint)
- Απροσδιόριστα διανύσματα κίνησης (unrestricted motion vector mode)
- Αριθμητική κωδικοποίηση βάσει σύνταξης (syntax-based arithmetic coding mode)
- Προηγμένη μέθοδος πρόβλεψης (advanced prediction mode)
- Μέθοδος PB-πλαισίων (PB-frames mode)
- Μέθοδος διόρθωση λαθών FEC (forward error correction)

Το πρότυπο H.263v2 ή αλλιώς H.263+ είναι η δεύτερη έκδοση του ITU-T H.263, όπου προστέθηκαν επιπλέον χαρακτηριστικά και δυνατότητες τα οποία μπορούν να βελτιώσουν την κωδικοποίηση. Κάποια από τα επιπλέον χαρακτηριστικά είναι:

- Προηγμένη εσωτερική κωδικοποίηση (advanced INTRA coding mode)
- Φίλτρο διαχωρισμού μπλοκ (deblocking filter mode)
- Διάρθρωση τεμαχίων (slice structured mode)
- Συμπληρωματικές πληροφορίες βελτίωσης προσδιορισμού (supplemental enhancement information specification)
- Βελτιωμένη λειτουργία PB-πλαισίων (improved PB-frames mode)
- Επιλογή εικόνας αναφοράς (reference picture selection mode)
- Χρονική, Σηματοθορυβική και Χωρική επεκτασιμότητα (temporal, SNR, Spatial Scalability mode)
- Επανα-δειγματοληψία της εικόνας αναφοράς (reference picture resampling)
- Ενημέρωση μειωμένης-ανάλυσης (reduced-resolution update mode)
- Ανεξάρτητη κωδικοποίηση τμήματος (independent segment coding mode)
- Εναλλακτικό INTER VLC (alternative INTER VLC mode)
- Τροποποιημένη μέθοδος κβαντοποίησης (modified quantization mode)
- Καθορισμός προφίλ και επιπέδων (profiles and levels definition)

Επίσης το H.263+ υποστηρίζει περισσότερα format εικόνων καθώς και εικόνες τροποποιημένης συχνότητας.

Στο πρότυπο H.263v3 ή H.263++ προστέθηκαν κάποια επιπλέον χαρακτηριστικά, μερικά από τα οποία είναι:

- Βελτιωμένη επιλογή εικόνας αναφοράς (enhanced reference picture selection mode)

- Τεμαχισμός δεδομένων σε ένα κομμάτι εικόνας (data-partitioning slice mode)
- Πρόσθετη συμπληρωματική πληροφορία βελτίωσης προσδιορισμού (additional supplemental enhancement information specification)
- Καθορισμός προφίλ και επιπέδων (profiles and levels definition)

	Έκδοση 1	Έκδοση 2	Έκδοση 3
Προηγμένη εσωτερική κωδικοποίηση	Ναι	Ναι	Ναι
Φίλτρο διαχωρισμού μπλοκ	Ναι	Ναι	Ναι
Συμπληρωματική πληροφορία βελτίωσης (πάγωμα πλήρους πλαισίου)	Ναι	Ναι	Ναι
Τροποποιημένη κβαντοποίηση	Ναι	Ναι	Ναι
Μη-αυστηρά διανύσματα κίνησης	Όχι	Ναι	Ναι
Δομή σε κομμάτια	Όχι	Ναι	Ναι
Επαναδειγματοληψία εικόνας αναφοράς	Όχι	Ναι	Ναι
Προηγμένη πρόβλεψη	Όχι	Ναι	Ναι
Βελτιωμένα ΡΒ-πλαίσια	Όχι	Όχι	Ναι
Ανεξάρτητη κωδικοποίηση τμήματος	Όχι	Όχι	Ναι
Εναλλακτικό INTER VLC	Όχι	Όχι	Ναι

Πίνακας 3 - Χαρακτηριστικά των προτύπων H.263, H.263+ και H.263++

2.3 Το πρότυπο MPEG-4

Το πρότυπο κωδικοποίησης MPEG-4 αναπτύχθηκε από την ομάδα ISO/IEC Moving Picture Experts (MPEG) και υιοθέτησε ορισμένα χαρακτηριστικά από τα προγενέστερα MPEG-1 και MPEG-2 αλλά προστέθηκαν και νέα, όπως η γλώσσα προτύπου εικονικής πραγματικότητας (virtual reality modeling language, VRML) η οποία υποστηρίζει τρισδιάστατη απόδοση γραφικών, αντικειμενοστραφή σύνθεση αρχείων (για τα αντικείμενα ήχου, εικόνας και VRML), υποστήριξη εξωτερικά προσδιορισμένων Δικαιωμάτων Ψηφιακής Διαχείρισης (Digital Rights Management) και διάφορους τύπους αλληλεπίδρασης.

Αρχικά το MPEG-4 είχε σκοπό να εξυπηρετεί εφαρμογές video χαμηλού ρυθμού bit (low bit-rate video communications) αλλά στη συνέχεια επεκτάθηκε και σε εφαρμογές πολυμέσων. Κάποια από τα χαρακτηριστικά του προτύπου MPEG-4 είναι:

- Βελτιωμένη απόδοση κωδικοποίησης
- Η δυνατότητα να κωδικοποιεί μικτά δεδομένα (video, ήχο και φωνή)
- Ικανότητα ανάκτησης του σήματος έπειτα από λάθος, βελτιστοποιώντας την μετάδοση (error resilience to enable robust transmission)
- Ικανότητα να αλληλεπιδρά με την οπτικοακουστική σκηνή που γεννάται στον δέκτη

2.4 Το πρότυπο MPEG-2

Το πρότυπο MPEG-2 είναι η βελτιωμένη έκδοση του MPEG-1 η οποία είχε αδυναμίες όπως χαμηλή ανάλυση, χαμηλή απόδοση συμπίεσης ήχου και μικρή ευελιξία καθώς υποστήριζε περιορισμένους τύπους πακέτων.

Το πρότυπο MPEG-2 ξεπερνάει αυτά τα μειονεκτήματα με την συμπίεση των πλαισίων (frames) σε τρία είδη: εσωτερικά κωδικοποιημένα πλαίσια (intra-coded frames, I-frames), κωδικοποιημένα πλαίσια με πρόβλεψη (predictive-coded frames, P-frames) και κωδικοποιημένα πλαίσια με αμφίδρομη πρόβλεψη βάσει των bits (bidirectionally-predictive-coded frames, B-frames) καθώς και με τη χρήση μακρο-μπλοκ.

Επίσης εισάγει νέες μεθόδους κωδικοποίησης ήχου όπως η χαμηλού ρυθμού bit με δείγματα ίσου αριθμού ψηφίων, πολυκάναλη κωδικοποίηση με μέχρι 5.1 κανάλια και την εξελιγμένη κωδικοποίηση ήχου (MPEG-2 Advanced Audio Coding, MPEG-2 AAC).

Συνομογραφία	Όνομα Προφίλ	Υποστηριζόμενοι τύποι εικόνων	Χρωματικό format	Αναλογία	Μεταβολή της κλίμακας
SP	Simple	I, P	4:2:0	Τετράγωνα pixels 4:3 ή 16:9	Καμία
MP	Main	I, P, B	4:2:0	Τετράγωνα pixels 4:3 ή 16:9	Καμία
SNR	SNR scalable	I, P, B	4:2:0	Τετράγωνα pixels 4:3 ή 16:9	SNR
Spatial	Spatial scalable	I, P, B	4:2:0	Τετράγωνα pixels 4:3 ή 16:9	SNR ή χωρική
HP	High	I, P, B	4:2:0 και 4:2:2	Τετράγωνα pixels 4:3 ή 16:9	SNR ή χωρική

Πίνακας 4 - Προφίλ του προτύπου MPEG-2

2.5 Το πρότυπο Motion JPEG-2000

Το πρότυπο Motion JPEG-2000 (γνωστό και ως MJ2 ή MJ2P) είναι κυρίαρχο στο ψηφιακό κινηματογράφο και υποστηρίζεται από την Πρωτοβουλία Ψηφιακού Κινηματογράφου (Digital Cinema Initiatives), μία ομάδα από τα μεγαλύτερα studio και πωλητές για την αποθήκευση, διανομή και έκθεση των ταινιών. Πρόκειται για βελτιωμένη έκδοση του MJPEG (MJ) το οποίο στηρίχθηκε στο πρότυπο JPEG.

Το MJ2 σε αντίθεση με τα κοινά πρότυπα κωδικοποίησης video δεν υποστηρίζει χρονική συμπίεση (temporal compression) ή εσωτερική συμπίεση πλαισίου (intra-frame compression) αλλά αντίθετα κάθε πλαίσιο είναι μία ξεχωριστή οντότητα, η οποία κωδικοποιείται ξεχωριστά. Η φυσική δομή τους δεν εξαρτάται από τη χρονική σειρά αλλά χρησιμοποιείται ένα ξεχωριστό προφίλ για τη συμπλήρωση των δεδομένων.

Για τη κωδικοποίηση του ήχου χρησιμοποιείται γραμμική παλμοκωδική διαμόρφωση (Linear pulse code modulation, LPCM) καθώς επίσης και κάποιες παραλλαγές του MPEG-4.

3. Συγκρίσεις αλγορίθμων κωδικοποίησης video

Σε αυτό το κεφάλαιο θα προσπαθήσουμε να κατατάξουμε ορισμένους αλγορίθμους κωδικοποίησης video ανάλογα με την απόδοσή τους υπό ορισμένες συνθήκες. Η κατάταξη αυτή της περισσότερες φορές γίνεται με βάση την ποιότητα της εικόνας (για τον υπολογισμό της οποίας μετράται συνήθως ο σηματοθορυβικός λόγος, PSNR) αλλά και με βάση τον ρυθμό bit που απαιτείται για την μετάδοση της πληροφορίας χωρίς λάθη. Η κατάταξη των αλγορίθμων δεν είναι πάντα εύκολη διαδικασία, καθώς όπως θα δούμε και στη συνέχεια πολλές φορές τα αποτελέσματα που προκύπτουν για τον ίδιο αλγόριθμο οδηγούν σε διαφορετικά συμπεράσματα όσο αναφορά την απόδοσή του.

3.1 Βελτιστοποίηση παραμόρφωσης (νέα μέθοδος ελέγχου ρυθμού bit)

Σε αυτό το πείραμα θα εξεταστεί μία νέα μέθοδος υπολογισμού της πολυπλοκότητας της κίνησης και ρύθμισης του ρυθμού bit η οποία παρουσιάζεται στο [5]. Γίνεται μία περιγραφή της λειτουργίας της νέας μεθόδου και παρουσιάζονται τα συγκριτικά αποτελέσματα της σύγκρισης μεταξύ της νέας και της «κλασικής» μεθόδου.

3.1.1 Βελτιστοποίηση παραμόρφωσης

Το πρότυπο H.264 σχεδιάστηκε ώστε να προσφέρει παρόμοια ποιότητα με αυτή των H.263+ και MPEG-4 αλλά με αρκετά καλύτερη επίδοση στη συμπίεση και πιο αξιόπιστη μετάδοση. Πειράματα έχουν δείξει ότι το H.264 έχει πετύχει καλύτερη ποιότητα βίντεο από αυτή των H.263, MPEG-2 και MPEG-4 με μέχρι και 50% οικονομία στο ρυθμό bit σε σύγκριση με το H.263 και το MPEG-4. Αυτό σημαίνει ότι το H.264 προσφέρει πολύ καλύτερη ποιότητα στον ίδιο ρυθμό bit.

Μία από τις σημαντικές καινοτομίες του προτύπου H.264 είναι η χρήση επτά διαφορετικών μεθόδων κωδικοποίησης μακρο-μπλοκ (macroblock coding modes) έτσι ώστε να γίνεται καλύτερη απεικόνιση των λεπτομερειών του χώρου (spatial) και του χρόνου (temporal), μέσα σε ένα μακρο-μπλοκ. Οι επτά τρόποι κωδικοποίησης (coding modes) είναι SKIP, INTER 16x16, INTER 16x8, INTER 8x16, INTER 8x8, INTRA 16x16 και INTRA 4x4. Στο INTER 8x8 κάθε μπλοκ μπορεί να χωριστεί περαιτέρω σε 8x8, 8x4, 4x8 ή 4x4 υπο-διαμερίσματα (subpartitions). Για την επιλογή του καλύτερου τρόπου κωδικοποίησης για κάθε μακρο-μπλοκ, δοκιμάζονται όλοι οι τρόποι και επιλέγεται αυτός που οδηγεί στην μικρότερο ρυθμό παραμόρφωσης (rate distortion cost). Αυτό γίνεται για πετύχουμε την καλύτερη ανταλλαγή μεταξύ του ρυθμού bit και του ρυθμού παραμόρφωσης (rate distortion, RD) και λέγεται βελτιστοποίηση του ρυθμού παραμόρφωσης (rate distortion optimization, RDO). Το RDO βρίσκεται επιλέγοντας την κωδικοποίηση με το μικρότερο RD κόστος χρησιμοποιώντας τον πολλαπλασιαστική Lagrangian, όπως φαίνεται στην **Εξίσωσή 1**.

$$\min J_i: J_i = D_i + \lambda_i \times R_i$$

Εξίσωση 1 - Ελαχιστοποίηση του ποσοστού παραμόρφωσης

Όπου το J_i είναι το RD κόστος ενός μακρο-μπλοκ χρησιμοποιώντας τον τρόπο i , R_i είναι τα bits δεδομένων που χρησιμοποιούνται και το D_i είναι η παραμόρφωση που προκαλείται. Η κωδικοποίηση με τη λιγότερη παραμόρφωση είναι αυτή που επιλέγεται σαν τη βέλτιστη για το συγκεκριμένο μακρο-μπλοκ.

Στις περισσότερες ιατρικές εικόνες και βίντεο υπάρχουν κάποιες περιοχές οι οποίες έχουν ειδικό ενδιαφέρον και άλλες οι οποίες δεν έχουν διαγνωστική αξία. Έτσι το RDO κάνει το H.264 κατάλληλο για την κωδικοποίηση ιατρικών εικόνων, αφού μπορούμε να μειώσουμε τον πολλαπλασιαστή Lagrangian λ_i έτσι ώστε να πετύχουμε μείωση της παραμόρφωσης στην περιοχή ενδιαφέροντος (region of interest, ROI) ή μπορούμε να αυξήσουμε το λ_i για να μειώσουμε τα bits που χρησιμοποιούνται σε περιοχές χωρίς ιατρικό ενδιαφέρον. Έτσι είναι δυνατό να επιλέξουμε τον καλύτερο τρόπο κωδικοποίησης έτσι ώστε να πετύχουμε ένα συνδυασμό υψηλής συμπίεσης και καλής ποιότητας video, υπολογίζοντας με ακρίβεια την περιοχή ενδιαφέροντος (ROI). Αν το H.264 χρησιμοποιηθεί για την κωδικοποίηση ιατρικών videos, μπορεί να επιτευχθεί η μεγαλύτερη απόδοση εάν επιλεγεί η κατάλληλη κωδικοποίηση η οποία απεικονίζει καλύτερα τα χαρακτηριστικά του ιατρικού περιεχομένου του videos.

3.1.2 Νέα μέθοδος για τη βελτιστοποίηση του ρυθμού παραμόρφωσης

Στη παραπάνω εξίσωση ο ρυθμός bit υπολογίζεται μόνο βάσει της πληρότητας του buffer ανεξάρτητα από το περιεχόμενο του πλαισίου (frame). Αυτό μπορεί να οδηγήσει σε μείωση του σηματοθορυβικού λόγου (peak signal-to-noise ratio, PSNR), ειδικά σε σκηνές με αυξημένη κινητικότητα ή στην αλλαγή σκηνών.

Στο [6] προτάθηκε το ποσοστό της απόλυτης διαφοράς (mean absolute difference ratio, MAD) σαν μονάδα μέτρησης της πολυπλοκότητας της κίνησης (motion complexity, MC) έτσι ώστε να βελτιωθεί η ποιότητα του video στις αλλαγές των σκηνών. Ωστόσο το MAD δεν είναι ο καλύτερος τρόπος για την αντιπροσώπευση της κίνησης, καθώς αντιπροσωπεύει μόνο την ομοιότητα μεταξύ του τρέχοντος πλαισίου και ενός πλαισίου αναφοράς. Επίσης το MAD μπορεί να χρησιμοποιηθεί για την εύρεση του βέλτιστου τρόπου κωδικοποίησης (coding mode). Ωστόσο γίνεται κατανοητό ότι ελάχιστο MAD δεν σημαίνει ελάχιστη παραμόρφωση.

Μια καλύτερη προσέγγιση είναι η χρήση των αισθητών χαρακτηριστικών του περιεχομένου του βίντεο. Στο [7] προτείνεται μία κωδικοποίηση βίντεο η οποία χρησιμοποιεί προσαρμοσμένο RDO (adaptive RDO) οδηγούμενο από αντιληπτούς παράγοντες (perceptual hints). Οι πολλαπλασιαστές Lagrange ρυθμίζονται σύμφωνα με προσεκτική οπτική ανάλυση.

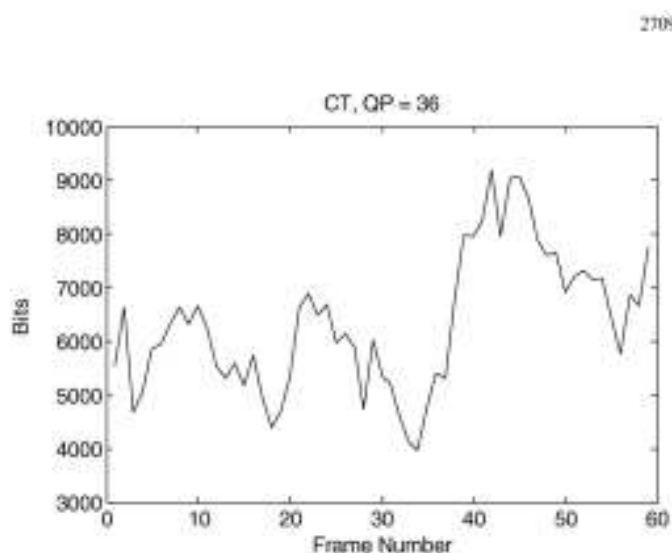
Στο [8] προτείνεται ένα αλγόριθμος όπου ο πολλαπλασιαστής Lagrangian ρυθμίζεται για κάθε εικόνα και καθορίζει την παράμετρο κβαντοποίησης (Quantization parameter, QP) για κάθε μακρο-μπλοκ. Η επιτυχία όλων των παραπάνω αισθητικών αλγορίθμων εξαρτάται αρκετά από τον σωστό υπολογισμό των οπτικών χαρακτηριστικών.

Στο [5] παρουσιάζεται μία αισθητική τεχνική κατανομής των bits (perceptual bit allocation scheme) για το H.264, έτσι ώστε να ξεπεραστούν τα παραπάνω προβλήματα του αλγορίθμου ελέγχου του ρυθμού bit (rate control algorithm). Το ενδιαφέρον συγκεντρώνεται στο εκτιμώμενο επιθυμητό ρυθμό bit (estimating target bit) στις αλλαγές των σκηνών και στην υψηλή κινητικότητα (high motions), όπου ο πολλαπλασιαστής Lagrangian προσαρμόζεται ανάλογα με τα αντιληπτά χαρακτηριστικά του περιεχομένου του video. Επιπλέον στη προτεινόμενη τεχνική περιλαμβάνονται το επίπεδο πλαισίου (frame level) και το επίπεδο μακρο-μπλοκ (MB level). Στο επίπεδο πλαισίου τα bits υπολογίζονται χρησιμοποιώντας μία νέα μέτρηση της πολυπλοκότητας της κίνησης (motion complexity (MC) measure) ή οποία αντιπροσωπεύει το ποσοστό των κινούμενων στοιχείων ανάμεσα σε δύο διαδοχικά πλαίσια. Στο επίπεδο μακρο-μπλοκ κατανέμουμε τα bits αισθητικά, υπολογίζοντας τον πολλαπλασιαστή Lagrangian βάσει του αρχικού μακρο-μπλοκ.

3.1.3 Νέα μέθοδος υπολογισμού της πολυπλοκότητας της κίνησης

Ο σκοπός του ελέγχου του ρυθμού bit είναι να έχουμε υψηλής ποιότητας video για ένα συγκεκριμένο ρυθμό bit χρησιμοποιώντας κατάλληλα τα bits για την υπολειπόμενη πληροφορία, τα διανύσματα κίνησης (vector motions) και κάποιες άλλες παραμέτρους. Για καλύτερη ποιότητα τα bits κατανέμονται σύμφωνα με την πολυπλοκότητα της κίνησης. Θυμίζουμε ότι η πολυπλοκότητα της κίνησης αναφέρεται στα κινούμενα στοιχεία δύο διαδοχικών πλαισίων. Αν τα στοιχεία σε ένα πλαίσιο είναι ακίνητα ή ένα πλαίσιο ταιριάζει αρκετά με το προηγούμενό του, τότε λέμε ότι το πλαίσιο είναι χαμηλής κίνησης (low motion frame). Αν υπάρχει αλλαγή σκηνής μεταξύ δύο πλαισίων ή ένα κινούμενο στοιχείο είναι δύσκολο να βρει ένα ταιριασμα στο προηγούμενο πλαίσιο τότε λέμε ότι το πλαίσιο είναι υψηλής κίνησης (high motion frame).

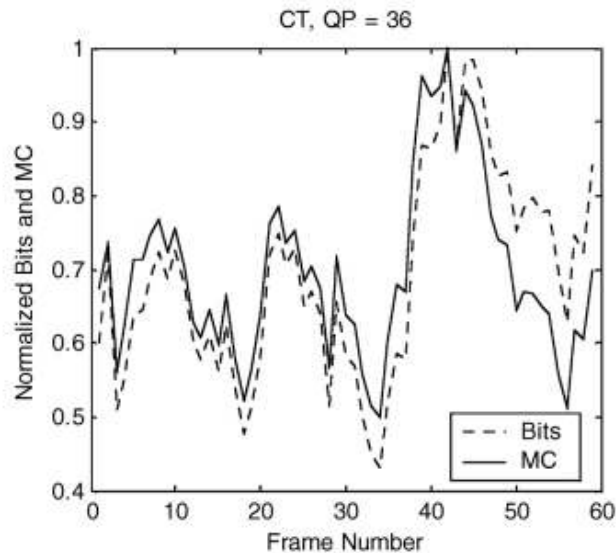
Για να καταλάβουμε καλύτερα πως η κατανομή των bits συνδέεται με την πολυπλοκότητα της κίνησης έχουμε το διάγραμμα της **Εικόνας 1**, όπου φαίνεται η σχέση κατανομής των bits και της πολυπλοκότητας της κίνησης.



Εικόνα 1 - Διάγραμμα σχέσης κατανομής των bits και της πολυπλοκότητας της κίνησης

Όπως βλέπουμε μεταξύ του 43^{ου} και του 44^{ου} πλαισίου έχουμε έντονη κίνηση. Τα bits που δεσμεύονται στο 44^ο πλαίσιο είναι περίπου 14% περισσότερα από αυτά που δεσμευτήκανε στο 43^ο. Αυτό μας οδηγεί στο συμπέρασμα ότι τα bits που δεσμεύονται είναι μία καλή μέτρηση για τον υπολογισμό της πολυπλοκότητας της κίνησης.

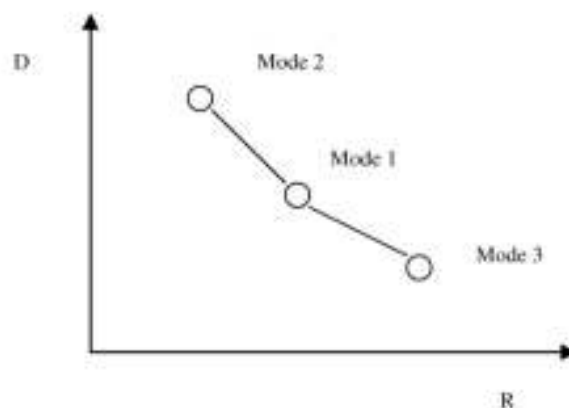
Σύμφωνα με τη νέα μέθοδο που προτείνεται στο [5], η πολυπλοκότητα της κίνησης δεν θα εξαρτάται μόνο από τα bits που έχουνε δεσμευτεί στο προηγούμενο πλαίσιο αλλά θα εξαρτάται από τα bits που έχουνε δεσμευτεί σε ένα αριθμό προηγούμενων κωδικοποιημένων πλαισίων. Στην **Εικόνα 2** φαίνεται ο υπολογισμός της πολυπλοκότητας της κίνησης και των δεσμευμένων bits (όμοια με την **Εικόνα 1**). Όπως παρατηρούμε η πολυπλοκότητα της κίνησης των στοιχείων υπολογίζεται περίπου ίδια και με τη νέα μέθοδο που προτείνεται.



Εικόνα 2 – Σχέση νέας μεθόδου υπολογισμού της πολυπλοκότητας της κίνησης και κατανομής των bits

3.1.3 Δέσμευση bits με χρήση αντιληπτών παραμέτρων

Όπως αναφέρθηκε και παραπάνω η επιλογή του πολλαπλασιαστή Lagrangian καθορίζει και την επιλογή του τρόπου κωδικοποίησης (coding mode).



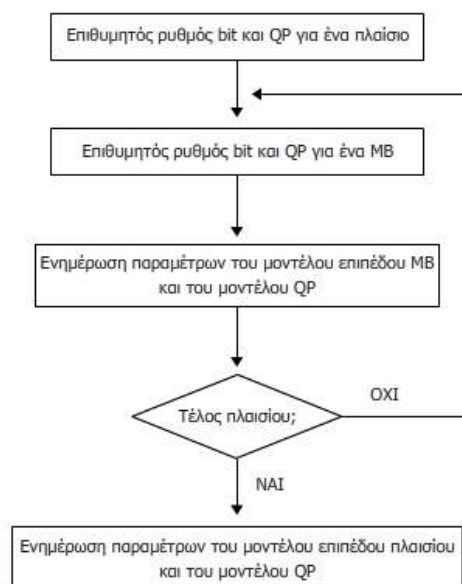
Εικόνα 3 - Καμπύλη του ποσοστού παραμόρφωσης και επιλογή της μεθόδου κωδικοποίησης

Κάθε σημείο της καμπύλης RD της **Εικόνας 3** αντιπροσωπεύει και έναν τρόπο κωδικοποίησης (coding mode). Ας υποθέσουμε ότι το mode 1 είναι το αρχικό, όπου ο πολλαπλασιαστής Lagrangian υπολογίζεται από την **Εξίσωση 1**. Αν αυξήσουμε τον πολλαπλασιαστή Lagrangian τότε θα μεταφερθούμε στο mode 2 που σημαίνει ότι θα δεσμεύονται λιγότερα bits άρα θα έχουμε μεγαλύτερη παραμόρφωση. Αν αντίστοιχα μειώσουμε τον πολλαπλασιαστή Lagrangian θα μεταφερθούμε στο mode 3 όπου θα δεσμεύονται περισσότερα bits άρα θα έχουμε μικρότερη παραμόρφωση.

Ένα από τα σημαντικότερα χαρακτηριστικά του ανθρώπινου οπτικού συστήματος (human vision system, HVS) ονομάζεται απόκρυψη ορίου (edge masking), όπου το HSV είναι πιο ευαίσθητο σε λάθη σε ένα εμφανές όριο μέσα σε μία εικόνα ενώ αντίθετα είναι λιγότερο ευαίσθητο σε λάθη δίπλα στο όριο αυτό. Έτσι πρέπει να δεσμεύονται περισσότερα bits για τη διαμόρφωση του ορίου και λιγότερα για την διαμόρφωση της υφής (texture pattern). Έτσι σύμφωνα με το παραπάνω διάγραμμα της **Εικόνας 3** μπορούμε να πούμε ότι αν το μακρο-μπλοκ ανήκει στη διαμόρφωση της υφής πρέπει να χρησιμοποιηθούν λιγότερα bits άρα μπορεί να χρησιμοποιηθεί το mode 2, ενώ αν το μακρο-μπλοκ ανήκει στη διαμόρφωση του ορίου χρειάζονται περισσότερα bits άρα μπορεί να χρησιμοποιηθεί το mode 3. Έτσι αλλάζοντας τον πολλαπλασιαστή Lagrangian σύμφωνα με τη διαμόρφωση στην οποία ανήκει το μακρο-μπλοκ, αλλάζουμε τη μέθοδο κωδικοποίησης άρα και τα δεσμευμένα bits.

3.1.4 Νέα μέθοδος ελέγχου του ρυθμού bit

Βασίζομενοι στη μέτρηση της πολυπλοκότητας και στην αισθητική επιλογή της μεθόδου κωδικοποίησης (coding mode), προτείνεται στο [5] μία μέθοδος ελέγχου του ρυθμού bit, η οποία περιλαμβάνει δύο επίπεδα ελέγχου (levels of rate control): έναν έλεγχο ρυθμού επιπέδου πλαισίου (frame-level rate control) και έναν έλεγχο ρυθμού επιπέδου μακρο-μπλοκ (MB-level rate control). Στο έλεγχο επιπέδου πλαισίου υπολογίζεται ο επιθυμητός ρυθμός bit (target bit rate) του τρέχοντος πλαισίου σύμφωνα με την πολυπλοκότητα της κίνησης (motion complexity). Στον έλεγχο επιπέδου μακρο-μπλοκ υπολογίζεται ο πολλαπλασιαστής Lagrangian σύμφωνα με τα οπτικά χαρακτηριστικά. Το διάγραμμα ροής του αλγορίθμου φαίνεται στην **Εικόνα 4**.



Εικόνα 4 - Διάγραμμα ροής του προτεινόμενου αλγορίθμου ελέγχου του ρυθμού bit

3.1.5 Πειραματική διαδικασία

Στη πειραματική διαδικασία πραγματοποιήθηκαν δύο συγκρίσεις: μία μεταξύ του MPEG-4 και του H.264 χωρίς έλεγχο ρυθμού bit και μία μεταξύ του H.264 με τον προκαθορισμένο έλεγχο του ρυθμού bit (H014) και του H.264 με τη νέα μέθοδο ελέγχου του ρυθμού bit. Το

πρότυπο MPEG-4 εφαρμόζεται με το MoMuSys codec ενώ το H.264 με το H.264 μοντέλο αναφοράς JM6.1e.

Για να προσδιορίσουμε τον επιθυμητό ρυθμό bit κωδικοποιούμε πρώτα τα δείγματα προς εξέταση χρησιμοποιώντας ένα σταθερό QP με το H.264 χωρίς έλεγχο του ρυθμού bit. Έπειτα ο πραγματικός αριθμός των bits που βρέθηκε χρησιμοποιείται σαν ο επιθυμητός ρυθμός bit στα ακόλουθα πειράματα. Θεωρώντας διαφορετικές συνθήκες δικτύου χρησιμοποιήθηκαν διαφορετικά QP (28, 32, 36 και 40) και έτσι υπάρχουν τέσσερις ρυθμοί bit για κάθε video προς εξέταση (test sequence).

Κατά την πειραματική διαδικασία είναι ενεργοποιημένα μόνο τα βασικά χαρακτηριστικά όπως αναφέρονται στο Βασικό προφίλ (Main Profile) του H.264.

Video	Format	Μέγεθος	Αριθμός πλαισίων
ηλεκτρονική τομογραφία	YUV	512x512	60
ηχοκαρδιογράφημα	YUV	640x480	60

Πίνακας 5 - Videos που χρησιμοποιήθηκαν για τη σύγκριση του MPEG-4 και του H.264

Τύπος πλαισίου (frame type)	IPP
Ρυθμός πλαισίου (frame rate)	30fps
Slice mode	OFF
RDO	ON
Έλεγχος bit rate (Rate control)	OFF
Hardmard	OFF
Εύρος αναζήτησης (Search range)	32
Όριο εύρους αναζήτησης (restrict search range)	Χωρίς περιορισμό (no restriction)
Ρύθμιση συμβόλου (symbol mode)	CABAC
Ρύθμιση τεμαχισμού (partition mode)	Χωρίς τεμαχισμό δεδομένων (no data partition)
Ρύθμιση εκτός αρχείου (Out file mode)	Annex B

Πίνακας 6 - Ρυθμίσεις του προτύπου H.264 για την διεξαγωγή του πειράματος

3.1.6 Αποτελέσματα σύγκρισης μεταξύ των H.264 και MPEG-4

Οι συνθήκες επιλέχθηκαν έτσι ώστε το MPEG-4 να ταιριάζει όσο το δυνατό περισσότερο με το H.264. Για παράδειγμα ο τύπος πλαισίου (frame type) είναι IPPP, ο ρυθμός πλαισίου (frame rate) 30fps, το εύρος αναζήτησης (search range) 32 κτλ. Στον **Πίνακα 7** φαίνονται τα αποτελέσματα για τέσσερεις διαφορετικούς ρυθμούς bit.

Video	Επιθυμητός ρυθμός bit (kbps)	PSNR (dB)		Κέρδος PSNR (dB)	Ρυθμός bit (kbps)		Εξοικονόμηση ρυθμού bit (%)
		MPEG-4	H.264		MPEG-4	H.264	
Ηλεκτρονική τομογραφία	438.19	36.12	43.07	6.95	450.93	438.19	2.83
	285.03	35.05	40.55	5.50	303.56	285.03	6.10
	197.55	35.10	37.73	2.63	214.89	197.55	8.07
	145.10	35.07	34.91	-0.16	162.38	145.10	10.64
Ηχοκαρδιογράφημα	2266.14	35.79	38.09	2.30	2266.51	2266.14	0.02
	1317.29	32.94	35.15	2.21	1292.83	1317.29	-1.89
	667.08	30.98	32.44	1.46	703.83	667.08	5.22
	308.26	30.41	29.81	-0.60	327.52	308.26	5.88

Πίνακας 7 - PSNR και ρυθμοί bit του MPEG-4 και του H.264 για τα videos ηλεκτρονικής τομογραφίας και ηλεκτροκαρδιογραφήματος

Το κέρδος σηματοθορυβικού λόγου (PSNR gain) είναι η διαφορά του σηματοθορυβικού λόγου σε dB (decibels) μεταξύ του H.264 και του MPEG-4. Αντίστοιχα η εξοικονόμηση ρυθμού bit είναι η ποσοστιαία (%) διαφορά ρυθμού bit μεταξύ του H.264 και του MPEG-4.

Όπως μπορούμε να παρατηρήσουμε το H264 αποδίδει καλύτερα από το MPEG-4. Στην ηλεκτρονική τομογραφία (computed tomography, CT) πετυχαίνει θετικό κέρδος σε όλες σχεδόν τις μετρήσεις, ενώ συγχρόνως έχουμε εξοικονόμηση στο ρυθμό bit. Ακόμα και σε υψηλό ρυθμό bit βλέπουμε ότι πετυχαίνει αρκετό κέρδος PSNR, περίπου 6.95dB.

Για το ηχοκαρδιογράφημα (Echocardiography) το H.264 πετυχαίνει και πάλι θετικό κέρδος PSNR σε όλες τις περιπτώσεις.

Για να έχουμε μια συνολική εικόνα της απόδοσης του H.264 σε σχέση με το MPEG-4, στον **Πίνακα 8** παρουσιάζεται η μέση διαφορά των σηματοθορυβικών λόγων (average PSNR difference) και η μέση διαφορά ρυθμού bit (average bit rate difference). Τα μεγέθη αυτά χρησιμοποιούνται για σύγκριση του ποσοστού παραμόρφωσης (Rate Distortion) μεταξύ των δύο μεθόδων.

Video	Μέση διαφορά PSNR για όλο το εύρος των bit rate (dB)	Μέση διαφορά bit rate για όλο το εύρος PSNR (%)
Ηλεκτρονική τομογραφία	4.35	<-100
Ηχοκαρδιογράφημα	1.51	-59.71

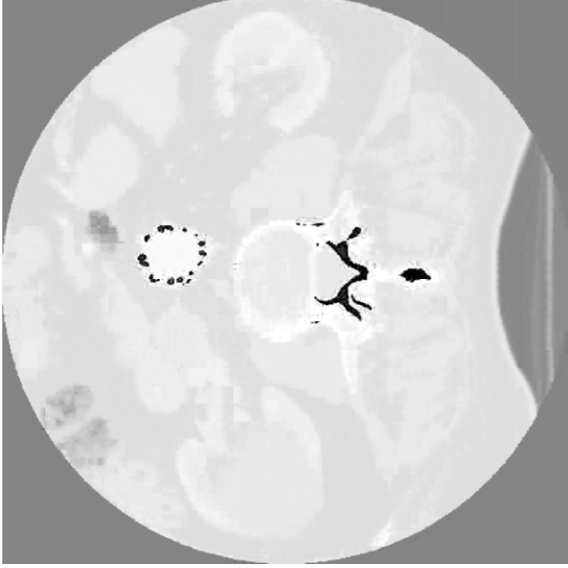
Πίνακας 8 - Μέση διαφορά PSNR και ρυθμού bit μεταξύ MPEG-4 και H.264

Όπως φαίνεται το H.264 αποδίδει αρκετά καλύτερα από το MPEG-4 όσο αναφορά τη μέση διαφορά PSNR και τη μέση διαφορά ρυθμού bit. Αυτό μας οδηγεί στο συμπέρασμα ότι το H.264 είναι πιο αποδοτική λύση από το MPEG-4 για συμπίεση ιατρικών βίντεο.

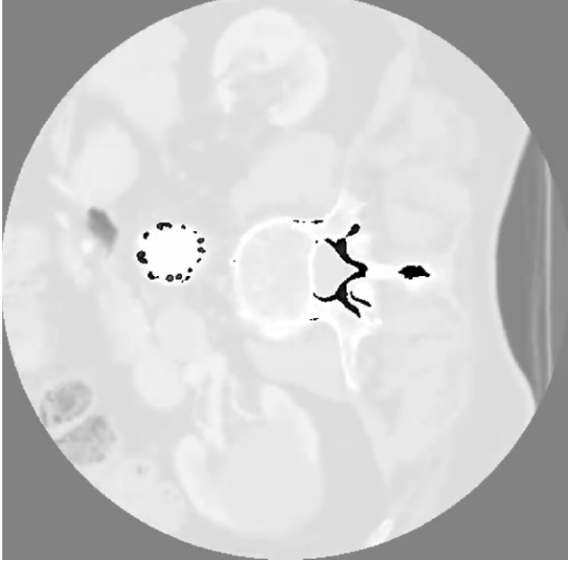
Στις παρακάτω εικόνες παρουσιάζεται το 44^ο πλαίσιο του CT για ρυθμό bit 438.19kbps όπως προέκυψε από το H.264 και το MPEG-4 καθώς επίσης παρουσιάζεται και η ασυμπίεστη εικόνα. Επίσης δίνεται το 32^ο πλαίσιο του ηχοκαρδιογραφήματος για ρυθμό bit 667.08kbps για το H.264 και το MPEG-4 καθώς επίσης και η ασυμπίεστη εικόνα.



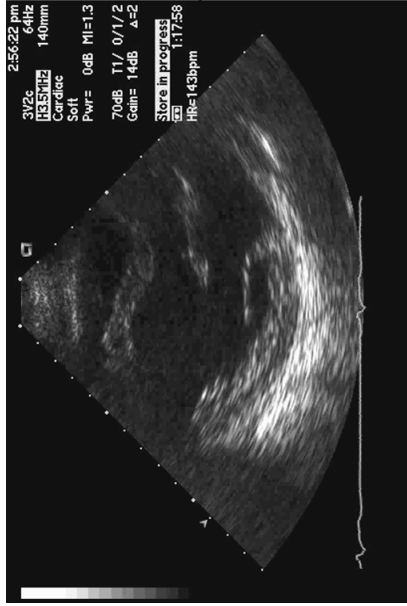
Εικόνα 5 - Ασυμπίεστη εικόνα ηλεκτρονικής τομογραφίας



Εικόνα 6 - Συμπιεσμένη εικόνα ηλεκτρονικής τομογραφίας με MPEG-4



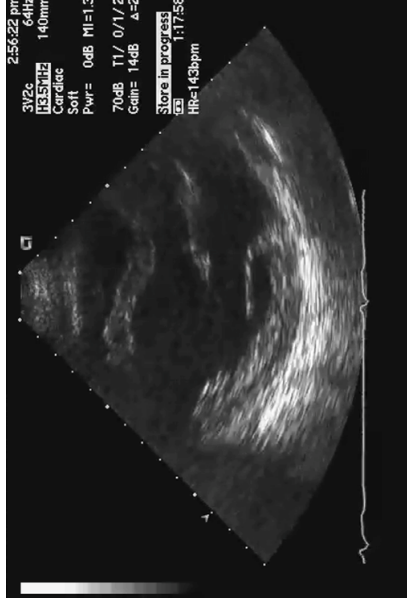
Εικόνα 7 - Συμπιεσμένη εικόνα ηλεκτρονικής τομογραφίας με H.264



Εικόνα 8 - Ασυμπίεστη εικόνα ηχοκαρδιογραφήματος



Εικόνα 9 - Συμπιεσμένη εικόνα ηχοκαρδιογραφήματος με MPEG-4



Εικόνα 10 - Συμπιεσμένη εικόνα ηχοκαρδιογραφήματος με H.264

3.1.7 Σύγκριση του ελέγχου του ρυθμού bit H014 και της νέας μεθόδου

3.1.7.1 Αντικειμενική αξιολόγηση ποιότητας

Στον Πίνακα 9 παρουσιάζονται τα αποτελέσματα του αλγορίθμου ελέγχου του ρυθμού bit H014 και της προτεινόμενης μεθόδου ελέγχου. Για να δούμε τη διαφορά μεταξύ του H014 και της προτεινόμενης μεθόδου, παρουσιάζεται το κέρδος PSNR και η εξοικονόμηση ρυθμού bit σαν μονάδες μέτρησης της απόδοσης. Το κέρδος PSNR είναι η διαφορά μεταξύ των σηματοθορυβικών λόγων μεταξύ της προτεινόμενης μεθόδου και του H014 και αντίστοιχα η εξοικονόμηση ρυθμού bit είναι η ποσοστιαία διαφορά (%) του ρυθμού bit της προτεινόμενης μεθόδου και του H014.

Video	Επιθυμητός ρυθμός bit (kbps)	PSNR (dB)		Κέρδος PSNR (dB)	Ρυθμός bit (kbps)		Εξοικονόμηση ρυθμού bit (%)
		H014	Προτεινόμενη		H014	Προτεινόμενη	
Ηλεκτρονική τομογραφία	438.19	42.89	42.97	0.08	439.45	439.80	+0.08
	285.03	40.28	40.43	0.15	286.37	286.92	+0.19
	197.55	37.43	37.69	0.26	198.00	197.91	0.05
	145.10	35.06	35.33	0.27	146.73	145.79	0.64
Ηχοκαρδιογράφημα	2266.14	38.18	38.16	-0.01	2268.22	2265.24	0.13
	1317.29	35.20	35.27	0.07	1319.49	1318.66	0.06
	667.08	32.43	32.56	0.13	671.47	670.46	0.15
	308.26	29.84	30.00	0.16	311.73	308.45	1.05

Πίνακας 9 - PSNR και ρυθμός bit των μεθόδων H014 και της προτεινόμενης για τα videos ηλεκτρονικής τομογραφίας και ηχοκαρδιογραφήματος

Όπως παρατηρούμε με την προτεινόμενη μέθοδο ελέγχου του ρυθμού bit έχουμε καλύτερη απόδοση και στο κέρδος PSNR και στην εξοικονόμηση ρυθμού bit στις περισσότερες περιπτώσεις. Όσο αυξάνεται ο QP (ή αλλιώς όσο μειώνεται ο επιθυμητός ρυθμός bit) βλέπουμε ότι το κέρδος PSNR με χρήση της προτεινόμενης μεθόδου σε σχέση με το H014 αυξάνεται. Αυτό μας οδηγεί στο συμπέρασμα ότι η προτεινόμενη μέθοδος λειτουργεί καλύτερα σε δίκτυα με χαμηλό ρυθμό bit (low bit rate network environment).

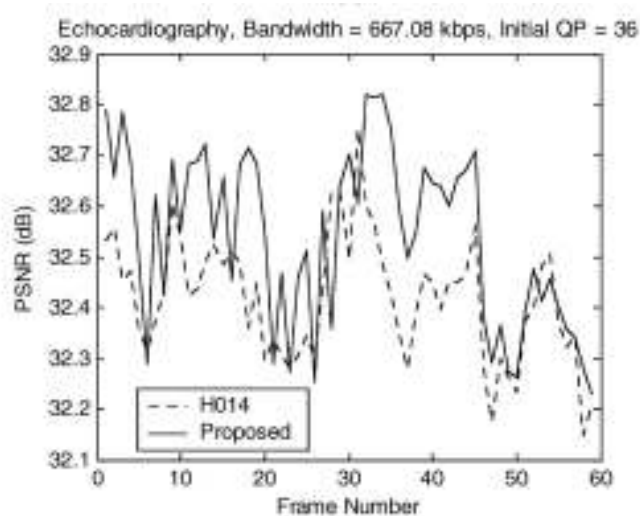
Συγκεκριμένα η προτεινόμενη μέθοδος ελέγχου αποδίδει καλύτερα στο υψηλής κίνησης ηλεκτρονικό τομογράφο (CT) όπου το κέρδος PSNR είναι θετικό για όλους τους επιθυμητούς ρυθμούς bit, ακόμα και αν η εξοικονόμηση ρυθμού bit είναι αρνητική, πράγμα που καθιστά την συνολική απόδοση της προτεινόμενης μεθόδου καλύτερη. Αλλά και στο ηχοκαρδιογράφημα (echocardiography) η προτεινόμενη μέθοδος έχει θετική εξοικονόμηση ρυθμού bit σε όλες τις περιπτώσεις και θετικό κέρδος PSNR στις περισσότερες.

Για να δούμε και την συμπεριφορά των codecs σε συνθήκες πολύ υψηλού ρυθμού bit έγιναν μετρήσεις με QP=24 και επιθυμητό ρυθμό bit 3402.4kbps. Στη συγκεκριμένη περίπτωση ο σηματοθορυβικός λόγος του H014 είναι 41.29dB και ο ρυθμός bit 3403.36kbps, ενώ με χρήση της προτεινόμενης μεθόδου είναι 41.26dB και 3403.38kbps αντίστοιχα. Έτσι το κέρδος PSNR προκύπτει -0,03dB και η εξοικονόμηση ρυθμού bit 0.00%. Άρα οδηγούμαστε στο συμπέρασμα ότι το H014 αποδίδει οριακά καλύτερα από την προτεινόμενη μέθοδο σε συνθήκες υψηλού ρυθμού bit.

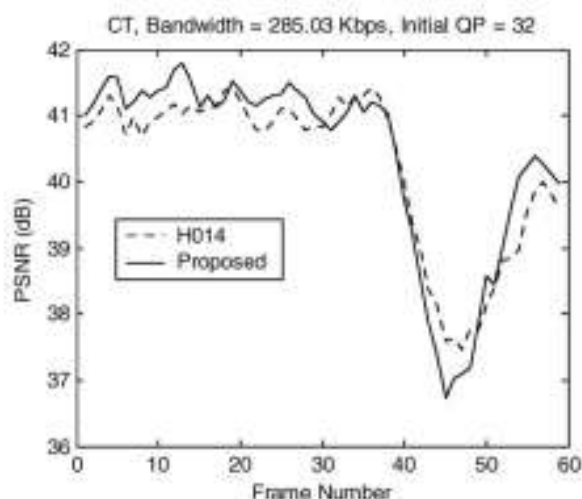
Ο Πίνακας 10 παρουσιάζει τη μέση διαφορά PSNR και τη μέση διαφορά ρυθμού bit μεταξύ της προτεινόμενης μεθόδου και του H014. Όπως παρατηρείται η προτεινόμενη μέθοδος ελέγχου αποδίδει καλύτερα σε συνολικό επίπεδο.

Video	Μέση διαφορά PSNR για όλο το εύρος των bit rate (dB)	Μέση διαφορά bit rate για όλο το εύρος PSNR (%)
Ηλεκτρονική τομογραφία	0.19	-2.65
Ηχοκαρδιογράφημα	0.11	-2.66

Πίνακας 10 - Μέση διαφορά PSNR και ρυθμού bit μεταξύ των μεθόδων H014 και της προτεινόμενης



Εικόνα 11 - Σχέση PSNR και αριθμού πλαισίου για το video ηχοκαρδιογραφήματος, με ρυθμό bit 667.08kbps



Εικόνα 12 - Σχέση PSNR και αριθμού πλαισίου για το video ηλεκτρονικής τομογραφίας, με ρυθμό bit 285.03kbps

3.1.7.2 Υποκειμενική αξιολόγηση ποιότητας

Για την υποκειμενική αξιολόγηση της ποιότητας χρησιμοποιήθηκε η μέθοδος διπλού συνεχούς ερεθίσματος (Double Stimulus Continuous Quality Scale, DSCQS), όπως αυτή περιγράφεται στο ITU-R BT.500-10. Η κλίμακα MOS² αξιολόγησης της ποιότητας για τους παρατηρητές ήταν εξαιρετική, καλή, ανεκτή, φτωχή και κακή με αντίστοιχη βαθμολογία 5, 4, 3, 2 και 1. Στο πείραμα πήραν μέρος πέντε παρατηρητές ένας εκ των οποίων ήταν λειτουργός υγείας (medical doctor) και όλοι τους είχαν κάποιες γνώσεις από επεξεργασία εικόνας. Στο Πίνακα 11 φαίνονται οι τιμές MOS που πέτυχε το κάθε codec.

Video	QP	Επιθυμητός ρυθμός bit (kbps)	Score		Κέρδος
			H014	Προτεινόμενη	
Ηλεκτρονική τομογραφία	28	438.19	4.20	4.20	0.00
	32	285.03	3.60	3.80	0.20
	36	197.55	3.60	4.00	0.40
	40	145.10	3.60	3.80	0.20
Ηχοκαρδιογράφημα	28	2266.14	4.20	4.20	0.00
	32	1317.29	3.80	4.20	0.40
	36	667.08	4.00	4.20	0.20
	40	308.26	3.60	3.80	0.20

Πίνακας 11 - Υποκειμενικό score για τη μέθοδο ελέγχου του ρυθμού bit H014 και την προτεινόμενη

Όπως προκύπτει και από τα αποτελέσματα η προτεινόμενη μέθοδος ελέγχου του ρυθμού bit αποδίδει καλύτερα από το H014. Είναι άξιο προσοχής ότι όλα τα κέρδη στο score

² Βλ. Παράρτημα A12

πετυχαίνονται για χαμηλούς ρυθμούς bit, το οποίο συμφωνεί με τις αντικειμενικές μετρήσεις.

3.2 Σύγκριση των codecs H.264, H.263 και H.263+

Στο [9] συγκρίνεται το πρότυπο H.264 με το πρότυπο H.263 και το βελτιωμένο H.263+. Η σύγκριση της ποιότητας γίνεται με βάση τον σηματοθορυβικό λόγο (PSNR) και το ρυθμό bit.

3.2.1 Πειραματική διαδικασία

Το σύστημά μας θεωρείται ότι είναι ένα σύστημα κινητής τηλεφωνίας το οποίο απαιτεί μικρό μέγεθος πλαισίου (frame size). Επίσης χρησιμοποιήθηκαν QCIF (176x144) videos. Τα videos αντιπροσωπεύουν έχουν τυπικό περιεχόμενο με χαμηλή και υψηλή καθυστέρηση. Τα videos είναι το Foreman, το Coast Guard, το Car Phone, το Miss America και το Tennis. Χρησιμοποιήθηκαν 150 πλαίσια για κωδικοποίηση και όλα ήταν κωδικοποιημένα με 15fps. Τα μεγέθη που συγκρίνονται είναι ο τελικός ρυθμός bit (output bit rate) και ο σηματοθορυβικός λόγος (PSNR) και λήφθηκε υπόψη μόνο η φωτεινότητα των στοιχείων (luminance component) καθώς το μάτι είναι πιο ευαίσθητο στη φωτεινότητα σε σχέση με τα χρώματα. Για το πρότυπο H.264 χρησιμοποιήθηκε το μοντέλο αναφοράς έκδοση 7.6 (model reference software version 7.6).

Το πρότυπο H.263 χρησιμοποιείται κυρίως για εφαρμογές με χαμηλή καθυστέρηση και χαμηλού έως μέτριου ρυθμού bit, όπως είναι η video-διάσκεψη και η παρακολούθηση χώρων.

Για να εξασφαλιστεί η χαμηλή καθυστέρηση δεν χρησιμοποιήθηκαν B-pictures.

Το πρότυπο H.263 που χρησιμοποιήθηκε ήταν η έκδοση 3 της UBC και ρυθμίστηκε για $\frac{1}{4}$ του pixel ανάλυση διανύσματος κίνησης (quarter pixel motion vector resolution), με χρήση πέντε πλαισίων για εσωτερική κίνηση (inter motion), με προσαρμοσμένη δυαδική κωδικοποίηση γενικευμένου πλαισίου (context-based adaptive binary coding, CABAC) για κωδικοποίηση συμβόλων και ενεργοποίηση της λήψης απόφασης ποσοστού παραμόρφωσης (rate distortion made decision).

Το πρότυπο H.263+ είναι μία επέκταση του H.263, το οποίο παρέχει ορισμένες επιπλέον δυνατότητες και χαρακτηριστικά για την βελτίωση της κωδικοποίησης και την ανάκτηση λαθών.

Για την εκτέλεση του πειράματος χρησιμοποιήθηκε το πρότυπο TMN H.263 έκδοση 3 από την UBC, με ενεργοποιημένη την προηγμένη μέθοδο εσωτερικής κωδικοποίησης (advanced intra coding mode), ενεργοποιημένο το φίλτρο διαχωρισμού των μπλοκ (deblocking filter mode), ενεργοποιημένο το χαρακτηριστικό συμπληρωματικής βοηθητικής πληροφορίας (supplemental enhancement information mode), ενεργοποιημένη τη προηγμένη μέθοδο πρόβλεψης (advance prediction mode) και τη αριθμητική κωδικοποίηση βάσει της σύνταξης (syntax-based arithmetic coding options), με παράθυρο αναζήτησης (search window) 15x15 και πέντε πλαίσια για εσωτερική αναζήτηση κίνησης (inter search motion).

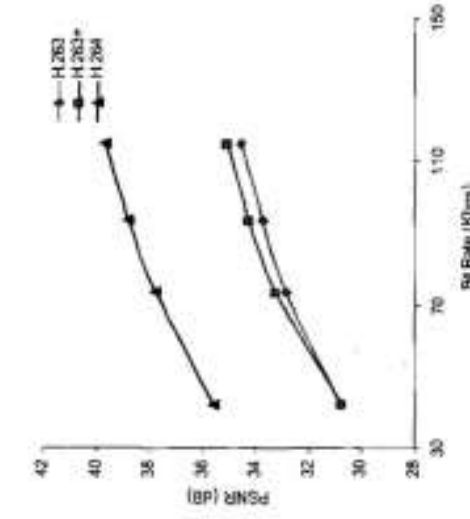
3.2.2 Αποτελέσματα σύγκρισης

Στον Πίνακα 12 και στις Εικόνες 13 μέχρι 17 παρουσιάζονται τα αποτελέσματα της σύγκρισης των τριών codecs. Όλα τα αποτελέσματα δείχνουν ότι το πρότυπο H.264 αποδίδει καλύτερα από το H.263 και το H.263+ για όλα τα video ελέγχθηκαν.

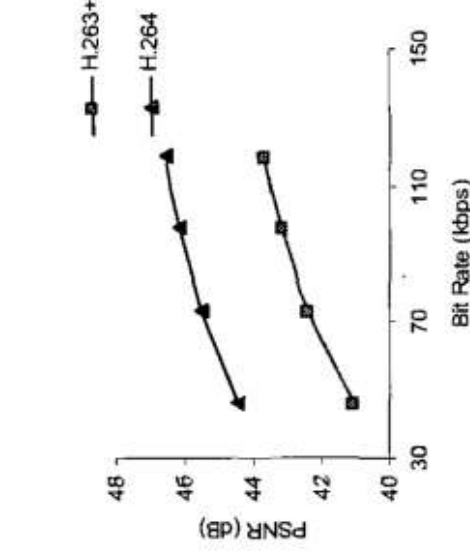
Video	Bit rate	PSNR Φωτεινότητας (Luminance PSNR) (dB)			Διαφορά του H.264 με το	
		H.263	H.263+	H.264	H.263	H.263+
	kbps					
Foreman	42	30.81	30.80	35.58	4.77	4.78
	74	32.85	33.29	37.78	4.93	4.49
	94	33.73	34.23	38.75	5.02	4.52
	115	34.53	35.03	39.62	5.09	4.59
Coast Guard	37	27.74	27.57	30.71	2.97	3.14
	68	29.86	30.01	32.97	3.11	2.96
	95	31.06	31.31	34.24	3.18	2.93
	116	31.82	32.08	35.03	3.21	2.95
Miss America	46	41.05	41.07	44.46	3.41	3.39
	73	42.28	42.40	45.52	3.24	3.12
	98	43.13	43.20	46.16	3.03	2.96
	119	43.67	43.73	46.57	2.90	2.84
Car phone	42	33.68	33.87	37.89	4.21	4.02
	62	35.24	35.47	39.75	4.51	4.28
	83	36.45	36.73	41.01	4.56	4.28
	103	37.39	37.70	42.02	4.63	4.32
Tennis	43	29.32	29.72	33.44	4.12	3.72
	68	31.30	32.05	35.70	4.40	3.65
	94	32.74	33.60	37.14	4.40	3.54
	120	33.90	34.76	38.36	4.46	3.60

Πίνακας 12 - Μέσο PSNR φωτεινότητας του H.264, H.263 και H.263+

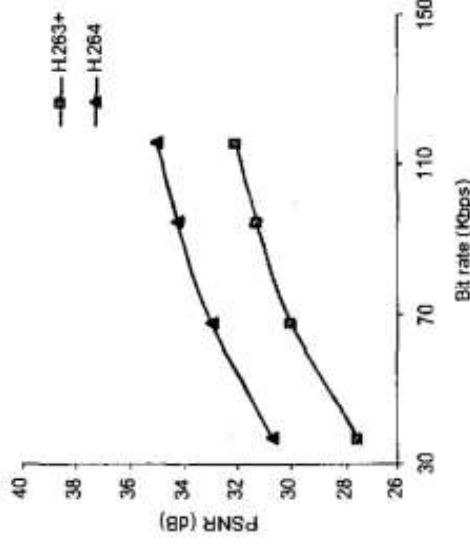
Από τον **Πίνακα 12** φαίνεται ότι το πρότυπο H.264 αποδίδει καλύτερα σε όλα τα videos που εξετάστηκαν και έχουμε ένα μέσο κέρδος PSNR περίπου 4%. Αντίστοιχα σε σχέση με το H.263+ το μέσο κέρδος PSNR που έχουμε με το H.264 είναι 3.7%.



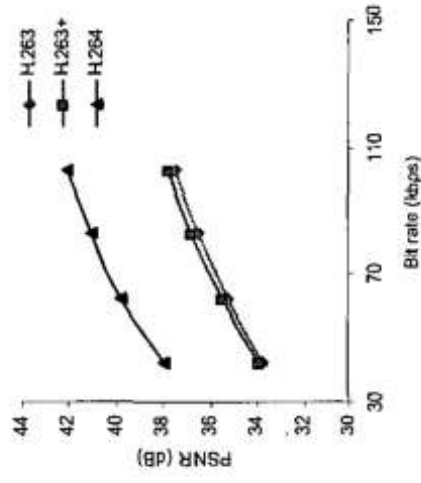
Εικόνα 13 - Διάγραμμα ρυθμού bit / PSNR για το video Foreman QCIF



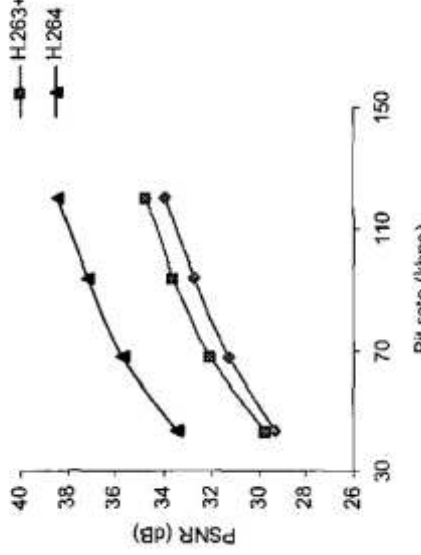
Εικόνα 14 - Διάγραμμα ρυθμού bit / PSNR για το video Miss America QCIF



Εικόνα 15 - Διάγραμμα ρυθμού bit / PSNR για το video Coast Guard QCIF



Εικόνα 16 - Διάγραμμα ρυθμού bit / PSNR για το video Car and Phone QCIF



Εικόνα 17 - Διάγραμμα ρυθμού bit / PSNR για το video Tennis QCIF

3.3 Σύγκριση του H.264, του MPEG-2 και του H.263

Στο πείραμα που παρουσιάζεται στο [10] συγκρίνεται η απόδοση του προτύπου H.264 με δύο προφίλ του προτύπου H.263, το βασικό προφίλ και το προφίλ υψηλής συμπίεσης και το πρότυπο MPEG-2.

Το πρότυπο H.264 έχει σχεδιαστεί σε δύο στρώματα (layers): ένα στρώμα κωδικοποίησης video (video coding layer, VCL) και ένα στρώμα δικτύου NAL (network adaptation layer, NAL). Το συγκεκριμένο πείραμα συγκεντρώνεται περισσότερο στο στρώμα VCL του H.264.

3.3.1 Πειραματική διαδικασία

Στο πείραμα χρησιμοποιήθηκαν τρία videos που αντιπροσωπεύουν ένα τυπικό περιεχόμενο για χαμηλής και υψηλής καθυστέρησης εφαρμογές. Χρησιμοποιήθηκαν 200 πλαίσια για κωδικοποίηση σε κάθε video. Τα videos χωρίζονται σε τρία σετ.

Το πρώτο σετ αποτελείται από τα: Carousel, Flower, Garden και Football. Οι εικόνες είναι CCIR-601 (720x480) με ρυθμό πλαισίων (frame rate) 30fps.

Το δεύτερο σετ αποτελείται από τα: Irene, Mobile and Calendar, Paris, Students και Tempete. Οι εικόνες εδώ είναι CIF format (352x288) με ρυθμό πλαισίων 30fps.

Το τρίτο σετ αποτελείται από τα: Akiyo, Carp Honec, Lairec, Coast Guard, Container, Foreman, Irene, Miss America, Mobile and calendar, News, Paris και Tempete. Οι εικόνες εδώ είναι QCIF format (176x144) με ρυθμό πλαισίων 15fps.

Σαν κριτήριο της απόδοσης των codecs χρησιμοποιήθηκε ο σηματοθορυβικός λόγος και ο ρυθμός bit. Για το υπολογισμό του σηματοθορυβικού λόγου λήφθηκε υπόψη μόνο το στοιχείο της φωτεινότητας (luminance component).

3.3.2 Αποτελέσματα σύγκρισης του H.264 με το MPEG-2

Η απόδοση των προτύπων H.264 και MPEG-2 συγκρίνεται με βάση τα δύο πρώτα σετ video. Η κωδικοποίηση με MPEG-2 έγινε με βάση του codec που αναπτύχθηκε από το Πανεπιστήμιο της Καλιφόρνια στο Berkley (University of California at Berkley), με ενεργοποιημένη τη μέθοδο απόφασης για βελτιστοποίηση του ποσοστού παραμόρφωσης στα μακρο-μπλοκ (rate-distortion optimized macroblock mode decision), την εκτίμηση κίνησης (motion estimation) και προηγμένο έλεγχο του ρυθμού bit (advanced rate control module).

Για τη κωδικοποίηση με H.264 χρησιμοποιήθηκε GOP δομή με μέγεθος 15. Ο κωδικοποιητής ρυθμίστηκε να έχει πέντε πλαίσια για εσωτερική αναζήτηση (inter motion search), $\frac{1}{4}$ του pixel ανάλυση διανύσματος κίνησης (1/4 pel motion vector resolution), προσαρμοσμένη δυαδική κωδικοποίηση γενικευμένου πλαισίου (context-based adaptive binary coding, CABAC) για κωδικοποίηση συμβόλων και ενεργοποίηση της λήψης απόφασης βελτιστοποίησης του ποσοστού παραμόρφωσης (rate distortion optimized mode decision). Και οι δύο κωδικοποιητές χρησιμοποιούν πλήρη αναζήτηση υπολογισμού της κίνησης (full search motion estimation) με εύρος αναζήτησης (search range) 16x16.

Στις προσομοιώσεις με το H.264 η παράμετρος κβαντοποίησης (quantization parameter, QP) ρυθμίστηκε ίση με 26, 28, 31, 36 και 41 έτσι ώστε να καλυφθεί το φάσμα από χαμηλούς μέχρι μέτρια υψηλούς ρυθμούς bit. Για το MPEG-2 οι ρυθμοί bit επιλέχθηκαν έτσι ώστε η ποιότητα του κωδικοποιημένου βίντεο (PSNR τιμές) να είναι κοντά στα αντίστοιχα videos

του H.264. Τα videos του MPEG-2 δημιουργήθηκαν σε λειτουργία συνεχούς ρυθμού bit (constant bit rate mode, CBR).

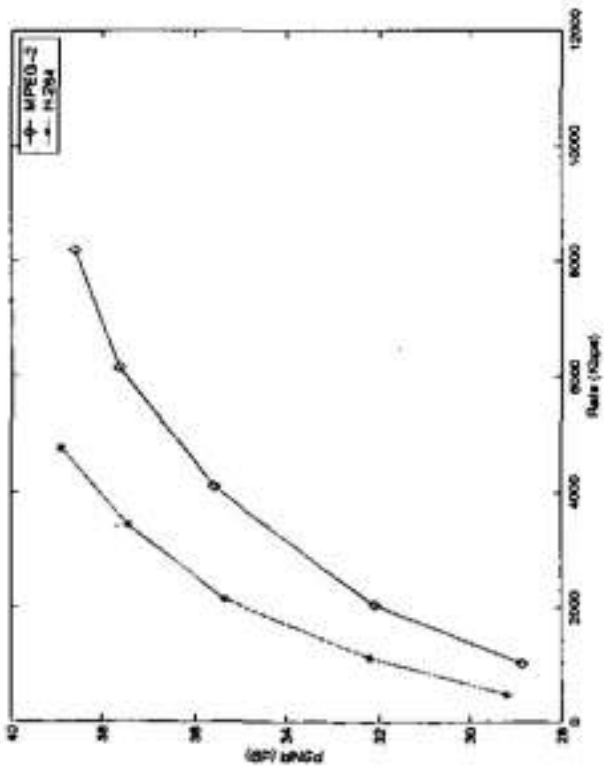
Για κάθε video το κέρδος κωδικοποίησης υπολογίστηκε από τη μέση τιμή των bits που εξοικονομούνται σε κάθε σημείο της προσομοίωσης στα διαγράμματα Ρυθμού bit / PSNR.

Video	Κέρδος κωδικοποίησης
Carousel	41%
Flower	55%
Football	52%
Μέσο κέρδος	49.3%

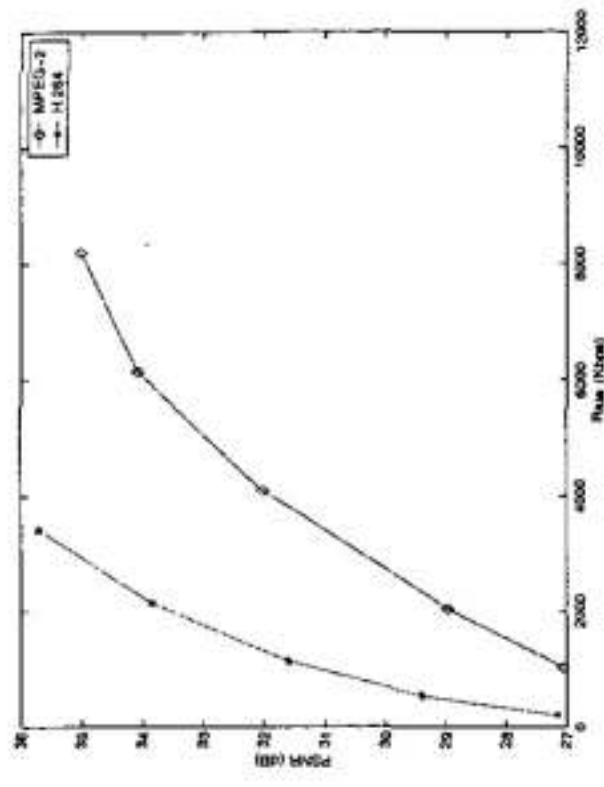
Πίνακας 13 - Κέρδος κωδικοποίησης του H.264 σε σχέση με το MPEG-2 για video CCIR

Video	Κέρδος κωδικοποίησης
Irene	53%
Mobile and Calendar	58%
Paris	49%
Students	45%
Tempete	41%
Μέσο κέρδος	49.2%

Πίνακας 14 - Κέρδος κωδικοποίησης του H.264 σε σχέση με το MPEG-2 για video CIF



Εικόνα 18 - Διάγραμμα ρυθμού bit / PSNR του H.264 και του MPEG-2 για το video Carousei



Εικόνα 19 - Διάγραμμα ρυθμού bit / PSNR του H.264 και του MPEG-2 για το video Flower

Όπως μπορούμε να δούμε με χρήση του προτύπου H.264 μπορούμε να έχουμε ένα μέσο κέρδος PSNR 50% σε σχέση με το MPEG-2 και μείωση του ρυθμού bit σχεδόν στο μισό σε σχέση με το MPEG-2.

3.3.3 Αποτελέσματα σύγκρισης του H.264 με το H.263 με βασικό προφίλ

Το H.263 χρησιμοποιείται συνήθως για χαμηλής και μέσης καθυστέρησης εφαρμογές όπως στη video-διάσκεψη. Στο πείραμα δεν χρησιμοποιήθηκαν εικόνες τύπου-B έτσι ώστε να καλύψουμε τις απαιτήσεις για χαμηλή καθυστέρηση.

Για την κωδικοποίηση με H.263 χρησιμοποιείται το codec H.263 της Telenor, με ενεργοποιημένη την προηγμένη πρόβλεψη (advanced prediction) και τη αριθμητική κωδικοποίηση βάσει της σύνταξης (syntax-based arithmetic coding options). Και για τις δύο κωδικοποιήσεις η πρώτη εικόνα κωδικοποιήθηκε σαν τύπου-I (I-type) και τα υπόλοιπα πλαίσια σαν τύπου-P (P-type).

Το H.264 ρυθμίστηκε να έχει πέντε πλαίσια για εσωτερική αναζήτηση κίνησης (inter motion search), $\frac{1}{4}$ του pixel ανάλυση διανύσματος κίνησης (1/4 pel motion vector resolution), προσαρμοσμένη δυαδική κωδικοποίηση γενικευμένου πλαισίου (context-based adaptive binary coding, CABAC) για κωδικοποίηση συμβόλων και ενεργοποιημένη τη λήψη απόφασης βελτιστοποίησης του ποσοστού παραμόρφωσης (rate distortion optimized mode decision).

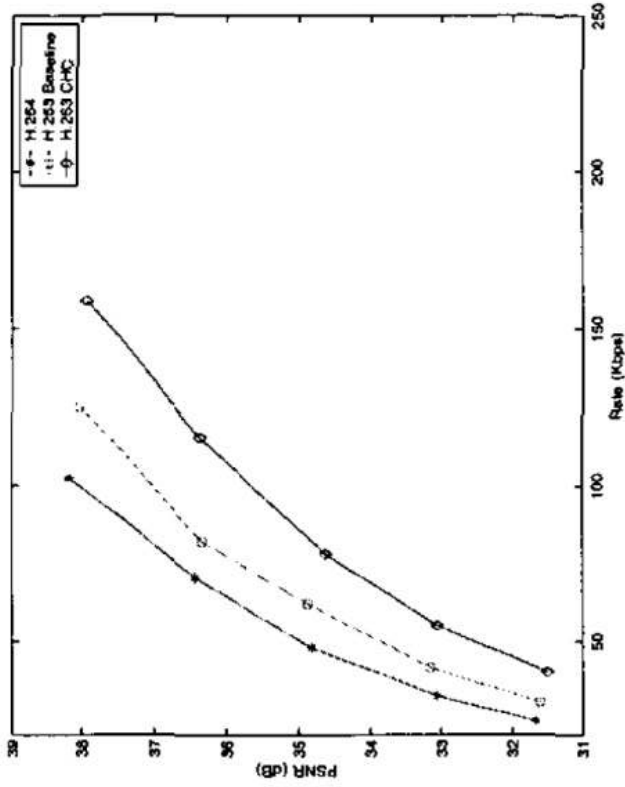
Τα σετ videos που επιλέχτηκαν για αυτή τη σύγκριση είναι το δεύτερο και το τρίτο. Ο σηματοθορυβικός λόγος και ο ρυθμός bit υπολογίστηκαν με χρήση διαφόρων τιμών του παράγοντα κβαντοποίησης: 21, 26, 31 και 41 για το H.264 και 8, 15, 20 και 33 για το H.263. Οι τιμές αυτές αντιπροσωπεύουν ένα εύρος από χαμηλούς μέχρι μέτρια υψηλούς ρυθμούς bit για εφαρμογές χαμηλής καθυστέρησης.

Video	Κέρδος κωδικοποίησης
Irene	53%
Mobile and Calendar	58%
Paris	42%
Students	45%
Tempete	41%
Μέσο κέρδος	49.2%

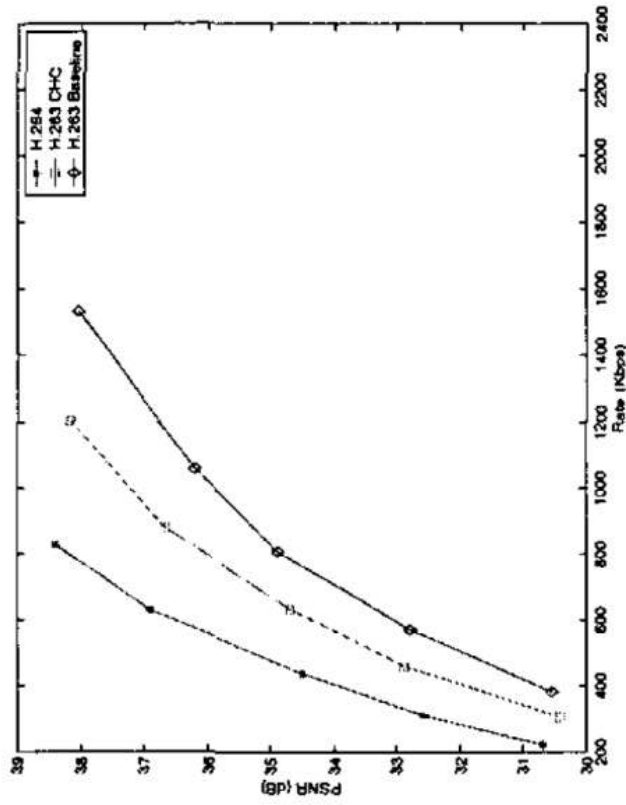
Πίνακας 15 - Κέρδος κωδικοποίησης του H.264 σε σχέση με το H.263 (βασικό προφίλ) για video CIF

Video	Κέρδος κωδικοποίησης
Akiyo	51%
Car Phone	40%
Claire	48%
Coast Guard	38%
Container	41%
Foreman	42%
Irene	50%
Miss America	44%
Mobile and Calendar	48%
News	56%
Paris	52%
Tempete	37%
Μέσο κέρδος	45.6%

Πίνακας 16 - Κέρδος κωδικοποίησης του H.264 σε σχέση με το H.263 (βασικό προφίλ) για video QCIF



Εικόνα 21 – Διάγραμμα ρυθμού bit / PSNR του H.264, του H.263 (βασικό προφίλ) και του H.263 (προφίλ υψηλής συμπίεσης) για το video Foreman QCIF



Εικόνα 20 – Διάγραμμα ρυθμού bit / PSNR του H.264, του H.263 (βασικό προφίλ) και του H.263 (προφίλ υψηλής συμπίεσης) για το video Paris CIF

Όπως παρατηρείται με τη χρήση του H.264 μπορούμε να έχουμε ένα μέσο κέρδος PSNR 49.2% για τα πέντε CIF video που ελέγχθηκαν στα 30fps και 45.6% για τα 12 QCIF στα 15fps, έναντι του H.263 στο βασικό προφίλ.

3.3.4 Αποτελέσματα σύγκρισης του H.264 με το H.263 με προφίλ υψηλής συμπίεσης

Εδώ συγκρίνεται η απόδοση του προτύπου H.264 με αυτή του H.263 με χρήση ενός προφίλ υψηλής συμπίεσης (Conversational High Compression, CHC).

Για την κωδικοποίηση με H.263 χρησιμοποιήθηκε το codec H.263 που αναπτύχθηκε από το Πανεπιστήμιο της Αγγλίας Columbia (University of British Columbia), με ενεργοποιημένη την προηγμένη πρόβλεψη (advanced prediction), τη αριθμητική κωδικοποίηση βάσει της σύνταξης (syntax-based arithmetic coding), προηγμένη εσωτερική κωδικοποίηση (advanced intra coding) και απεριόριστα διανύσματα κίνησης (unrestricted motion vector modes). Ο κωδικοποιητής ρυθμίστηκε ώστε να χρησιμοποιεί φίλτρο διαχωρισμού των μπλοκ (deblocking filter) και πέντε πλαίσια αναφοράς (reference frames) για εσωτερική κωδικοποίηση (inter coding).

Για την κωδικοποίηση με H.264 ο κωδικοποιητής ρυθμίστηκε να έχει πέντε πλαίσια για εσωτερική αναζήτηση κίνησης (inter motion search), $\frac{1}{4}$ του pixel ανάλυση διανύσματος κίνησης (1/4 pel motion vector resolution), προσαρμοσμένη δυαδική κωδικοποίηση γενικευμένου πλαισίου (context-based adaptive binary coding, CABAC) για κωδικοποίηση συμβόλων και ενεργοποίηση της λήψης απόφασης βελτιστοποίησης του ποσοστού παραμόρφωσης (rate distortion optimized mode decision).

Και στις δύο κωδικοποιήσεις χρησιμοποιήθηκε GOP δομή με μέγεθος 15. Ο σηματοθορυβικός λόγος (PSNR) και ο ρυθμός bit μετρήθηκε για διάφορες τιμές του παράγοντα κβαντοποίησης: 21, 26, 31 και 41 για το H.264 και 8, 15, 20 και 33 για το H.263 CHC. Τα σετ videos που χρησιμοποιήθηκαν είναι το δεύτερο και το τρίτο.

Video	Κέρδος κωδικοποίησης
Irene	29%
Mobile and Calendar	31%
Paris	22%
Students	24%
Tempete	20%
Μέσο κέρδος	25.4%

Πίνακας 17 - Κέρδος κωδικοποίησης του H.264 σε σχέση με το H.263 (προφίλ υψηλής συμπίεσης) για video CIF

Video	Κέρδος κωδικοποίησης
Akiyo	25%
Car Phone	20%
Claire	34%
Coast Guard	19%
Container	22%
Foreman	20%
Irene	26%
Miss America	25%
Mobile and Calendar	22%
News	28%
Paris	22%
Tempete	17%
Μέσο κέρδος	23.3%

Πίνακας 18 - Κέρδος κωδικοποίησης του H.264 σε σχέση με το H.263 (προφίλ υψηλής συμπίεσης) για video QCIF

Όπως παρατηρείται το πρότυπο H.264 πετυχαίνει ένα μέσο κέρδος 25.4% έναντι του H.263 για τα CIF videos στα 30fps και 23.3% για τις 12 QCIF videos στα 15fps.

3.4 Σύγκριση του προτύπου H.264, του M-JPEG2000 και του PACC

Στο [11] γίνεται σύγκριση μεταξύ των προτύπων κωδικοποίησης H.264 με το Motion-JPEG2000 και το PACC, από όπου προκύπτουν ορισμένα πολύ ενδιαφέροντα συμπεράσματα τα οποία αναφέρονται στη συνέχεια του κεφαλαίου.

Το πρότυπο Motion-JPEG2000, σαν μέρος της σουίτας JPEG2000 επεκτείνει τη χρήση του JPEG2000 σε εφαρμογές που χρησιμοποιούν κατεχοχόν κωδικοποιητές video όπως το H.264/AVC. Εδώ θα ελέγξουμε την απόδοση του Motion-JPEG2000 και του H.264/AVC όσο αναφορά την κωδικοποίηση βίντεο, δηλαδή σε εφαρμογές που απαιτούν τυχαία προσπέλαση σε ανεξάρτητες εικόνες.

Ορισμένες εφαρμογές όπως η επεξεργασία ιατρικού βίντεο ή η επαγγελματική παραγωγή βίντεο, απαιτούν ένα εργαλείο κωδικοποίησης το οποίο να επιτρέπει την τυχαία προσπέλαση σε ανεξάρτητες εικόνες. Κάποια επιπλέον χαρακτηριστικά, τα οποία μπορεί να είναι πολύ σημαντικά για κάποιες εφαρμογές είναι η χωρική επεκτασιμότητα (spatial scalability), η επεκτασιμότητα του σήματος προς θόρυβο (SNR scalability) και η κωδικοποίηση περιοχών ενδιαφέροντος (region-of-interest coding).

3.4.1 PACC

Ο αλγόριθμος κωδικοποίησης PACC είναι ένας εναλλακτικός κωδικοποιητής εικόνας βάσει κυματομορφών (wave-let based images), ο οποίος παρ' όλο που εκδόθηκε το 1997 μπορεί να θεωρηθεί σαν ένας από τους πιο αντιπροσωπευτικούς της τεχνολογίας των κωδικοποιητών εικόνας βάσει κυματομορφής. Έχει εφαρμοστεί επιτυχώς σε κωδικοποιήσεις βίντεο καθώς επίσης έχει συνδυάσει απωλεστική / μη-απωλεστική (lossy / lossless) κωδικοποίηση στατικών εικόνων.

Το PACC είναι μία διαδικασία κωδικοποίησης τριών σταδίων. Τα τρία βήματα της διαδικασίας είναι: (1) τεμαχισμός (partitioning) (2) συσσωμάτωση (aggregation) και (3) κωδικοποίηση υπό συνθήκες (conditional coding).

Η αρχική διαδικασία τεμαχισμού χωρίζει την κβαντισμένη απεικόνιση της κυματομορφής (wavelet) σε τρεις υπό-πηγές (sub-sources) ανάλογα με τη σπουδαιότητα, το μέγεθος και το σήμα του κάθε κβαντισμένου συντελεστή κυματομορφής.

Στη δεύτερη προαιρετική διαδικασία γίνεται ενσωμάτωση των ασήμαντων συντελεστών σε διαφορετικές κλίμακες (επίπεδα ανάλυσης, resolution levels) με χρήση του γνωστού εργαλείου μηδενικών δέντρων (instrument of zerotrees).

Στη τελική διαδικασία αντιστοιχίζεται ένα κατάλληλο μοντέλο γενικευμένου πλαισίου (context) για αυτή κάθε αυτή την κωδικοποίηση, σε κάθε ένα στοιχείο των τριών υπό-πηγών. Για κάθε μία από τις τρεις υπό-πηγές έχουν αναπτυχθεί εμπειρικά κάποια πρότυπα βελτιστοποίησης (optimized prototype templates).

Σημειωτέον ότι σε αντίθεση με το JPEG2000, το PACC δεν λειτουργεί σε επίπεδο κβαντισμένων τιμών bits (bitplanes of quantizer indices) και έτσι οι συνεχείς ροές bits (bitstreams) που προκύπτουν δεν είναι ταξινομημένες με έναν ενσωματωμένο τρόπο ο οποίος ενεργοποιεί την πρόοδο του SNR, ωστόσο υποστηρίζει προοδευτική μετάδοση (progressive transmission) με χωρική ανάλυση και φασματικές συνιστώσες (spatial resolution and spectral component).

3.4.2 Πειραματική διαδικασία

Για να γίνει όσο το δυνατόν πιο σωστή και δίκαιη σύγκριση μεταξύ του Motion-JPEG2000 και του H.264/AVC, στο δεύτερο χρησιμοποιήθηκε η ρύθμιση καθαρής εσωτερικής κωδικοποίησης (pure intra coding mode). Για τη συλλογή επιπλέον πληροφοριών σχετικά με τη δυνατότητα κωδικοποίησης εικόνων κυματομορφής (wavelet-based images) γίνεται σύγκριση του codec PACC με τα δύο προηγούμενα. Τα αποτελέσματα προήλθαν από ένα πλήθος video, το οποίο καλύπτει διάφορες αναλύσεις και άλλα χαρακτηριστικά, όπως υπόστρωμα σταδιακού συγχρονισμού (underlying progressive timing) η πεπλεγμένου συγχρονισμού (interlaced timing) της αρχικής εικόνας.

Στο πείραμα εξετάζονται τέσσερις περιπτώσεις, κάθε μία για διαφορετική ανάλυση video. Η πρώτη αναφέρεται σε χαμηλής ανάλυσης CIF video (352x288 pixels), η οποία χρησιμοποιείται συνήθως σε video-διάσκεψη και εφαρμογές συνεχούς ροής video (video streaming).

Στη δεύτερη περίπτωση συγκρίνεται η απόδοση όσο αναφορά το ρυθμό παραμόρφωσης (Rate Distortion) για πεπλεγμένη σάρωση πρότυπης ευκρίνειας video (interlaced-scan standard definition television (ITU-R 601) video sequences) σε ανάλυση 720x576 pixels (25Hz).

Στη τρίτη περίπτωση συγκρίνεται πάλι η απόδοση όσο αναφορά το ρυθμό παραμόρφωσης στα 60Hz για σταδιακή σάρωση υψηλής ευκρίνειας (progressive-scan high-definition, HD) σε ανάλυση 1280x720 pixels (720p).

Τέλος στην τέταρτη περίπτωση τα videos είναι ακόμα υψηλότερης ανάλυσης 1920x720 pixels σταδιακά σαρωμένα (progressively scanned) με συχνότητα 25Hz.

Στο **Πίνακα 19** συνοψίζονται τα δεδομένα εισαγωγής για όλες τις περιπτώσεις. Σημειωτέον ότι όλες οι σκηνές είναι σε YUV 4:2:0 χρωματικό (color) format, όπου τα δύο χρωματικά

στοιχεία (chrominance components) (U, V) έχουν δειγματοληφθεί με συντελεστή 2 διαστάσεων (down-sampled by factor of two in each spatial dimension)

Video	Ανάλυση	Αριθμός πλαισίων	Χαρακτηριστικά περιεχομένου
Paris	CIF	100	Μέτρια χωρική και χρωματική λεπτομέρεια
Bus	CIF	260	Μέτρια χωρική λεπτομέρεια
Mobile and Calendar	720x576i	100	Υψηλή χωρική και χρωματική λεπτομέρεια
Canoe	720x576i	100	Μέτρια προς υψηλή χωρική και χρωματική λεπτομέρεια
Crew	1280x720p	300	Μέτρια χωρική λεπτομέρεια
Harbour	1280x720p	300	Υψηλή χωρική λεπτομέρεια
Vintage Car	1920x1080p	250	Μέτρια προς υψηλή χωρική λεπτομέρεια με θόρυβο
Book	1920x1080p	220	Μέτρια χωρική λεπτομέρεια με θόρυβο

Πίνακας 19 - Videos που χρησιμοποιήθηκαν για τον έλεγχο του ρυθμού παραμόρφωσης (RD)

Σαν αντικειμενική μέτρηση της απόδοσης επιλέχθηκε ο σηματοθορυβικός λόγος των φωτεινών στοιχείων (luminance component) σε όλα τα πλαίσια σε ένα video, για την απεικόνιση της οπτικής παραμόρφωσης. Ο σηματοθορυβικός λόγος των φωτεινών στοιχείων (U, V) έχει ρυθμιστεί για τους διάφορους κωδικοποιητές, έτσι ώστε οι γραφικές παραστάσεις R-D που απεικονίζουν την μέση παραμόρφωση (U, V) σε σχέση με τον μέσο συνολικό ρυθμό bit, να είναι σχεδόν όμοιες και για τους τρεις κωδικοποιητές, με κάποια όρια ανοχής. Αυτό πραγματοποιήθηκε βγάζοντας πρώτα τη γραφική παράσταση R-D του Motion-JPEG2000 κωδικοποιητή και εν συνεχεία η παράμετρος κβαντοποίησης άλλαζε επαναληπτικά τόσο για την κωδικοποίηση με H.264/AVC τόσο και με PACC, μέχρι οι γραφικές R-D των τριών κωδικοποιητών να είναι όσο το δυνατό πιο όμοιες.

Απαιτώντας να είναι σχεδόν ίδιες οι γραφικές παραστάσεις R-D είναι επαρκές, για κάθε τεστ και video, να παρουσιάζεται μόνο μία γραφική R-D που δείχνει το μέσο σηματοθορυβικό λόγο (luminance PSNR) συναρτήσε του μέσου ρυθμού bit.

3.4.2.1 Ρυθμίσεις του Motion-JPEG2000

Χρησιμοποιήθηκε το μοντέλο (verification model, VM) του Motion-JPEG2000, έκδοση 2.1 αναπτυγμένο πάνω στην έκδοση 8.6 του JPEG2000 VM. Οι ρυθμίσεις που χρησιμοποιήθηκαν για την μεγιστοποίηση της R-D απόδοσης του JPEG2000 Part 1 είναι:

- Ένα στοιχείο (tile) ανά εικόνα (no tiling)
- Πέντε επίπεδα διάσπασης κυματομορφής (wavelet decomposition)
- Βιο-ορθογώνιο φίλτρο κυματομορφής πυρήνα με βήμα 9/7 (9/7-tap default biorthogonal wavelet filter kernel)
- Σαφή κβαντισμό με έλεγχο μεγέθους βήματος (explicit quantization by step size control)
- Μέγεθος κωδικού-μπλοκ (codeblock) 64x64 δείγματα (samples)

- Κωδικοποίηση αλληλουχίας (Sequential coding mode) και κωδικοποίηση πλήρους επίδοσης (full effort coding mode) (απενεργοποιημένη η παράλληλη κωδικοποίηση (parallel coding mode) και το lazy coding)
- Ένα στρώμα ποιότητας (one single quality layer)

Επιπλέον για πιο δίκαιη σύγκριση της R-D απόδοσης του Motion-JPEG2000 ένα μεγάλο μέρος των δεδομένων επικεφαλίδας (header data) (συνήθως περισσότερα από 170 bytes ανά εικόνα) έχει απορριφθεί για τον υπολογισμό του ρυθμού bit.

3.4.2.2 Ρυθμίσεις του H.264/AVC

Για την κωδικοποίηση των δεδομένων της πηγής σταδιακής σάρωσης (progressive-scan source) χρησιμοποιήθηκε το Joint Model (JM) της Joint Video Team (JVT), έκδοση 7.1, ενώ για την κωδικοποίηση του πεπλεγμένου (interlaced) υλικού μία εφαρμογή του Main Profile H.264/AVC, η οποία αναπτύχθηκε από το Fraunhofer-Inst. HHI. Και οι δύο κωδικοποιητές ακολουθούσαν την ίδια Lagrangian στρατηγική για την βελτιστοποίηση του ρυθμού παραμόρφωσης (όπως αναφέρετε στο [25]). Επίσης χρησιμοποιήθηκε κωδικοποίηση CABAC. Τέλος επιτρεπόταν μόνο ένα κομμάτι (slice) ανά εικόνα, ενώ για το πεπλεγμένο υλικό επιλέχθηκε η κωδικοποίηση μακρο-μπλοκ προσαρμοσμένη στο πλαίσιο/πεδίο (macroblock adaptive frame/field (MBAFF) coding mode).

3.4.2.3 Ρυθμίσεις του PACC

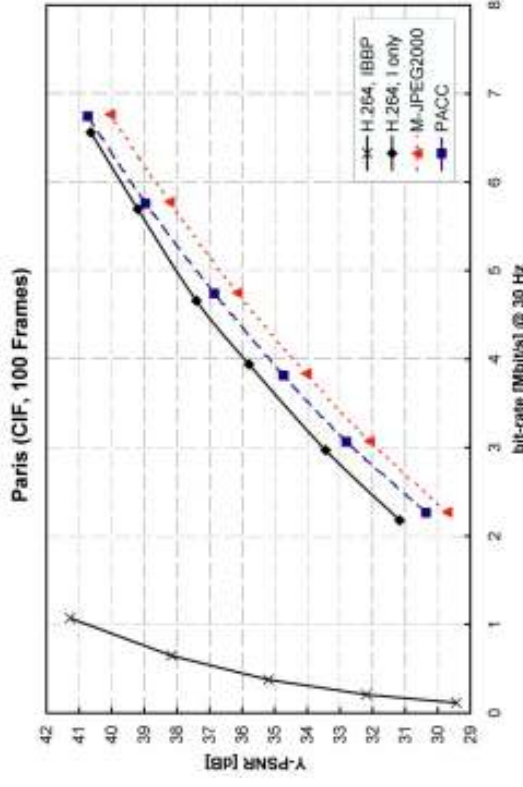
Για τον έλεγχο του PACC χρησιμοποιήθηκε η έκδοση που χρησιμοποιήθηκε και στις προσομοιώσεις όπως αναφέρονται στο [26]. Η εφαρμογή αυτή περιλαμβάνει διακριτή διάσπαση κυματομορφής 4 επιπέδων (four-level standard discrete wavelet decomposition) με χρήση φίλτρου με βήμα 9/7 (9/7-tap filter). Τέλος λόγω έλλειψης υποστήριξης πολύπλεξης (interlace) το PACC εφαρμόστηκε μόνο στα δεδομένα σταδιακής σάρωσης (progressive-scan source).

3.4.3 Αποτελέσματα σύγκρισης

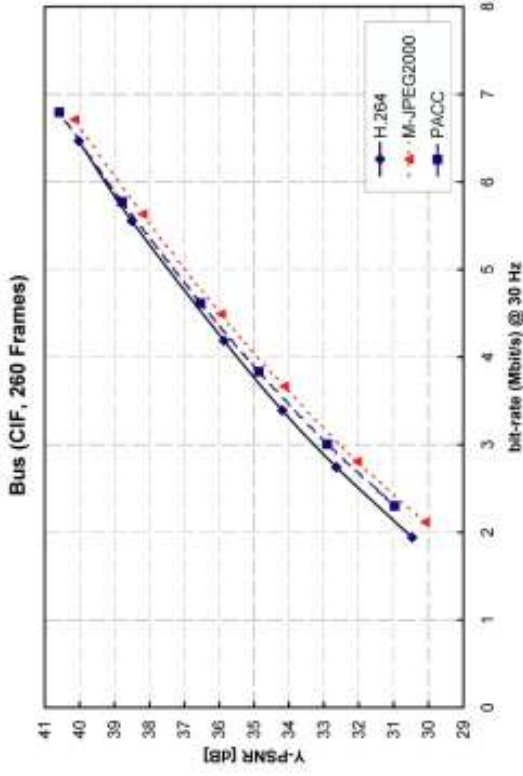
Στις **Εικόνες 22 μέχρι και 25** παρουσιάζονται οι γραφικές παραστάσεις R-D για δύο videos της πρώτης (CIF) και της δεύτερης περίπτωσης (ITU-R 601, 720x576i).

Ως ένδειξη του μεγάλου κενού παρατήρησης μεταξύ ενός εσωτερικού κωδικοποιητή (pure intra coder) και ενός codec πλήρους λειτουργίας (full operating video codec), ο οποίος χρησιμοποιεί όλα τα είδη των χρονικών προβλέψεων (temporal prediction modes) (σε P και B εικόνες) το R-D διάγραμμα για την εικόνα "Paris" περιλαμβάνει επίσης τη γραφική R-D όπως προέκυψε από το H.264/AVC με χρήση μόνο μίας I-εικόνας στην αρχή της ακολουθίας, ακολουθούμενη από δύο B-εικόνες, οι οποίες εισήχθησαν ανάμεσα σε κάθε P-εικόνα.

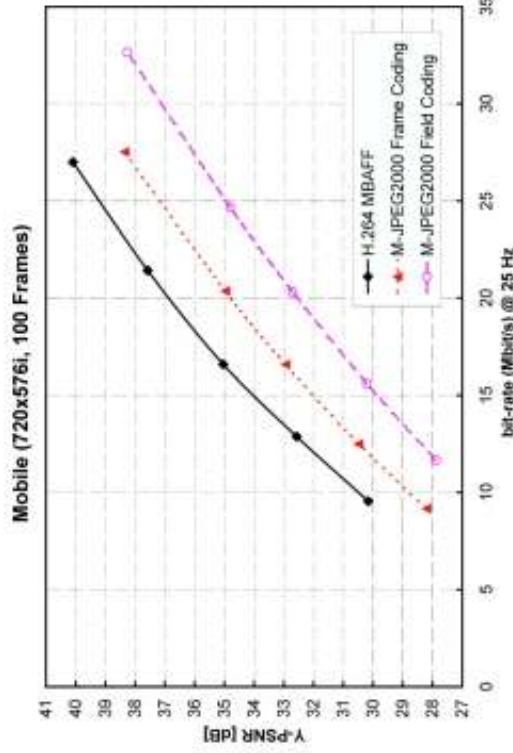
Παρατηρείται μία αύξηση του ρυθμού bit της τάξης 7/10 (by a factor of 7-10) για τον εσωτερικό κωδικοποιητή (intra coder) σε αυτό το τεστ, μιας τυπικής video-διάσκεψης, το οποίο είναι υψηλό τίμημα στην απόδοση της κωδικοποίησης, δεδομένης της τυχαίας προσπέλασης των πλαισίων.



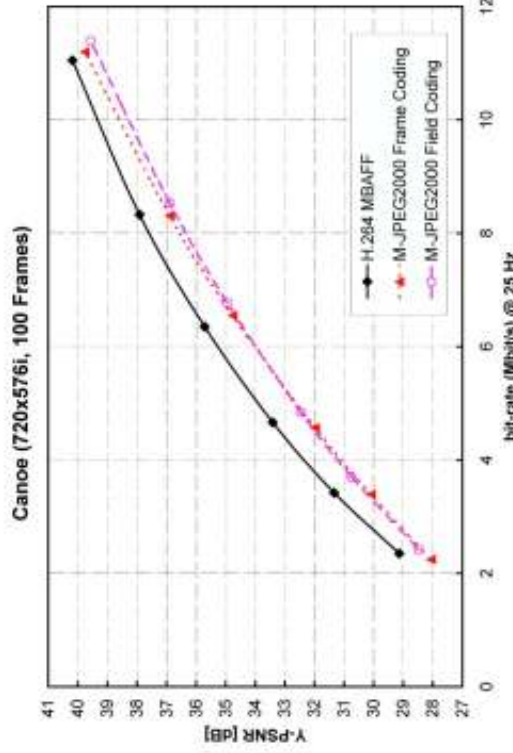
Εικόνα 22 - Διάγραμμα ρυθμού bit / PSNR των H.264, MJPEG2000 και PACC για το video Paris



Εικόνα 23 - Διάγραμμα ρυθμού bit / PSNR των H.264, MJPEG2000 και PACC για το video Bus



Εικόνα 24 - Διάγραμμα ρυθμού bit / PSNR των H.264, MJPEG2000 και PACC για το video Mobile



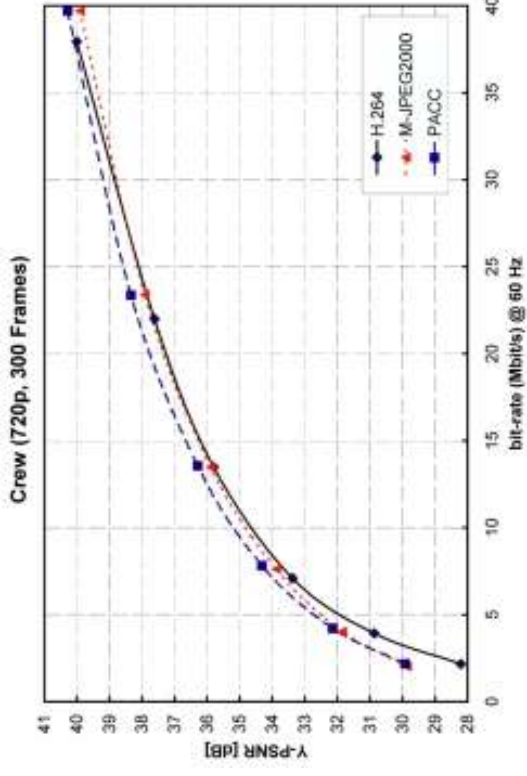
Εικόνα 25 - Διάγραμμα ρυθμού bit / PSNR των H.264, MJPEG2000 και PACC για το video Canoe

Με βάση ότι η υποστήριξη μιας τέτοιας δυνατότητας αξίζει την απώλεια R-D απόδοσης, το πείραμα δείχνει ότι το H.264/AVC αποδίδει καλύτερα από το Motion-JPEG2000 στην περίπτωση της εσωτερικής κωδικοποίησης (intra coding) στα πρώτα δύο τεστ. Για τα videos ανάλυσης CIF παρατηρείται ένα κέρδος PSNR περίπου 0.5 – 1.5 dB υπέρ του H.264/AVC, ενώ το PACC δείχνει μία απόδοση περίπου ανάμεσα από τα δύο άλλα codecs.

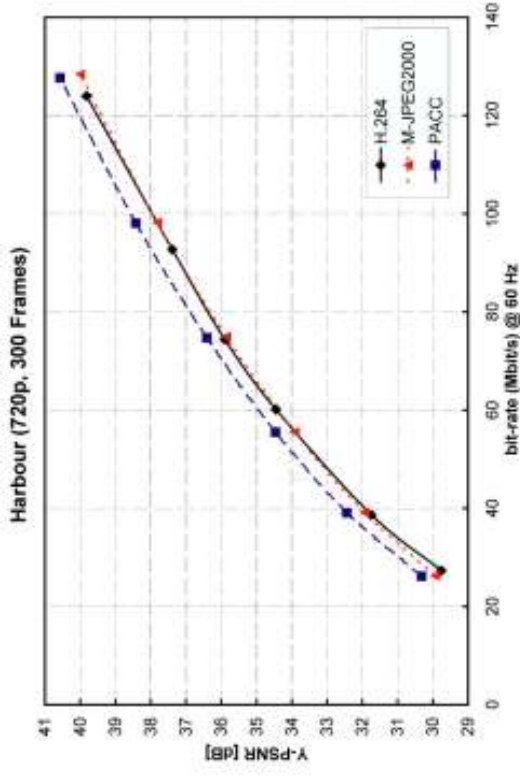
Στην περίπτωση του ITU-R 601 πεπλεγμένου υλικού, έχουμε δύο γραφικές R-D για το Motion-JPEG2000: η μία για κωδικοποίηση κάθε εικόνας σαν ένα ολόκληρο πλαίσιο (frame coding) και η άλλη γραφική για κωδικοποίηση των δύο περιοχών (fields) της εικόνας ξεχωριστά (field coding). Όπως φαίνεται και από τις γραφικές παραστάσεις η R-D απόδοση και τον δύο codecs εξαρτάται από το είδος του υλικού. Για το “Mobile and Calendar” μία εικόνα έχει μεγάλη χωρική πυκνότητα κατά τις διπλανές σειρές, η οποία συλλαμβάνεται (captured) σε διαφορετικές χρονικές στιγμές. Καθώς σε αυτή την ακολουθία η κάμερα κάνει zoom αργά τα κινούμενα στοιχεία καταλαμβάνουν ένα μικρό κομμάτι της κάθε εικόνας. Αυτό, παρόλο που είναι αντίθετο με τα χαρακτηριστικά του “Canoe” το οποίο περιλαμβάνει μία γρήγορη μετακίνηση (pan) της κάμερας με γρήγορα κινούμενα στοιχεία στη σκηνή, οδηγεί σε μείωση της εξάρτησης μεταξύ δύο διπλανών γραμμών που ανήκουν στην ίδια περιοχή.

Στη τελευταία περίπτωση η κωδικοποίηση βάσει περιοχής (field-based coding) είναι καλύτερη ενώ στις άλλες περιπτώσεις η κωδικοποίηση βάσει πλαισίου (frame-based coding) οδηγεί σε καλύτερα αποτελέσματα. Αυτό φαίνεται και καθαρά από τις γραφικές παραστάσεις. Παρόλα αυτά πρέπει να αναφερθεί ότι η κωδικοποίηση βάσει πλαισίου για το “Canoe” αποδίδει υποκειμενικά χειρότερα από ότι η κωδικοποίηση βάσει περιοχής, παρ’ όλο που οι γραφικές τους είναι αρκετά ίδιες. Η κωδικοποίηση προσαρμοσμένου πλαισίου/περιοχής (Picture-adaptive frame/field (PAFF) coding) ίσως θα μπορούσε να βελτιώσει την R-D απόδοση του Motion-JPEG2000 αλλά αυτή η επιλογή δεν υποστηρίζεται από το Motion-JPEG2000 VM 2.1.

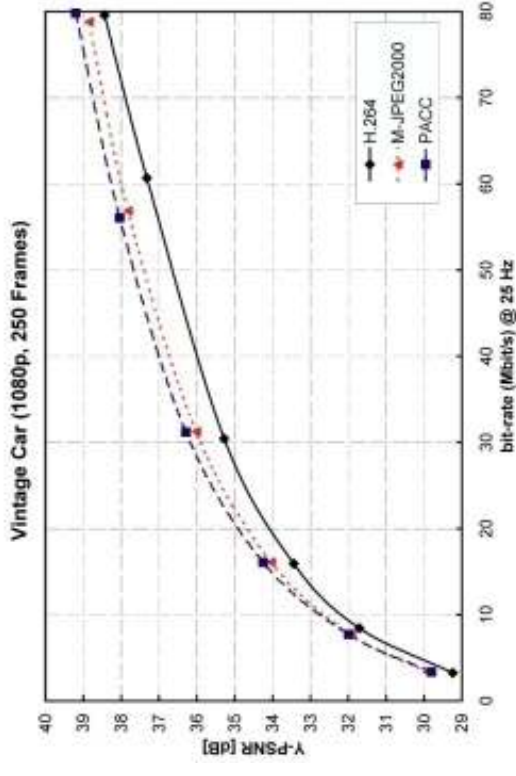
Το H.264/AVC, MBAFF συνήθως παρέχει τον πιο αποδοτικό τρόπο κωδικοποίησης του πεπλεγμένου υλικού. Η μέθοδος αυτή της τοπικής κωδικοποίησης προσαρμοσμένης σε πλαίσιο/περιοχή (frame/field) είναι ένα μοναδικό χαρακτηριστικό του H.264/AVC και σκοπό έχει να μειώσει το ρυθμό bit μέχρι και 15% σε σχέση με το PAFF, ειδικά σε υλικό με ανάμικτες περιοχές. Αν συγκρίνουμε το H.264/AVC MBAFF με την καλύτερη κωδικοποίηση του Motion-JPEG2000 παρατηρείται ένα σημαντικό κέρδος PSNR 0.5 – 2 db, υπέρ του H.264/AVC όπως φαίνεται και από τις γραφικές.



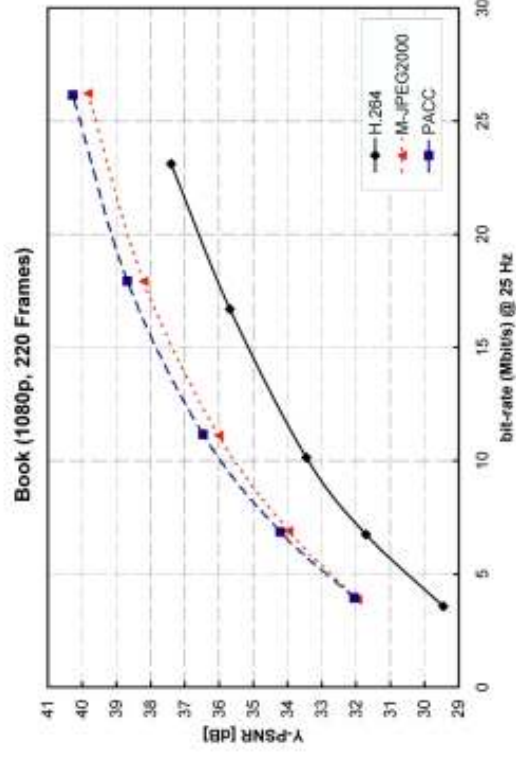
Εικόνα 26 - Διάγραμμα ρυθμού bit / PSNR των H.264, MJPEG2000 και PACC για το video Crew



Εικόνα 27 - Διάγραμμα ρυθμού bit / PSNR των H.264, MJPEG2000 και PACC για το video Harbour



Εικόνα 28 - Διάγραμμα ρυθμού bit / PSNR των H.264, MJPEG2000 και PACC για το video Vintage Car



Εικόνα 29 - Διάγραμμα ρυθμού bit / PSNR των H.264, MJPEG2000 και PACC για το video Book

Ένα καθαρό πλεονέκτημα του Motion-JPEG2000 έναντι του H.264/AVC μπορεί να παρατηρηθεί στα παραπάνω διαγράμματα για την περίπτωση των 1080p. Σε τόσο υψηλές αναλύσεις το Motion-JPEG2000 πετυχαίνει ένα κέρδος 0.5 – 2 dB σε σχέση με το H.264/AVC. Επιπλέον κέρδος περίπου 0.5 dB παρατηρείται για το PACC codec, ειδικά στα υψηλά bit-rates.

Η υπεροχή των κωδικοποιητών βάσει κυματομορφής (wavelet-based) όσο αναφορά την R-D απόδοση για αυτού του είδους το υλικό, μπορεί να αποδοθεί σε δύο φαινόμενα. Πρώτον, η συσχέτιση των εσωτερικών pixels (inter-pixel) αυξάνεται συνήθως με την αύξηση της ανάλυσης, τόσο ώστε οι καλύτερες ιδιότητες αποσυσχέτισης (decorrelating) του μετασχηματισμού κυματομορφής σε μεγάλες λείες επιφάνειες μπορεί να είναι πολύ αποτελεσματικές. Δεύτερον, το βήμα χωρικής πρόβλεψης (spatial prediction step) στον κωδικοποιητή H.264/AVC μπορεί να επηρεαστεί από το ίχνος θορύβου του film (film grain noise), ο οποίος συνήθως παρατηρείται στο 1080p και ο οποίος στις περιπτώσεις των κωδικοποιητών βάσει κυματομορφής (wavelet-based coders) διαχειρίζεται καλύτερα στα χαμηλότερης σημασίας πεδία των bits (bitplanes).

3.4.4 Πολυπλοκότητα, Λειτουργικότητα και Εφαρμογές

Σε αυτό το σημείο θα εξεταστεί η απόδοση από πλευράς της πολυπλοκότητας των συγκεκριμένων αλγορίθμων. Επίσης προσδιορίζονται κάποιες συγκεκριμένες λειτουργίες που παρέχονται από αυτά τα τρία codecs πέραν της συμπίεσης και τέλος πως μπορούν να χρησιμοποιηθούν σε εφαρμογές αυτές οι λειτουργίες.

Ο υπολογισμός της πολυπλοκότητας σε όρους εύρους/προσπέλασης μνήμης (memory bandwidth/access) και/ή υπολογιστικού κόστους είναι μία πολύ δύσκολη διαδικασία, καθώς εξαρτάται σε μεγάλο βαθμό από τον τρόπο υλοποίησης του πειράματος. Από την οπτική μεριά του αποκωδικοποιητή (decoder centric point of view) ο σημαντικότερος παράγοντας για την επιβράδυνση της διαδικασίας (bottleneck) και στους τρεις αλγορίθμους φαίνεται να βρίσκεται στη μονάδα κωδικοποίησης (entropy encoding unit), λόγω της έμφυτης διαδοχικά οργανωμένης αριθμητικής (sequentially organized arithmetic) διαδικασίας αποκωδικοποίησης. Σε μελέτες σχετικές με υλοποιήσεις υλικού (hardware-based implementations) του JPEG2000 προέκυψε το συμπέρασμα ότι ο αποκωδικοποιητής διαζώματος 1 (T1 entropy decoder) βασίζεται περισσότερο από 60% στην πολυπλοκότητα του υλικού (hardware) παρά στην υλοποίηση του αποκωδικοποιητή. Στις περιπτώσεις των κωδικοποιήσεων που δεν προσανατολίζονται από το επίπεδο των bits (non-bitplane oriented entropy coding) όπως των CABAC και PACC το πρόβλημα αυτό είναι μικρότερο, όχι τόσο όμως ώστε να μπορεί να αγνοηθεί. Ένας άλλος σημαντικός παράγοντας της πολυπλοκότητας είναι το μέγεθος της μεταφερόμενης μνήμης (amount of memory transfer), το οποίο μπορούμε να πούμε ότι έχει μικρότερη σημασία στην υβριδική κωδικοποίηση βάσει του μπλοκ (block-based hybrid coder) του H.264/AVC.

Πρέπει να αναφερθεί ότι τα codecs έχουν συνήθως ένα σετ εργαλείων το οποίο μπορεί να βάλει κάποιους περιορισμούς στους παράγοντες κωδικοποίησης έτσι ώστε να έχουμε μία κλίμακα πολυπλοκότητας. Κάποια παραδείγματα αυτών των παραμέτρων είναι η επιλογή του πυρήνα κυματομορφής (wavelet kernel), το μέγεθος του στοιχείου (tile) ή του κωδικού-μπλοκ (codeblock) στο JPEG2000 ή αντίστοιχα το μέγεθος του κομματιού (slice) ή ο αριθμός των τρόπων πρόβλεψης (prediction modes) στο H.264/AVC. Ωστόσο πρέπει να αναφερθεί ότι η εισαγωγή αυτών των περιορισμών έχει συνήθως αρνητική επίπτωση στην R-D απόδοση.

Πέραν της απλής συμπίεσης, οι αλγόριθμοι κωδικοποίησης για εικόνα και video πρέπει να υποστηρίζουν και κάποιες άλλες λειτουργίες που απαιτούνται από κάποιες εφαρμογές. Εδώ πρέπει να γίνει και ένας διαχωρισμός μεταξύ των βασικών λειτουργιών (basic functionalities) και των προαιρετικών λειτουργιών (optional value-added features). Το πιο σημαντικό βασικό τεχνικό χαρακτηριστικό που πρέπει να παρέχεται από όλα τα codecs πολυμέσων είναι μία ευέλικτη διεπαφή (flexible interface) όσο αναφορά τα μέσα αποθήκευσης (storage media) και τα επίπεδα μετάδοσης (transmission layers). Για το σκοπό αυτό το Motion-JPEG2000 έχει το δικό του MJ2 file format, το οποίο προέρχεται από το JPEG2000 JP2 file format και είναι στενά συνδεδεμένο με το MPEG-4 MP4 file format. Σε σύγκριση με αυτό το H.264/AVC προσφέρει ένα πιο ευέλικτο σχεδιασμό το λεγόμενο στρώμα επιπέδου NAL, το οποίο προσαρμόζει το επίπεδο κωδικοποίησης video (video coding layer).

Ένα άλλο σημαντικό χαρακτηριστικό που πρέπει να παρέχεται από τους αλγορίθμους κωδικοποίησης είναι η ευρωστία (robustness) στη απώλεια και στα λάθη δεδομένων (data errors/losses). Το H.264/AVC παρέχει μία πληθώρα εργαλείων για αυτό το σκοπό. Το Motion-JPEG2000 παρέχει ένα δύο μεθόδους κωδικοποίησης (coding modes) για τη βελτίωση της αντιμετώπισης των λαθών, τα οποία προέρχονται από το JPEG2000.

Όσο αναφορά τα προαιρετικά (value-added) χαρακτηριστικά, η επεκτασιμότητα (scalability) είναι το πιο σημαντικό χαρακτηριστικό του Motion-JPEG2000. Συνήθως με αυτό το χαρακτηριστικό αναφερόμαστε στην δυνατότητα να εξάγει διαφορετικές αναλύσεις (resolutions), ποιότητες (fidelities), στοιχεία (components) ή χωρικές τοποθεσίες (spatial locations) από μία συμπιεσμένη συνεχή ροή bits. Αυτή είναι μία μοναδική λειτουργία του JPEG2000 / Motion-JPEG2000 που δεν έχει αντίστοιχη το H.264/AVC πέρα της ευελιξίας της διάταξης των μακρο-μπλοκ (flexible macroblock ordering, FMO) ή οποία επιτρέπει ένα είδος υποστήριξης περιοχών ενδιαφέροντος (region-of-interest support). Η επεκτασιμότητα έχει και επιπτώσεις στην R-D απόδοση. Για παράδειγμα, ο ρυθμός bit ή η επεκτασιμότητα SNR (SNR-scalability) μπορεί να προκαλέσει πτώση του PSNR μέχρι και 0.5 dB σε σύγκριση με την περίπτωση του μονού επιπέδου (single layer) που ελέγχθηκε στο παρόν πείραμα. Μια άλλη δυνατότητα που παρέχεται από το Motion-JPEG2000 είναι η δυνατότητα παραγωγής μίας μόνο ροής bits (single bitstream) από την οποία μπορούν να εξαχθούν και οι μη-απωλεστικές αναπαραστάσεις (lossless representations) και ένας αριθμός απωλεστικών αναπαραστάσεων (lossy representations). Αλλά και στην περίπτωση αυτή υπάρχει ένα κόστος το οποίο είναι η σημαντική απώλεια στην R-D απόδοση 0.5 – 1 dB PSNR για τις απωλεστικές ανακατασκευές (lossy reconstructions), λόγω του λιγότερο αποδοτικού φίλτρου αντίστροφης κυματομορφής με βήμα 5/3 (5/3-trap reversible wavelet filter).

Το εύρος των εφαρμογών στο οποίο μπορούν να συναγωνιστούν το Motion-JPEG2000 και το H.264/AVC είναι πολύ μεγάλο. Στο μέλλον μπορεί να υπάρξει μία ομάδα εφαρμογών όπου το codec θα είναι ενσωματωμένο (built-in) για συγκεκριμένους σκοπούς. Για παράδειγμα, ψηφιακές κάμερες ήπιας λήψης (digital still cameras) με ρύθμιση για καταγραφή έκρηξης κίνησης (burst capturing mode) κατάλληλο για το JPEG2000/Motion-JPEG2000 ή ψηφιακές φορητές κάμερες (digital camcorders) με ρύθμιση λήψης ήπιας κίνησης (still image capturing mode) καταλληλότερο για το H.264/AVC. Η λήψη video από υπολογιστή (PC-based video capturing) είναι μια άλλη δυνατότητα των μελλοντικών εφαρμογών αν και εκεί μπορεί να έχουμε μεγάλο ανταγωνισμό μεταξύ των codecs.

Κρίνοντας από τα αποτελέσματα της R-D απόδοσης και των χαρακτηριστικών συμπεραίνουμε ότι το Motion-JPEG2000 είναι καταλληλότερο για υψηλής ποιότητας και ανάλυσης κωδικοποίησης video. Επίσης θα μπορούσε να εφαρμοστεί και στο πεδίο της

υψηλής ποιότητας ιατρικών και δορυφορικών εικόνων (high-resolution medical or satellite imaging).

Το H.264/AVC από την άλλη μεριά μπορεί να βρει εφαρμογή σε μετάδοση video πάνω από κανάλια με μεγάλη πιθανότητα λάθους ή δίκτυα με περιορισμένες δυνατότητες όπου η τυχαία προσπέλαση των εικόνων είναι δευτερεύουσας σημασίας.

3.4.5 Σύνοψη

Σαν αποτέλεσμα παρατηρήσαμε την μεγάλη υπεροχή του H.264/AVC στις χαμηλές αναλύσεις και για πεπλεγμένο υλικό πρότυπης ευκρίνειας (standard definition television interlaced). Στο πεδίο της σταδιακής σάρωσης (progressive-scan), μεσαίας και υψηλής ανάλυσης παρατηρείται μία ισορροπία μεταξύ των εξεταζόμενων codecs. Στις υψηλές αναλύσεις και ρυθμούς bit το Motion-JPEG2000 υπερέχει ξεκάθαρα του H.264/AVC σε όρους PSNR σε σχέση με τον ρυθμό bit.

3.5 Συνδυασμός MPEG-2 και H.264

Στο [12] ελέγχεται η απόδοση του συνδυασμού ενός πλαισίου χαμηλής ανάλυσης MPEG2 και ενός πλαισίου υψηλής ανάλυσης H.264 έτσι ώστε να έχουμε ένα βελτιωμένο πλαίσιο.

Το H.264 απαιτεί σχεδόν 50% λιγότερο ρυθμό bit σε σχέση με το MPEG2, χωρίς όμως να έχουμε μείωση στην ποιότητα. Καθώς όμως πολλά συστήματα χρησιμοποιούν ήδη το ευρέως διαδεδομένο MPEG2 η μετάβαση στο H.264 είναι δύσκολη.

Η περισσότερη μελέτη γύρω από τη συμπίεση video αφορά την επεκτασιμότητα. Η βασική ιδέα της επεκτασιμότητας είναι ότι το video στέλνεται στο δέκτη σε ένα πλαίσιο βάση (base layer). Αυτό το πλαίσιο αποτελείται από μία έκδοση μειωμένης ποιότητας (reduced quality version) (για επεκτασιμότητα του SNR), μία έκδοση μειωμένου πλαισίου (frame reduced version) (για χρονική επεκτασιμότητα) ή μία χωρικά μειωμένη έκδοση (για χωρική επεκτασιμότητα). Το πλαίσιο βάση κωδικοποιείται συνήθως με το μικρότερο αριθμό bits για να εξασφαλίζουμε ότι θα υπάρχει αρκετό εύρος ζώνης για τη μετάδοση.

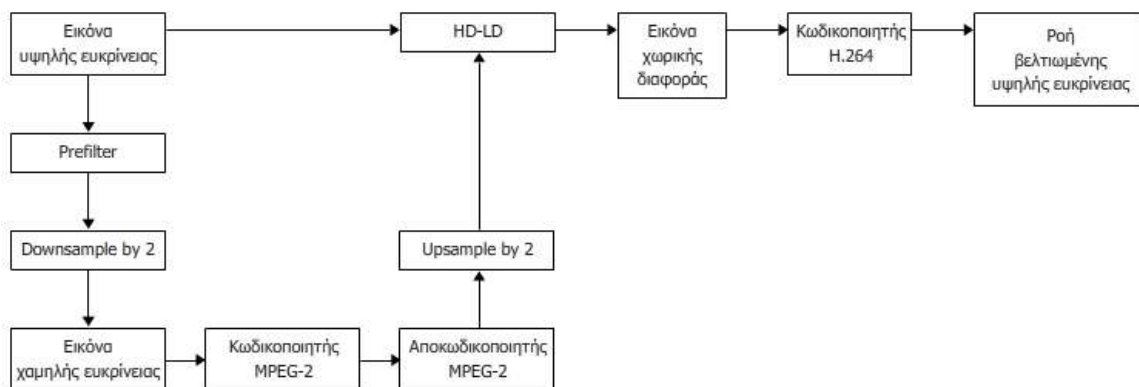
Στη συνέχεια δημιουργείται ένα πλαίσιο ενίσχυσης (enhancement layer) το οποίο περιέχει επιπλέον πληροφορία πρόσθετη από αυτή του πλαισίου βάση. Το πλαίσιο ενίσχυσης αυξάνει το μέγεθος του πλαισίου βάση βελτιώνοντας είτε την ποιότητα, είτε τη χρονική ανάλυση, είτε τη χωρική ανάλυση του τελικού video. Αν το επιτρέπει το εύρος ζώνης το πλαίσιο ενίσχυσης, ή μέρος αυτού στέλνεται στο δέκτη. Αν το εύρος ζώνης μειωθεί ξαφνικά κατά τη διάρκεια της μετάδοσης, ο εξυπηρετητής ο οποίος δημιουργεί τη ροή δεδομένων μπορεί να αμελήσει ένα μέρος ή ολόκληρο το πλαίσιο ενίσχυσης. Αφού ο δέκτης θα λαμβάνει σε κάθε περίπτωση ολόκληρο το πλαίσιο βάση, μπορεί να εξασφαλιστεί μια ελάχιστη ποιότητα video ακόμα και αν λείπει ολόκληρο ή ένα μέρος του πλαισίου ενίσχυσης.

Στο πείραμα αυτό δοκιμάζεται μία νέα μέθοδος συνδυασμού του MPEG2 και του H.264. Συγκεκριμένα ένα χωρικά-μειωμένο πλαίσιο βάση του MPEG2 επεκτείνεται με ένα πλαίσιο ενίσχυσης H.264. Αυτό το πλαίσιο ενίσχυσης αυξάνει τη χωρική ανάλυση και την ποιότητα της τελικής εικόνας, καθώς επίσης έχει εφαρμογή και στο HD video (High Definition video) και στη ψηφιακή μετάδοση.

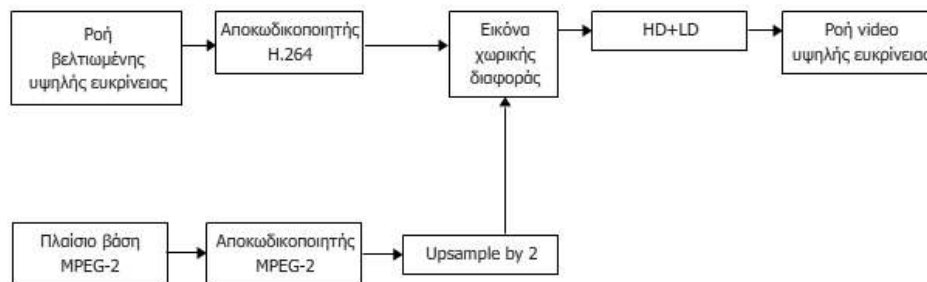
3.5.1 Πειραματική διαδικασία

Στο [13] μελετάται ένα πλαίσιο βελτίωσης υψηλής ευκρίνειας (HD spatial enhancement layer), από το οποίο προέκυψαν κάποια υποσχόμενα αποτελέσματα. Εδώ εξετάζεται η προσφορά ενός υψηλής ποιότητας H.264 χωρικής βελτίωσης πλαισίου (H.264 spatial enhancement layer) σε συνδυασμό με ένα MPEG2 πλαίσιο βάση.

Παρατίθενται τα μπλοκ διαγράμματα κωδικοποίησης και αποκωδικοποίησης χωρίς να αναφερθούμε σε περισσότερες λεπτομέρειες σχετικά με τη διαδικασία κωδικοποίησης και αποκωδικοποίησης.



Εικόνα 30 - Μπλοκ διάγραμμα λειτουργίας ενός κωδικοποιητή ενός συστήματος που χρησιμοποιεί ένα πλαίσιο χαμηλής ανάλυσης MPEG-2 και ένα πλαίσιο υψηλής ανάλυσης H.264



Εικόνα 31 - Μπλοκ διάγραμμα λειτουργίας ενός αποκωδικοποιητή ενός συστήματος που χρησιμοποιεί ένα πλαίσιο χαμηλής ανάλυσης MPEG-2 και ένα πλαίσιο υψηλής ανάλυσης H.264

Για το πείραμα χρησιμοποιήθηκε το πρότυπο H.264 έκδοση 7.3 της Joint Video Team (JVT), ενώ έγιναν κάποιες αλλαγές έτσι ώστε να μπορεί να δεχθεί αποκωδικοποίηση μαζί με το MPEG2 από το πλαίσιο βάση.

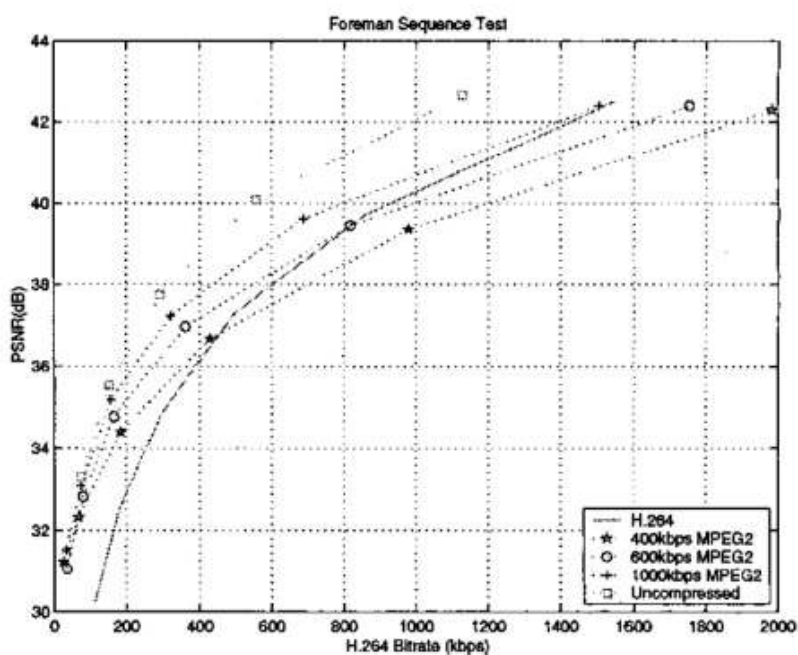
Για την πραγματοποίηση της σύγκρισης χρησιμοποιήθηκαν τα videos Foreman CIF και Coastguard CIF και κωδικοποιήθηκαν στα 30Hz. Οι τιμές σηματοθορυβικού λόγου (PSNR values) αναφέρονται στα στοιχεία φωτεινότητας (luminance component) και είναι μέσος όρος ανάμεσα σε όλα τα πλαίσια. Η μείωση του ρυθμού bit αναφέρεται στα bits που

εξοικονομούνται στο πλαίσιο βελτίωσης σε σύγκριση με την κωδικοποίηση μόνο με το H.264.

3.5.2 Αποτελέσματα και συμπεράσματα

3.5.2.1 Foreman

Στην **Εικόνα 32** φαίνεται η καμπύλη ρυθμού παραμόρφωσης (Rate Distortion, RD) για 30 πλαίσια του video Foreman CIF. Η συνεχόμενη γραμμή δείχνει τα bits που χρειάζονται για την κωδικοποίηση μόνο με H.264. Ελέγχθηκαν τρεις ρυθμοί bit για το πλαίσιο βάση MPEG2: 400kbps, 600kbps και 1000kbps. Η **Εικόνα 22** δείχνει τα bits που μπορούν να εξοικονομηθούν στο πλαίσιο βάση. Για ορισμένους ρυθμούς bit η καμπύλη RD MPEG2 τέμνει την καμπύλη του H.264, υποδηλώνοντας ότι από εκείνο το σημείο δεν χρειάζεται να χρησιμοποιείται ο συνδυασμός MPEG2 και H.264, αφού η κωδικοποίηση μόνο με H.264 αποδίδει καλύτερα. Για ρυθμούς bit κάτω από αυτό το σημείο είναι καλύτερα να επεκτείνουμε το πλαίσιο βάση του MPEG2 με ένα πλαίσιο βελτίωσης H.264.



Εικόνα 32 - Διάγραμμα ρυθμού bit / PSNR του προτεινόμενου συνδυασμού πλαισίων H.264 και MPEG-2 για το video Foreman

Η γραμμή “uncompressed” στην γραφική παράσταση της **Εικόνας 22** αντιπροσωπεύει το άνω όριο της μείωσης του ρυθμού bit χρησιμοποιώντας τη συγκεκριμένη μέθοδο. Υπολογίστηκε με τη χρήση μιας ακολουθίας πλαισίου βάση πριν από την κωδικοποίηση με MPEG2. Αν μία ακολουθία κωδικοποιηθεί με MPEG2, σίγουρα θα έχουμε κάποια αλλοίωση (λόγω της κβαντοποίησης των DCT συντελεστών) ενώ το ασυμπιεστο πλαίσιο βάση θα έχει πάντα τη καλύτερη ποιότητα.

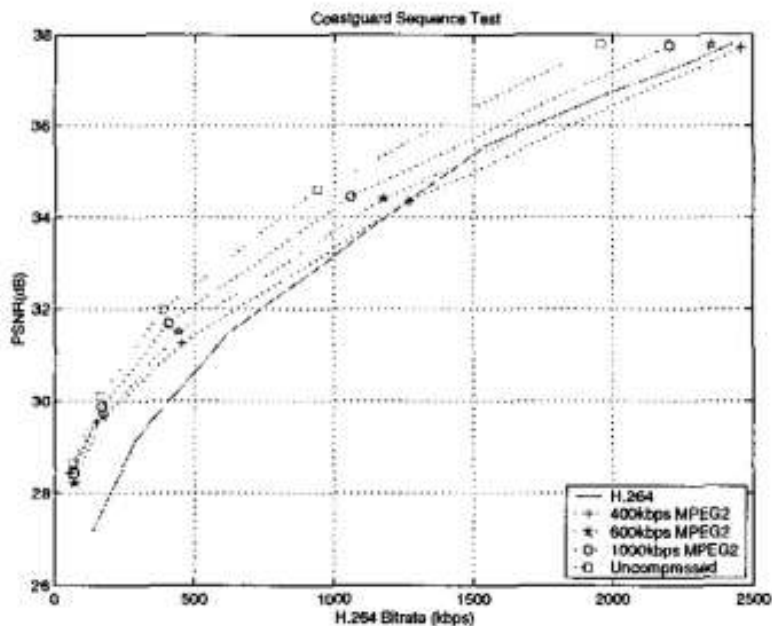
Ο **Πίνακας 20** δείχνει τη μείωση στο ρυθμό bit για το video Foreman αν χρησιμοποιηθεί η προτεινόμενη μέθοδος επεκτασιμότητας (scalability scheme). Όσο αυξάνεται ο ρυθμός bit της ροής δεδομένων του MPEG2, η εξοικονόμηση ρυθμού bit που εξασφαλίζουμε με το πλαίσιο βελτίωσης H.264 γίνεται όλο και πιο σημαντική, φτάνοντας το 50% για PSNR 38dB.

Ρυθμός MPEG2 (kbps)	PSNR	Μείωση ρυθμού bit
400	34	25%
600	34	50%
1000	34	54%
400	36	6%
600	36	28%
1000	36	43%
400	38	None
600	38	8%
1000	38	25%

Πίνακας 20 - Ποσοστό μείωσης του ρυθμού bit του προτεινόμενου συνδυασμού πλαισίων H.264 και MPEG-2 έναντι του MPEG-2 για το video Foreman

3.5.2.2 Coast Guard

Το video Coastguard έχει πληροφορία υψηλότερης συχνότητας από αυτή του video Foreman (λόγω των κυμάτων του ωκεανού) και έτσι απαιτούνται επιπλέον bits για τη συμπίεση. Οι ρυθμοί bit του πλαισίου βάση MPEG2 που ελέχθησαν είναι 400kbps, 600kbps και 1000kbps.



Εικόνα 33 – Διάγραμμα ρυθμού bit / PSNR του προτεινόμενου συνδυασμού πλαισίων H.264 και MPEG-2 για το video Coast Guard

Ο Πίνακας 21 δείχνει τη μείωση στο ρυθμό bit που έχουμε με χρήση της προτεινόμενης μεθόδου επεκτασιμότητας για το video Coastguard. Καθώς ο ρυθμός bit της ροής δεδομένων MPEG2 αυξάνει η εξοικονόμηση σε ρυθμό bit από το πλαίσιο ενίσχυσης H.264 φτάνει μέχρι και 20% για PSNR 38dB.

MPEG2 Rate (kbps)	PSNR	Μείωση ρυθμού bit
400	32	7%
600	32	20%
1000	32	46%
400	34	1%
600	34	8%
1000	34	23%
400	36	None
600	36	4%
1000	36	12%
400	38	None
600	38	3%
1000	38	10%

Πίνακας 21 - Ποσοστό μείωσης του ρυθμού bit του προτεινόμενου συνδυασμού πλαισίων H.264 και PEG-2 έναντι του MPEG-2 για το video Coast Guard

Αν και με το πλαίσιο ενίσχυσης H.264 εξοικονομούμε ρυθμό bit για video υψηλής συχνότητας (όπως το Coastguard), σε video μικρότερης πολυπλοκότητας (όπως το Foreman) η εξοικονόμηση είναι ακόμα πιο θεαματική.

3.5.3 Σύνοψη

Από τα αποτελέσματα του πειράματος μπορούμε να συμπεράνουμε ότι για ρυθμό bit MPEG2 600kbps και τελικό PSNR 36dB για το video Foreman, μπορούμε να μειώσουμε το ρυθμό bit μέχρι και 28% με τη χρήση του πλαισίου ενίσχυσης H.264. Τυπικά η μείωση για το video Foreman για PSNR μεταξύ 34dB και 38dB είναι περίπου 54%, αναλόγως την ποιότητα του πλαισίου βάση MPEG2.

Αξιοσημείωτα ωστόσο είναι και τα αποτελέσματα για το video Coastguard, το οποίο περιέχει πληροφορία υψηλότερης συχνότητας από αυτή του Foreman. Για ρυθμό bit του πλαισίου βάση MPEG2 600kbps και τελικό PSNR 34dB, η μείωση στο ρυθμό bit είναι περίπου 8%. Η συνηθισμένη μείωση στο ρυθμό bit για αυτό το video είναι περίπου 46%, αναλόγως την ποιότητα του πλαισίου MPEG2 και την επιθυμητή τελική ποιότητα.

3.6 Σύγκριση του H.264/AVC με τα DivX, WMV, XviD και MC-EZBC

Σε [14] συγκρίνονται οι σηματοθορυβικοί λόγοι (peak signal-to-noise ration, PSNR) και η ελάχιστη διαφορά (just noticeable difference, JNR) διαφόρων μεθόδων συμπίεσης video. Τα codecs που εξετάζονται είναι το XviD 0.9.1 (σύμφωνα με το MPEG-4 Visual Simple Profile), το DivX 5.1 (σύμφωνα με το MPEG-4 Visual Advanced Simple Profile), το Windows Media Video 9, το MC-EZBC και το H.264/AVC AHM 2.0 (η έκδοση JM 6.1 με επέκταση στον έλεγχο ρυθμού bit (rate control)).

Το πρώτο codec είναι το Ad Hoc Model 2.0 (AHM 2.0) μία εφαρμογή του προτύπου H.264/AVC η οποία επεκτείνει το JM 6.1 με έναν αλγόριθμο ελέγχου του ρυθμού bit (rate control). Δύο codecs στηρίζονται στο πρότυπο MPEG-4, το XviD 0.9.1 και το DivX 5.1. Το τέταρτο codec είναι το Windows Media Video 9 (WMV9) και το τελευταίο είναι ένα επεκτάσιμο codec κωδικοποίησης video βάσει κυματομορφής (scalable wavelet-based video codec) το οποίο αναπτύχθηκε από τον J. Woods et al. (motion compensated embedded zero block coding, MC-EZBC).

3.6.1 Πειραματική διαδικασία και ρύθμιση των codecs

Σε αυτό το κεφάλαιο περιγράφονται οι ρυθμίσεις των συγκρινόμενων codecs. Σαν είσοδος χρησιμοποιήθηκαν έξι σειρές σταδιακού video (progressive video) σε σειρά YCbCr 4:2:0. Μία επισκόπηση των κομματιών video (sequences) δίνεται στον **Πίνακα 22**. Οι κλάσεις A,B και C είναι μία επισκόπηση της ποσότητας της κίνησης και / ή της χωρικής λεπτομέρειας της αντίστοιχης ακολουθίας.

Video	Κλάση	Αριθμός πλαισίων	Συντομογραφία
Head with glasses	A	300	HG
Mother and daughter	A	300	MD
Silent	B	300	Si
Foreman	B	300	Fo
Stefan	C	300	St
Mobile and Calendar	C	300	MC

Πίνακας 22 - Videos που χρησιμοποιήθηκαν για τη σύγκριση των H.264, DivX, WMV, XviD και MC-EZBC

Χρησιμοποιήθηκαν δύο διαφορετικές αναλύσεις: η QCIF (Quarter Common Intermediate Format, QCIF, 176x144) και η CIF (Common Intermediate Format, CIF, 352x288). Έτσι προκύπτουν δώδεκα videos στην είσοδο. Τα videos αυτά κωδικοποιούνται με κωδικοποίηση συνεχόμενου ρυθμού bit (constant bit rate coding, CBR). Χρησιμοποιήθηκαν τριάντα διαφορετικοί ρυθμοί bit, από πολύ χαμηλούς μέχρι πολύ υψηλούς (20, 40, 60, 80, 100, 200, 300, ..., 2500 και 2600kbps) και σε κάθε ρυθμό bit η κωδικοποίηση γινόταν στα 30 fps (frames per second).

Η ροή δεδομένων του H.264 ήταν σύμφωνη με το Βασικό προφίλ (Main Profile). Η δομή GOP είναι IBBBP και το μέγεθος του GOP είναι 16. Για το XviD 0.9.1 χρησιμοποιήθηκε η βιβλιοθήκη πυρήνα (core library), η οποία παρέχει ένα API για την κωδικοποίηση και αποκωδικοποίηση των videos. Η μόνη τροποποίηση που έγινε ήταν στο διάστημα πλαισίου κλειδί (key frame interval) το οποίο ρυθμίστηκε στο 16. Κάθε video κωδικοποιήθηκε μία φορά και εν συνεχεία συλλέχθηκε αρκετές φορές έτσι ώστε να γίνει αποκωδικοποίηση στους διάφορους ρυθμούς bit.

Όσο αναφορά τα WMV9 και DivX 5.1 τα πράγματα είναι πιο πολύπλοκα καθώς αποτελούν μέρος ενός ολόκληρου πλαισίου (framework). Στο WMV9 χρησιμοποιήθηκε ο κωδικοποιητής VCM (Video Compression Manager). Καθώς το WMV9 και το DivX αναμένουν ότι η ροή των δεδομένων του video είναι σε αρχείο AVI, χρησιμοποιήθηκε ένα εργαλείο το οποίο προσθέτει στα YCbCr 4:2:0 δεδομένα ένα περίβλημα και μεταφέρονται ως AVI. Επίσης τροποποιήθηκε το VirtualDub 1.5.10 έτσι ώστε η διαδικασία κωδικοποίησης του WMV9 και DivX να γίνει γρηγορότερη και να διατηρηθεί η σχετική πληροφορία που παράγεται από τους δύο κωδικοποιητές (όπως ο τύπος και το μέγεθος των συμπιεσμένων πλαισίων). Για την αποκωδικοποίηση χρησιμοποιήθηκε ένα εργαλείο βασισμένο στο DirectShow API το οποίο παράγει δεδομένα YUV. Έτσι για όλα τα codecs προέκυψαν μία μόνο YUV γραφικά.

Για τα WMV9 και DivX 5.1 επιλέχθηκαν ροές bit οι οποίες είναι σύμφωνες με το Βασικό προφίλ του WMV9. Στη περίπτωση του DivX ήταν ενεργοποιημένα τα χαρακτηριστικά $\frac{1}{4}$ pixel (quarter-pel) και αντιστάθμιση γενική κίνησης (global motion compensation). Συνήθως χρησιμοποιούνται δύο περάσματα κωδικοποίησης (two-pass encoding) για τη βελτίωση της ποιότητας, αλλά στην περίπτωση των WMV9 και DivX 5.1 ενεργοποιήθηκε η κωδικοποίηση ενός περάσματος (one-pass constant bit rate).

Για τον έλεγχο της απόδοσης ενός codec μπορεί να χρησιμοποιηθεί μόνο η κωδικοποίηση μεταβλητού ρυθμού bit (variable bit rate coding, VBR) και να συγκριθούν τα διάφορα codecs σε αυτό τον τομέα. Σε αυτή την περίπτωση ο έλεγχος του ρυθμού bit (rate control) δεν αποτελεί κομμάτι της σύγκρισης και η απόδοση του codec μπορεί να υπολογιστεί με μεγαλύτερη ακρίβεια. Παρ' όλα αυτά πέρα από το VBR είναι αρκετοί ακόμα παράγοντες που πρέπει να ληφθούν υπόψη. Για παράδειγμα λόγω της διαφορετικής φύσης των I, P και B εικόνων, πρέπει να χρησιμοποιείται διαφορετική μέθοδος κβαντοποίησης για κάθε τύπο. Επίσης λόγω της υψηλής «εξυπνάδας» της κωδικοποίησης CBR, η απόδοση ενός codec είναι μεγαλύτερη σε σχέση με αυτή όταν χρησιμοποιείται κωδικοποίηση VBR. Η χρήση της κωδικοποίησης VBR ακυρώνει (πολύ ή λίγο) τις αρνητικές επιπτώσεις του αλγορίθμου ελέγχου του ρυθμού bit στην ολική απόδοση του codec, ενώ συγχρόνως δεν έχει επίπτωση σε άλλους αλγορίθμους που χρησιμοποιούνται. Έτσι στο πείραμα ο αλγόριθμος ελέγχου του ρυθμού bit μπορεί να θεωρηθεί απών.

3.6.2 Μέτρηση της ποιότητας

Στην αξιολόγηση των videos ο σηματοθορυβικός λόγος (PSNR-Y) είναι η πιο συνηθισμένη μονάδα μέτρησης της ποιότητας. Ο PSNR-Y είναι ο από κορυφή σε κορυφή λόγος του σήματος προς το θόρυβο των φωτεινών στοιχείων (luma component). Η χρωματική πληροφορία (γνωστή και ως χρωματικά στοιχεία - chroma components) δεν λαμβάνεται υπόψη. Για να πάρουμε τον PSNR ολόκληρης του video, υπολογίζεται ο μέσος όρος του PSNR-Y κάθε πλαισίου ξεχωριστά. Αυτός είναι ένας τρόπος για να πάρουμε το PSNR ολόκληρου του video. Ένας άλλος τρόπος θα ήταν, να παίρναμε τη μικρότερη τιμή PSNR-Y που εμφανίζεται στα πλαίσια καθώς σε ορισμένες περιπτώσεις το video αξιολογείται βάσει του χειρότερου κομματιού του.

Μία ακόμα μονάδα μέτρησης της ποιότητας είναι η ελάχιστη διαφορά (just noticeable difference, JND). Το JND αναπτύχθηκε από την Sarnoff Corporation και βασίζεται σε ένα οπτικό μοντέλο το οποίο παρέχει εκτιμήσεις στην ορατότητα των διαφορών μεταξύ της γνήσιας και της παραμορφωμένης εικόνας. Το μοντέλο αυτό βασίζεται σε γνωστές ψυχολογικές και ψυχοφυσικές αρχές του ανθρώπινου οπτικού διαχωρισμού (human visual discrimination performance). Η μονάδα JND ορίζεται ως εξής: ένα JND αντιπροσωπεύει το 75% της πιθανότητας ένας παρατηρητής βλέποντας της δύο εικόνες να δει τη διαφορά. Όμοια οι δύο βαθμοί JND αντιπροσωπεύουν $0.75 + 0.75 \times (1 - 0.75) = 93.75\%$ πιθανότητα να παρατηρήσει τη διαφορά.

Εδώ πρέπει να σημειωθεί ότι αύξηση του JND σημαίνει μείωση της ποιότητας, ενώ αύξηση του PSNR σημαίνει αύξηση της ποιότητας.

Η έλλειψη χρονικής διάστασης και στις δύο μονάδες μέτρησης της ποιότητας μπορεί να θεωρηθεί σαν μειονέκτημα (χρήσιμο όταν θέλουμε να μετρήσουμε την ποιότητα κωδικοποιημένων video πολλαπλών πλαισίων, multilayered video coding schemes).

3.6.3 Αποτελέσματα

Μερικά αποτελέσματα (11 από τους 30 συνολικά ρυθμούς bit και 4 από τα 6 video CIF) παρουσιάζονται στους Πίνακες 23 και 24. Τα videos "Silent" και "Mother & Daughter" δεν

παρουσιάζονται καθώς περιέχουν μία λωρίδα από μαύρα pixels και το codec WM9 εφήρμοσε ένα φίλτρο κατά προσέγγιση (extrapolation filter) έτσι ώστε να απαλλαγεί από αυτή τη μαύρη γραμμή, προκαλώντας έτσι λανθασμένες μετρήσεις στην ποιότητα. Σύμφωνα με τα δεδομένα του **Πίνακα 23** το σύνολο των αποτελεσμάτων συγκεντρώθηκε και προέκυψε ένας μέσος όρος, έτσι ώστε να πάρουμε μία καλύτερη εικόνα των αποτελεσμάτων. Για κάθε video και μονάδα μέτρησης (quality metric) δίνεται η μέση τιμή της διαφοράς μεταξύ των τιμών της ποιότητας του H.264/AVC και των αντίστοιχων τιμών του αντίστοιχου codec. Αν το H.264/AVC έχει καλύτερη απόδοση στη κωδικοποίηση η διαφορά των PSNR είναι θετική ενώ η διαφορά του JND είναι αρνητική.

	DivX 5.1		WMV9		MC-EZBC		XviD 0.9.1	
	PSNR	JND	PSNR	JND	PSNR	JND	PSNR	JND
HG	1.64	0.00	2.18	0.00	-0.22	0.10	1.63	-0.04
MD	1.52	-0.11	12.88	-1.93	-0.24	0.10	1.74	-0.15
Si	1.93	-0.13	9.87	-1.45	0.23	-0.21	2.22	-0.19
Fo	2.35	0.31	2.64	-0.04	1.56	-0.10	2.80	-0.36
St	2.84	-0.21	4.08	-2.14	2.43	-0.07	1.81	0.43
MC	4.00	-0.90	4.79	-0.86	0.74	0.02	5.73	-1.24

Πίνακας 23 - Συνολικό κέρδος κωδικοποίησης του H.264/AVC έναντι των υπολοίπων codecs

target rate	H.264/AVC AHM 2.0			DivX 5.1			WMV9			MC-EZBC			XviD 0.9.1		
	rate	PSNR	JND	rate	PSNR	JND	rate	PSNR	JND	rate	PSNR	JND	rate	PSNR	JND
foreman, CIF, 30Hz															
60	60.1	26.0	7.98	-	-	-	-	-	-	-	-	-	-	-	-
100	100.5	28.3	6.99	-	-	-	-	-	-	-	-	-	-	-	-
300	297.0	33.4	5.22	362.1	32.5	4.05	416.2	33.7	3.57	300.0	32.8	4.34	307.3	31.9	4.13
500	493.2	34.1	5.90	480.1	34.0	3.64	593.1	35.1	3.13	500.0	35.1	3.63	512.4	34.1	3.54
700	650.9	35.3	5.65	755.3	36.1	2.98	784.8	36.2	2.82	700.0	36.5	3.18	715.4	35.5	3.20
900	892.0	39.2	2.71	902.8	36.9	2.79	979.8	37.1	2.65	900.0	37.6	2.85	916.9	36.5	2.96
1100	1089.5	39.9	2.68	1058.4	37.6	2.61	1173.1	37.8	2.52	1100.0	38.5	2.65	1118.9	37.3	2.78
1300	1287.7	40.9	2.21	1225.5	38.2	2.48	1368.1	38.4	2.41	1300.0	39.2	2.48	1319.5	38.0	2.65
1600	1548.2	40.7	3.01	1574.5	39.3	2.26	1663.0	39.4	2.31	1600.0	40.3	2.22	1621.6	39.0	2.46
2000	1982.2	42.9	1.81	1971.0	40.3	2.07	2056.2	40.4	2.23	2000.0	41.4	2.02	2033.3	40.1	2.27
2500	2476.7	44.0	1.67	2610.0	41.6	1.89	2569.0	41.4	2.13	2500.0	42.5	1.78	2524.8	41.2	2.04
head with glasses, CIF, 30Hz															
60	59.3	26.1	7.04	-	-	-	-	-	-	-	-	-	-	-	-
100	99.0	28.8	5.93	-	-	-	-	-	-	-	-	-	-	-	-
300	296.8	34.4	3.84	254.0	30.9	4.06	434.2	34.4	3.19	300.0	34.1	3.58	309.8	32.5	3.71
500	495.3	36.8	3.22	483.3	34.9	3.05	640.6	35.9	2.79	500.0	36.9	2.88	512.3	35.2	3.07
700	693.5	38.3	2.72	740.6	36.8	2.65	829.0	36.9	2.61	700.0	38.5	2.60	713.6	36.7	2.75
900	892.0	39.2	2.51	876.5	37.5	2.54	1001.1	37.5	2.49	900.0	39.5	2.42	913.9	37.6	2.56
1100	1089.0	40.0	2.41	1178.0	38.6	2.39	1186.2	37.9	2.41	1100.0	40.2	2.33	1112.3	38.3	2.47
1300	1286.9	40.5	2.34	1345.1	39.1	2.32	1359.9	38.4	2.34	1300.0	40.9	2.23	1312.0	39.0	2.39
1600	1583.9	41.3	2.24	1527.0	39.5	2.26	1648.2	39.2	2.25	1600.0	41.6	2.13	1613.0	39.7	2.28
2000	1979.7	42.1	2.15	2100.4	40.7	2.12	2057.8	40.1	2.12	2000.0	42.4	2.02	2008.6	40.7	2.16
2500	2474.6	43.1	2.02	2513.3	41.4	2.05	2567.4	41.1	2.01	2500.0	43.3	1.81	2512.0	41.8	2.04
mobile, CIF, 30Hz															
60	62.3	18.5	7.72	-	-	-	-	-	-	-	-	-	-	-	-
100	99.0	20.1	7.00	-	-	-	-	-	-	-	-	-	-	-	-
300	297.0	25.6	5.08	-	-	-	-	-	-	-	-	-	-	-	-
500	494.8	28.5	3.96	524.3	26.2	4.33	738.9	26.8	3.97	500.0	28.7	3.86	514.8	24.1	5.02
700	693.2	30.5	3.32	712.7	27.3	4.06	950.3	27.8	3.70	700.0	30.4	3.35	719.3	25.4	4.58
900	891.4	31.9	3.00	986.3	28.6	3.77	1143.6	28.6	3.54	900.0	31.5	2.96	925.9	26.5	4.30
1100	1090.0	32.9	2.82	1150.2	29.2	3.65	1364.8	29.4	3.40	1100.0	32.5	2.65	1127.8	27.5	4.07
1300	1287.4	33.8	2.60	1335.3	29.8	3.56	1554.8	30.0	3.25	1300.0	33.2	2.53	1332.3	28.3	3.87
1600	1583.9	34.9	2.42	1751.4	31.1	3.35	1777.7	30.6	3.20	1600.0	34.1	2.44	1633.0	29.3	3.67
2000	1980.0	36.2	2.20	1984.1	31.7	3.26	2111.9	31.5	3.08	2000.0	35.3	2.16	2035.4	30.5	3.33
2500	2474.8	37.5	2.02	2495.3	33.1	2.94	2580.5	32.7	2.97	2500.0	36.6	2.03	2534.0	31.8	3.09
stefan, CIF, 30Hz															
60	61.1	19.8	9.17	-	-	-	-	-	-	-	-	-	-	-	-
100	98.7	21.3	8.28	-	-	-	-	-	-	-	-	-	-	-	-
300	304.9	26.0	6.87	-	-	-	-	-	-	-	-	-	-	-	-
500	487.9	28.1	6.40	-	-	-	-	-	-	-	-	-	-	-	-
700	695.2	30.9	4.52	822.5	28.6	4.32	899.5	29.5	5.55	700.0	28.5	4.49	721.0	29.2	4.20
900	1041.4	31.4	6.06	883.1	29.1	4.24	1084.6	30.1	5.51	900.0	29.7	4.18	924.2	30.4	3.94
1100	1183.9	32.5	5.74	1107.6	30.6	3.95	1276.3	30.6	5.52	1100.0	31.0	3.89	1127.6	31.6	3.72
1300	1412.0	33.6	5.33	1255.0	31.4	3.86	1445.1	31.0	5.50	1300.0	31.8	3.72	1329.7	32.6	3.56
1600	1584.8	35.6	3.23	1582.7	32.7	3.66	1707.2	31.5	5.48	1600.0	33.0	3.45	1631.3	33.7	3.35
2000	1980.1	36.8	3.04	1975.9	34.0	3.43	2072.6	32.0	5.47	2000.0	34.5	3.20	2034.4	35.1	3.14
2500	2475.1	38.3	2.78	2439.6	35.3	3.26	2570.6	32.5	5.46	2500.0	36.0	2.87	2536.7	36.6	2.88

Πίνακας 24 - Μετρήσεις ποιότητας των H.264/AVC, DivX, WMV9, XviD και MC-EZBC

3.6.3.1 H.264/AVC και MC-EZBC

Σε αυτό το κομμάτι συγκρίνεται η απόδοση της κωδικοποίησης του H.264/AVC AHM 2.0 και του MC-EZBC.

Μία πρώτη παρατήρηση είναι ότι για ορισμένους ρυθμούς bit δεν υπάρχουν μετρήσεις του MC-EZBC (το οποίο συμβαίνει και σε άλλα codecs). Για χαμηλούς ρυθμούς bit αυτό σημαίνει ότι το MC-EZBC δεν μπορεί να κωδικοποιήσει το συγκεκριμένο video σε τόσο χαμηλό ρυθμό bit. Από τη άλλη τα άλλα codecs παράγουν μία ροή δεδομένων έχοντας το μικρότερο δυνατό ρυθμό bit (για το συγκεκριμένο codec). Αυτό οδηγεί σε παραγωγή όμοιων ροών bit (identical bit streams) για μία ομάδα διαφορετικών ρυθμών bit. Το ίδιο μπορεί να συμβεί και για υψηλούς ρυθμούς bit. Στη περίπτωση των χαμηλών ρυθμών bit επίσης το codec μπορεί να αποφασίσει να αγνοήσει κάποια πλαίσια.

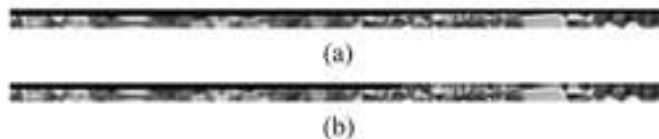
Για video με χαμηλή κίνηση και/ή μικρή χωρική λεπτομέρεια (Head with Glasses, Mother & Daughter) παρατηρούμε ότι το MC-EZBC αποδίδει λίγο καλύτερα από το H.264/AVC. Αυτό δείχνει ότι για το συγκεκριμένο είδος video η κωδικοποίηση βάσει της κυματομορφής (wavelet-based) μπορεί να έχει συγκρίσιμη απόδοση με αυτή του H.264/AVC.

Για ακολουθίες με υψηλή κινητικότητα (Foreman και Stefan) βλέπουμε ότι το H.264/AVC AHM2.0 αποδίδει καλύτερα από το MC-EZBC όσο αναφορά τη μονάδα μέτρησης PSNR-Y, ενώ στο JND βλέπουμε ότι η διαφορά αυτή είναι πολύ μικρότερη. Παρατηρούμε ότι το MC-EZBC σε αντίθεση με το H.264/AVC δεν μπορεί να κωδικοποιήσει αυτά τα δύσκολα videos σε ρυθμούς bit κάτω από 200kbps. Ένας λόγος που συμβαίνει αυτό είναι ότι οι κωδικοποιημένες ροές bit (coded bit streams) έχουν τρία επίπεδα χωρικής επεκτασιμότητας (spatial scalability) και τέσσερα επίπεδα χρονικής επεκτασιμότητας (temporal scalability), το οποίο σημαίνει μία επιβάρυνση η οποία είναι κρίσιμη στους χαμηλούς ρυθμούς bit. Στο video Stefan ακόμα και ο μέσος PSNR-Y του XviD (MPEG-4 Visual SP) είναι μεγαλύτερος από του MC-EZBC για όλους τους ρυθμούς bit. Η ανισότητα στη ποιότητα περιορίζεται μόνο στα videos με υψηλή χωρική λεπτομέρεια και χαμηλά ποσοστά κίνησης (για παράδειγμα Mobile and Calendar: 0.74dB και 0.02 JND).

Ένα ενδιαφέρον φαινόμενο παρατηρήθηκε όταν κωδικοποιήθηκε το video Mother and Daughter σε QCIF ανάλυση. Σε αυτή τη περίπτωση, το MC-EZBC φτάνει τη μέγιστη δυνατή ποιότητα (53.1 dB και 0.78 JND) στα 1600kbps. Από την άλλη το H.264/AVC AHM 2.0 μπορεί να χρησιμοποιήσει περισσότερα bits (higher bit budget) για να πετύχει υψηλότερο PSNR-Y και χαμηλότερες JND τιμές (φτάνοντας τα 59.9dB και 0.50 JND) στα 2600kbps. Αν και είναι δύσκολο να δούμε τη διαφορά ανάμεσα στα 53dB και 60dB, αυτό δείχνει ότι το H.264/AVC έχει μία πιο ευέλικτη μέθοδο κβαντοποίησης ή οποία μπορεί να οδηγήσει σε καταπληκτικά PSNR.

3.6.3.2 H.264/AVC, WMV9 και DivX

Σε αυτό το σημείο γίνεται μία επισκόπηση της απόδοσης κωδικοποίησης των δύο εμπορικών codecs WMV9 και DivX και συγκρίνονται με το H.264/AVC AHM 2.0.



Εικόνα 34 - Το WMV9 εξάλειψε τη μαύρη λωρίδα στην κορυφή του video Mother and Daughter (a) Αρχική (b) Συμπιεσμένη με WMV9

Όπως αναφέρθηκε και προηγουμένως τα αποτελέσματα του WMV9 υπήρχαν κάποιες αλλοιώσεις. Όσο αναφορά το video Mother & Daughter, οι δύο πάνω σειρές pixels έχουν μία μαύρη λωρίδα την οποία εξάλειψε το WMV9 κατά τη κωδικοποίηση. Αυτό είχε ως αποτέλεσμα τον λανθασμένο υπολογισμό του PSNR και του JND. Επίσης για το video Silent οι δύο πάνω σειρές pixels του video αλλοιώθηκαν τόσο πολύ από το WMV9 έτσι ώστε δημιουργήθηκε μία γκριζα λωρίδα. Αυτό ενδεχομένως να οφείλεται στο μοντέλο (psychovisual model) που χρησιμοποιεί το WMV9 κατά την κωδικοποίηση. Αυτού του είδους οι αλλοιώσεις παρατηρούνται (στη πάνω δεξιά μεριά της εικόνας) και στην ακολουθία Stefan όταν κωδικοποιείται με WMV9, παρ' όλα αυτά όμως δεν οδηγούν σε πολύ χαμηλές μετρήσεις της ποιότητας. Στην **Εικόνα 34** φαίνεται η αλλοίωση στην ακολουθία Mother & Daughter από το WMV9.

Βάσει των τιμών του PSNR του **Πίνακα 23** το H.264/AVC αποδίδει καλύτερα και από τα δύο codecs, το DivX και το WMV9. Το H.264/AVC αποδίδει καλύτερα όταν κωδικοποιούνται «πολύπλοκες» ακολουθίες. Δηλαδή, το κέρδος είναι μεγαλύτερο όταν έχουμε video κλάσης C (video sequences class C). Σε αυτό το σημείο (και σύμφωνα με τα προηγούμενα) βλέπουμε ότι η επιλογή της ακολουθίας είναι ένας πολύ σημαντικός παράγοντας όταν συγκρίνεται η απόδοση της κωδικοποίησης διαφόρων τεχνολογιών codecs.

Παρατηρώντας τον **Πίνακα 23** βλέπουμε ότι υπάρχουν κάποιες ιδιομορφίες σχετικά με τις δύο διαφορετικές μονάδες μετρήσεις. Το PSNR δείχνει ότι το H.264/AVC έχει καλύτερη απόδοση για τα videos Foreman και Head with Glasses. Σε αντίθεση το JND δείχνει ότι δεν υπάρχει διαφορά στη Head with Glasses για τα τρία codecs. Επιπλέον το JND δείχνει ότι στο video Foreman το DivX αποδίδει καλύτερα από το H.264/AVC και η διαφορά με το WMV9 είναι οριακή. Αν προσέξουμε το video Head with Glasses κωδικοποιημένο με H.264/AVC, DivX και WMV9 σε ρυθμό bit 1000kbps βλέπουμε ότι είναι δύσκολο να διακρίνουμε διαφορά μεταξύ των τριών αυτών κωδικοποιημένων videos (βάσει των μηδενικών διαφορών του JND) ενώ το PSNR δείχνει μία διαφορά 1.64dB και 2.18dB.

Σε περιπτώσεις υψηλού ρυθμού bit, υψηλής ποιότητας και «απλών» σκηνών, το JND αντιπροσωπεύει την αντιληπτή ποιότητα, ενώ το PSNR αντιπροσωπεύει τη θεωρητική ποιότητα. Για ρυθμούς bit κάτω από 1000kbps παρατηρείται μία καθαρή διαφοροποίηση μεταξύ της ποιότητας των codecs. Το H.264/AVC αποδίδει καλύτερα όπως δείχνει και το PSNR, ενώ σε περιπτώσεις χαμηλού ρυθμού bit και χαμηλής ποιότητας το JND δεν είναι καλή μονάδα μέτρησης.

Μία δεύτερη παρατήρηση έχει να κάνει με το video Mobile and Calendar. Ενώ ο PSNR δείχνει ότι το DivX έχει καλύτερη κωδικοποίηση από το WMV9 για όλους του ρυθμούς bit, το JND δείχνει το αντίθετο. Για το λόγο αυτό έγινε ένα ακόμα αντικειμενικό οπτικό πείραμα (subjective viewing test), το οποίο έδειξε ότι τα videos που κωδικοποιήθηκαν με WMV9 έχουν καλύτερη οπτική ποιότητα, σε σχέση με αυτές του DivX. Το αποτέλεσμα αυτό συμφωνεί με τα αποτελέσματα του JND. Η κωδικοποίηση με DivX περιέχει περισσότερες οπτικές αλλοιώσεις (blocking artifacts) και έτσι οδηγούμαστε σε μειωμένη ποιότητα. Οι κωδικοποιήσεις με H.264/AVC έχουν υψηλότερη ποιότητα και από αυτές του DivX και από του WMV9, όπως προκύπτει και από το PSNR και το JND.

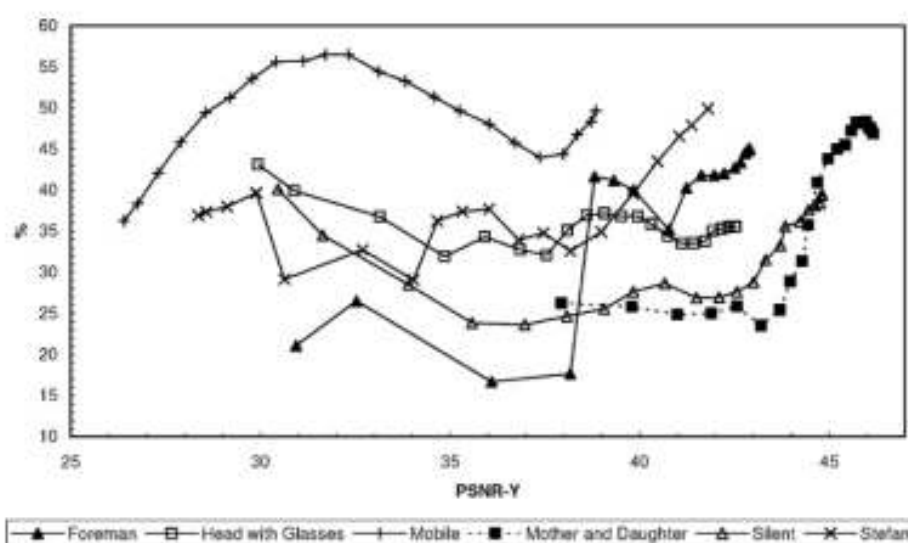
Μία ακόμα παρατήρηση είναι ότι το WMV9 και DivX δεν κωδικοποιούν όλα τα πλαίσια του video αν ο ρυθμός bit είναι χαμηλός (κάτω από 300kbps). Ένα συμπιεσμένο πλαίσιο προηγείται των πλαισίων που δεν κωδικοποιήθηκαν, το οποίο επαναλαμβάνεται έτσι ώστε να ισοσταθμιστούν τα πλαίσια που δεν κωδικοποιήθηκαν. Έτσι τα πλαίσια τα οποία κωδικοποιήθηκαν είχαν καλή ποιότητα, αλλά η ομαλότητα της κίνησης χειροτέρευε. Τα codecs αυτά κάνουν μία ανταλλαγή μεταξύ της οξύτητας (sharpness) των πλαισίων και της ομαλότητας (smoothness) της όλης ακολουθίας. Η συμπεριφορά αυτή μπορεί να ελεγχθεί από τις παραμέτρους του κωδικοποιητή (στο πείραμα χρησιμοποιήθηκαν οι προεπιλεγμένες ρυθμίσεις). Επειδή τα κωδικοποιημένα videos περιλαμβάνουν ένα ή περισσότερα μη-κωδικοποιημένα πλαίσια (skipped frames) δεν μπορεί να γίνει σύγκριση με το πρωτότυπο video χρησιμοποιώντας το PSNR και το JND. Τα videos αυτά δεν περιλαμβάνονται στα αποτελέσματα και στον πίνακα είναι – (παύλες).

3.6.3.3 H.264/AVC και MPEG-4

Η εφαρμογή XviD του MPEG-4 Visual SP έχει απόδοση κωδικοποίησης που είναι γενικά παρόμοια με αυτή του DivX, με την εξαίρεση ότι παρουσιάζεται μεγάλη διαφορά όταν κωδικοποιούνται πολύπλοκα videos κλάσης C (“complex” sequences, class C). Παρατηρώντας τον **Πίνακα 24** βλέπουμε ότι το XviD μπορεί να κωδικοποιήσει τα videos σε χαμηλότερους ρυθμούς bit σε σχέση με τα άλλα codecs (με εξαίρεση το H.264/AVC) χωρίς να παραλείψει πλαίσια, το οποίο είναι σύμφωνο με το Απλό προφίλ (Simple Profile) του MPEG-4 Visual. Στον **Πίνακα 23** το PSNR και το JND δείχνουν αντίθετα αποτελέσματα όσο αναφορά το video Stefan όταν συγκρίνεται το XviD με το H.264/AVC (αυτή η αντίθεση παρατηρείται και σε ρυθμούς bit μεγαλύτερους από 1400kbps). Με μία οπτική επιθεώρηση (στους αντίστοιχους ρυθμούς bit δίνουν αντίθετα αποτελέσματα) παρατηρούμε ότι τα videos που είναι κωδικοποιημένα με H.264/AVC περιέχουν ορισμένα άσχημα κομμάτια (bad parts) (περίπου 70 πλαίσια συνολικά) με κάποια ίχνη (ghosts) και ατέλειες στα μπλοκ (blocking artifacts), ενώ τα videos κωδικοποιημένα με XviD δίνουν την αίσθηση μίας πιο σταθερής αλλά χαμηλότερης ποιότητας, καθ’ όλη τη διάρκεια του video. Η διαφορά στις δύο μονάδες μέτρησης (PSNR και JND) οφείλεται στο ότι το JND επηρεάζεται περισσότερο από αυτά τα άσχημα κομμάτια, ενώ ο PSNR εξισορροπείται στο σύνολο της ακολουθίας. Ποια από τις δύο μονάδες είναι σωστή εξαρτάται από τις προτεραιότητες και προτιμήσεις του χρήστη.

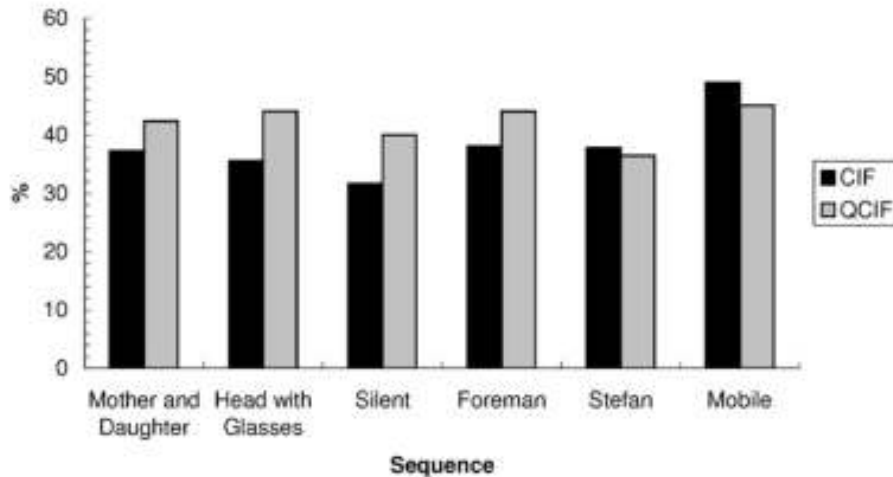
Στην **Εικόνα 35** φαίνεται η εξοικονόμηση ρυθμού bit (bit rate savings) του H.264/AVC AHM 2.0 σε σχέση με το DivX ανά μονάδα ποιότητας (σε τιμές PSNR) για όλες της ακολουθίες CIF.

Όπως φαίνεται η εξοικονόμηση ρυθμού bit εξαρτάται αρκετά από το περιεχόμενο του video. Επίσης videos ίδιας κλάσης δεν έχουν όμοια αποτελέσματα (για παράδειγμα η Mobile and Calendar και η Stefan). Σημαντική εξοικονόμηση ρυθμού bit έχουμε για πολύ χαμηλές και πολύ υψηλές ποιότητες. Αυτό επιβεβαιώνει το γεγονός ότι το H.264/AVC έχει σχεδιαστεί για ένα μεγάλο εύρος ρυθμών bit. Σημαντική παρατήρηση είναι ότι η εξοικονόμηση bit στο video Mobile and Calendar είναι πολύ μεγαλύτερη σε σχέση με τα άλλα videos. Το video Mobile and Calendar είναι το πιο πολύπλοκο που χρησιμοποιήθηκε και έχει τη φήμη του video αναφοράς για τους αλγορίθμους κωδικοποίησης video. Πολλά από τα χαρακτηριστικά του H.264/AVC φαίνεται να κωδικοποίησαν το video Mobile and Calendar πολύ αποτελεσματικά.



Εικόνα 35 – Εξοικονόμηση ρυθμού bit / PSNR με χρήση του H.264/AVC σε σύγκριση με το DivX 5.1 για όλα τα videos

Στην **Εικόνα 36** φαίνεται η εξοικονόμηση ρυθμού bit για κάθε video. Στο σύνολο το H.264/AVC AHM 2.0 έχει ένα εξοικονόμηση ρυθμού bit 42% στις QCIF αναλύσεις και 38% στις CIF αναλύσεις, με μονάδα μέτρησης της ποιότητας το PSNR. Βλέπουμε ότι μπορούμε να έχουμε σημαντικό κέρδος κάνοντας χρήση των βελτιωμένων χαρακτηριστικών του H.264. Παρόμοια αποτελέσματα προκύπτουν και αν ληφθεί υπόψη σαν μονάδα μέτρησης της ποιότητα το JND.



Εικόνα 36 - Εξοικονόμηση ρυθμού bit με χρήση του H.264/AVC σε σύγκριση με το DivX 5.1 ανά video

3.6.4 Σύνοψη

Σε αυτό το πείραμα έγινε μία επισκόπηση της απόδοσης πέντε τεχνολογιών κωδικοποίησης video. Τα αποτελέσματα δείχνουν ότι το H.264/AVC μπορεί να αποδώσει καλύτερα σχεδόν σε όλες τις περιπτώσεις.

Σε μερικές περιπτώσεις το PSNR και το JND σαν μονάδες μέτρησης της ποιότητας (quality metrics) οδηγούν σε διαφορετικά συμπεράσματα, όσο αναφορά την απόδοση των αλγορίθμων. Τα αποτελέσματα δείχνουν ότι το JND δεν πρέπει να χρησιμοποιείται στις χαμηλές ποιότητες, αλλά πρέπει να λαμβάνεται μόνο το PSNR υπόψη (αν και ούτε αυτό είναι 100% ορθό). Αντίθετα στις υψηλές ποιότητες το JND είναι πιο έγκυρο σε σύγκριση με το PSNR.

4. Συγκρίσεις του προτύπου H.264 διαφορετικών κατασκευαστών και του MPEG-4

Σε αυτό το κεφάλαιο θα ελεγχθεί η απόδοση μεταξύ του προτύπου H.264 διαφορετικών κατασκευαστών. Τα πειράματα πραγματοποιήθηκαν από το Graphics & Media Lab Video Group το οποίο είναι τμήμα του Computer Science Department of Moscow State University και παρουσιάζονται ξεκινώντας από το πιο πρόσφατο και προχωρώντας στο πιο παλιό.

4.1 Σύγκριση XviD, MainConcept H.264, Intel H.264, x264, AMD H.264 και Artemis H.264

Το πείραμα αυτό πραγματοποιήθηκε τον Δεκέμβριο του 2007 και παρουσιάζεται στο [15]. Συγκρίνονται τα codecs XviD (MPEG-4 ASP codec), MainConcept H.264, Intel H.264, x264, AMD H.264 και Artemis H.264. Μία σύντομη περιγραφή των codecs φαίνεται στον Πίνακα 25. Το codec XviD θεωρείται σαν codec αναφοράς.

Codec	Κατασκευαστής	Έκδοση
MainConcept H.264/AVC	MainConcept AG	Build 7.3.0 at 2007/07/25 rev. 18090
AMD H.264/AVC	Adanced Micro Devices, Inc.	
Intel H.264	Intel Corp.	Dev. Version for 07.09.2007
Raw H.264	Artemis	07.2007
X264	X264 Development Team	X264 core:56 svn-671
XviD raw MPEG4	XviD	Version for 24.08.2007

Πίνακας 25 - Codecs που χρησιμοποιήθηκαν για τη σύγκριση XviD, MainConcept H.264, Intel H.264, x264, AMD H.264 και Artemis H.264

Τα κομμάτια video που χρησιμοποιήθηκαν στο πείραμα φαίνονται στον Πίνακα 26.

Video	Αριθμός πλαισίων	Ρυθμός πλαισίων	Ανάλυση και χρώματα
Salesman	449	30	172x144 (YV12)
Foreman	300	30	352x288 (YV12)
News	300	30	352x288 (YV12)
Battle	1599	24	704x288 (YV12)
Smith	772	24	720x432 (YV12)
Ice Age	491	24	720x576 (YV12)
Lord of the rings	292	24	720x416 (YV12)
Troy	300	24	1920x1072 (YV12)
Matrix (HDTV)	250	30	1920x1072 (YV12)

Πίνακας 26 - Videos που χρησιμοποιήθηκαν για τη σύγκριση XviD, MainConcept H.264, Intel H.264, x264, AMD H.264 και Artemis H.264

Σκοπός του πειράματος είναι η παρουσίαση των αποτελεσμάτων της σύγκρισης μεταξύ του codec H.264 διαφόρων εταιριών. Οι ρυθμίσεις που χρησιμοποιήθηκαν ήταν σύμφωνες με τον κατασκευαστή του codec.

4.1.1 Πειραματική διαδικασία

Το πείραμα διαιρείται σε τρεις βασικούς τύπους εφαρμογών, οι οποίοι διαφέρουν όσο αναφορά την ανάλυση, το ρυθμό bit και τις απαιτήσεις σε ταχύτητα κωδικοποίησης. Έτσι έχουμε:

- Video-διάσκεψη (videoconferences)
 - ρυθμός bit 50 – 400 kbps
 - απαιτήσεις ταχύτητας για video CIF 200kbps: ελάχιστο 60fps για ρύθμιση υψηλής ταχύτητας (High Speed) και 30fps για ρύθμιση υψηλής ποιότητας (High Quality)
- Ταινίες (movies)
 - ρυθμός bit 500 – 1500 kbps
 - απαιτήσεις ταχύτητας για video 4CIF 750kbps: ελάχιστο 15fps για ρύθμιση υψηλής ταχύτητας και 4fps για ρύθμιση υψηλής ποιότητας
- Τηλεόραση υψηλής ευκρίνειας (high-definition television, HDTV)
 - ρυθμός bit 1 – 10 Mbps
 - απαιτήσεις ταχύτητας για video 1280x720 3Mbps: ελάχιστο 4fps για ρύθμιση υψηλής ταχύτητας και 1 fps για ρύθμιση υψηλής ποιότητας

Οι ρυθμίσεις των codecs παρήχθησαν από τους κατασκευαστές του κάθε codec. Κάθε codec ελέγχθηκε τρεις φορές και χρησιμοποιήθηκε η μέση τιμή των τριών μετρήσεων. Το video πηγή ήταν σε YV12 format (κατάληξη αρχείου .yuv) και για τις μετρήσεις χρησιμοποιήθηκαν τα μεγέθη PSNR (Y, U, V components) και SSIM (Y, U, V components).

4.1.2 Αποτελέσματα

4.1.2.1 Video-διάσκεψη

Οι **Πίνακες 27 και 28** παρουσιάζουν το σχετικό ρυθμό bit κάθε codec για την ίδια ποιότητα. Το MainConcept αποδίδει καλύτερα από όλα τα υπόλοιπα, σύμφωνα με τις αντικειμενικές μετρήσεις (objective quality metrics) και ακολουθεί το x264, ενώ τρίτο είναι το Intel IPP codec. Η διαφορά μεταξύ αυτών των τριών codecs δεν είναι τρομερή. Το AMD είναι το μόνο codec το οποίο είναι χειρότερο από το XviD (το οποίο είναι και το codec αναφοράς). Το Artemis x264 φαίνεται να αποδίδει χειρότερα σύμφωνα με το Y-PSNR αλλά δείχνει να είναι καλύτερο σύμφωνα με το Y-SSIM.

Εδώ πρέπει να σημειωθεί ότι στους **Πίνακες 27 και 28** κάθε αριθμός αντιπροσωπεύει το ίδιο τμήμα ρυθμού bit. Αυτά τα τμήματα μπορεί να είναι αρκετά διαφορετικά λόγω της διαφορετικής ποιότητας των συγκρινόμενων codecs. Αυτό μπορεί να οδηγήσει σε ανεπαρκή συμπεράσματα όταν συγκρίνονται πάνω από τρία codecs.

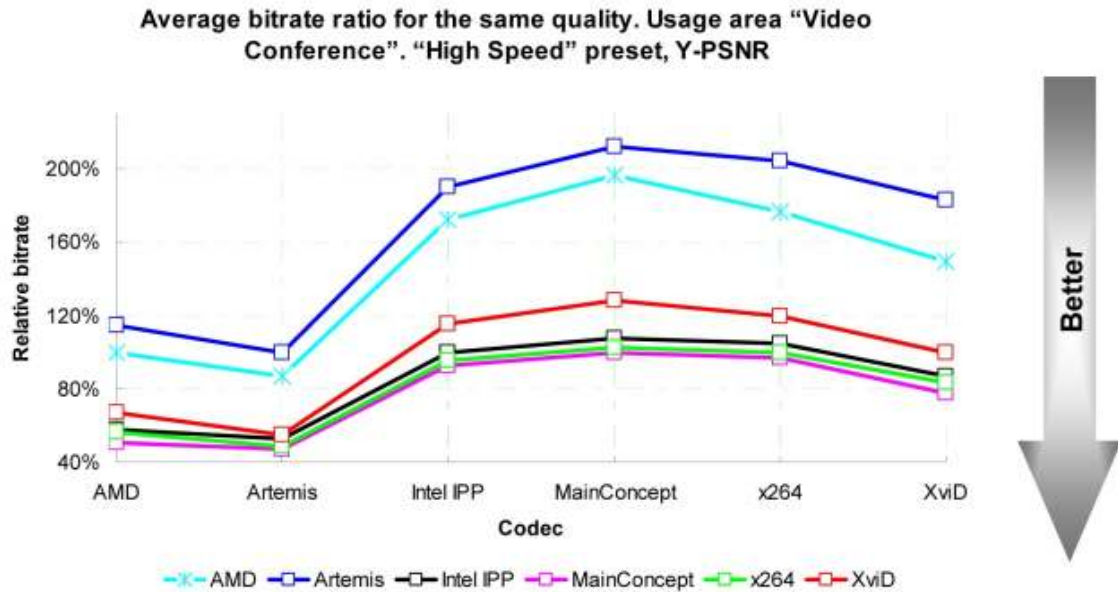
	AMD	Artemis x264	Intel IPP	MainConcept	X264	XviD
AMD	100.00%	114.61%	58.00%	50.89%	56.70%	66.70%
Artemis x264	87.25%	100.00%	52.53%	47.12%	48.89%	54.60%
Intel IPP	172.42%	190.37%	100.00%	92.81%	95.57%	115.30%
MainConcept	196.49%	212.20%	107.75%	100.00%	102.95%	128.19%
X264	176.37%	204.55%	104.63%	97.13%	100.00%	119.53%
XviD	149.93%	183.13%	86.73%	78.01%	73.66%	100.00%

Πίνακας 27 – Μέσο ποσοστό ρυθμού bit για συγκεκριμένη ποιότητα για τα videos video-διάσκεψης για ρύθμιση Υψηλής ταχύτητας

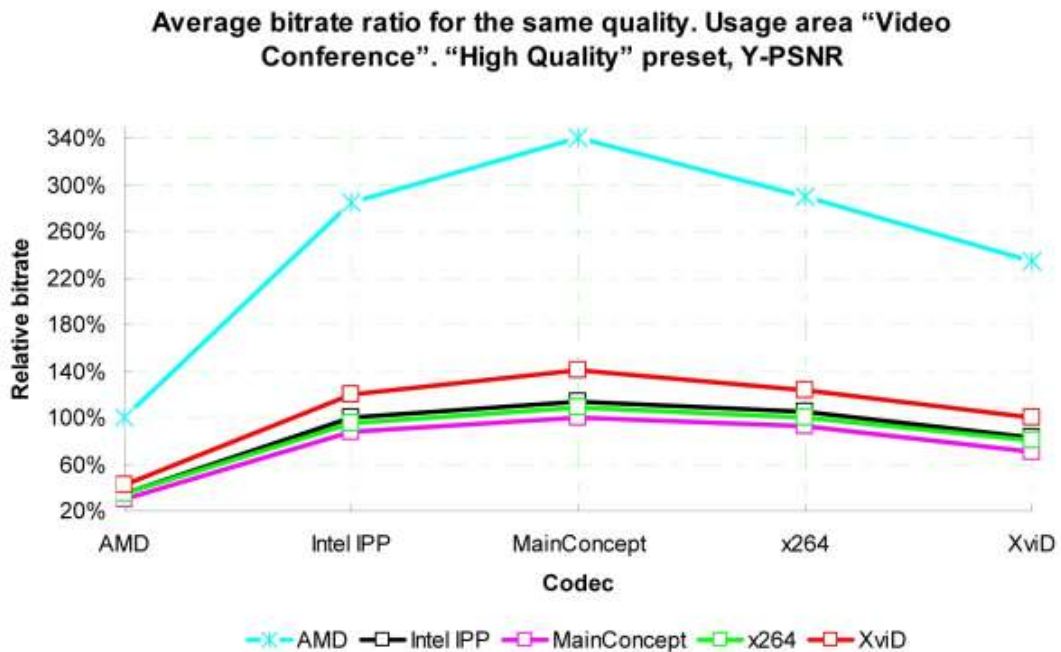
	AMD	Intel IPP	MainConcept	X264	XviD
AMD	100.00%	35.17%	29.36%	34.53%	42.73%
Intel IPP	284.30%	100.00%	88.08%	95.52%	120.20%
MainConcept	340.60%	113.53%	100.00%	108.39%	140.98%
X264	289.60%	104.69%	92.26%	100.00%	123.56%
XviD	234.05%	83.19%	70.93%	80.93%	100.00%

Πίνακας 28 - Μέσο ποσοστό ρυθμού bit για συγκεκριμένη ποιότητα για τα videos video-διάσκεψης για ρύθμιση Υψηλής ποιότητας

Οι **Εικόνες 37 και 38** αντιστοιχούν στους **Πίνακες 27 και 28**. Κάθε γραμμή αντιπροσωπεύει ένα codec. Οι τιμές στον κάθετο άξονα είναι ο μέσος ρυθμός bit που συγκρίνεται με τα codecs στον οριζόντιο άξονα.



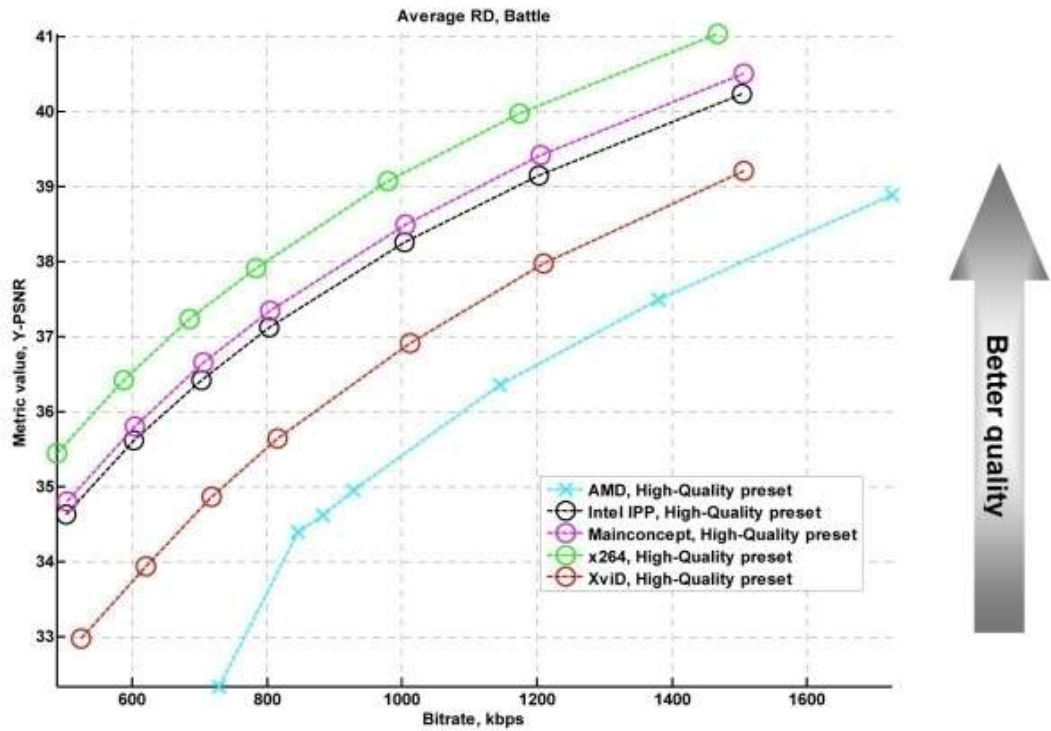
Εικόνα 37 - Μέσος ποσοστό ρυθμού bit για συγκεκριμένη ποιότητα για τα videos video-διάσκεψης για ρύθμιση Υψηλής ταχύτητας



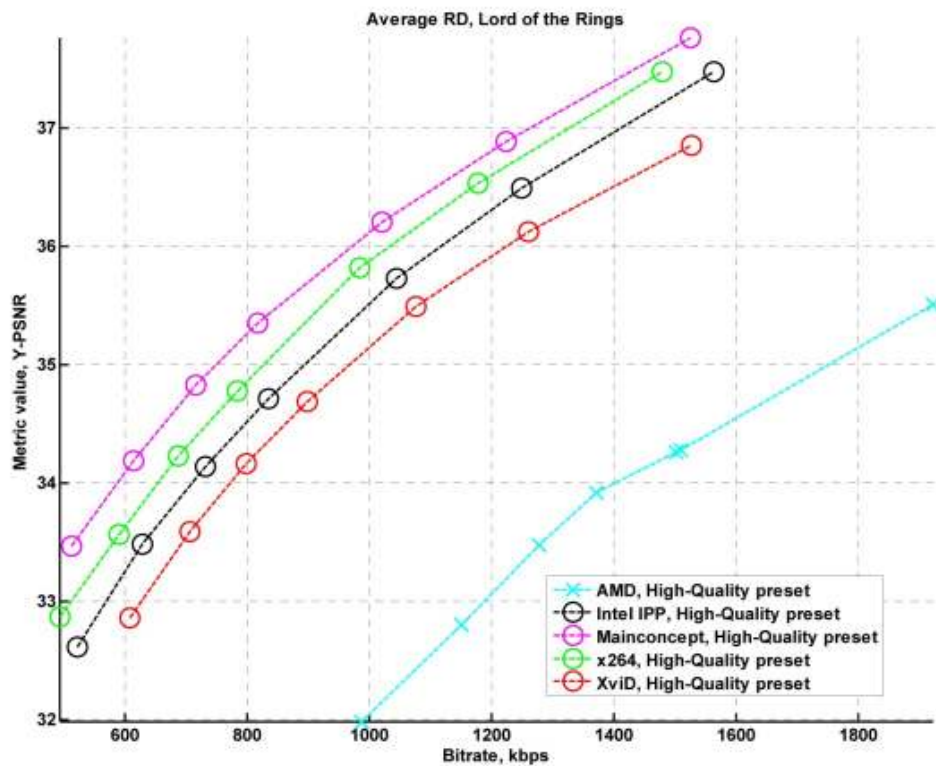
Εικόνα 38 - Μέσο ποσοστό ρυθμού bit για συγκεκριμένη ποιότητα για τα videos video-διάσκεψης για ρύθμιση Υψηλής ποιότητας

4.1.2.2 Ταινίες (movies)

Στις **Εικόνες 39 και 40** παρουσιάζονται οι γραφικές παραστάσεις των αποτελεσμάτων για ρύθμιση υψηλής ποιότητας. Το x264 φαίνεται να αποδίδει καλύτερα από όλα τα υπόλοιπα εκτός από το “Lord of the rings” όπου το MainConcept φαίνεται να έχει τα καλύτερα αποτελέσματα. Το AMD φαίνεται να αποδίδει χειρότερα από όλα ενώ το Intel IPP έρχεται τρίτο.



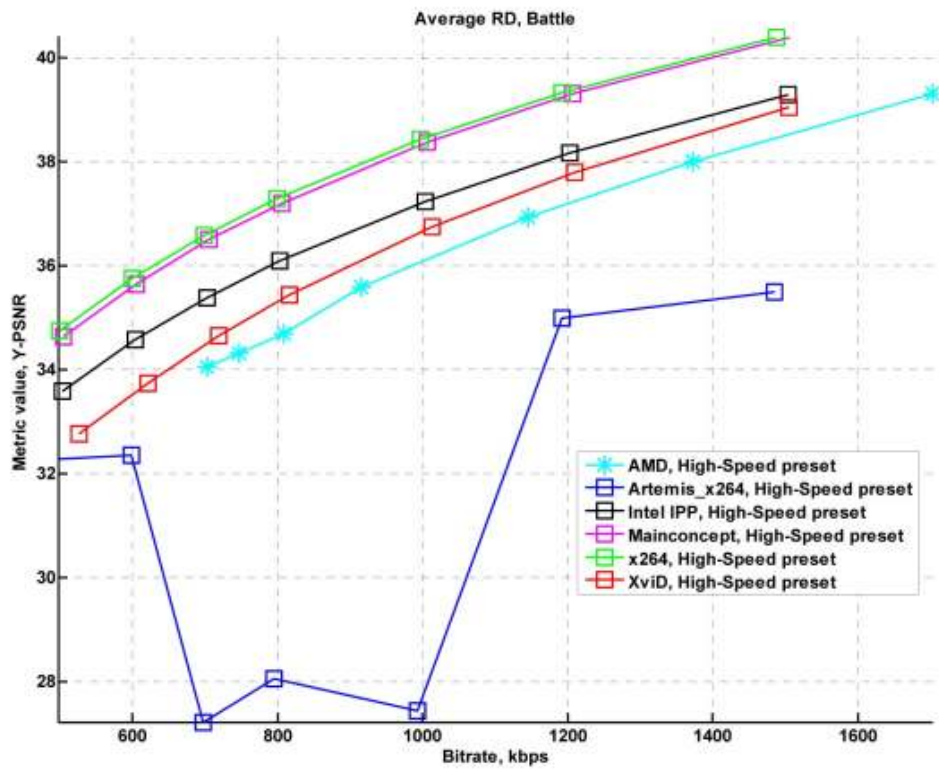
Εικόνα 39 – Διάγραμμα ρυθμού bit / Y-PSNR για το video Battle για ρύθμιση Υψηλής ποιότητας



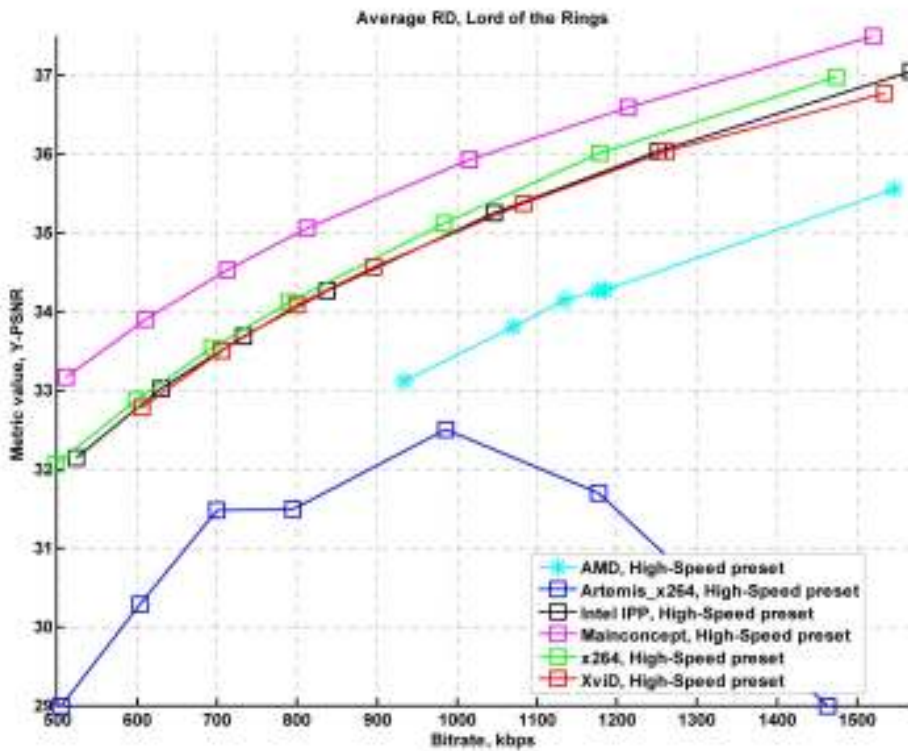
Εικόνα 40 - Διάγραμμα ρυθμού bit / Y-PSNR για το video Lord of the Rings για ρύθμιση Υψηλής ποιότητας

Στις **Εικόνες 41 και 42** παρουσιάζονται τα αποτελέσματα για τη ρύθμιση υψηλής ταχύτητας. Αυτό που παρατηρείται είναι τα ασταθή αποτελέσματα για το Artemis x264, όπου τα Y-PSNR αποτελέσματα είναι πολύ χαμηλά για όλα τα videos. Για τα Y-SSIM αποτελέσματα είναι σχετικά πιο σταθερά αλλά παραμένει αρκετά πίσω σε σχέση με τα υπόλοιπα codecs. Μόνο για το video “Ice Age” μπορούμε να πούμε ότι τα αποτελέσματά του είναι συγκρίσιμα των υπολοίπων.

Το x264 και το MainConcept φαίνεται να αποδίδουν καλύτερα σε αυτή τη περίπτωση. Το MainConcept είναι καλύτερο στο “Lord of the rings” και τα αποτελέσματά του είναι πολύ κοντά σε αυτά του x264 για τα υπόλοιπα videos. Το AMD φαίνεται να έχει τα γρηγορότερα αποτελέσματα αλλά με μειωμένη ποιότητα.



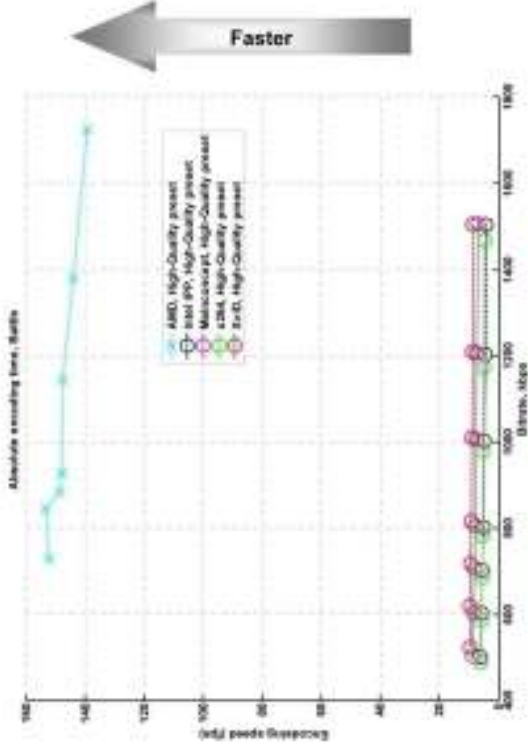
Εικόνα 41 – Διάγραμμα ρυθμού bit / Y-PSNR για το video Battle για ρύθμιση Υψηλής ταχύτητας



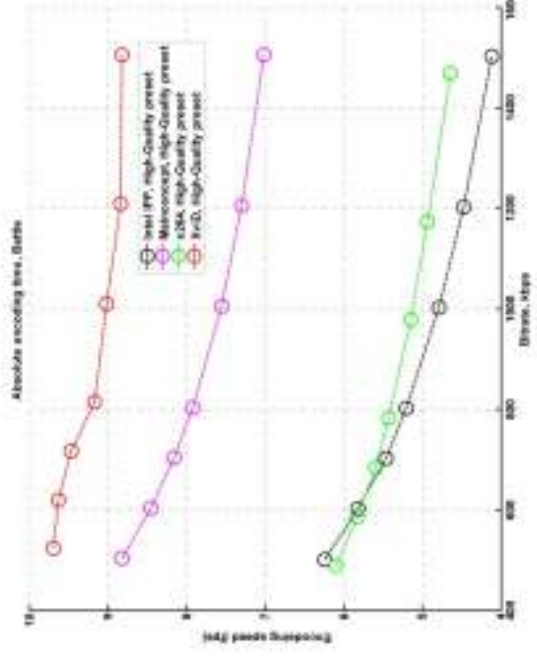
Εικόνα 42 - Διάγραμμα ρυθμού bit / Y-PSNR για το video Lord of the Rings για ρύθμιση Υψηλής ταχύτητας

4.1.2.2.1 Ταχύτητα κωδικοποίησης

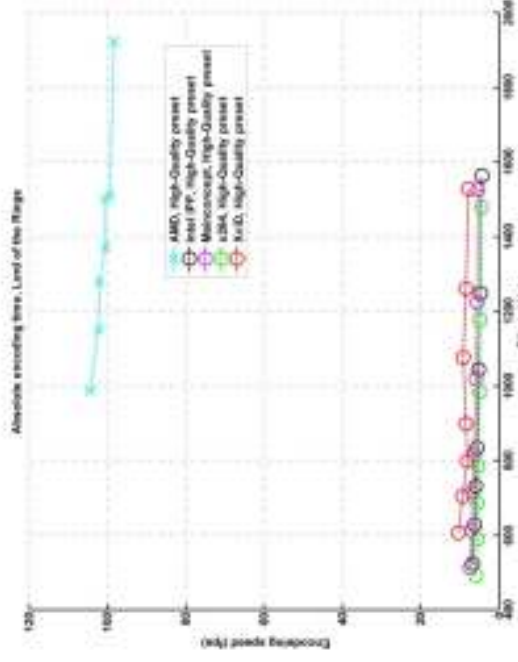
Στις **Εικόνες 43 μέχρι και 50** φαίνονται τα αποτελέσματα για την ταχύτητα κωδικοποίησης. Αξίζει να παρατηρηθεί η εξάρτηση της ταχύτητας κωδικοποίησης από το ρυθμό bit. Το Intel IPP H.264 φαίνεται να παρουσιάζει την πιο γρήγορη αύξηση στην ταχύτητα κωδικοποίησης με την αύξηση του ρυθμού bit. Τα αποτελέσματα για το XviD όπως παρατηρείται είναι ασταθή, ενώ το AMD παρουσιάζει αρκετά υψηλή ταχύτητα λόγω μιας ειδικής ρύθμισης του codec που προσανατολίζεται στην αύξησης της ταχύτητας κωδικοποίησης.



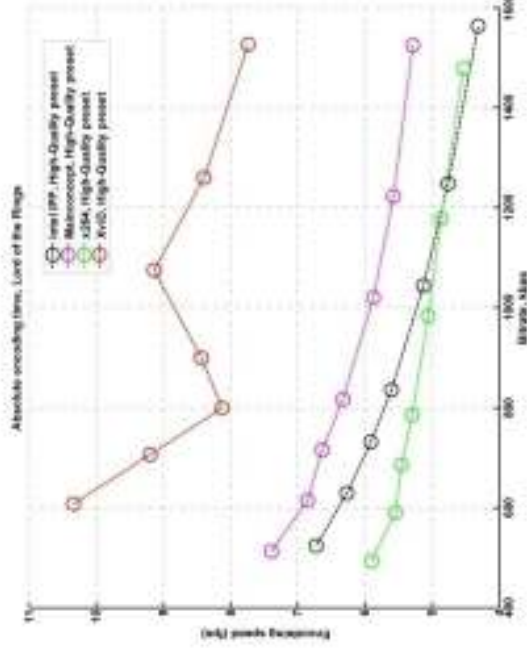
Εικόνα 43 - Ταχύτητα κωδικοποίησης για το video Battle για ρύθμιση Υψηλής ποιότητας



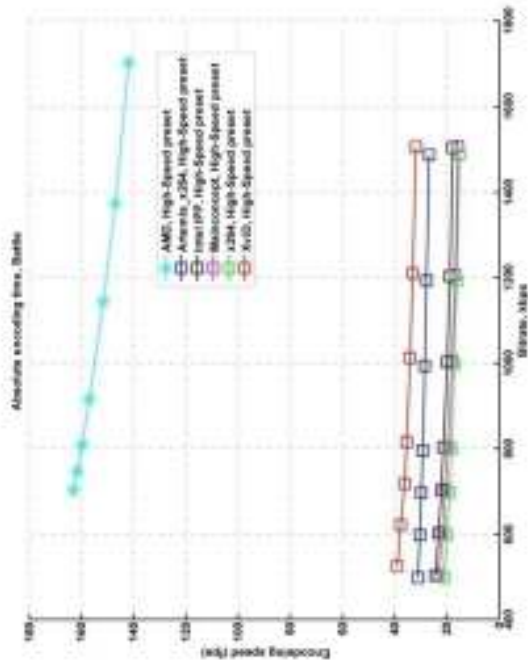
Εικόνα 44 - Ταχύτητα κωδικοποίησης για το video Battle για ρύθμιση Υψηλής ποιότητας, εκτός του codec AMD



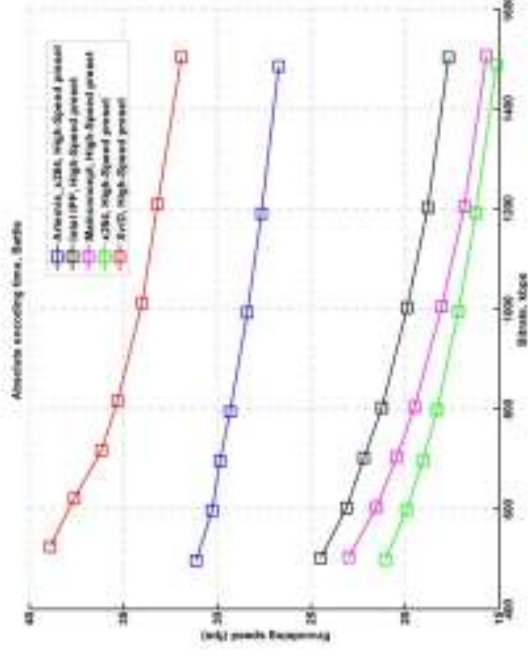
Εικόνα 45 - Ταχύτητα κωδικοποίησης για το video Lord of the Rings για ρύθμιση Υψηλής ποιότητας



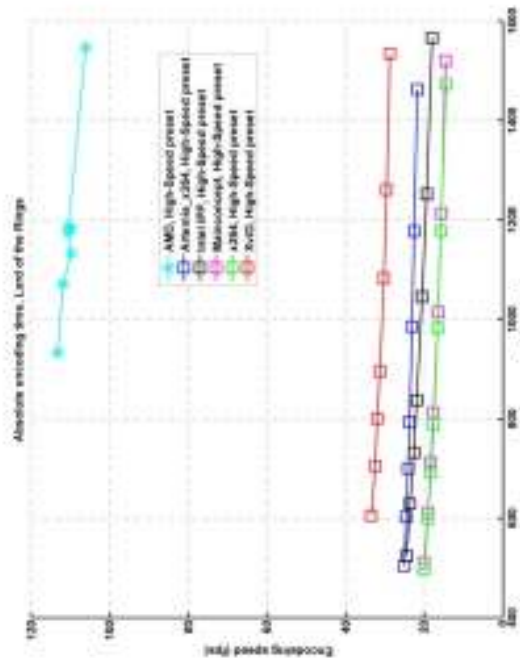
Εικόνα 46 - Ταχύτητα κωδικοποίησης για το video Lord of the Rings για ρύθμιση Υψηλής ποιότητας, εκτός του codec AMD



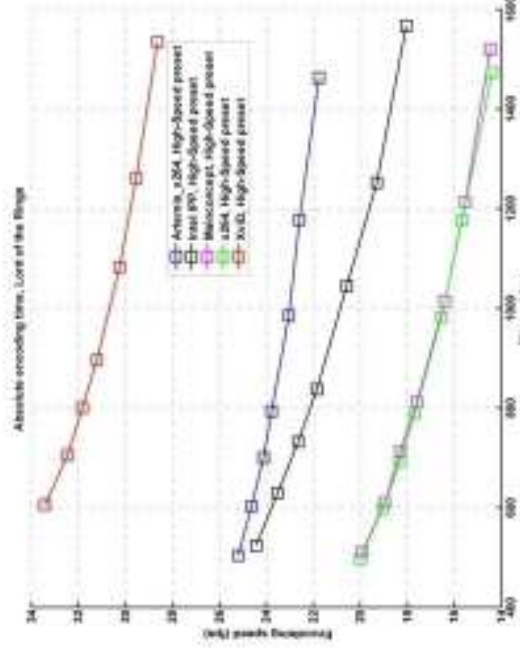
Εικόνα 47 - Ταχύτητα κωδικοποίησης για το video Battle για ρύθμιση Υψηλής ταχύτητας



Εικόνα 48 - Ταχύτητα κωδικοποίησης για το video Battle για ρύθμιση Υψηλής ταχύτητας, εκτός του codec AMD



Εικόνα 49 - Ταχύτητα κωδικοποίησης για το video Lord of the Rings για ρύθμιση Υψηλής ταχύτητας

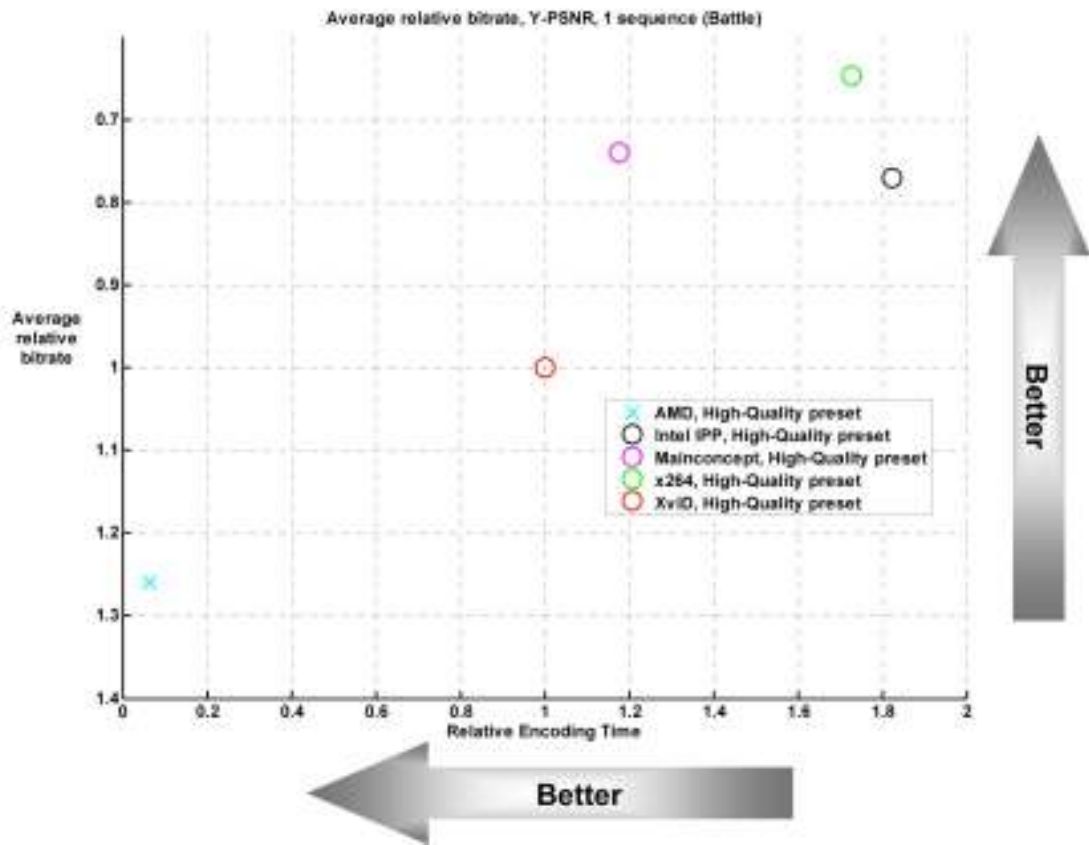


Εικόνα 50 - Ταχύτητα κωδικοποίησης για το video Lord of the Rings για ρύθμιση Υψηλής ταχύτητας, εκτός του codec AMD

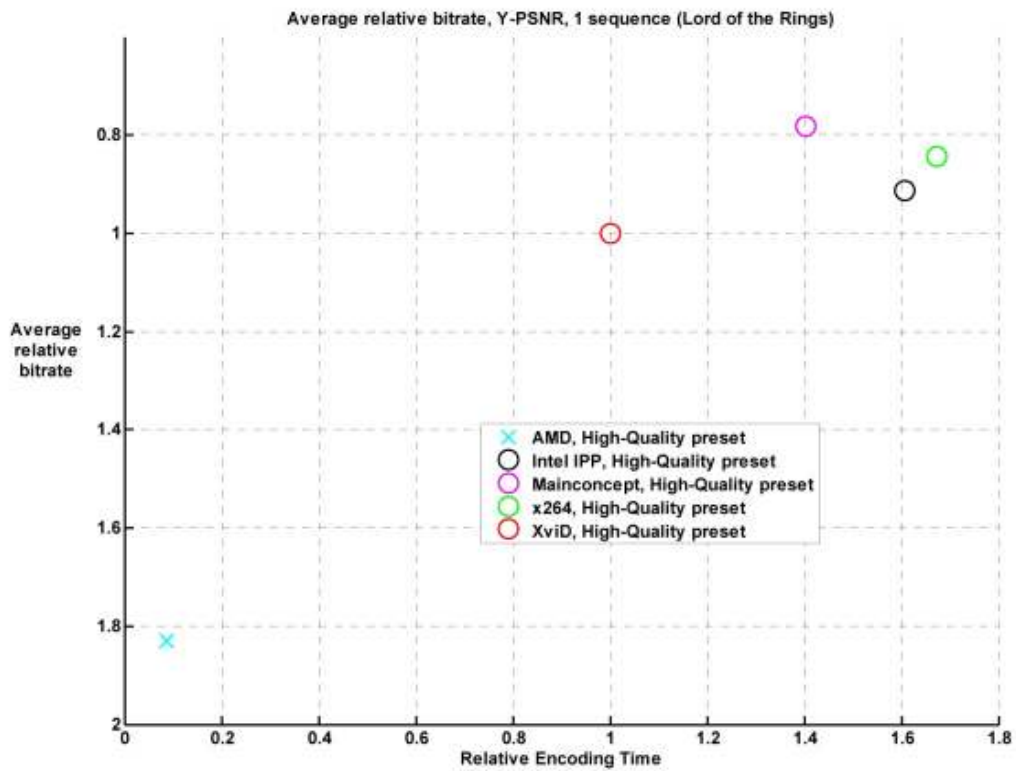
4.1.2.2.2 Ανταλλαγή ταχύτητας / ποιότητας

Τα διαγράμματα ανταλλαγής ταχύτητας/ποιότητας (Speed/Quality Tradeoff) δείχνουν τη σχέση της ταχύτητας και της ποιότητας για τα συγκεκριμένα codecs που ελέγχονται. Και εδώ το codec αναφοράς είναι το XviD και όσο αναφορά την ποιότητα και όσο αναφορά τη ταχύτητα. Έχει γίνει κανονικοποίηση σε όλες οι τιμές των codecs.

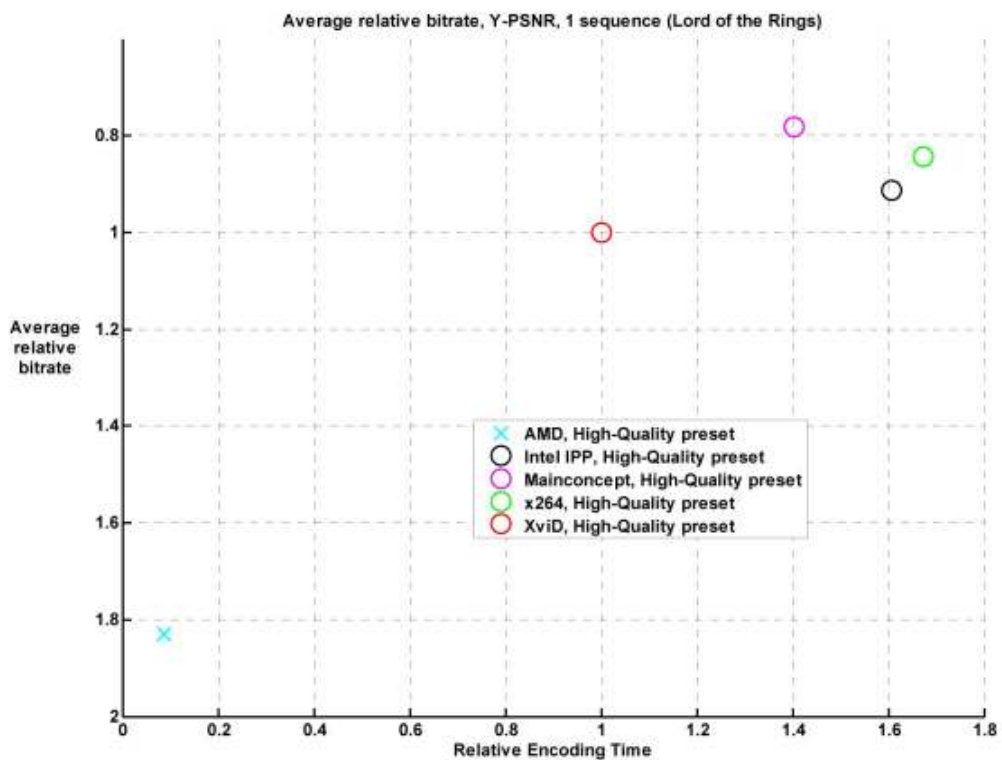
Στις **Εικόνες 51 μέχρι και 53** φαίνονται τα αποτελέσματα για ρύθμιση υψηλής ποιότητας. Το MainConcept codec φαίνεται να έχει καλύτερα αποτελέσματα από το Intel IPP για όλα τα videos. Επίσης το MainConcept φαίνεται να είναι καλύτερο από το x264 για το video “Lord of the rings”. Επίσης παρατηρείται ότι τα αποτελέσματα Y-PSNR και Y-SSIM είναι αρκετά όμοια.



Εικόνα 51 –Διάγραμμα ταχύτητας κωδικοποίησης / Y-PSNR για το video Battle για ρύθμιση Υψηλής ποιότητας

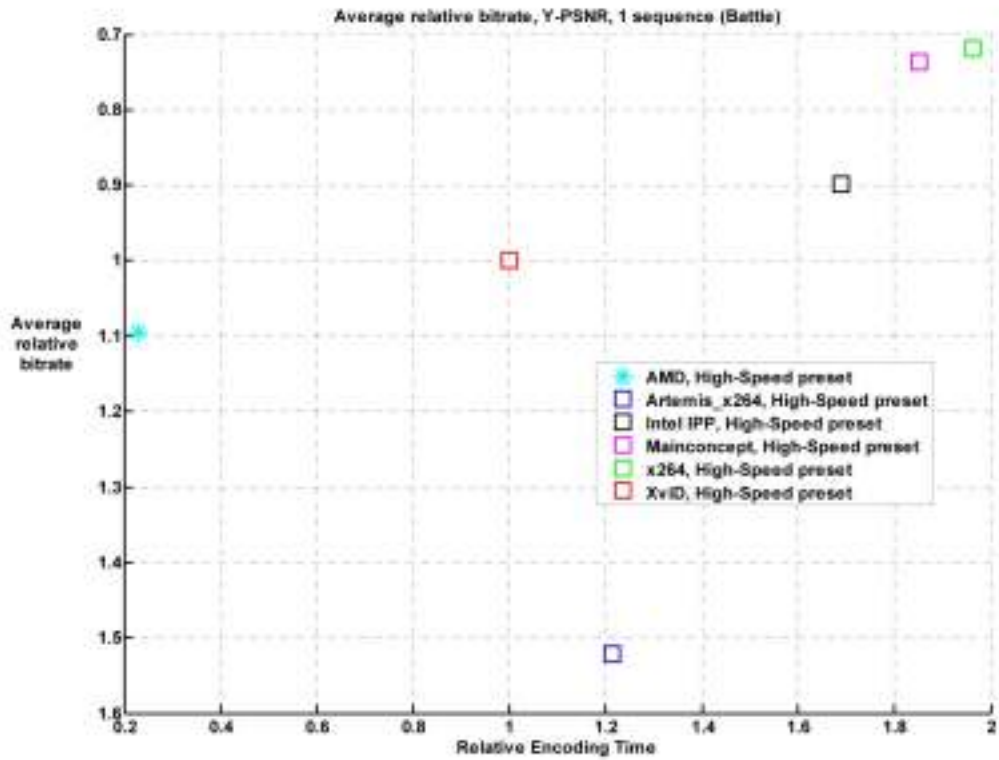


Εικόνα 52 - Διάγραμμα ταχύτητας κωδικοποίησης / Y-PSNR για το video Lord of the Rings για ρύθμιση Υψηλής ποιότητας

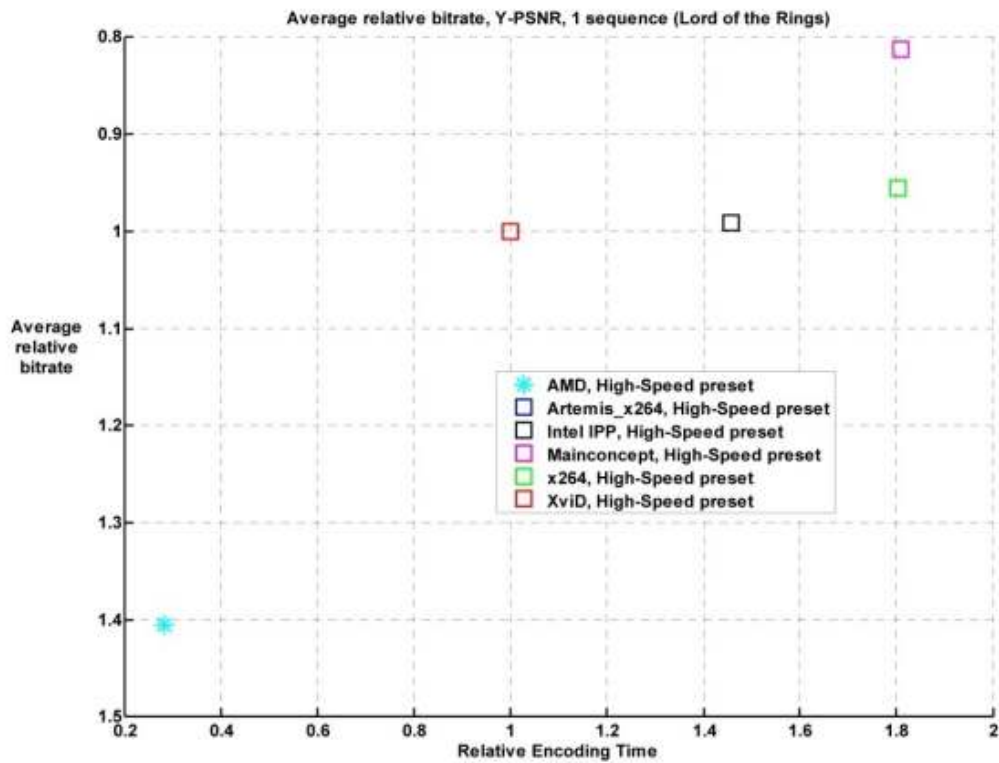


Εικόνα 53 - Διάγραμμα ταχύτητας κωδικοποίησης / Y-PSNR για όλα τα videos της ομάδας Ταινίες για ρύθμιση Υψηλής ποιότητας

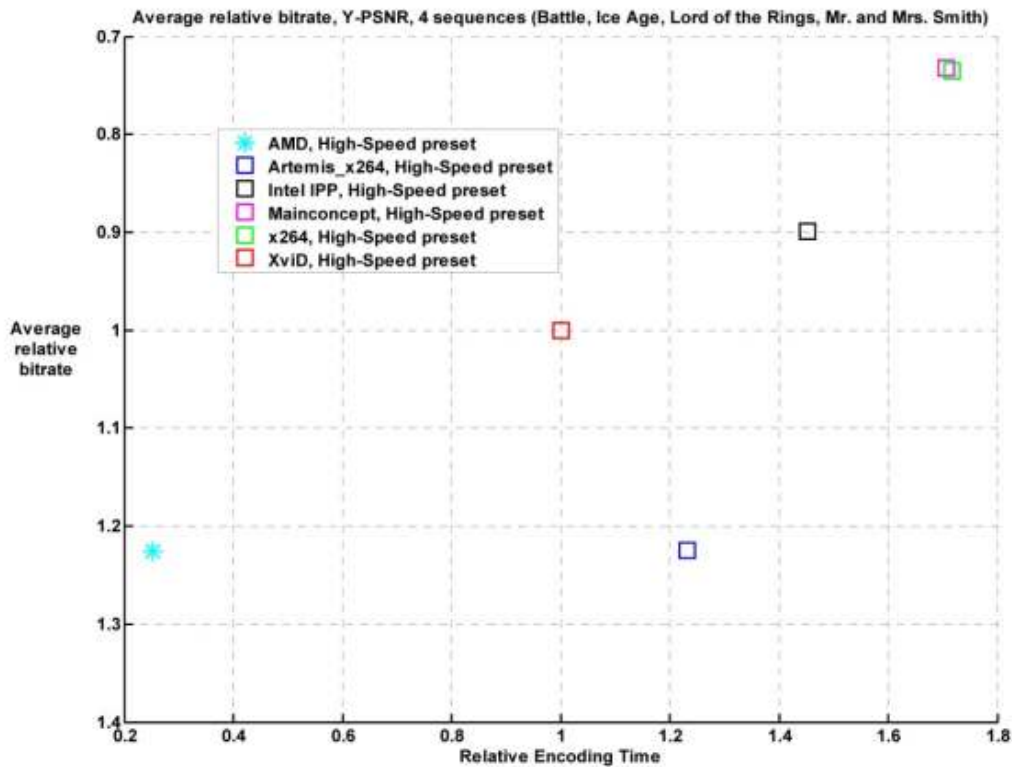
Στις **Εικόνες 54 μέχρι και 56** φαίνονται τα αποτελέσματα για ρύθμιση υψηλής ταχύτητας. Στο σύνολο των αποτελεσμάτων φαίνεται ότι το XviD είναι καλύτερο από το Artemis για όλα τα videos. Το x264 και το MainConcept έχουν αρκετά παρόμοια αποτελέσματα. Η ανάλυση κάθε video ξεχωριστά παρουσιάζει αρκετές διακυμάνσεις αποτελεσμάτων. Για παράδειγμα το MainConcept είναι καλύτερο από το x264 για το “Lord of the rings” αλλά χειρότερο για το “Ice age”. Επίσης για το Intel IPP και το video “Mr. And Mrs. Smith” το αποτέλεσμα εξαρτάται από τη μέτρηση που χρησιμοποιείται (PSNR ή SSIM).



Εικόνα 54 - Διάγραμμα ταχύτητας κωδικοποίησης / Y-PSNR για το video Battle για ρύθμιση Υψηλής ταχύτητας



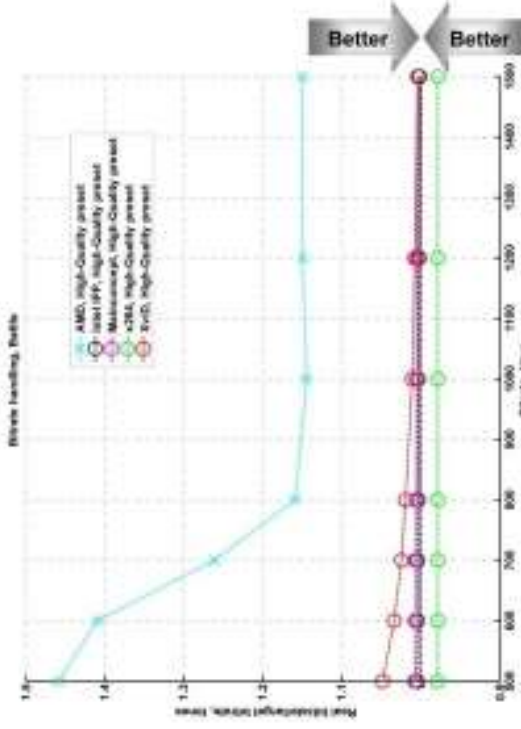
Εικόνα 55 - Διάγραμμα ταχύτητας κωδικοποίησης / Y-PSNR για το video Lord of the Rings για ρύθμιση Υψηλής ταχύτητας



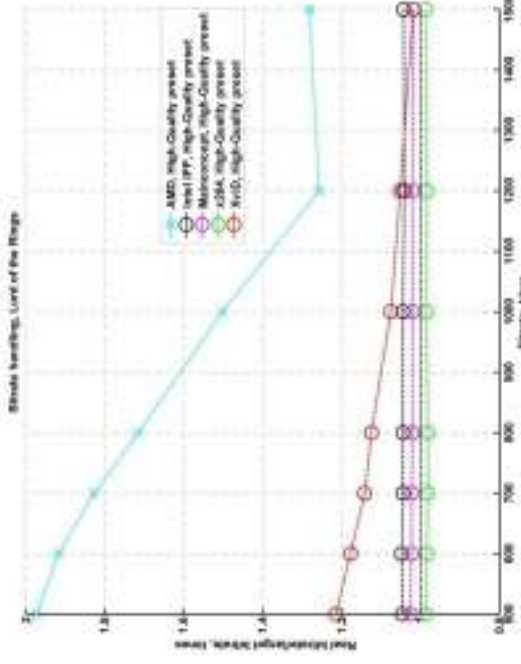
Εικόνα 56 - Διάγραμμα ταχύτητας κωδικοποίησης / Y-PSNR για όλα τα videos της ομάδας Ταινίες για ρύθμιση Υψηλής ταχύτητας

4.1.2.2.3 Διαχείριση ρυθμού bit

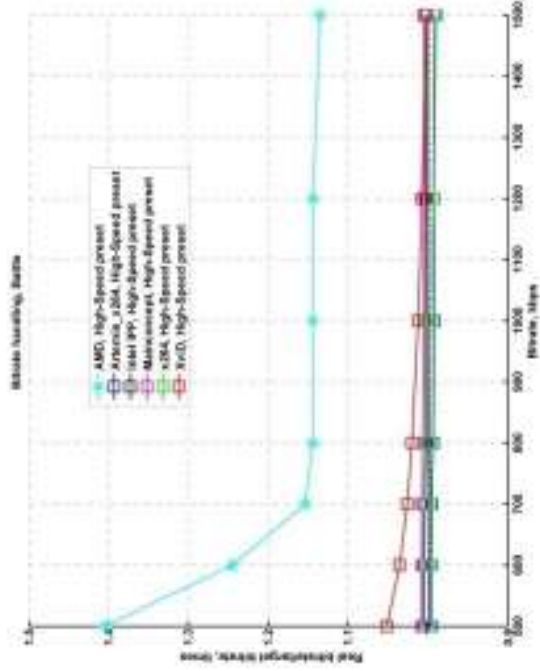
Στις Εικόνες 57 μέχρι και 60 φαίνονται τα διαγράμματα διαχείρισης του ρυθμού bit (bit rate handling) για τα διάφορα codecs για τα διάφορα videos. Το AMD δεν αποδίδει καθόλου καλά καθώς αυξάνει πάρα πολύ το ρυθμό bit (μέχρι και το διπλάσιο). Το XviD στους χαμηλούς ρυθμούς bit φαίνεται να τον αυξάνει, στους υπόλοιπους φαίνεται ότι τον διαχειρίζεται καλά αλλά όχι τόσο καλά όσο το MainConcept, το x264 και το Intel IPP.



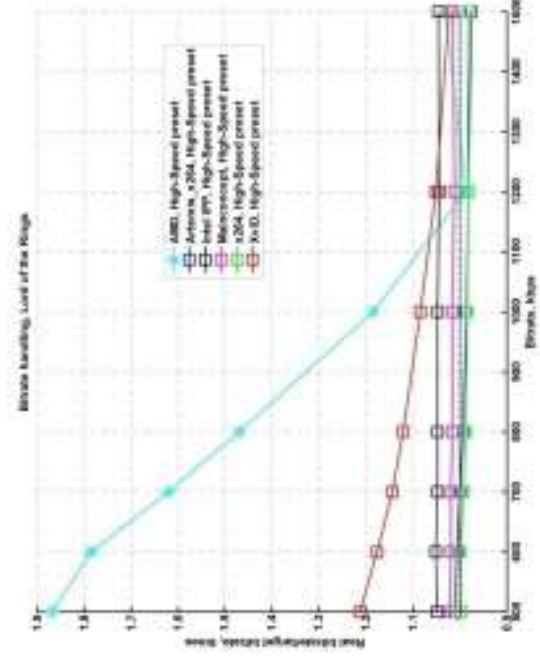
Εικόνα 57 – Διαχείριση ρυθμού bit για το video Battle για ρύθμιση Υψηλής ποιότητας



Εικόνα 58 – Διαχείριση ρυθμού bit για το video Lord of the Rings για ρύθμιση Υψηλής ποιότητας



Εικόνα 59 – Διαχείριση ρυθμού bit για το video Battle για ρύθμιση Υψηλής ταχύτητας



Εικόνα 60 – Διαχείριση ρυθμού bit για το video Lord of the Rings για ρύθμιση Υψηλής ταχύτητας

Οι Πίνακες 29 και 30 δείχνουν τους σχετικούς ρυθμούς bit για μία σταθερή ποιότητα για όλα τα codecs. Οι πίνακες αυτοί δεν περιέχουν καμιά πληροφορία όσο αναφορά την ταχύτητα κωδικοποίησης.

Για τη ρύθμιση υψηλής ταχύτητας παρατηρούμε ότι τα αποτελέσματα για το Artemis x264 εξαρτώνται κατά πολύ από τη μονάδα μέτρηση της ποιότητας (τα αποτελέσματα σε Y-SSIM είναι καλύτερα από τα αποτελέσματα σε Y-PSNR). Πέρα από αυτό το Artemis x264 φαίνεται να αποδίδει χειρότερα από το XviD (codec αναφοράς). Επίσης το AMD αποδίδει χειρότερα από το XviD. Τα codecs που αποδίδουν καλύτερα για τη ρύθμιση υψηλής ταχύτητας είναι το MainConcept και το x264. Το Intel IPP είναι λίγο πιο κάτω από τα τελευταία.

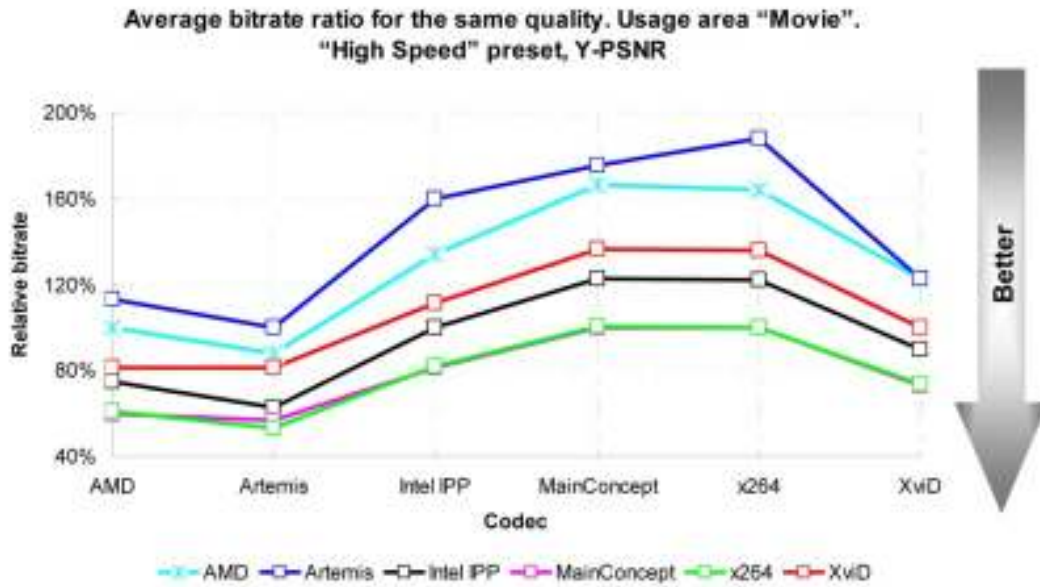
Όσο αναφορά τη ρύθμιση υψηλής ποιότητας τα codecs που αποδίδουν καλύτερα είναι και πάλι το x264 και το MainConcept με ένα μικρό προβάδισμα για το x264. Τρίτο έρχεται το Intel IPP και τελευταίο το AMD.

	AMD	Artemis x264	Intel IPP	MainConcept	X264	XviD
AMD	100.0%	113.4%	74.5%	60.0%	61.0%	81.6%
Artemis x264	88.2%	100.0%	62.7%	56.9%	53.2%	81.6%
Intel IPP	134.1%	159.6%	100.0%	81.6%	81.8%	111.2%
MainConcept	166.7%	175.7%	122.5%	100.0%	100.3%	136.6%
X264	164.1%	188.0%	122.2%	99.8%	100.0%	136.0%
XviD	122.6%	122.5%	89.9%	73.2%	73.5%	100.0%

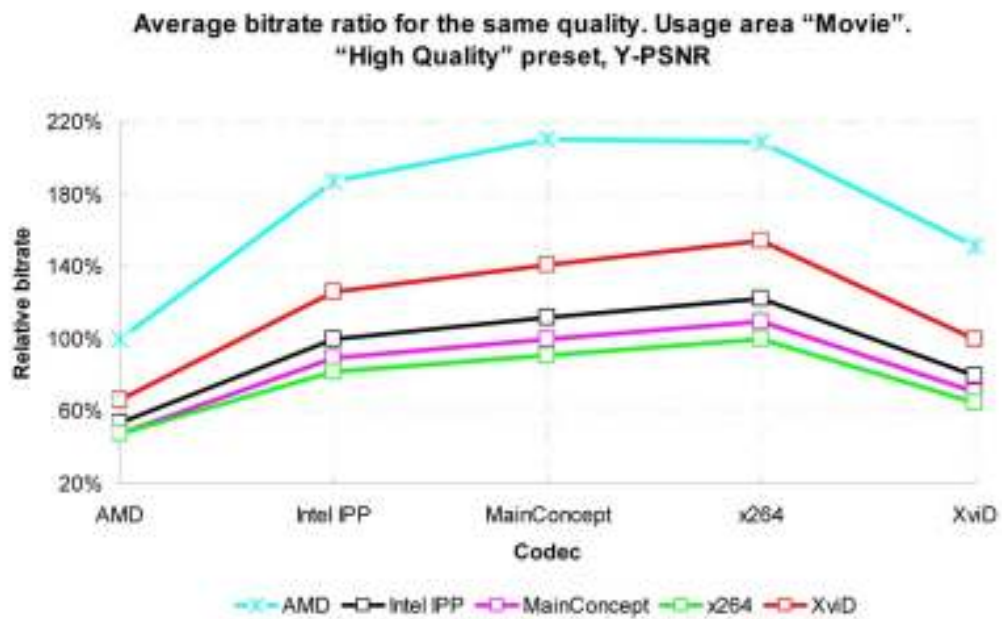
Πίνακας 29 – Μέσο ποσοστό ρυθμού bit για σταθερή ποιότητα για όλα τα videos της ομάδας Τανίες για ρύθμιση Υψηλής ταχύτητας

	AMD	Intel IPP	MainConcept	X264	XviD
AMD	100.0%	53.3%	47.5%	47.9%	66.0%
Intel IPP	187.5%	100.0%	89.7%	81.7%	125.7%
MainConcept	210.6%	111.5%	100.0%	91.2%	141.0%
X264	208.7%	122.4%	109.6%	100.0%	154.1%
XviD	151.6%	79.5%	70.9%	64.9%	100.0%

Πίνακας 30 - Μέσο ποσοστό ρυθμού bit για σταθερή ποιότητα για όλα τα videos της ομάδας Τανίες για ρύθμιση Υψηλής ποιότητας



Εικόνα 61 - Μέσο ποσοστό ρυθμού bit για σταθερή ποιότητα για όλα τα video της ομάδας Τανίες για ρύθμιση Υψηλής ταχύτητας



Εικόνα 62 - Μέσο ποσοστό ρυθμού bit για σταθερή ποιότητα για όλα τα videos της ομάδας Τανίες για ρύθμιση Υψηλής ποιότητας

4.1.2.3 HDTV

Οι Πίνακες 31 και 32 δείχνουν το σχετικό ρυθμό bit ανάμεσα στα codecs για συγκεκριμένη και σταθερή ποιότητα. Το MainConcept φαίνεται να αποδίδει καλύτερα από όλα τα υπόλοιπα για τη ρύθμιση υψηλής ταχύτητας, ακολουθούμενο από το x264. Οι διαφορές ανάμεσα σε αυτά τα δύο codecs εξαρτώνται αρκετά από τη μονάδα μέτρησης (PSNR ή SSIM). Έχουμε μία διαφορά 3% για το Y-PSNR και 16% για το Y-SSIM. Το AMD codec είναι το μοναδικό το οποίο αποδίδει χειρότερα από το XviD MPEG-4 (το οποίο οφείλεται στο σχεδιασμό του για βελτιστοποίηση της ταχύτητας).

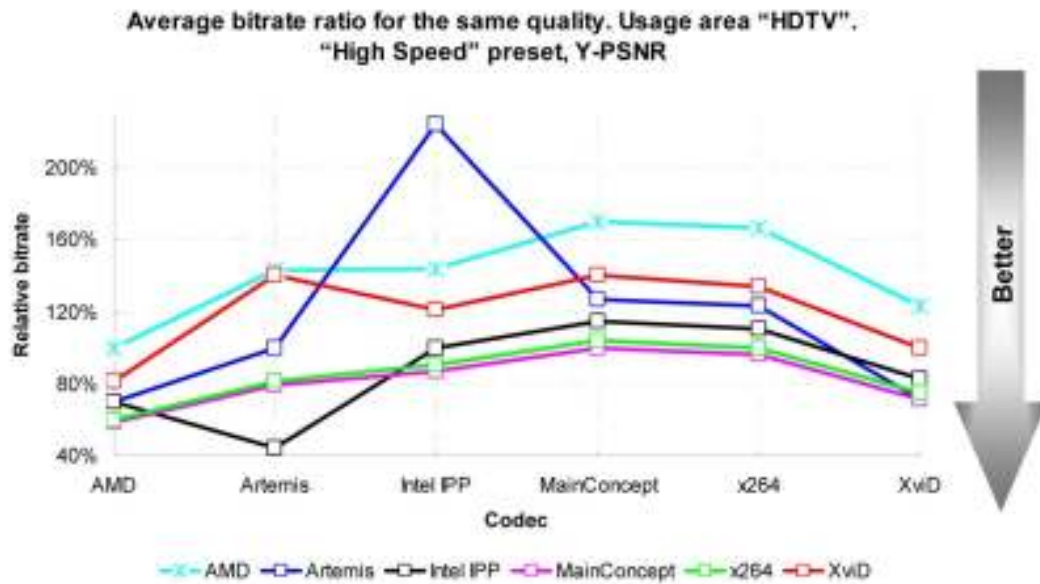
Στη ρύθμιση υψηλής ποιότητας τα αποτελέσματα για το MainConcept και x264 είναι αντίστροφα, αφού το x264 φαίνεται να αποδίδει καλύτερα.

	AMD	Artemis x264	Intel IPP	MainConcept	X264	XviD
AMD	100.0%	69.9%	69.7%	58.7%	60.1%	81.3%
Artemis x264	143.0%	100.0%	44.6%	79.1%	81.3%	140.2%
Intel IPP	143.6%	224.0%	100.0%	87.2%	90.7%	120.9%
MainConcept	170.3%	126.5%	114.7%	100.0%	103.9%	140.1%
X264	166.4%	123.0%	110.2%	96.3%	100.0%	133.6%
XviD	123.1%	71.3%	82.7%	71.4%	74.8%	100.0%

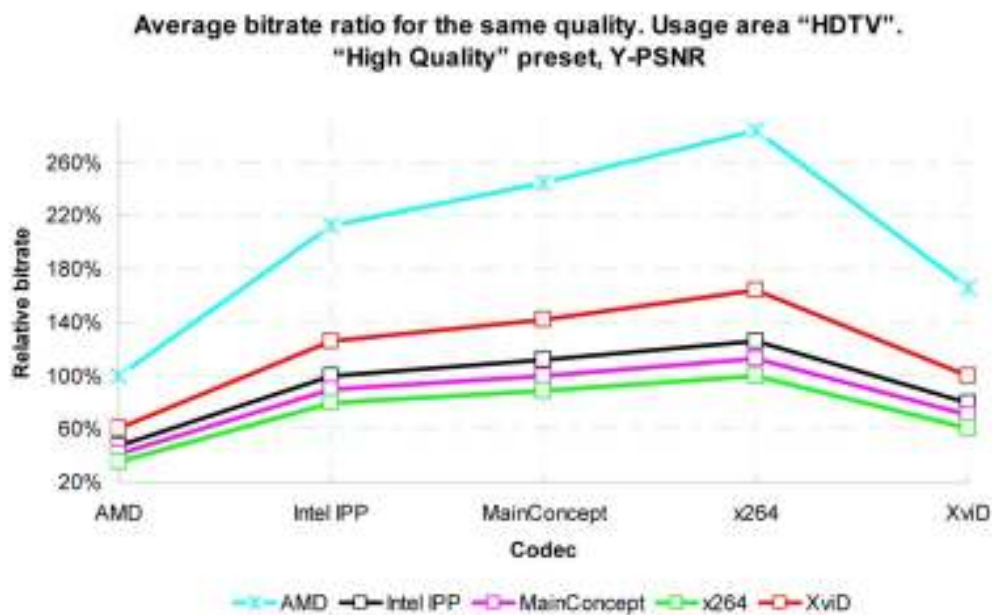
Πίνακας 31 - Μέσο ποσοστό ρυθμού bit για σταθερή ποιότητα για όλα τα videos της ομάδας HDTV για ρύθμιση Υψηλής ταχύτητας

	AMD	Intel IPP	MainConcept	X264	XviD
AMD	100.0%	47.1%	40.9%	35.3%	60.1%
Intel IPP	212.4%	100.0%	89.3%	79.2%	125.7%
MainConcept	244.5%	112.0%	100.0%	88.6%	142.4%
X264	283.6%	126.2%	112.9%	100.0%	164.5%
XviD	166.4%	79.6%	70.2%	60.8%	100.0%

Πίνακας 32 - Μέσο ποσοστό ρυθμού bit για σταθερή ποιότητα για όλα τα videos της ομάδας Τανίες για ρύθμιση Υψηλής ποιότητας



Εικόνα 63 - Μέσο ποσοστό ρυθμού bit για σταθερή ποιότητα για όλα τα videos της ομάδας HDTV για ρύθμιση Υψηλής ταχύτητας



Εικόνα 64 - Μέσο ποσοστό ρυθμού bit για σταθερή ποιότητα για όλα τα videos της ομάδας HDTV για ρύθμιση Υψηλής ποιότητας

4.1.3 Συμπεράσματα

4.1.3.1 Video-διάσκεψη

Την καλύτερη απόδοση στη video-διάσκεψη την έχουν το x264 και το MainConcept, με το MainConcept να είναι ελαφρώς καλύτερο. Τη χειρότερη επίδοση την παρουσίασε το AMD codec, το οποίο οφείλεται στο σχεδιασμό του για γρήγορη κωδικοποίηση (σχεδόν πέντε φορές γρηγορότερη από αυτή του XviD).

4.1.3.1.1 Υψηλή ποιότητα

Το MainConcept έχει την καλύτερη ποιότητα για όλα τα videos. Τα τρία καλύτερα codecs για αυτή τη ρύθμιση είναι:

1. MainConcept
2. X264
3. Intel IPP

Τα τρία αυτά codecs παρουσιάζουν και σχετικά καλή διαχείριση του ρυθμού bit.

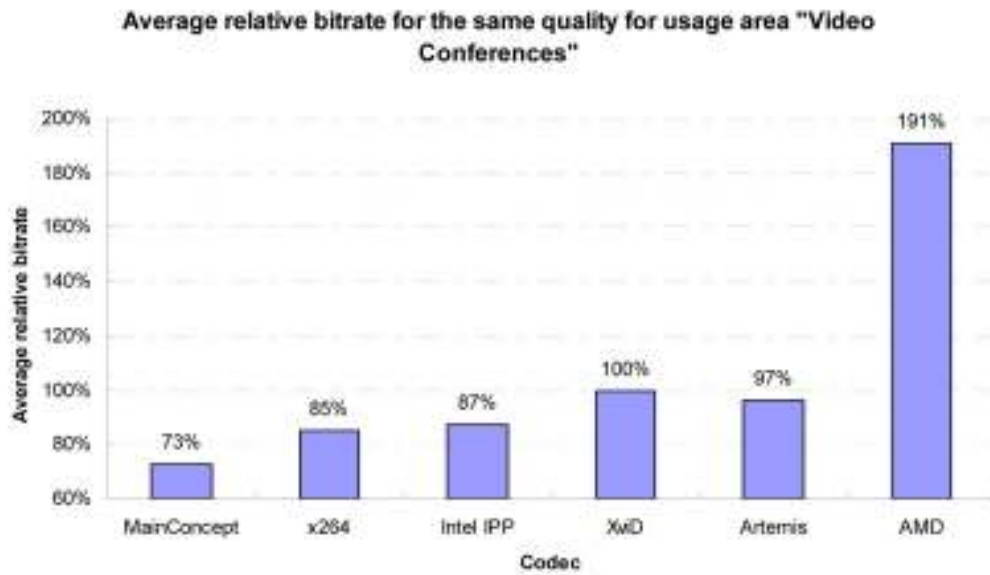
4.1.3.1.2 Υψηλή ταχύτητα

Και εδώ το MainConcept παρουσιάζει την καλύτερη ποιότητα για όλα τα videos. Η σειρά κατάταξης για αυτή τη ρύθμιση είναι:

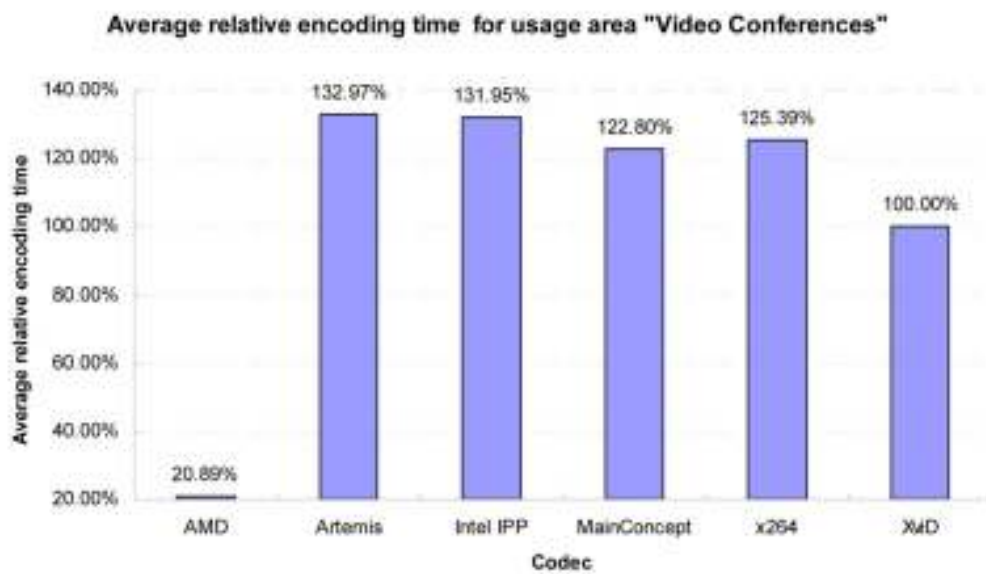
1. MainConcept
2. X264
3. Intel IPP
4. XviD

Το AMD και Artemis x264 βρίσκονται στην 5 ή 6 ανάλογα με τη μονάδα μέτρησης, Y-PSNR ή Y-SSIM.

Τα τρία πρώτα codecs παρουσιάζουν και σχετικά καλή διαχείριση ρυθμού bit.



Εικόνα 65 - Μέσος σχετικός ρυθμός bit για σταθερή ποιότητα για όλα τα videos της ομάδας video-διάσκεψη, για όλες τις ρυθμίσεις - YSSIM



Εικόνα 66 - Μέσος σχετικός χρόνος κωδικοποίησης για σταθερή ποιότητα για όλα τα videos της ομάδας video-διάσκεψη, για όλες τις ρυθμίσεις

4.1.3.2 Ταινίες

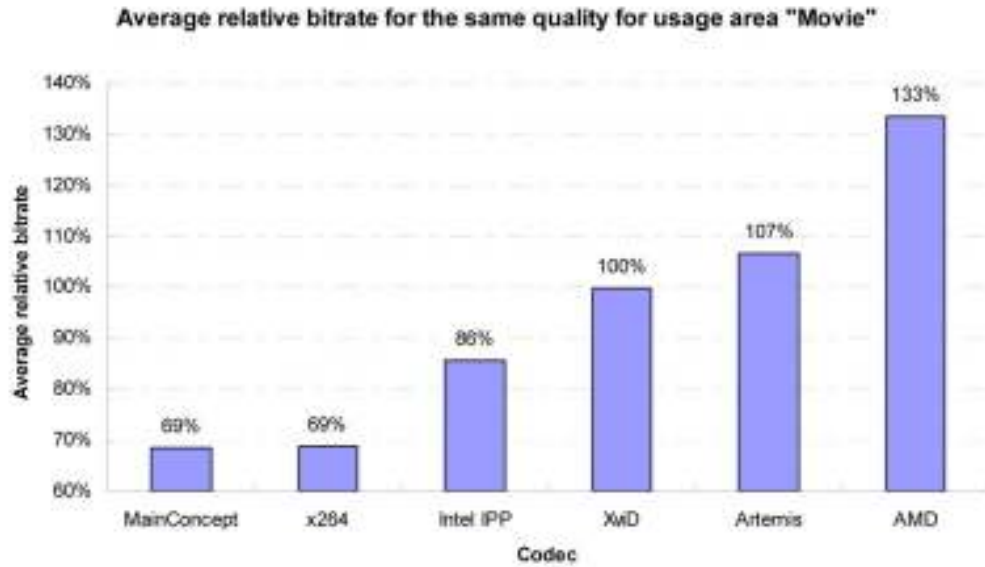
Σε αυτή την κατηγορία την καλύτερη απόδοση την παρουσιάζει το x264 και το MainConcept. Το AMD και εδώ παρουσιάζει μειωμένη απόδοση.

4.1.3.2.1 Υψηλή ποιότητα

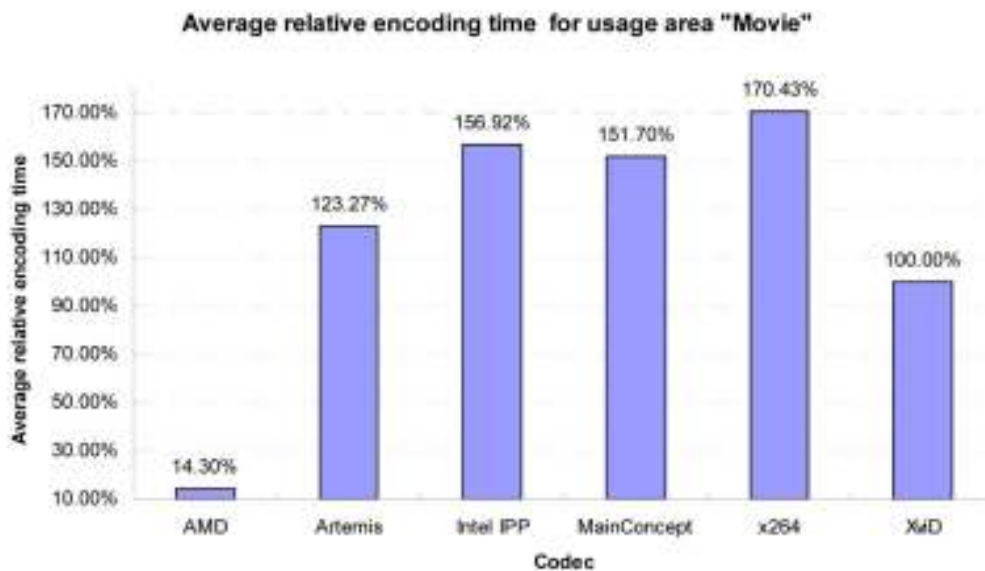
Σε αυτή τη ρύθμιση το x264 αποδίδει καλύτερα σε όλα τα videos εκτός από το “Lord of the rings” (για περίπου 10% ρυθμό bit για συγκεκριμένη – σταθερή ποιότητα) αλλά έχει πιο αργή απόδοση για ίδιο ρυθμό bit περίπου 30%. Για το video “Lord of the rings” το MainConcept είναι ταχύτερο και έχει καλύτερη ποιότητα. Και για τα δύο codecs η διαχείριση του ρυθμού bit είναι καλή. Το Intel IPP codec και εδώ καταλαμβάνει την τρίτη θέση. Σε κάποιες περιπτώσεις το Intel IPP αποδίδει ελαφρώς χειρότερα από το MainConcept, αλλά παρόλα αυτά αποδίδει σταθερά. Η σύγκριση των XviD και AMD με άλλα codecs είναι δύσκολη καθώς είναι γρηγορότερα και παρουσιάζουν μειωμένη ποιότητα για συγκεκριμένο – σταθερό ρυθμό bit. Η αντικειμενική ποιότητα του AMD είναι χαμηλότερη από αυτή του XviD αλλά η κωδικοποίηση του AMD codec είναι περίπου δέκα φορές γρηγορότερη. Επίσης το AMD codec παρουσιάζει αρκετά προβλήματα στη διαχείριση του ρυθμού bit (σε κάποια videos έχουμε μέχρι και διπλασιασμό του ρυθμού bit).

4.1.3.2.2 Υψηλή ταχύτητα

Τα αποτελέσματα για τη ρύθμιση υψηλής ταχύτητας είναι παρόμοια με αυτά της υψηλής ποιότητας. Την καλύτερη απόδοση παρουσιάζουν το x264 και το MainConcept και στην τρίτη θέση βρίσκεται πάλι το Intel IPP. Τα αποτελέσματα για την ανταλλαγή ταχύτητας / ποιότητας του Intel IPP σε αυτή τη ρύθμιση είναι βελτιωμένα καθώς είναι γρηγορότερο από το MainConcept. Το codec Artemis x264 είναι αρκετά ασταθές και είναι μόλις 20% ταχύτερο από το Intel IPP, αλλά με χαμηλότερη ποιότητα από το XviD. Το AMD codec και εδώ είναι αρκετά γρήγορο αλλά με χαμηλή ποιότητα.



Εικόνα 67 - Μέσος σχετικός ρυθμός bit για σταθερή ποιότητα για όλα τα videos της ομάδας Ταινίες, για όλες τις ρυθμίσεις - YSSIM



Εικόνα 68 - Μέσος σχετικός χρόνος κωδικοποίησης για σταθερή ποιότητα για όλα τα videos της ομάδας Ταινίες, για όλες τις ρυθμίσεις

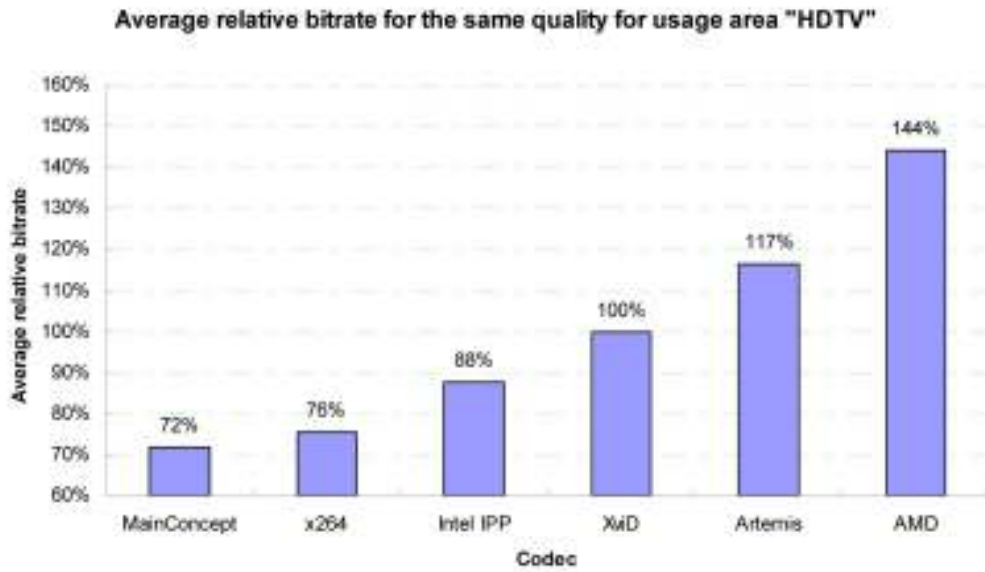
4.1.3.3 HDTV

4.1.3.3.1 Υψηλή ποιότητα

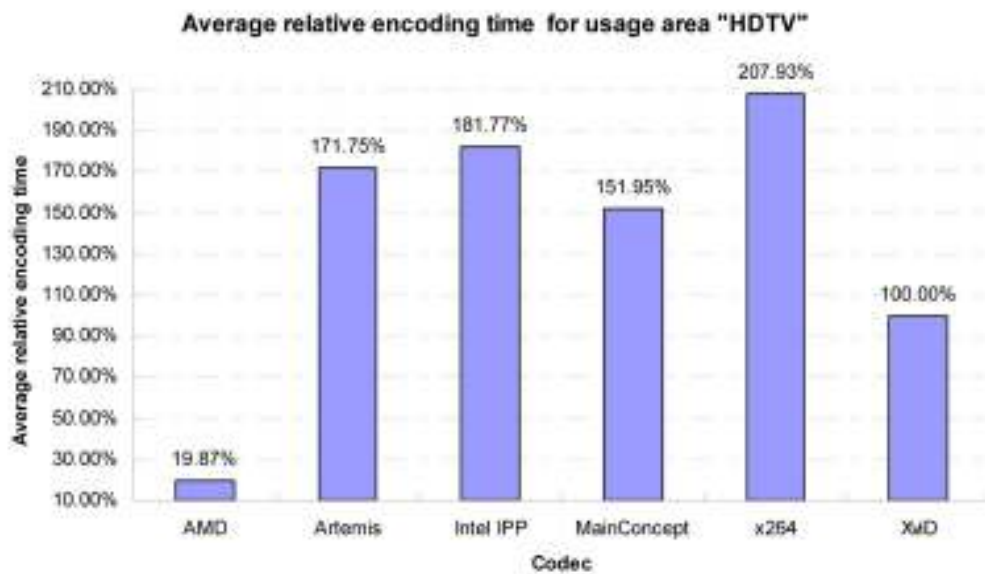
Σε αυτή τη σύγκριση το x264 και το MainConcept παρουσιάζουν την καλύτερη ποιότητα. Η ποιότητα κωδικοποίησης του x264 είναι καλύτερη από αυτή του MainConcept αλλά πιο αργή. Στη τρίτη θέση ποιοτικά έρχεται το Intel IPP παρόλο που είναι πιο αργό από το MainConcept. Και σε εδώ το AMD και το XviD είναι τα ταχύτερα από όλα τα codecs.

4.1.3.3.2 Υψηλή ταχύτητα

Σε αυτή τη περίπτωση το MainConcept αποδίδει καλύτερα και στην ποιότητα και στην ταχύτητα από το x264, το Intel IPP και το Artemis x264. Η ποιότητα του Artemis x264 είναι πολύ ασταθής και φαίνεται να είναι το χειρότερο codec σε αυτή τη κατηγορία.



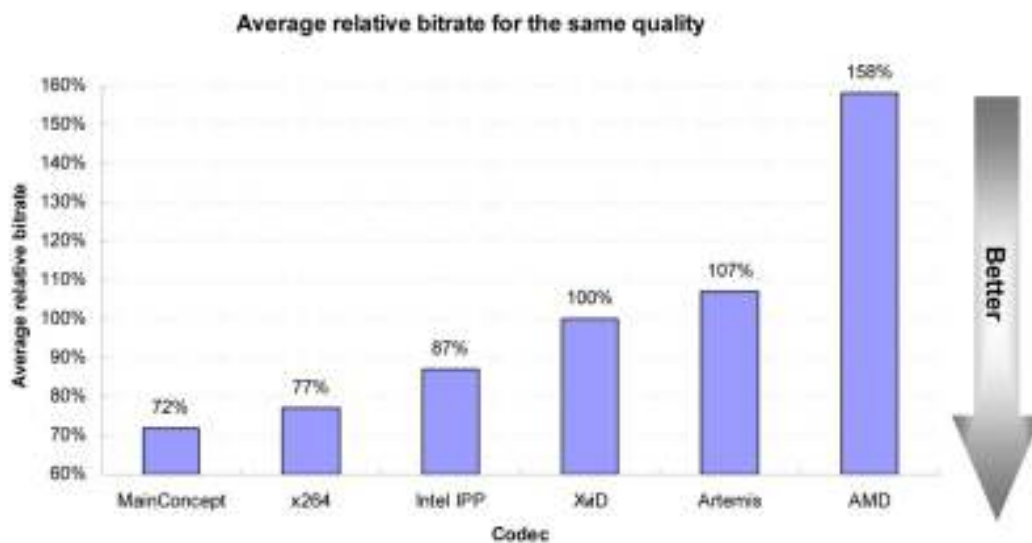
Εικόνα 69 - Μέσος σχετικός ρυθμός bit για σταθερή ποιότητα για όλα τα videos της ομάδας HDTV, για όλες τις ρυθμίσεις - YSSIM



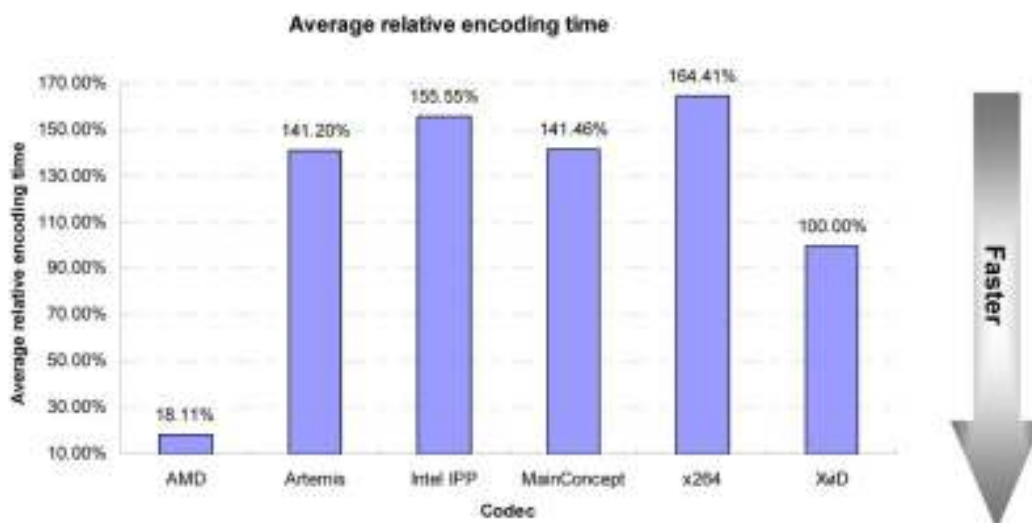
Εικόνα 70 - Μέσος σχετικός χρόνος κωδικοποίησης για σταθερή ποιότητα για όλα τα videos της ομάδας HDTV, για όλες τις ρυθμίσεις

4.1.3.4 Συνολικά συμπεράσματα

Στο σύνολο των αποτελεσμάτων μπορούμε να πούμε ότι την καλύτερη απόδοση την έχουν το MainConcept και το x264 και το Intel IPP να ακολουθεί. Το XviD (MPEG-4 ASP) codec είναι στο σύνολο καλύτερο από το AMD codec και το Artemis x264, πράγμα που αποδεικνύει ότι το AMD και το Artemis δεν χρησιμοποιούν όλα τα χαρακτηριστικά του προτύπου H.264. Ο βασικός λόγος που το AMD codec παρουσιάζει χαμηλή ποιότητα είναι η γρήγορη ταχύτητα κωδικοποίησης. Το XviD όπως και το AMD παρουσιάζουν προβλήματα στη διαχείριση του ρυθμού bit.



Εικόνα 71 - Μέσος σχετικός ρυθμός bit για σταθερή ποιότητα για όλα τα videos όλων των ομάδων, για όλες τις ρυθμίσεις - YSSIM



Εικόνα 72 - Μέσος σχετικός χρόνος κωδικοποίησης για σταθερή ποιότητα για όλα τα videos όλων των ομάδων, για όλες τις ρυθμίσεις

Η κατάταξη των codecs στο σύνολο των αποτελεσμάτων είναι:

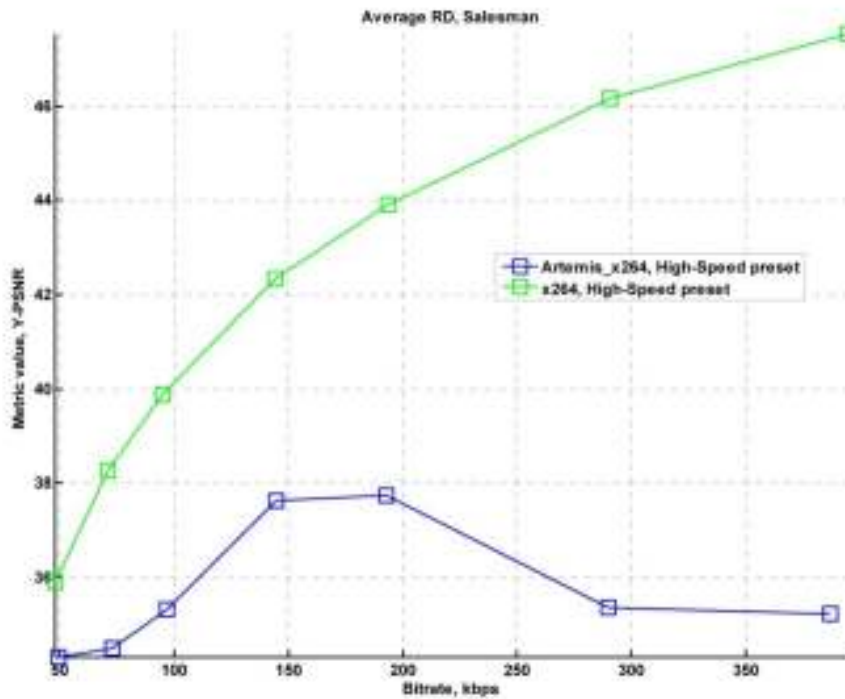
1. MainConcept
2. X264
3. Intel IPP
4. XviD
5. Artemis x264
6. AMD

Η κατάταξη αυτή βασίζεται μόνο στα αποτελέσματα της ποιότητας (**Εικόνα 71**) και δεν λαμβάνεται υπόψη η ταχύτητα κωδικοποίησης.

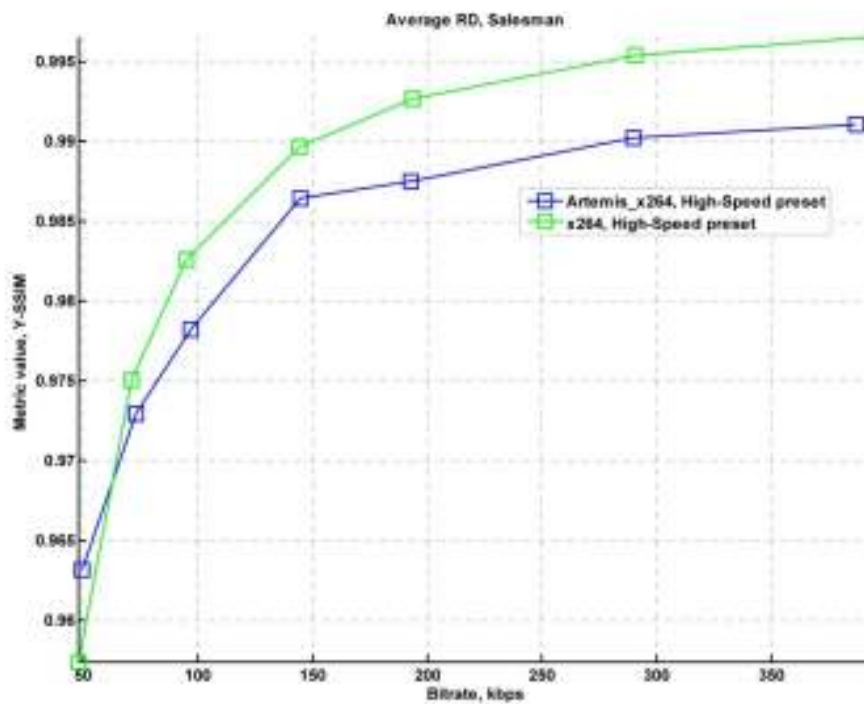
Η διαφορά του MainConcept και του x264 είναι πολύ μικρή και έτσι μπορεί να θεωρηθεί ότι είναι τα επικρατέστερα codecs σε αυτό το πείραμα. Για το Artemis x264 πρέπει να σημειωθεί ότι δεν έγινε έλεγχος υψηλής ποιότητας (High Quality test) και έτσι η βαθμολογία του βασίζεται μόνο στα αποτελέσματα για υψηλή ταχύτητα (High Speed). Η ποιότητα του Artemis x264 (H.264) είναι χαμηλότερη από αυτή του XviD (MPEG-4 ASP) και αυτό σημαίνει ότι το Artemis δεν χρησιμοποιεί όλες τις δυνατότητες του προτύπου H.264. Η χαμηλή ποιότητα του AMD codec μπορεί να δικαιολογηθεί από την υψηλή ταχύτητα κωδικοποίησης.

4.1.4 Σύγκριση του Artemis x264 και του x264

Το Artemis x264 είναι μία μετατροπή της προηγούμενης έκδοσης του x264 και τα αποτελέσματα της σύγκρισης των δύο έχουν ενδιαφέρον. Στην **Εικόνα 73** φαίνονται τα διαγράμματα του ρυθμού παραμόρφωσης για το video Salesman με ρύθμιση υψηλής ταχύτητας.



Εικόνα 73 – Διάγραμμα ρυθμού bit / Y-PSNR για το video Salesman, για τα codecs Artemis H.264 και x264 για ρύθμιση Υψηλής ταχύτητας



Εικόνα 74 – Διάγραμμα ρυθμού bit / Y-SSIM για το video Salesman, για τα codecs Artemis H.264 και x264 για ρύθμιση Υψηλής ταχύτητας

Παρατηρείται ότι η καμπύλη για το Artemis x264 έχει διαφορετική κλίση για τις δύο διαφορετικές μονάδες μέτρησης. Η Y-PSNR δεν είναι μονότονη (monotonic) ενώ η Y-SSIM είναι μονότονη με μικρή μείωση στα 200kbps.

4.2 Σύγκριση DivX 6.2.5, MainConcept H.264, Intel H.264, VSS H.263, x264 και Apple H.264

Το πείραμα αυτό που παρουσιάζεται στο [16], πραγματοποιήθηκε τον Νοέμβριο του 2006 και συγκρίνονται τα codecs DivX 6.2.5 (MPEG-4 ASP codec), MainConcept H.264, Intel H.264, VSS H.263, x264, Apple H.264 (μερικώς ελεγχόμενο) (partial tested) και Sorenson H.264 (μερικώς ελεγχόμενο). Μία σύντομη περιγραφή των ελεγχόμενων codecs φαίνεται στον Πίνακα 33. Το codec MPEG-4 ASP είναι το codec αναφοράς.

Codec	Κατασκευαστής	Έκδοση
DivX	DivX, nc	6.2.5
VSS H.264 Codec Pron 3.0	Vanguard Software Solutions, Inc	3.0.7.5
MainConcept H.264/AVC	MainConcept AG	2.1.5217
Intel H.264	Intel Corp.	Dev. Version for 10.08.2006
X264	X264 Development Team	Version 544
Apple	Apple Computers Inc.	QuickTime 7.1.3 for Windows
Sorenson	Sorenson Media Inc.	Build 2.00.106.00

Πίνακας 33 - Codecs που χρησιμοποιήθηκαν για τη σύγκριση DivX 6.2.5, MainConcept H.264, Intel H.264, VSS H.263, x264 και Apple H.264

Τα videos που χρησιμοποιήθηκαν για την εξαγωγή των αποτελεσμάτων παρουσιάζονται στον Πίνακα 34.

Videos	Αριθμός πλαισίων	Ρυθμός πλαισίων	Ανάλυση και χρώματα
Foreman	300	30	352x288 (YV12)
Akiyo	300	25	352x288 (YV12)
Carphone	382	25	176x144 (YV12)
Battle	1599	24	704x288 (YV12)
Rancho	1237	24	704x288 (YV12)
Matrix	239	25	720x416 (YV12)
Futurama	292	25	720x576 (YV12)
Concert	390	25	1664x1088 (YV12)

Πίνακας 34 - Videos που χρησιμοποιήθηκαν για τη σύγκριση DivX 6.2.5, MainConcept H.264, Intel H.264, VSS H.263, x264 και Apple H.264

4.2.1 Πειραματική διαδικασία

Σκοπός της εξέτασης αυτής είναι η αξιολόγηση της ποιότητας του νέου προτύπου H.264 μέσα από αντικειμενικές μετρήσεις. Οι ρυθμίσεις του κάθε codec έγιναν από την ομάδα που το ανέπτυξε.

Η εξέταση διαιρείται σε τρεις τύπους εφαρμογών, οι οποίοι διαφέρουν όσο αναφορά τις απαιτήσεις στην ανάλυση, σε ρυθμό bit και στην ταχύτητα κωδικοποίησης.

- Video-διάσκεψη (videoconference)
 - ρυθμός bit 30 – 300 kbps
 - απαιτήσεις σε ταχύτητα για video CIF 200kbps: ελάχιστο 70fps για ρύθμιση υψηλής ταχύτητας και 30fps για υψηλή ποιότητα
- Ταινίες (movies)
 - ρυθμός bit 500 – 2000 kbps
 - απαιτήσεις σε ταχύτητα για video 4CIF 750kbps: ελάχιστο 10fps για ρύθμιση υψηλής ταχύτητας, 4fps για υψηλή ποιότητα (ενός περάσματος) και 1 fps για υψηλή ποιότητα δύο περασμάτων
- Τηλεόραση υψηλής ευκρίνειας (high definition television, HDTV)
 - ρυθμός bit 1 – 10 Mbps
 - απαιτήσεις σε ταχύτητα για video 1920x1080 3Mbps: ελάχιστο 0.4fps

Για κάθε codec οι ρυθμίσεις προέρχονται από την ομάδα ανάπτυξής του, εκτός από το DivX. Κάθε codec ελέγχθηκε τρεις φορές και το αποτέλεσμα είναι ο μέσος όρος των τριών αυτών μετρήσεων. Χρησιμοποιήθηκαν δύο τύπου video για την εξαγωγή των αποτελεσμάτων:

- Πηγαίο video (.yuv κατάληξη αρχείου) σε YV12 format
- Videos σε YV12 format (.avi κατάληξη αρχείου) Τα video αυτά χρησιμοποιήθηκαν για το DivX 6.2.5 codec και διαφέρουν από τα .yuv videos μόνο στην επικεφαλίδα

Κατά τη διάρκεια της εξέτασης μετρήθηκαν τα παρακάτω μεγέθη:

- PSNR για τα στοιχεία Y,U και V (Y, U, V components)
- SSIM για το στοιχείο Y
- VQM για το στοιχείο Y
- MSU ανεξαρτήτου φωτεινότητας PSNR για το στοιχείο Y (μόνο για το codec της Apple)

Εδώ παρουσιάζονται μόνο οι μετρήσεις του SSIM καθώς αποτελούν την πιο ακριβή μέτρηση σε αυτό που αντιλαμβάνεται το ανθρώπινο μάτι. Σημειώνεται ότι μερικά αποτελέσματα διαφέρουν από αυτά του SSIM, ανάλογα τη μέτρηση που λαμβάνεται υπόψη.

4.2.2 Video-διάσκεψη

Σε αυτό το κομμάτι εξετάζεται η συμπεριφορά των codecs για κωδικοποίηση videos video-διάσκεψης. Τα videos είναι σχετικά απλής κίνησης και χαμηλής ανάλυσης. Οι ρυθμοί bit

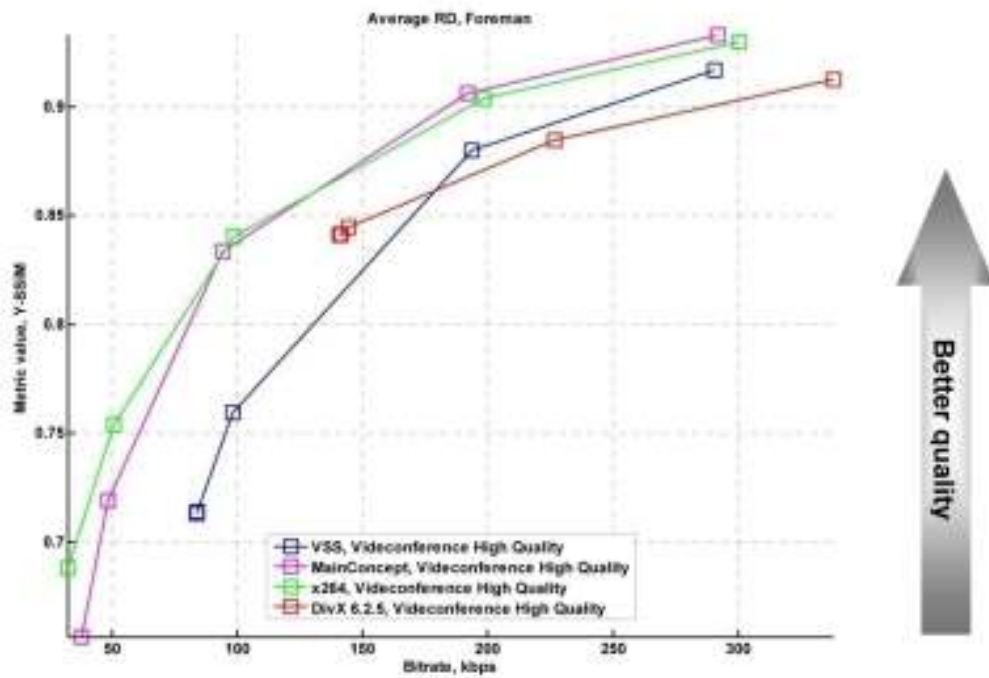
που επιλέχθηκαν είναι 30, 50, 100, 200 και 300kbps και απευθύνονται σε διαύλους περιορισμένων δυνατοτήτων, όπως ISDN ή xDSL διαύλους και δίκτυα κινητών.

Τα codecs που ελέγχθηκαν σε αυτό το κομμάτι είναι:

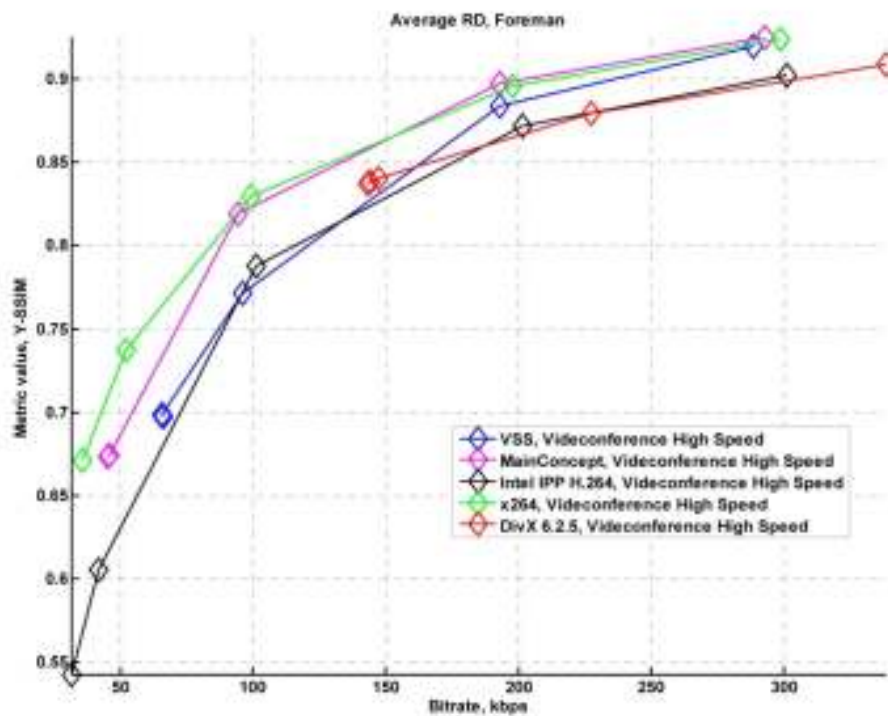
- DivX 6.2.1 (2 presets)
- MainConcept (2 presets)
- Intel H.264 (1 preset)
- VSS (2 presets)
- X264 (2 presets)

4.2.2.1 Αποτελέσματα

Όπως παρατηρείται στις **Εικόνες 75 και 76** το MainConcept codec και το x264 παρουσιάζουν παρόμοια ποιότητα για τη ρύθμιση υψηλής ποιότητας (με ένα μικρό προβάδισμα του x264) και είναι αρκετά καλύτερα από τα υπόλοιπα codecs. Παρόμοια αποτελέσματα παρουσιάζουν και για υψηλή ταχύτητα. Το VSS codec και το Intel codec παρουσιάζουν όμοια αποτελέσματα, με χειρότερη όμως ποιότητα video το οποίο μπορεί να δικαιολογηθεί από την μεγάλη ταχύτητα κωδικοποίησης. Επίσης παρατηρείται ότι το DivX δεν μπορεί να συμπίεσει video με πολύ χαμηλό ρυθμό bit.

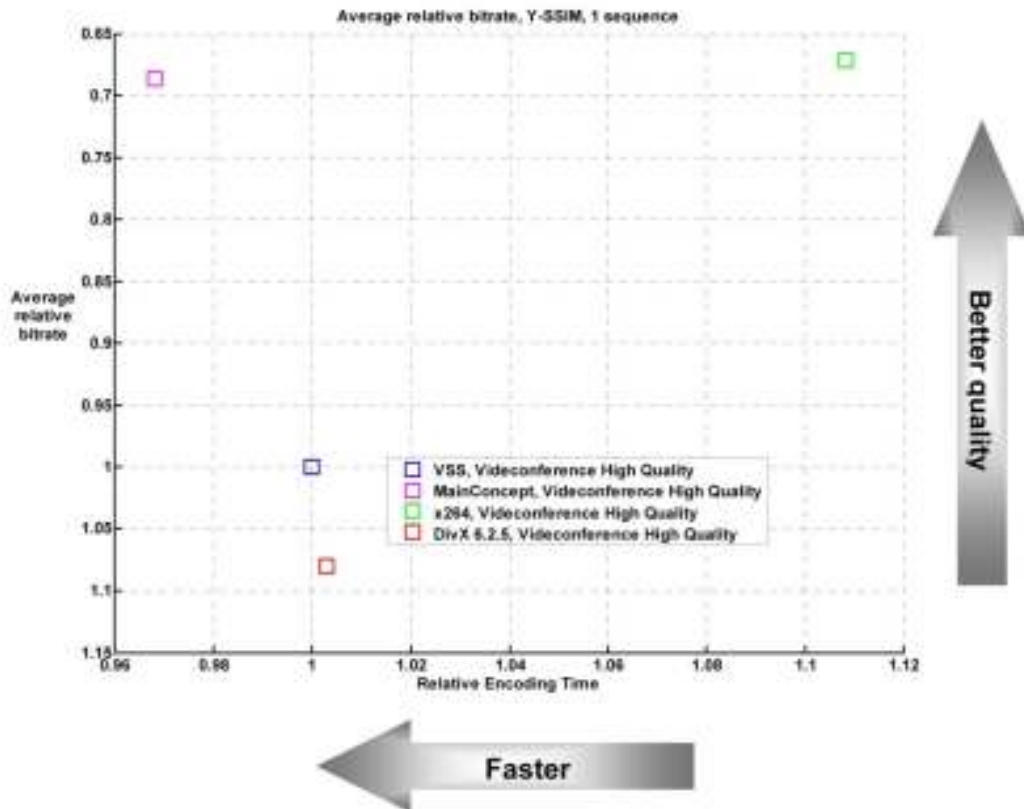


Εικόνα 75 – Διάγραμμα ρυθμού bit / Y-SSIM για το video Foreman για ρύθμιση Υψηλής ποιότητας



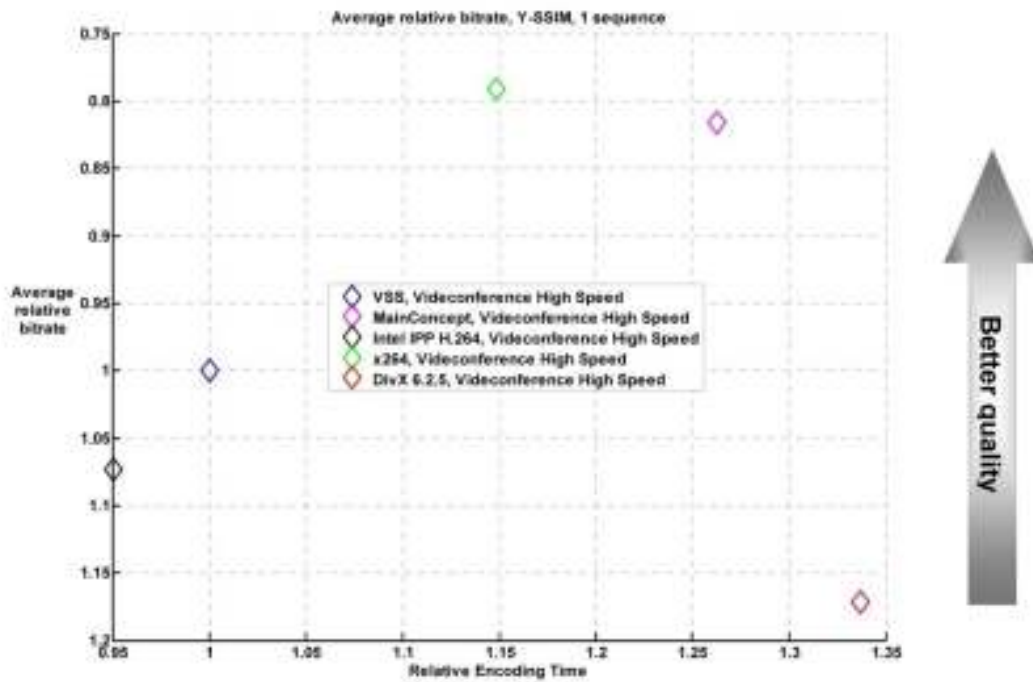
Εικόνα 76 – Διάγραμμα ρυθμού bit / Y-SSIM για το video Foreman για ρύθμιση Υψηλής ταχύτητας

Στην **Εικόνα 77** παρατηρούμε ότι το codec MainConcept έχει πολύ υψηλή ταχύτητα κωδικοποίησης ενώ ταυτόχρονα παρουσιάζει υψηλή ποιότητα, ενώ το x264 δείχνει λίγο καλύτερη ποιότητα αλλά με κόστος την μείωση της ταχύτητας. Το MainConcept παρουσιάζει καλύτερη ποιότητα από το VSS codec και το DivX codec με γρηγορότερη ταχύτητα κωδικοποίησης.



Εικόνα 77 – Σχετικός ρυθμός bit / σχετική ταχύτητα κωδικοποίησης για όλα τα videos της ομάδας video-διάσκεψη για ρύθμιση Υψηλής ποιότητας

Στην **Εικόνα 78** παρατηρούμαι ότι το codec της Intel παρουσιάζει αρκετά υψηλή ταχύτητα κωδικοποίησης. Επίσης το x264 παρουσιάζει καλύτερη ποιότητα από το MainConcept codec ενώ παράλληλα έχει υψηλότερη ταχύτητα κωδικοποίησης.



Εικόνα 78 – Σχετικός ρυθμός bit / σχετική ταχύτητα κωδικοποίησης για όλα τα videos της ομάδας video-διάσκεψη για ρύθμιση Υψηλής ταχύτητας

Στους **Πίνακες 35 και 36** παρουσιάζονται οι μέσες τιμές ρυθμού bit για μία σταθερή – συγκεκριμένη ποιότητα για όλα τα codecs. Οι μέσες τιμές προήλθαν από όλα τα video για αυτού του είδους την εφαρμογή (τηλεδιάσκεψη) (Foreman, Akiyo και Carphone).

	DivX	VSS	MainConcept	Intel H.264	X264
Μέσο ποσοστό μεγέθους αρχείου σε σχέση με το DivX	100.00%	78.67%	52.70%	-	62.58%

Πίνακας 35 – Μέσο μέγεθος αρχείου για σταθερή ποιότητα για όλα τα videos της ομάδας video-διάσκεψη, για ρύθμιση Υψηλής ποιότητας και για ρυθμούς bit 100 – 300 kbps

	DivX	VSS	MainConcept	Intel H.264	X264
Μέσο ποσοστό μεγέθους αρχείου σε σχέση με το DivX	100.00%	75.59%	58.4%	90.86%	62.37%

Πίνακας 36 - Μέσο μέγεθος αρχείου για σταθερή ποιότητα για όλα τα videos της ομάδας video-διάσκεψη, για ρύθμιση Υψηλής ταχύτητας και για ρυθμούς bit 100 – 300 kbps

4.2.2.2 Συμπεράσματα

Η κατάταξη των codecs σύμφωνα με τις μέσες τιμές του ρυθμού bit για σταθερή ποιότητα, όπως αυτές προέκυψαν από όλα τα videos για τη video-διάσκεψη είναι:

1. MainConcept
2. X264 (με μικρή καθυστέρηση)
3. VSS
4. Intel H.24
5. DivX (MPEG-4 ASP)

Όπως παρατηρείται το MPEG-4 ASP για εφαρμογές video-διάσκεψης αποδίδει καλύτερα από όλα τα codecs χρησιμοποιούν το πρότυπο H.264 (MPEG-4 AVC).

4.2.3 Ταινίες

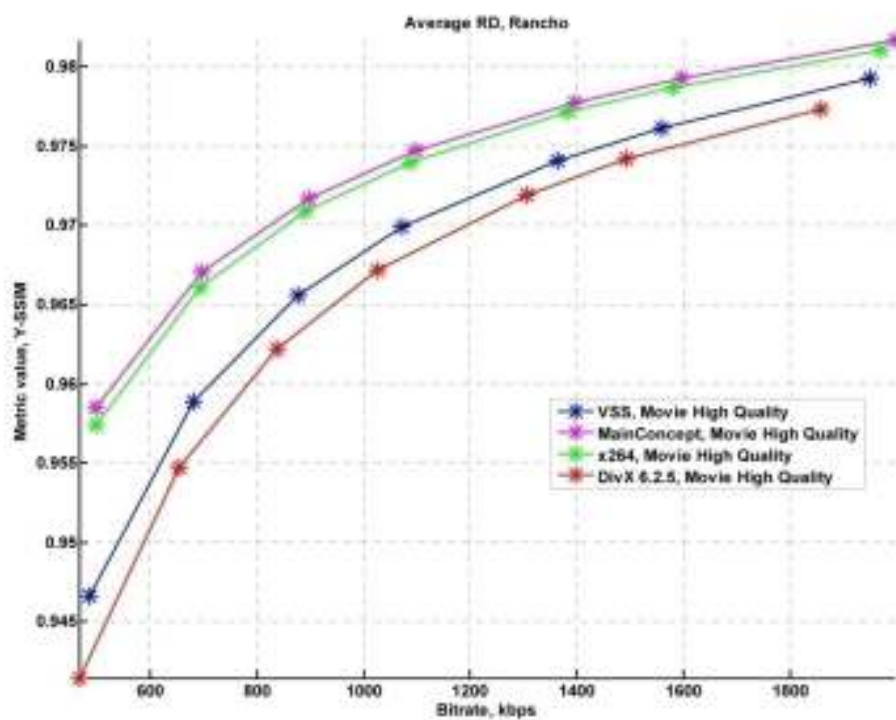
Σε αυτό το κομμάτι ελέγχεται η συμπεριφορά των codecs για την κωδικοποίηση ταινιών με πρότυπη ανάλυση (standard resolution, SDTV). Χρησιμοποιούνται διάφορα videos με διαφορετική πολυπλοκότητα κωδικοποίησης για διάφορους ρυθμούς bit (500, 700, 900, 1100, 1400, 1600 και 2000 kbps) τα οποία είναι συνηθισμένα στη κωδικοποίηση για CD-ROM, καλωδιακή τηλεόραση και τη ψηφιακή δορυφορική μετάδοση.

Τα codecs που ελέγχθηκαν είναι:

- DivX 6.2.1 (2 ρυθμίσεις)
- MainConcept (3 ρυθμίσεις)
- Intel H.264 (1 ρύθμιση)
- VSS (2 ρυθμίσεις)
- x264 (3 ρυθμίσεις)

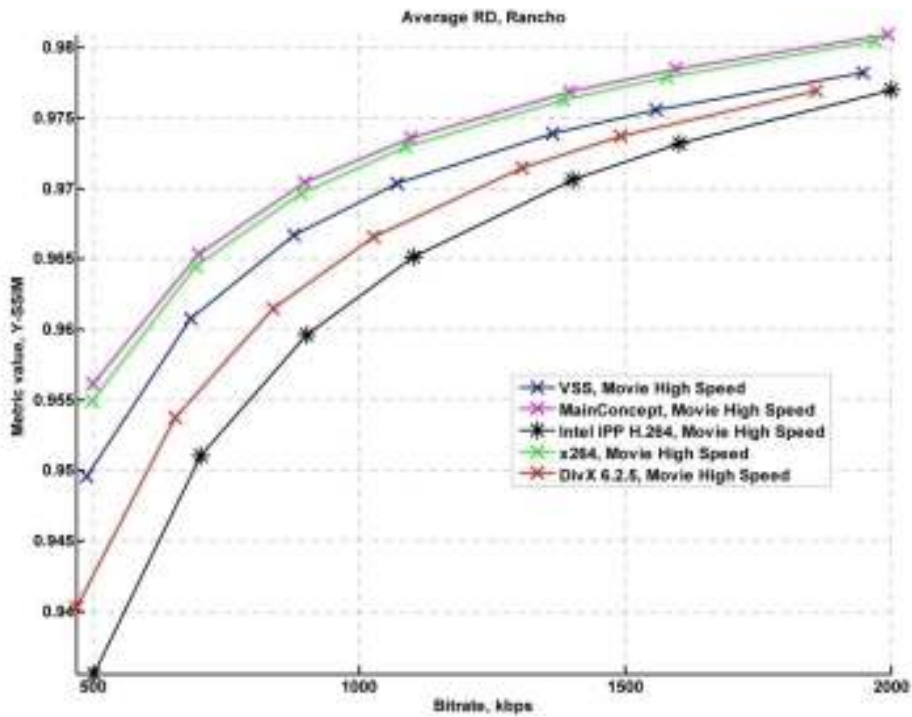
4.2.3.1 Αποτελέσματα

Στην **Εικόνα 79** παρατηρείται ότι όλα τα codecs που χρησιμοποιούν το πρότυπο H.264 παρουσιάζουν καλύτερη ποιότητα από το DivX το οποίο χρησιμοποιεί το πρότυπο MPEG-4 ASP. Ωστόσο αυτό δεν παρατηρείται σε όλα τα videos που χρησιμοποιήθηκαν σε αυτό το κομμάτι (ταινίες).



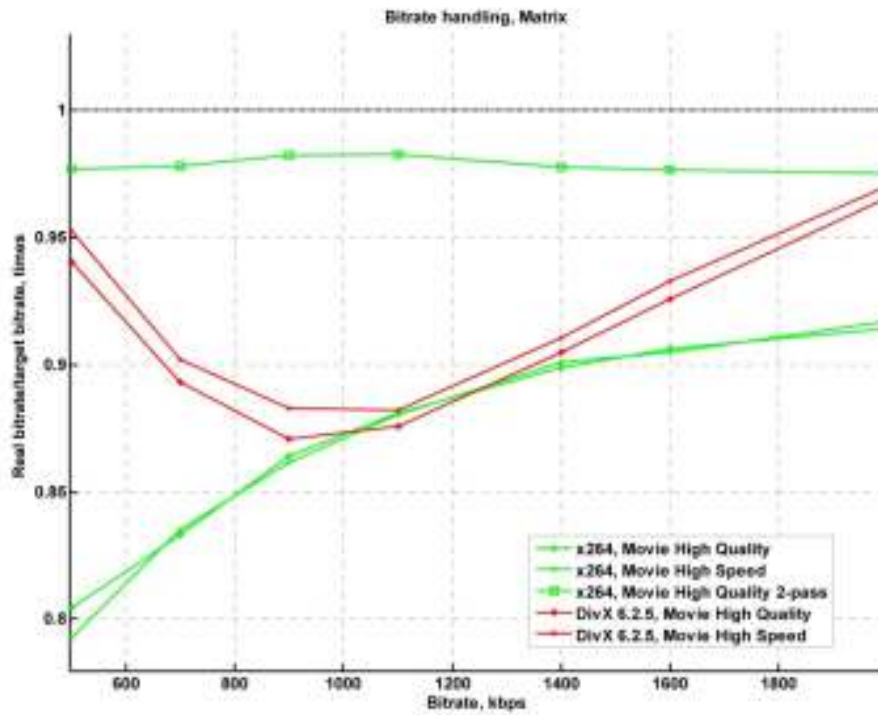
Εικόνα 79 – Διάγραμμα ρυθμού bit / Y-SSIM για το video Rancho για ρύθμιση Υψηλής ποιότητας

Για το ίδιο video για τη ρύθμιση υψηλής ταχύτητας τα αποτελέσματα φαίνονται στην **Εικόνα 80**. Η χαμηλότερη ποιότητα παρουσιάζεται από το codec της Intel. Αυτό μπορεί να δικαιολογηθεί από τη διαφορά στη ταχύτητα κωδικοποίησης. Το codec Intel H.264 κωδικοποιεί το συγκεκριμένο video 1.4 φορές γρηγορότερα από το DivX.



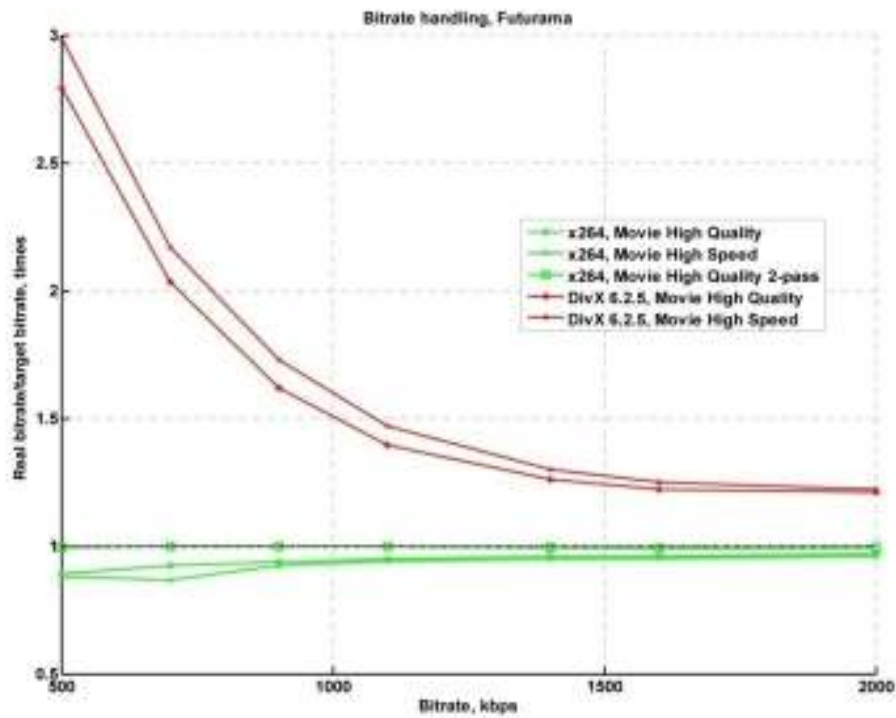
Εικόνα 80 - Διάγραμμα ρυθμού bit / Y-SSIM για το video Rancho για ρύθμιση Υψηλής ταχύτητας

Στην **Εικόνα 81** παρουσιάζεται ένα διάγραμμα διαχείρισης του ρυθμού bit από τα διάφορα codecs. Παρατηρείται ότι το 2-pass x264 έχει αρκετά καλό ρυθμό bit για το video “Matrix”. Η μέγιστη διαφορά είναι περίπου 20% για τη ρύθμιση ενός περάσματος (1-pass).



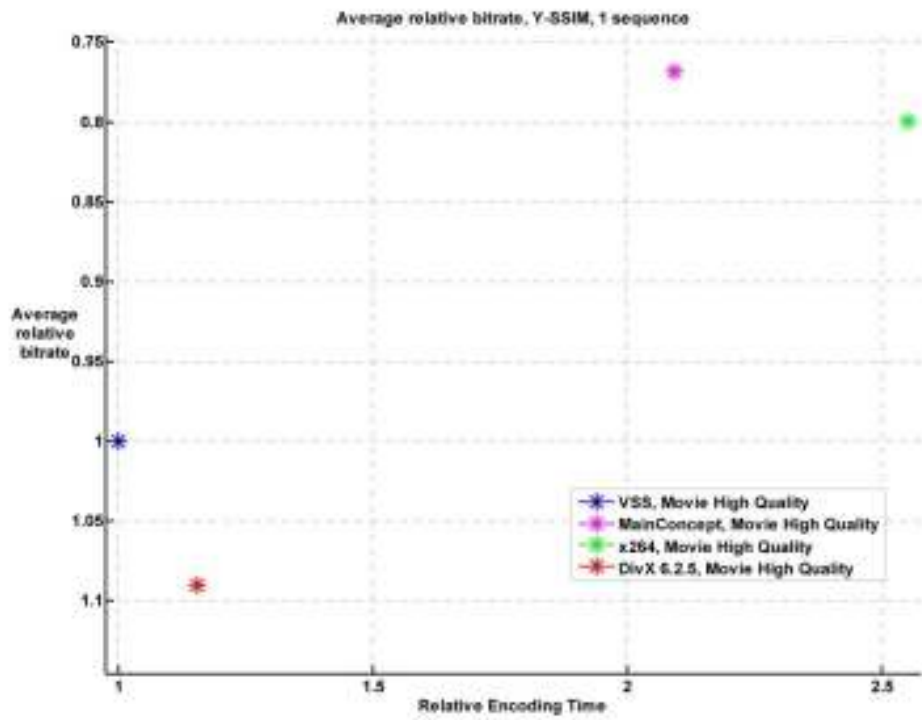
Εικόνα 81 – Διάγραμμα διαχείρισης ρυθμού bit των codecs DivX και x264 για το video Matrix

Στην **Εικόνα 82** παρουσιάζεται ένα ακόμα διάγραμμα διαχείρισης του ρυθμού bit, το οποίο αποκαλύπτει ότι το DivX έχει προβλήματα στη διατήρηση σταθερού ρυθμού bit για το video “Futurama” (ο πραγματικός ρυθμός bit είναι περίπου τριπλάσιος από τον επιθυμητό ρυθμό bit).



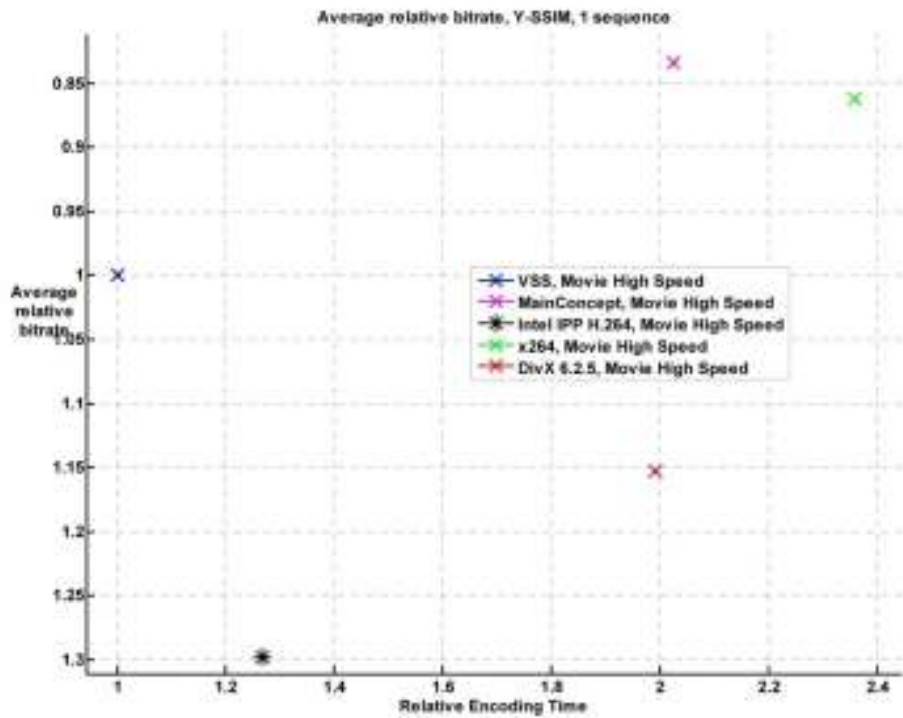
Εικόνα 82 - Διάγραμμα διαχείρισης ρυθμού bit των codecs DivX και x264 για το video Futurama

Στην **Εικόνα 83** παρατηρείται ότι το VSS codec υπερσχύει του DivX codec, καθώς παρουσιάζει καλύτερη ποιότητα ενώ ταυτόχρονα διατηρεί υψηλότερη ταχύτητα κωδικοποίησης.



Εικόνα 83 – Διάγραμμα σχετικού ρυθμού bit / σχετικής ταχύτητας κωδικοποίησης για το video Rancho για ρύθμιση Υψηλής ποιότητας

Στην **Εικόνα 84** παρατηρείται ότι τα codecs DivX και Intel αποδίδουν χειρότερα από το VSS, καθώς έχουν χαμηλότερη ποιότητα και πιο αργή κωδικοποίηση από το VSS για το video “Rancho” (με μέτρηση του SSIM).



Εικόνα 84 - Διάγραμμα σχετικού ρυθμού bit / σχετικής ταχύτητας κωδικοποίησης για το video Rancho για ρύθμιση Υψηλής ταχύτητας

Οι **Πίνακες 37, 38 και 39** παρουσιάζουν τις μέσες τιμές του ρυθμού bit για σταθερή ποιότητα. Η μέση τιμή προέκυψε από όλα τα videos για αυτό το είδος εφαρμογής (ταινίες) (δηλαδή τα videos Battle, Rancho, Matrix και Futurama).

	DivX	VSS	MainConcept	Intel H.264	X264
Μέσο ποσοστό μεγέθους αρχείου σε σχέση με το DivX	100.00%	105.00%	78.00%	107.00%	76.00%

Πίνακας 37 - Μέσο μέγεθος αρχείου για σταθερή ποιότητα για όλα τα videos της ομάδας Ταινίες, για ρύθμιση Υψηλής ταχύτητας και για ρυθμούς bit 600-1800kbps

	DivX	VSS	MainConcept	Intel H.264	X264
Μέσο ποσοστό μεγέθους αρχείου σε σχέση με το DivX	100.00%	96.00%	77.00%	-	72.00%

Πίνακας 38 - Μέσο μέγεθος αρχείου για σταθερή ποιότητα για όλα τα videos της ομάδας Ταινίες, για ρύθμιση Υψηλής ποιότητας και για ρυθμούς bit 600-1800kbps

	DivX	VSS	MainConcept	Intel H.264	X264
Μέσο ποσοστό μεγέθους αρχείου σε σχέση με το MainConcept	-	-	100.00%	-	90.00%

Πίνακας 39 - Μέσο μέγεθος αρχείου για σταθερή ποιότητα για όλα τα videos της ομάδας Ταινίες, για ρύθμιση Υψηλής ποιότητας - 2 περασμάτων και για ρυθμούς bit 600-1800kbps

4.2.3.2 Συμπεράσματα

Με βάση τη μέση τιμή του ρυθμού bit η κατάταξη των codecs ελέγχθηκαν είναι:

1. x264
2. MainConcept
3. DivX (MPEG-4 ASP)
4. Intel H.264
5. VSS

Αξίζει να παρατηρηθεί ότι για video που ανήκουν στο σύνολο εφαρμογών «ταινίες» το MPEG-4 ASP αποδίδει καλύτερα από κάποια codecs που χρησιμοποιούν το πιο πρόσφατο codec H.264. Επίσης πρέπει να σημειωθεί ότι η καλύτερη απόδοση παρατηρείται από το μη-εμπορικό codec x264.

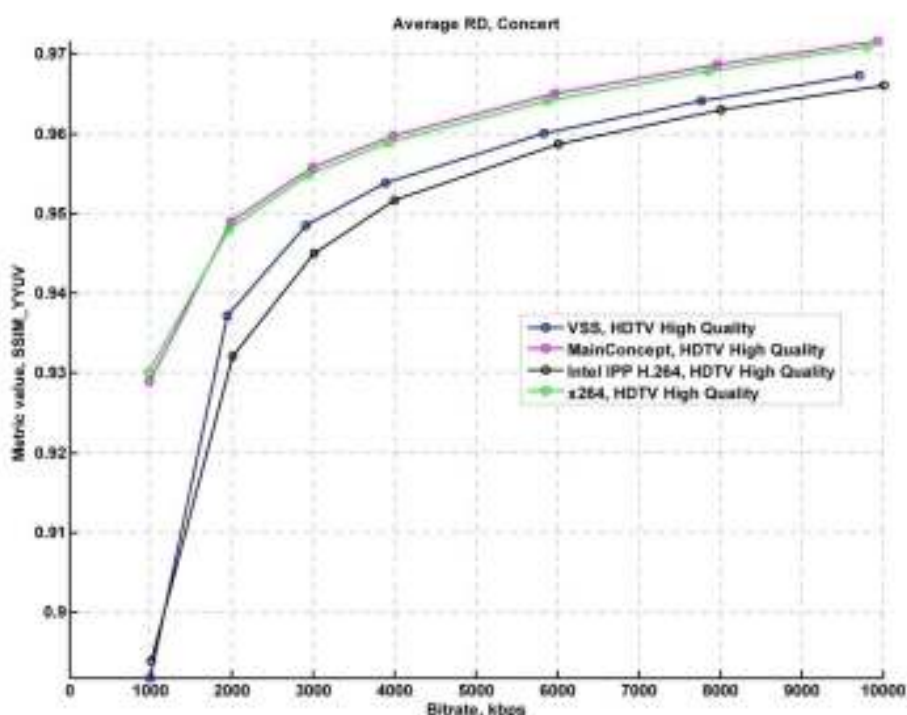
4.2.4 HDTV

Σε αυτό το κομμάτι ελέγχεται η συμπεριφορά των codecs όσο αναφορά την κωδικοποίηση video με ανάλυση υψηλής ευκρίνειας (high definition resolution, HDTV). Εδώ χρησιμοποιήθηκαν video υψηλής ανάλυσης και επιλέχθηκαν bit rates 1, 2, 3, 4, 6, 8 και 10 Mbps. Τα codecs που ελέγχθηκαν είναι:

- MainConcept
- Intel H.264
- VSS
- X264

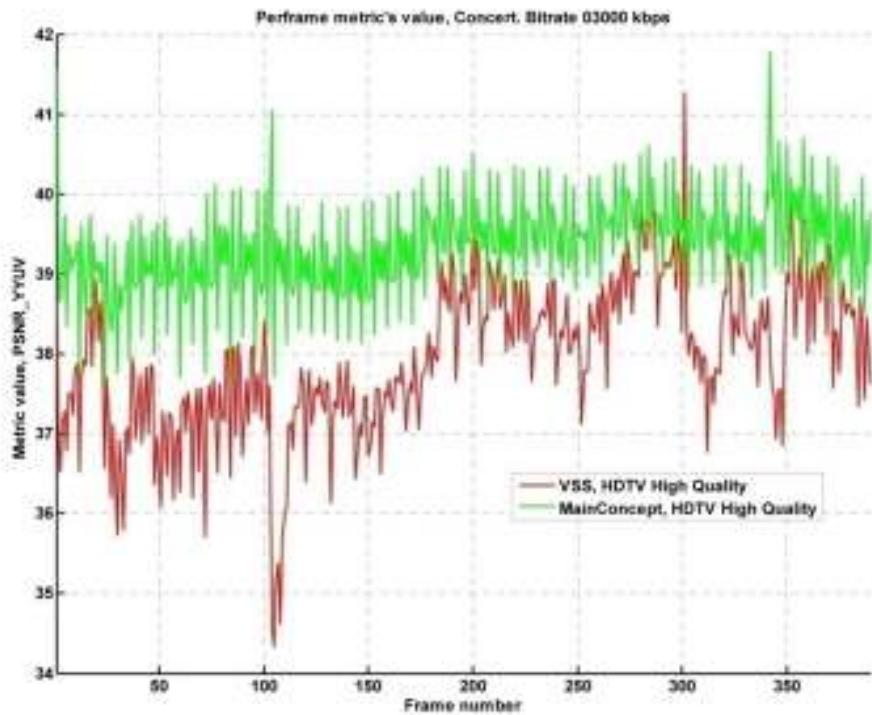
4.2.4.1 Αποτελέσματα

Στην **Εικόνα 85** βλέπουμε την υπεροχή του x264 και του MainConcept codec. Η μειωμένη ποιότητα των codecs Intel και VSS οφείλεται στη μεγαλύτερη ταχύτητα που παρουσιάζουν σε σχέση με το x264. Αξίζει να παρατηρηθεί ότι ενώ η διαφορά στη ταχύτητα ανάμεσα στο x264 και το MainConcept είναι μεγάλη, η διαφορά στη ποιότητα δεν είναι τόσο μεγάλη και το γρηγορότερο codec αποδίδει καλύτερη ποιότητα.



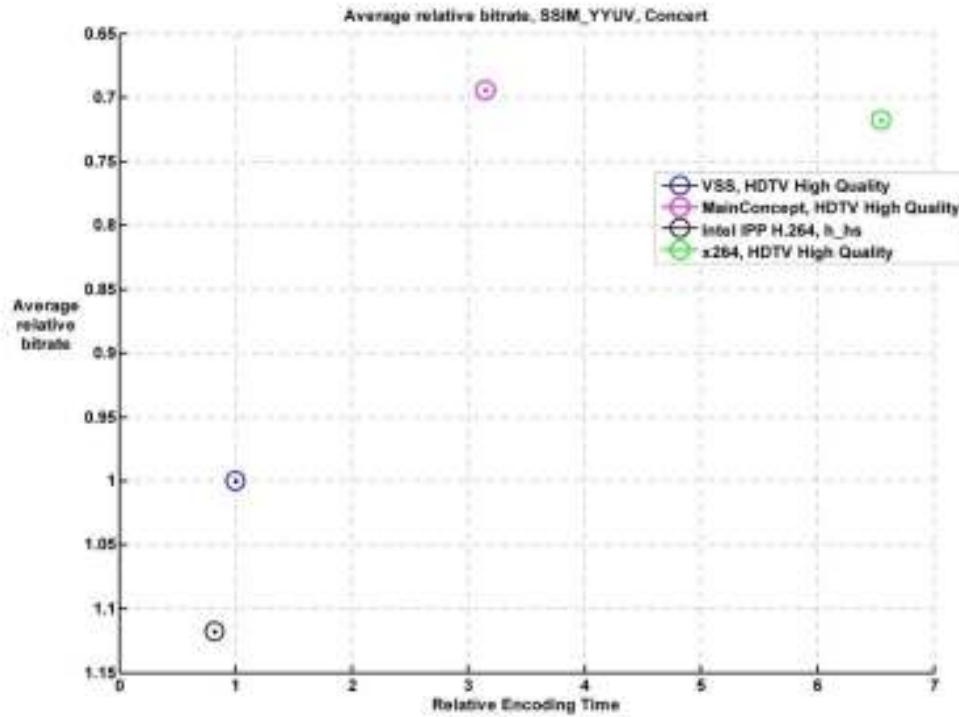
Εικόνα 85 – Διάγραμμα ρυθμού bit / ποιότητας για το video Concert για ρύθμιση Υψηλής ποιότητας

Στην **Εικόνα 86** φαίνεται η ποιότητα ανά πλαίσιο (per-frame quality) για τα codecs VSS και MainConcept. Παρατηρείται μία ύφεση στη ποιότητα στην αρχή και αύξηση στο τέλος του video. Επίσης παρατηρείται ότι το VSS έχει περισσότερες διακυμάνσεις από το MainConcept.



Εικόνα 86 - Διάγραμμα ποιότητας ανά πλαίσιο για το video Concert για ρύθμιση Υψηλής ποιότητας

Στην **Εικόνα 87** παρατηρείται ότι το MainConcept λειτουργεί γρηγορότερα και αποδίδει λίγο καλύτερα. Ωστόσο όταν χρησιμοποιείται το PSNR αντί του SSIM δεν παρατηρούνται τα ίδια αποτελέσματα.



Εικόνα 87 – Διάγραμμα σχετικού ρυθμού bit / σχετικής ταχύτητας κωδικοποίησης για το video Concert για ρύθμιση Υψηλής ποιότητας

Ο **Πίνακας 40** παρουσιάζει τις μέσες τιμές ρυθμού bit για σταθερή ποιότητα για όλα τα codecs. Η μέση τιμή προήλθε μόνο από το video “Concert”.

	VSS	MainConcept	Intel H.264	X264
Μέσο ποσοστό μεγέθους αρχείου σε σχέση με το VSS	100.00%	69.43%	111.80%	71.76%

Πίνακας 40 - Μέσο μέγεθος αρχείου για σταθερή ποιότητα για όλα τα videos της ομάδας HDTV, για ρυθμούς bit 2-10Mbps

4.2.4.2 Συμπεράσματα

Στο σύνολο των μετρήσεων για τις εφαρμογές που ανήκουν στη κατηγορία “HDTV” η κατάταξη των codecs που ελέγχθηκαν είναι:

1. MainConcept
2. X264
3. VSS
4. Intel H.264

Τα codecs MainConcept και x264 είναι τα καλύτερα για την κωδικοποίηση HDTV περιεχομένου με κριτήριο την αναλογία ποιότητας / ταχύτητας.

4.2.5 Συγκεντρωτικά αποτελέσματα και συμπεράσματα

Σε αυτό το κομμάτι παρουσιάζονται το σύνολο των αποτελεσμάτων για κάθε τύπο εφαρμογής (video-διάσκεψη, ταινίες και HDTV). Οι τιμές που παρουσιάζονται είναι οι μέσες τιμές από όλες τις μετρήσεις για κάθε codec.

Ο **Πίνακας 41** παρουσιάζει τη μέση τιμή ρυθμού bit που εξοικονομείται. Οι τιμές είναι σχετικές για σταθερή ποιότητα για όλα τα codecs και έτσι η μικρότερη τιμή είναι η καλύτερη.

	MainConcept	X264	VSS	DivX	Intel H.264
Video-διάσκεψη (υψηλή ποιότητα)	56%	63%	79%	100%	
Video-διάσκεψη (υψηλή ταχύτητα)	62%	62%	76%	100%	91%
Video-διάσκεψη	55%	62%	77%	100%	91%
Ταινίες (υψηλή ταχύτητα)	78%	76%	105%	100%	107%
Ταινίες (ένα πέρασμα)	77%	72%	96%	100%	-
Ταινίες (δύο περάσματα)	100%	90%	-	-	-
Ταινίες	84%	79%	100%	100%	107%
HDTV	69%	72%	100%		112%
HDTV	69%	72%	100%		112%
Σύνολο	69%	71%	92%	100%	103%

Πίνακας 41 – Αποτελέσματα του 2006 για το μέσο μέγεθος αρχείου για σταθερή ποιότητα



Εικόνα 88 – Μέσο μέγεθος video για στο σύνολο του πειράματος

Στον **Πίνακα 42** παρουσιάζει τη μέση τιμή του χρόνου κωδικοποίησης για όλα τα codecs για όλα τα video για όλους τους τύπους εφαρμογών. Αρχικά υπολογίστηκαν οι μέσες τιμές για κάθε ρύθμιση κάθε τύπου εφαρμογής και μετά υπολογίστηκαν οι μέσες τιμές για κάθε τύπο εφαρμογής. Για κάθε ρύθμιση (υψηλής ποιότητας ή υψηλής ταχύτητας) η μεγαλύτερη είναι το 100% και οι υπόλοιπες είναι σχετικές με αυτή.

	VSS	Intel H.264	MainConcept	DivX	x264
Video-διάσκεψη (υψηλή ποιότητα)	91%	-	84%	93%	100%
Video-διάσκεψη (υψηλή ταχύτητα)	71%	74%	91%	100%	88%
Video-διάσκεψη	81%	74%	88%	97%	94%
Ταινίες (υψηλή ταχύτητα)	48%	56%	84%	84%	100%
Ταινίες (ένα πέρασμα)	33%	-	64%	44%	100%
Ταινίες (δύο περάσματα)	-	-	36%	-	100%
Ταινίες	40%	56%	61%	64%	100%
HDTV	15%	13%	48%	-	100%
HDTV	15%	13%	48%	-	100%
Σύνολο	46%	47%	66%	80%	98%

Πίνακας 42 - Αποτελέσματα του 2006 για τη μέση ταχύτητα κωδικοποίησης

Στο σύνολο των αποτελεσμάτων για όλους τους τύπους των εφαρμογών η κατάταξη των codecs είναι:

1. MainConcept
2. X264
3. VSS
4. DivX (MPEG-4 ASP)
5. Intel H.264

Έχει ενδιαφέρον να πούμε ότι για κάθε τύπο εφαρμογής κάθε codec δείχνει ξεχωριστή συμπεριφορά και αποτελέσματα.

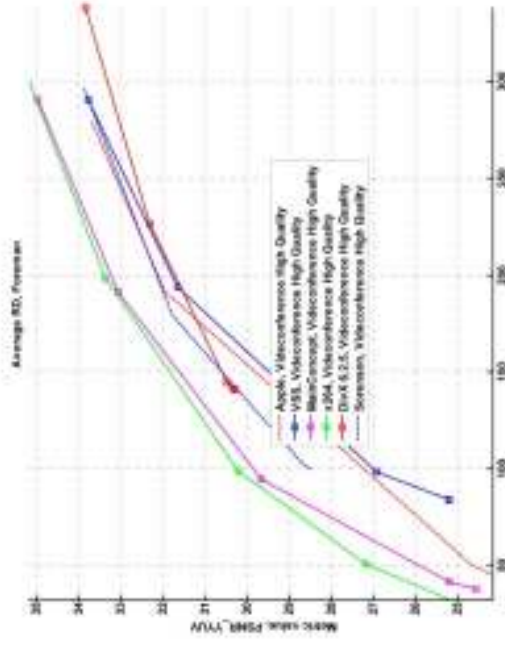
Βάση της ποιότητας τα codecs MainConcept και x264 καταλαμβάνουν τις πρώτες θέσεις για τους τύπους “video-διάσκεψη” και “HDTV”. Η διαφορά μεταξύ τους όμως δεν είναι σημαντική. Για εφαρμογές video-διάσκεψης τα codecs που χρησιμοποιούν το πρότυπο H.264 αποδίδουν καλύτερα έναντι αυτόν που χρησιμοποιούν το MPEG-4 ASP, όπως το DivX. Για τις εφαρμογές τύπου ταινίας το DivX (MPEG-4 ASP) παρουσιάζει αρκετά καλά αποτελέσματα σε σύγκριση με τα υπόλοιπα codecs που χρησιμοποιούν το πιο πρόσφατο H.264. Για τον τύπο εφαρμογής “HDTV” δεν είναι δυνατό να ελεγχθεί το DivX λόγω τεχνικών δυσκολιών, καθώς τα νέα codecs έχουν μεγαλύτερο εύρος όσο αναφορά το χρόνο κωδικοποίησης και την ποιότητα.

4.2.6 Σύγκριση των codecs Apple και Sorenson

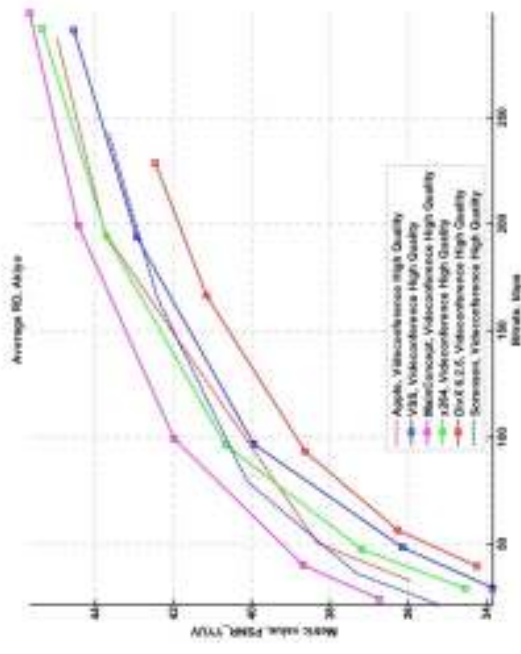
Λόγω τεχνικών προβλημάτων τα codecs της Apple και της Sorenson δεν πήραν μέρος στην προηγούμενη εξέταση και έτσι η ανάλυσή του παρουσιάζεται σε αυτό το κομμάτι. Τα δύο codecs ελέγχθηκαν για ρύθμιση υψηλής ποιότητας.

Μεγάλη διαφορά στα αποτελέσματα παρατηρείται αν χρησιμοποιηθεί τη μέτρηση του MSU σημαθορυβικού λόγου ανεξαρτήτου φωτεινότητας (MSU Brightness Independent PSNR,), καθώς το codec της Apple μεταβάλλει τη φωτεινότητα.

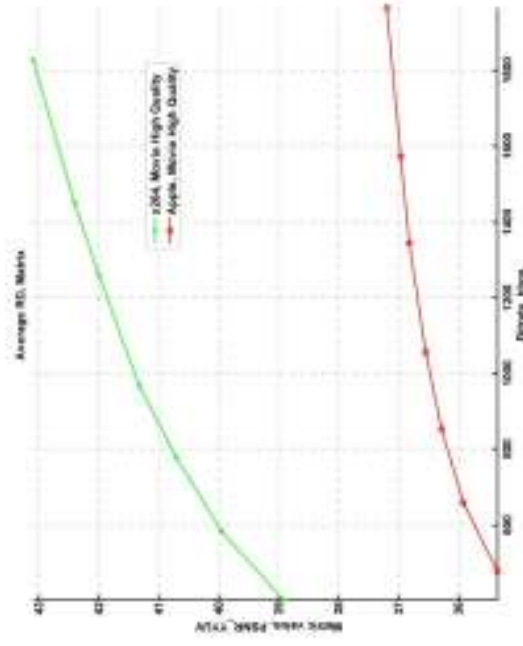
Η συμπεριφορά των codecs διαφέρει για άλλα video πέραν των δύο που παρουσιάζονται.



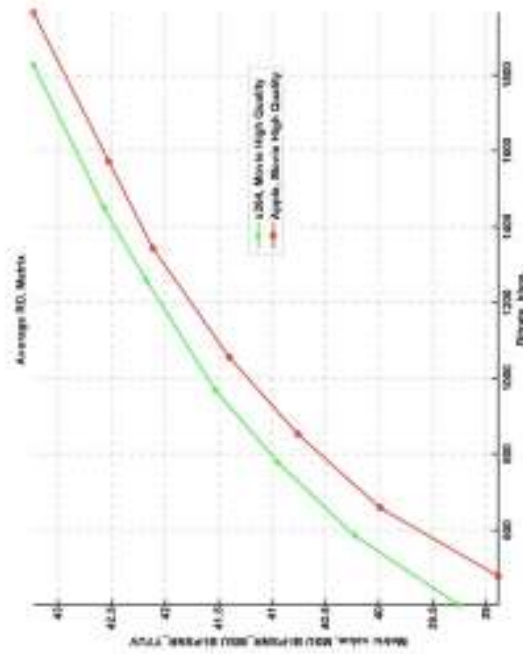
Εικόνα 89 – Διάγραμμα ρυθμού bit / ποιότητας για το video Foreman για ρύθμιση Υψηλής ποιότητας



Εικόνα 90 – Διάγραμμα ρυθμού bit / ποιότητας για το video Akigo για ρύθμιση Υψηλής ποιότητας



Εικόνα 91 – Διάγραμμα ρυθμού bit / ποιότητας των codecs x264 και Apple για το video Matrix για ρύθμιση Υψηλής ποιότητας



Εικόνα 92 – Διάγραμμα ρυθμού bit / ποιότητας των codecs x264 και Apple για το video Matrix για ρύθμιση Υψηλής ποιότητας, με μέτρηση του Y-BI-PSNR

4.2.6.1 Συμπεράσματα

Τα codecs της Apple και της Sorenson έχουν μέτρια απόδοση σε σύγκριση με τα άλλα codecs. Το codec της Apple παρουσιάζει χαμηλή ποιότητα σε ορισμένους τύπους video (όπως τις ταινίες) όταν υπολογίζεται με μέτρηση του Y-PSNR. Αυτό οφείλεται στη μεταβολή του εφαρμόζει το codec στη φωτεινότητα του κάθε video. Αν χρησιμοποιήσουμε το PSNR ανεξαρτήτως φωτεινότητας (MSU BI-PSNR) τα αποτελέσματα του codec της Apple είναι συγκρίσιμα με αυτά των υπολοίπων.

4.3 Σύγκριση των Mpegable AVC, Moonlight H.264, MainConcept H.264, Fraunhofer IIS, AteME MPEG-4 AVC / H.264, Videosoft H.264 και DivX Pro 5.1.1

Στο πείραμα αυτό, το οποίο πραγματοποιήθηκε τον Ιανουάριο του 2005 και παρουσιάζεται στο [17], συγκρίνονται τα codecs Mpegable AVC, Moonlight H.264 Video, MainConcept H.264, Fraunhofer IIS, AteME MPEG-4 AVC/H.264, Videosoft H.264 και το DivX Pro 5.1.1. Στον Πίνακα 43 παρουσιάζονται τα codecs που χρησιμοποιήθηκαν.

Codec	Κατασκευαστής	Έκδοση
Mpegable AVC	Dicas digital image coding GmbH	0.10
Moonlight H.264	Moonlight Cordless LTD	0.1.2546
MainConcept H.264	MainConcept AG	1.04.02.00
MPEG-4/AVC	Fraunhofer Institute for Integrated circuits	Date 25.11.2004
AteME MPEG-4 AVC/H.264	AteME	1.0.3.2
Videosoft H.264	Videosoft Inc.	2.1.0.2
DivX Pro 5.1.1	DivX Networks Inc.	5.1.1

Πίνακας 43 - Codecs που χρησιμοποιήθηκαν για τη σύγκριση των Mpegable AVC, Moonlight H.264, MainConcept H.264, Fraunhofer IIS, AteME MPEG-4 AVC / H.264, Videosoft H.264 και DivX Pro 5.1.1

Τα videos που χρησιμοποιήθηκαν για την εξαγωγή των αποτελεσμάτων κατά τη διάρκεια του πειράματος παρουσιάζονται στον Πίνακα 44.

Videos	Αριθμός πλαισίων	Ρυθμός πλαισίων	Ανάλυση και χρώματα
Bankomatdi	376	30	704x352 (RGB)
Battle	1599	24	704x288 (RGB)
Bbc3di	374	25	704x576 (RGB)
Foreman	300	15	352x288 (RGB)
Susidi	374	25	704x576 (RGB)

Πίνακας 44 - Videos που χρησιμοποιήθηκαν για τη σύγκριση των Mpegable AVC, Moonlight H.264, MainConcept H.264, Fraunhofer IIS, AteME MPEG-4 AVC / H.264, Videosoft H.264 και DivX Pro 5.1.1

4.3.1 Πειραματική διαδικασία

Σκοπός του πειράματος είναι η εκτίμηση των codecs όσο αναφορά την ποιότητα και την συμπίεση για δεδομένα video. Τα videos που χρησιμοποιήθηκαν πέρασαν από μία διαδικασία όπου τους εφαρμόστηκε ένα φίλτρο απόπλεξης (deinterlacing filter). Τα codecs είχαν τις προεπιλεγμένες ρυθμίσεις τους.

Το μέγεθος πλαισίου υπολογίστηκε ως το ποσοστό του μεγέθους του video προς τον αριθμό των πλαισίων.

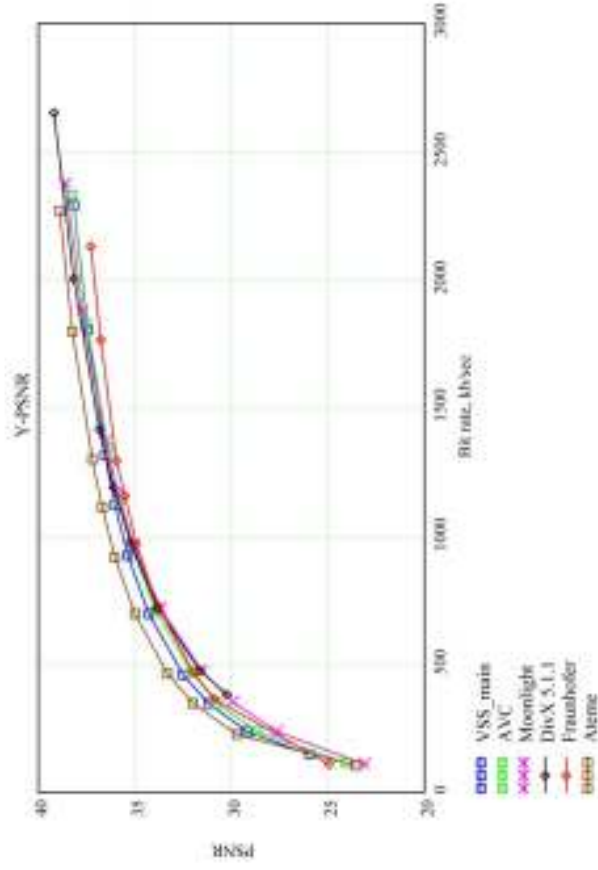
Ο Υ-άξονας στο διάγραμμα Δέλτα (Delta diagram) υπολογίστηκε ως η διαφορά μεταξύ του PSNR του ελεγχόμενου codec και του PSNR του DivX codec.

Το MainCocnert codec εισήγαγε επιπλέον πλαίσια στο κωδικοποιημένο video τα οποία αφαιρέθηκαν για να γίνει η σύγκριση. Αν το τελευταίο πλαίσιο του διορθωμένου video διέφερε από το τελευταίο πλαίσιο του ασυμπίεστου video τότε δεν γινόταν σύγκριση για αυτό το video.

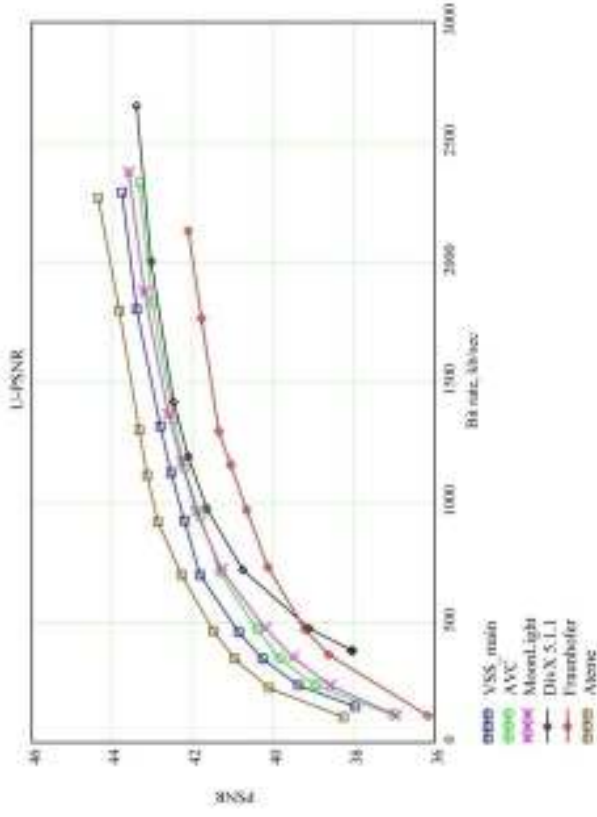
4.3.2 Διαγράμματα PSNR / ρυθμού bit

4.3.2.1 Video Bankomatdi

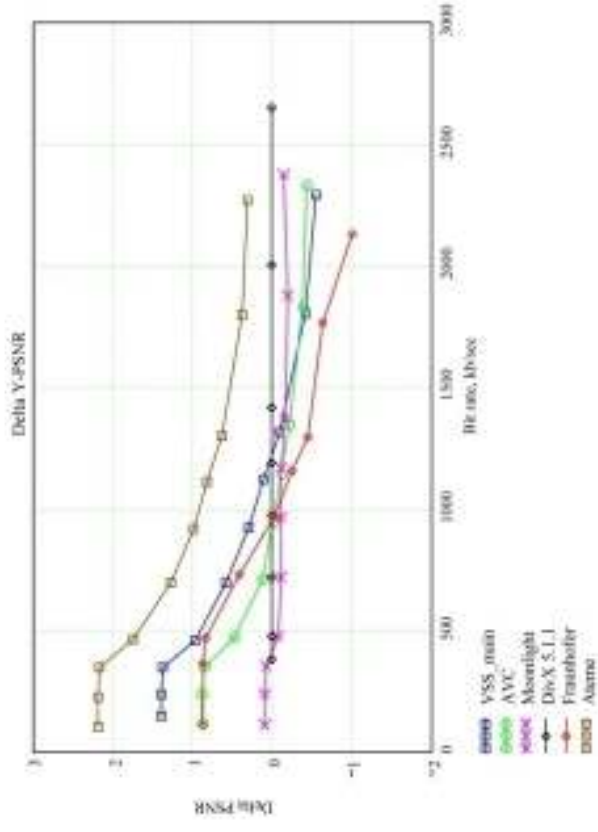
Το DivX codec στους χαμηλούς ρυθμούς bit αποδίδει χειρότερα από όλα τα codecs που χρησιμοποιούν το πρότυπο H.264. Το Ateme codec έχει τα καλύτερα αποτελέσματα στους μεσαίους και υψηλούς ρυθμούς bit. Τέλος πρέπει να αναφερθεί ότι τα διαγράμματα για τα U και V στοιχεία (components) είναι παρόμοια.



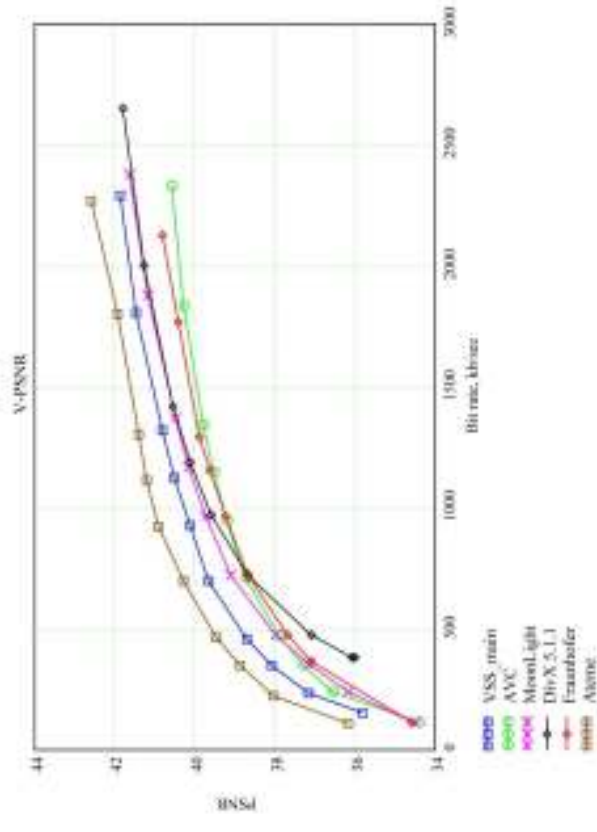
Εικόνα 93 - Διάγραμμα ρυθμού bit rate / Y-PSNR για το video Bankomatdi



Εικόνα 95 – Διάγραμμα ρυθμού bit / U-PSNR για το video Bankomatdi



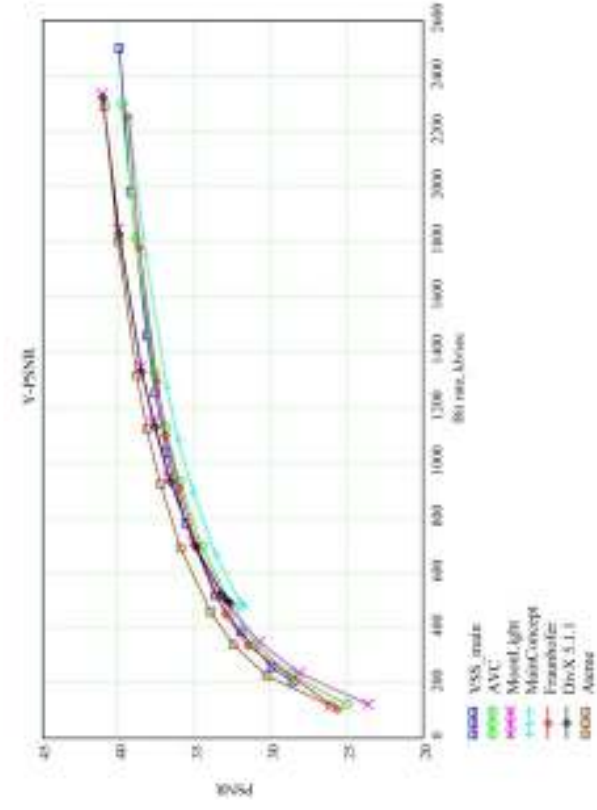
Εικόνα 94 - Διάγραμμα ρυθμού bit / Delta Y-PSNR για το video Bankomatdi



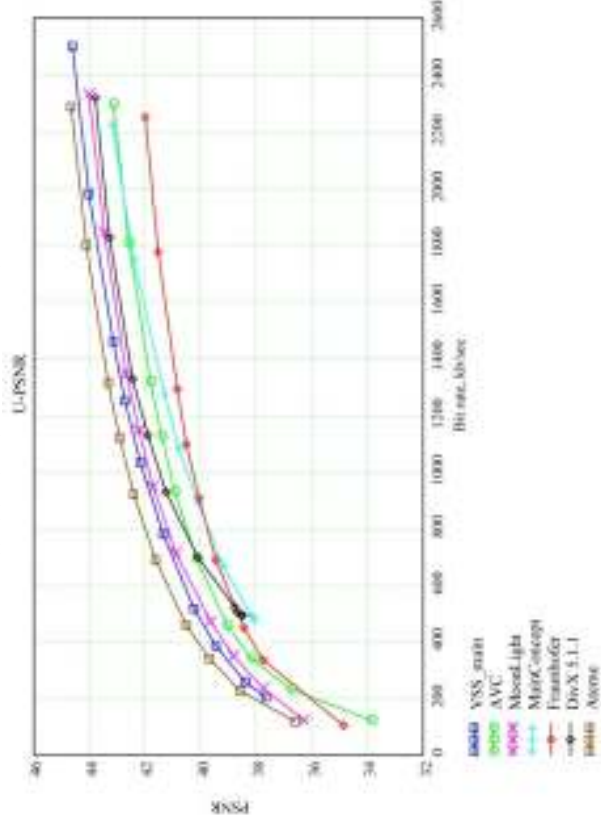
Εικόνα 96 – Διάγραμμα ρυθμού bit / V-PSNR για το video Bankomatdi

4.3.2.2 Video Battle

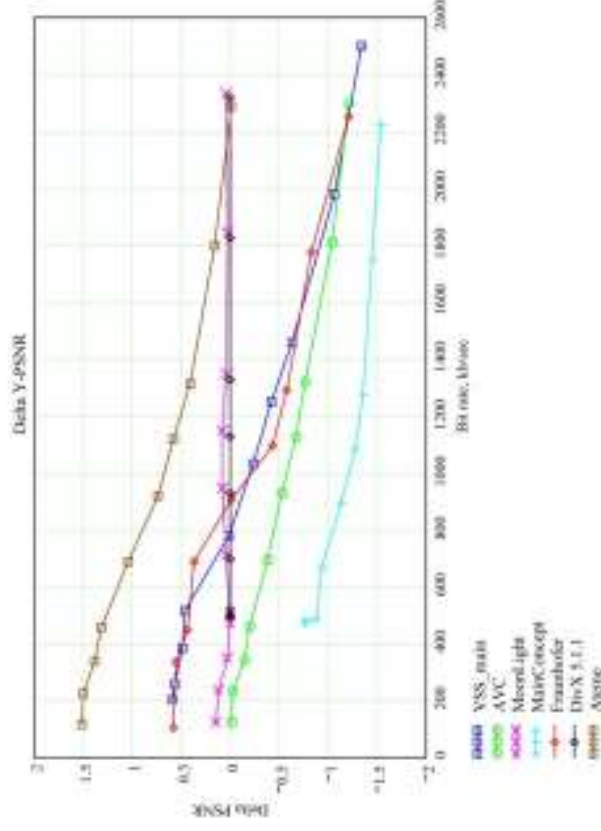
Το DivX codec αποδίδει πολύ χειρότερα από το Videosoft, Fraunhofer και το Ateme για τους χαμηλούς ρυθμούς bit. Το Ateme codec έχει τα καλύτερα αποτελέσματα στους μεσαίους και υψηλούς ρυθμούς bit. Και εδώ τα διαγράμματα για τα U και V στοιχεία (components) είναι παρόμοια.



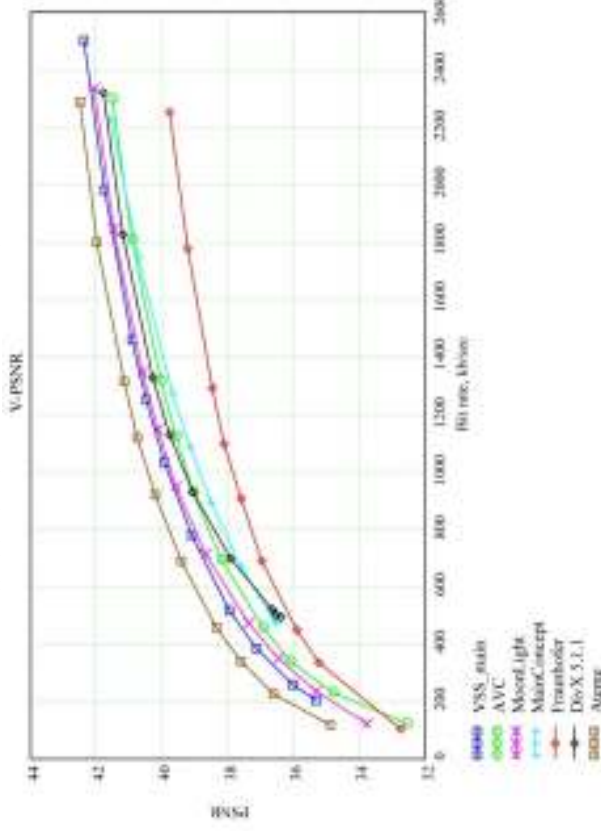
Εικόνα 97 – Διάγραμμα ρυθμού bit / Y-PSNR για το video Battle



Εικόνα 99 – Διάγραμμα ρυθμού bit / U-PSNR για το video Battle



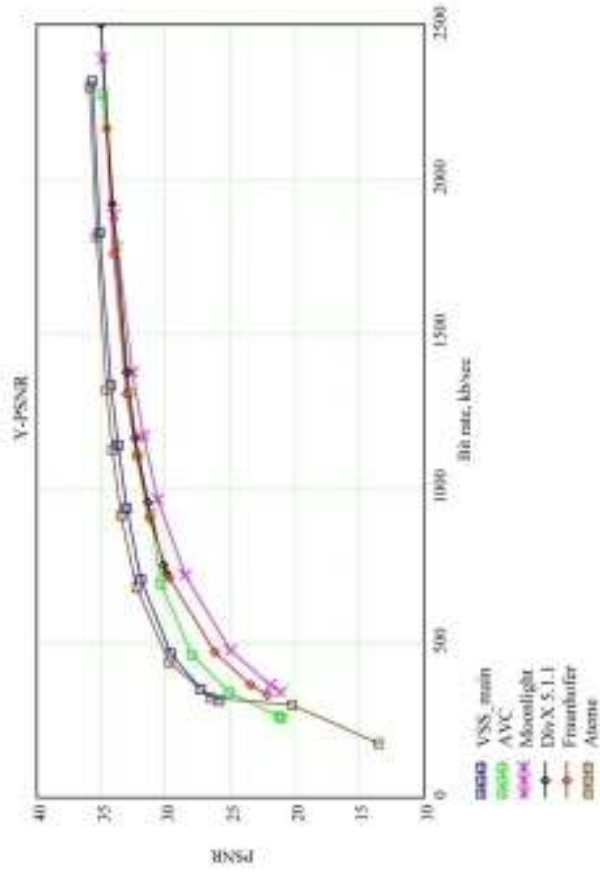
Εικόνα 98 – Διάγραμμα ρυθμού bit / Delta Y-PSNR για το video Battle



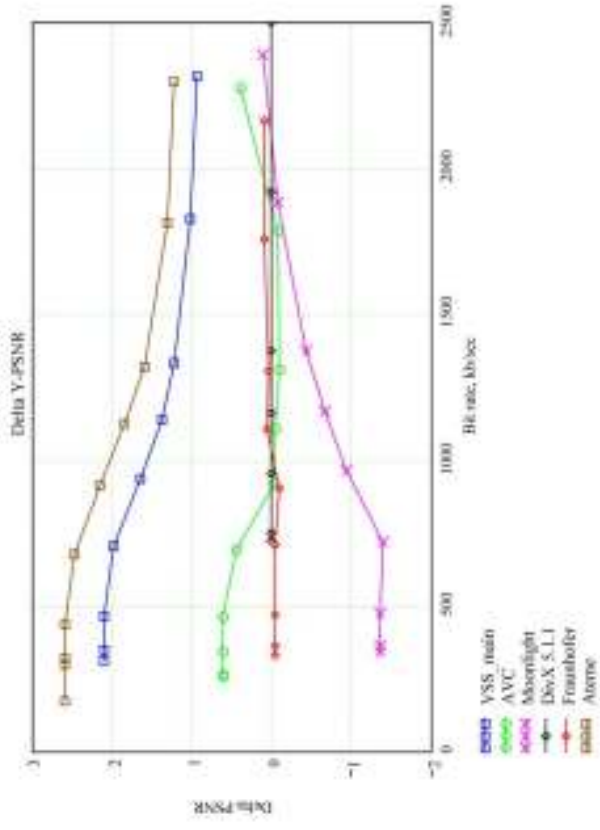
Εικόνα 100 – Διάγραμμα ρυθμού bit / V-PSNR για το video Battle

4.3.2.3 Video Bbc3di

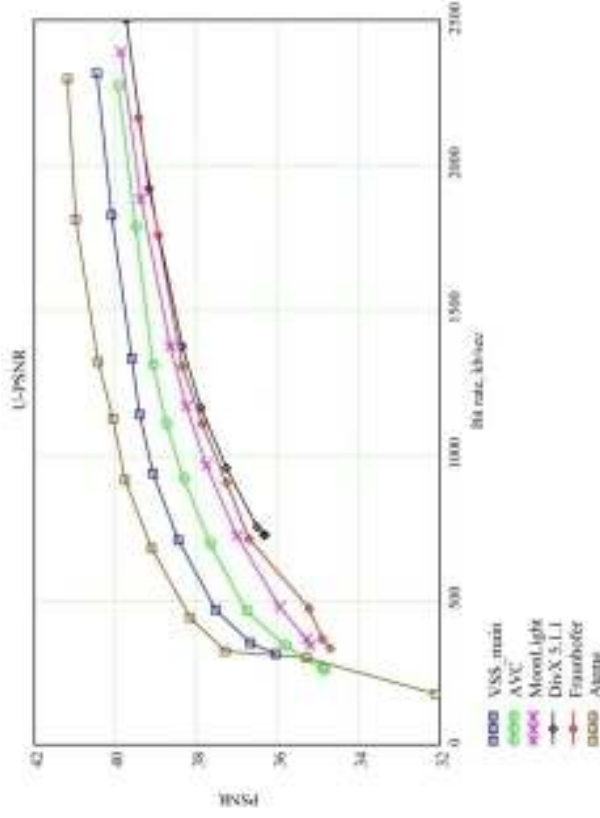
Το DivX codec αποδίδει χειρότερα στους χαμηλούς ρυθμούς bit από όλα τα codecs εκτός από το Moonlight. Τα codecs Atome και Videosoft έχουν τα καλύτερα αποτελέσματα στους μεσαίους και υψηλούς ρυθμούς bit. Και σε αυτή τη περίπτωση τα διαγράμματα για τα U και V στοιχεία είναι παρόμοια.



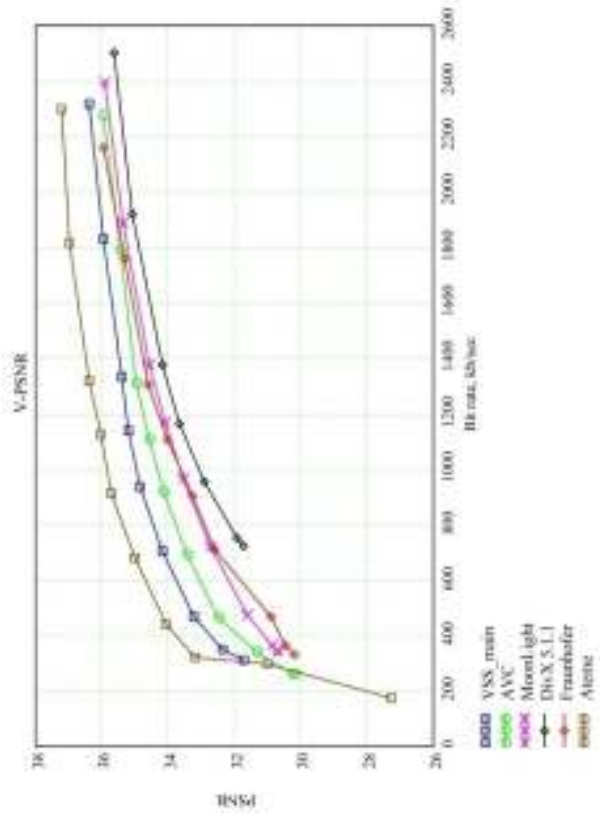
Εικόνα 101 – Διάγραμμα ρυθμού bit / Y-PSNR για το video Bbc3di



Εικόνα 102 – Διάγραμμα ρυθμού bit / Delta Y-PSNR για το video Bbc3di



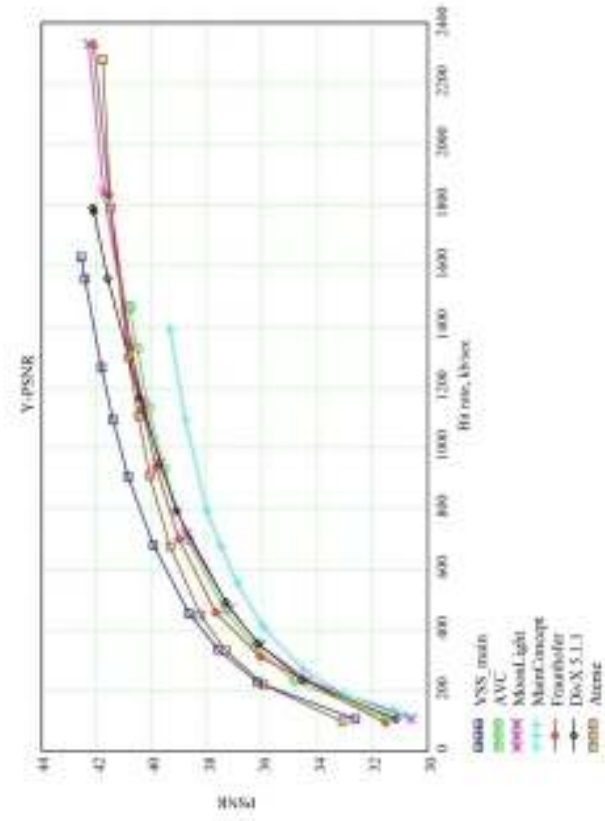
Εικόνα 103 – Διάγραμμα ρυθμού bit / U-PSNR για το video Bbc3di



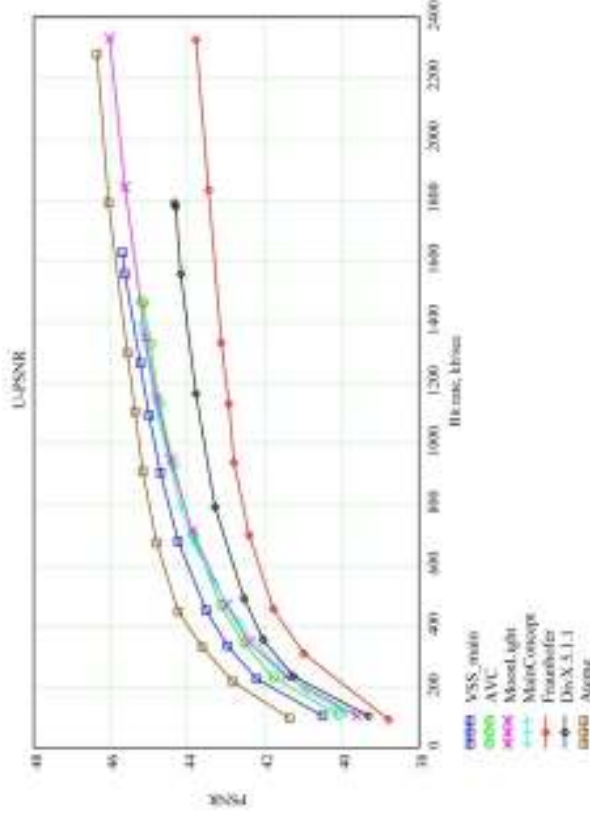
Εικόνα 104 – Διάγραμμα ρυθμού bit / V-PSNR για το video Bbc3di

4.3.2.4 Video Foreman

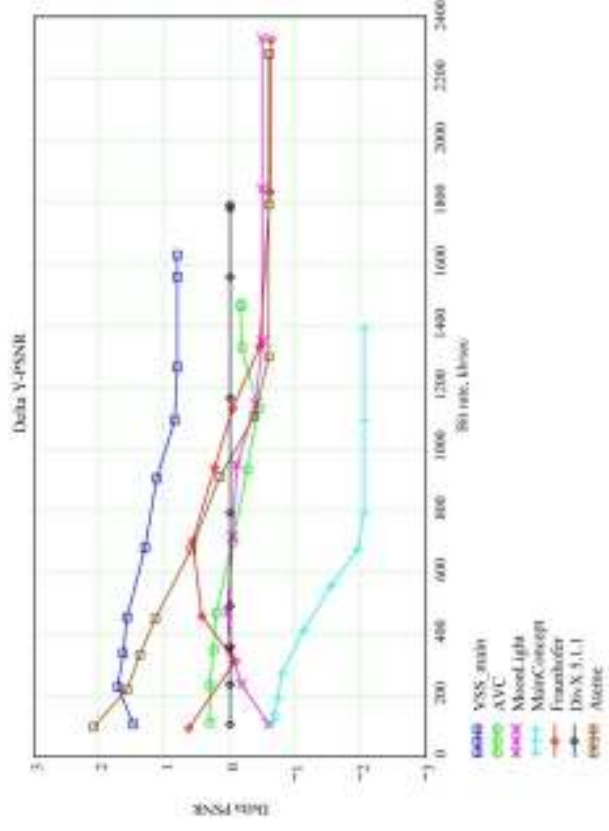
Στους χαμηλούς ρυθμούς bit το DivX αποδίδει χειρότερα από όλα τα codecs εκτός από το MainConcept. Το codec Videosoft έχει τα καλύτερα αποτελέσματα στους μεσαίους και χαμηλούς ρυθμούς bit. Τα codecs Moonlight, Fraunhofer και Atome έχουν παρόμοια αποτελέσματα στα υψηλούς ρυθμούς bit. Τα αποτελέσματα για τα U και V στοιχεία είναι παρόμοια εκτός από το Fraunhofer το οποίο έδειξε καλύτερα αποτελέσματα στα V στοιχεία και το AVC το οποίο έδειξε καλύτερα αποτελέσματα στα U στοιχεία.



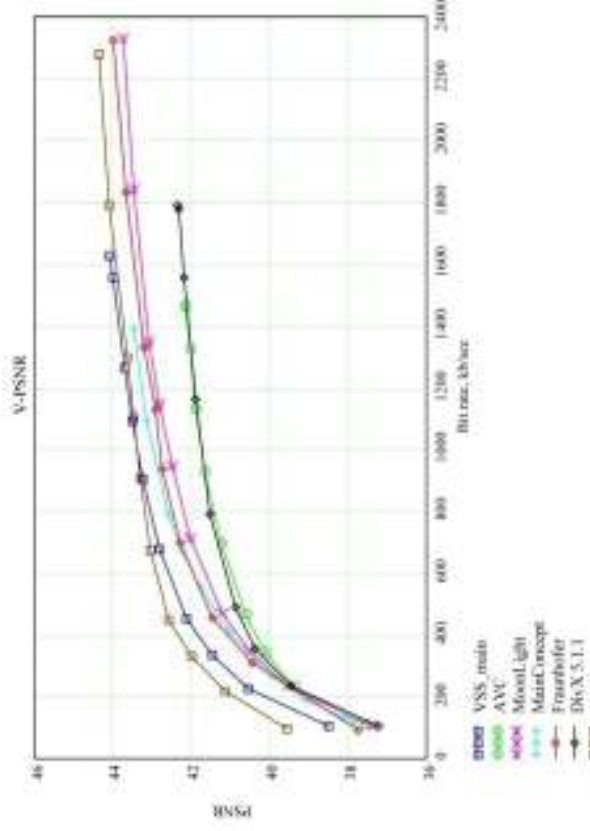
Εικόνα 105 – Διάγραμμα ρυθμού bit / Y-PSNR για το video Foreman



Εικόνα 107 – Διάγραμμα ρυθμού bit / U-PSNR για το video Foreman



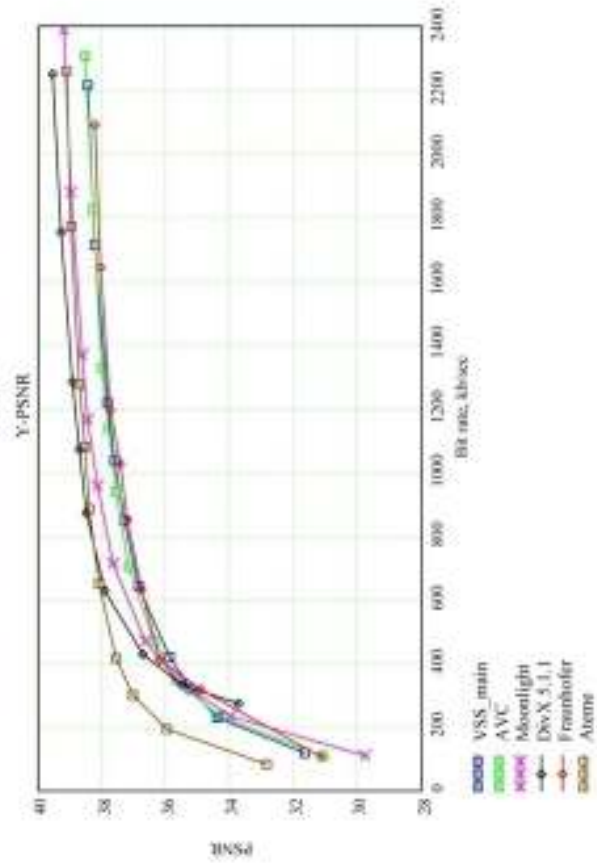
Εικόνα 106 - Διάγραμμα ρυθμού bit / Delta Y-PSNR για το video Foreman



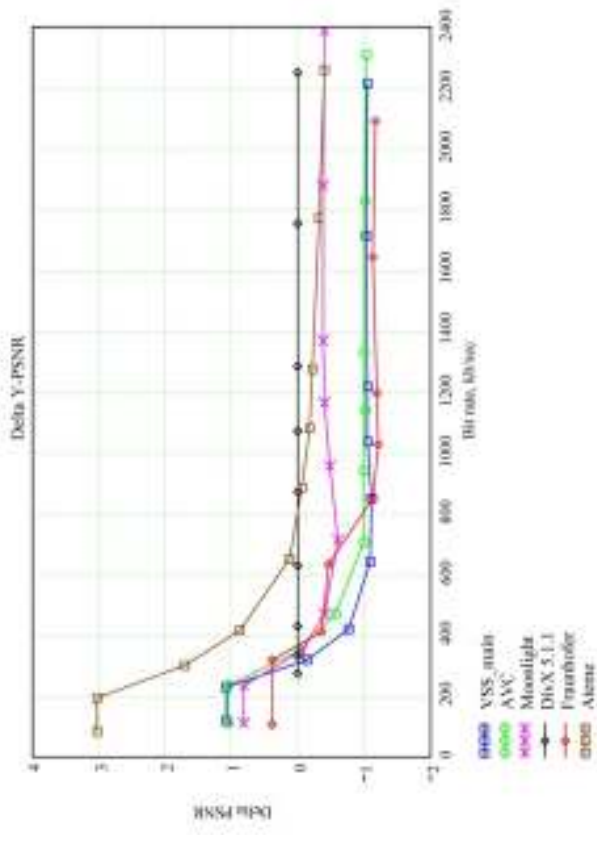
Εικόνα 108 – Διάγραμμα ρυθμού bit / V-PSNR για το video Foreman

4.3.2.5 Video Susidi

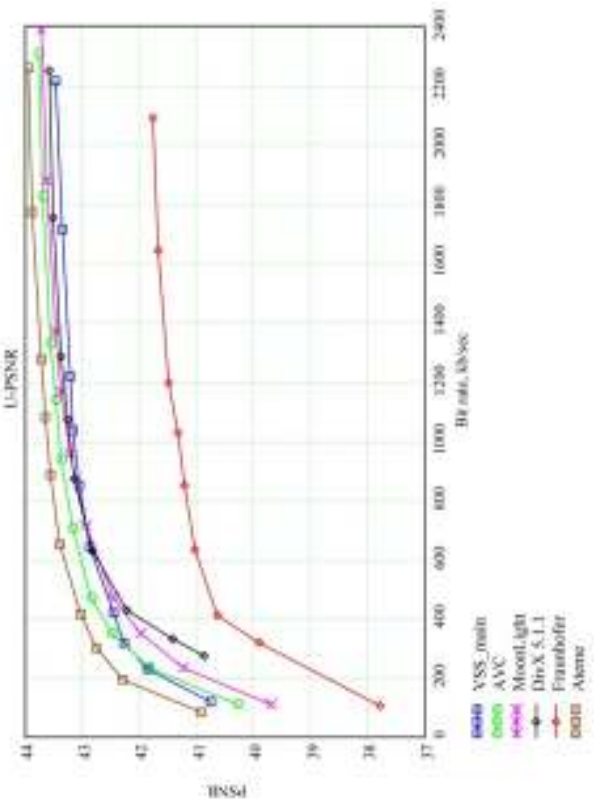
Το DivX αποδίδει αρκετά χειρότερα από το Atme στους χαμηλούς ρυθμούς bit και είναι το καλύτερο στους μεσαίους και υψηλούς ρυθμούς bit. Τα διαγράμματα για τα U και V στοιχεία είναι παρόμοια εκτός από το Fraunhofer όπου έχει καλύτερα αποτελέσματα στα V στοιχεία και το AVC που φαίνεται να έχει καλύτερα αποτελέσματα στα U στοιχεία.



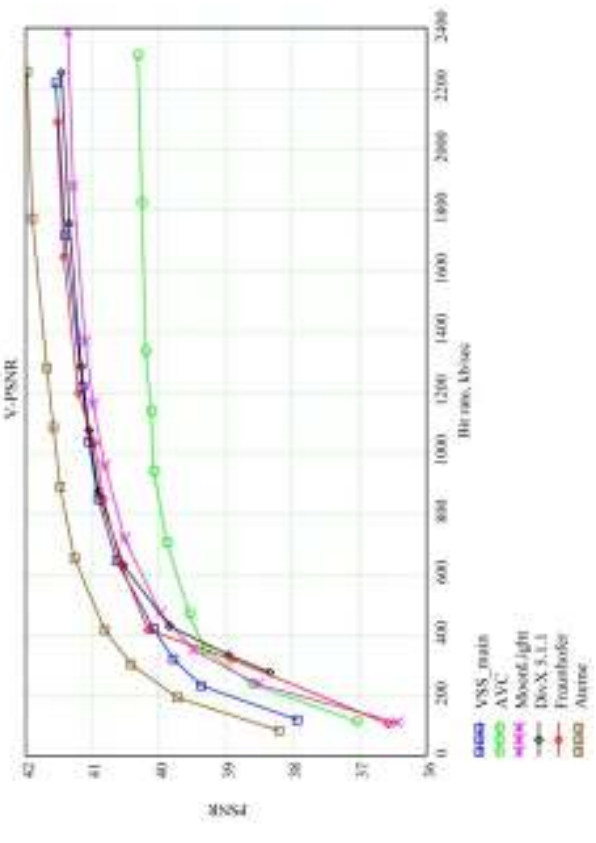
Εικόνα 109 – Διάγραμμα ρυθμού bit / Y-PSNR για το video Susidi



Εικόνα 110 – Διάγραμμα ρυθμού bit / Delta Y-PSNR για το video Susidi



Εικόνα 111 – Διάγραμμα ρυθμού bit / U-PSNR για το video Susidi

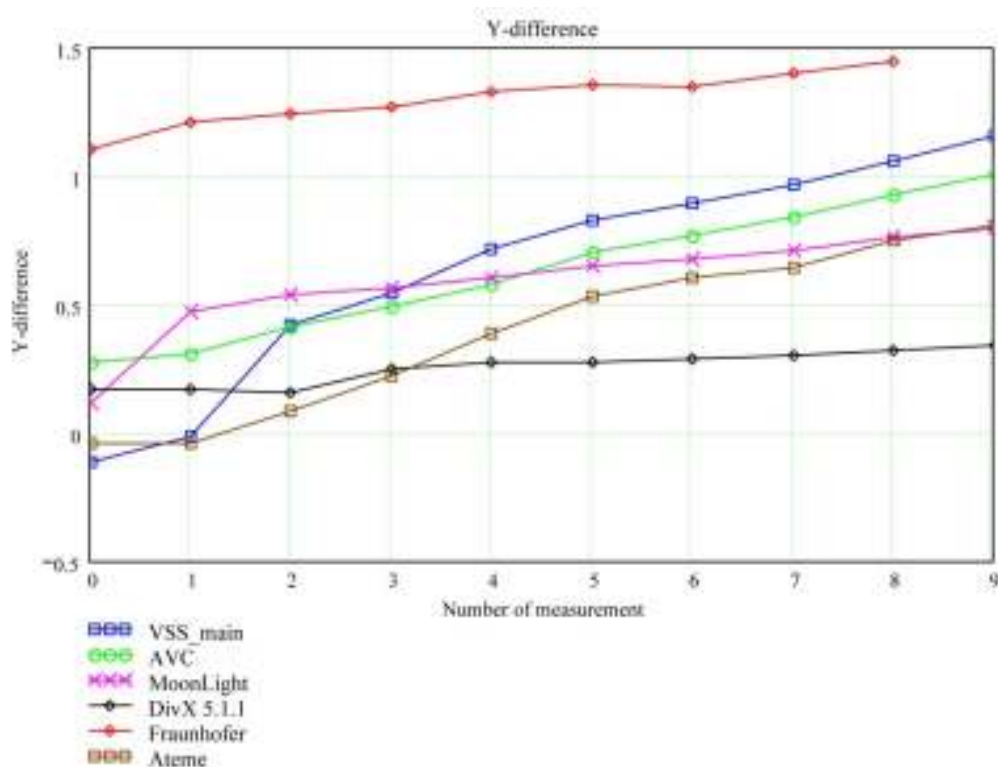


Εικόνα 112 – Διάγραμμα ρυθμού bit / V-PSNR για το video Susidi

4.3.3 Διαγράμματα μεταβολής της φωτεινότητας

Τα διαγράμματα μεταβολής της φωτεινότητας (Average brightness shift diagrams) δείχνουν τη μέση μεταβολή της φωτεινότητας που γίνεται από τα codecs. Κάθε σημείο αντιπροσωπεύει τη μέση φωτεινότητα για όλο το video και τον αριθμό των μετρήσεων σε διαφορετικούς ρυθμούς bit. Έτσι έχουμε κάθε γραμμή έχει δέκα σημεία δηλαδή δέκα ρυθμούς bit.

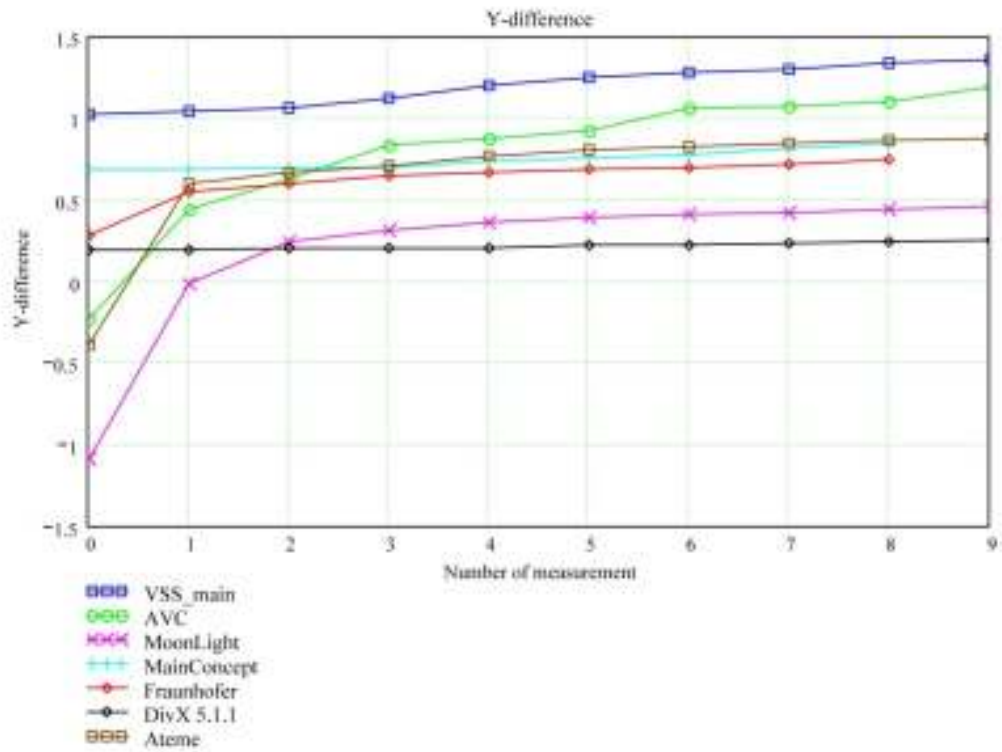
4.3.3.1 Video Bankomatdi



Εικόνα 113 – Διάγραμμα μεταβολής φωτεινότητας για το video Bankomatdi

Όλα τα codecs που χρησιμοποιούν το πρότυπο H.264 αυξάνουν τη φωτεινότητα. Όσο αυξάνεται ο ρυθμός bit τόσο αυξάνεται η φωτεινότητα. Το codec Fraunhofer αυξάνει τη φωτεινότητα πιο πολύ από τα άλλα codecs.

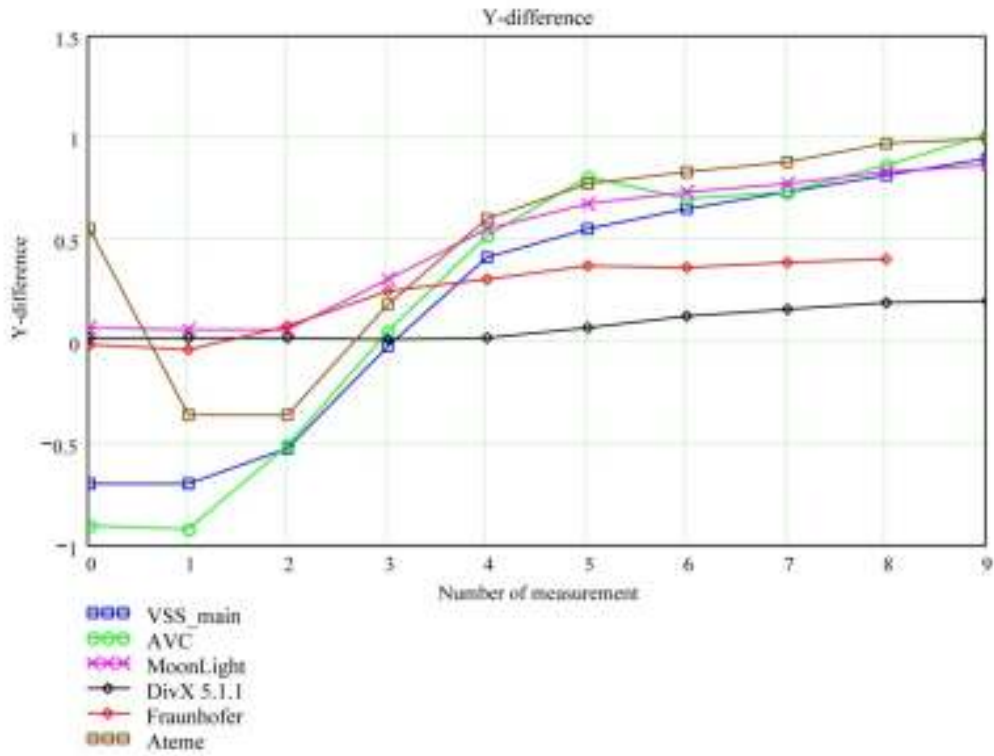
4.3.3.2 Video Battle



Εικόνα 114 - Διάγραμμα μεταβολής φωτεινότητας για το video Battle

Όλα τα codecs που χρησιμοποιούν το H.264 αυξάνουν τη φωτεινότητα. Όσο αυξάνεται ο ρυθμός bit τόσο αυξάνεται η φωτεινότητα. Παρατηρείται ότι το codec Videosoft αυξάνει τη φωτεινότητα πιο πολύ από τα άλλα codecs.

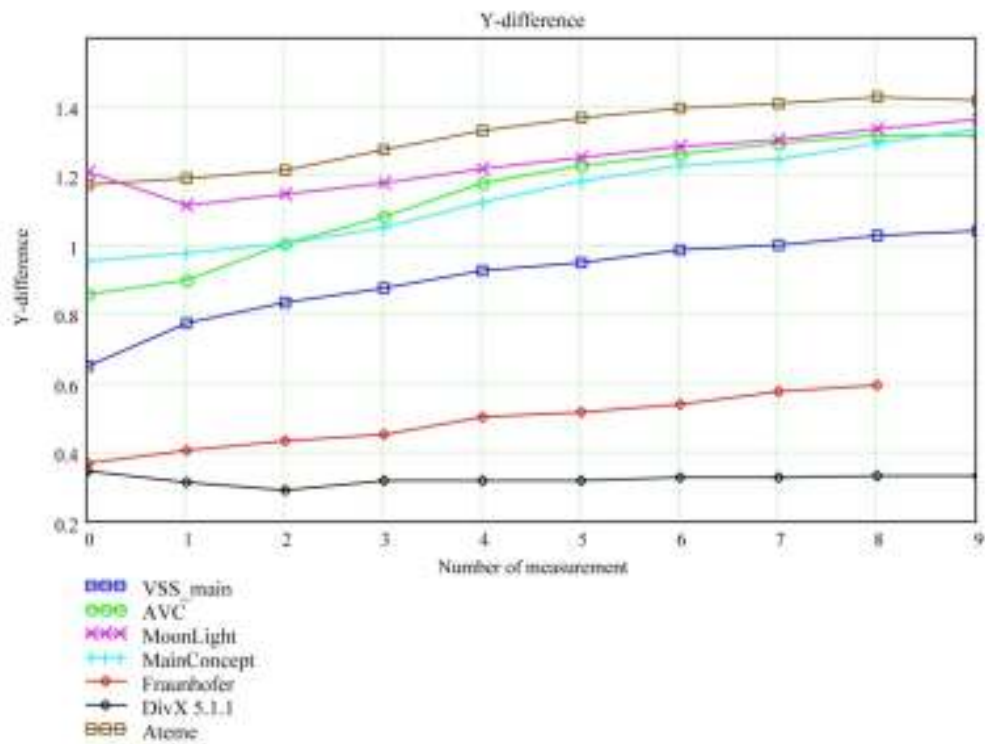
4.3.3.3 Video Bbc3di



Εικόνα 115 - Διάγραμμα μεταβολής φωτεινότητας για το video Bbc3di

Όλα τα codecs που χρησιμοποιούν το H.264 αυξάνουν τη φωτεινότητα. Όσο αυξάνεται ο ρυθμός bit τόσο αυξάνεται η φωτεινότητα. Η αλλαγή της φωτεινότητας για χαμηλούς ρυθμούς bit είναι μικρή.

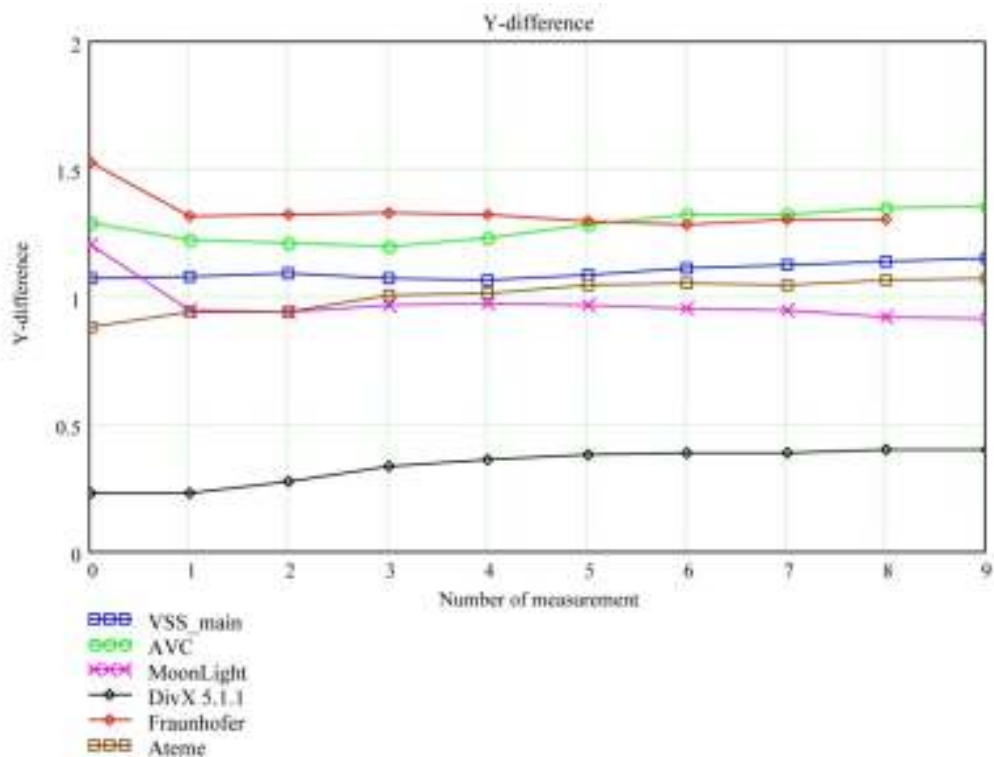
4.3.3.4 Video Foreman



Εικόνα 116 - Διάγραμμα μεταβολής φωτεινότητας για το video Foreman

Όλα τα codecs που χρησιμοποιούν το H.264 αυξάνουν τη φωτεινότητα. Όσο αυξάνεται ο ρυθμός bit τόσο αυξάνεται η φωτεινότητα.

4.3.3.5 Video Susidi



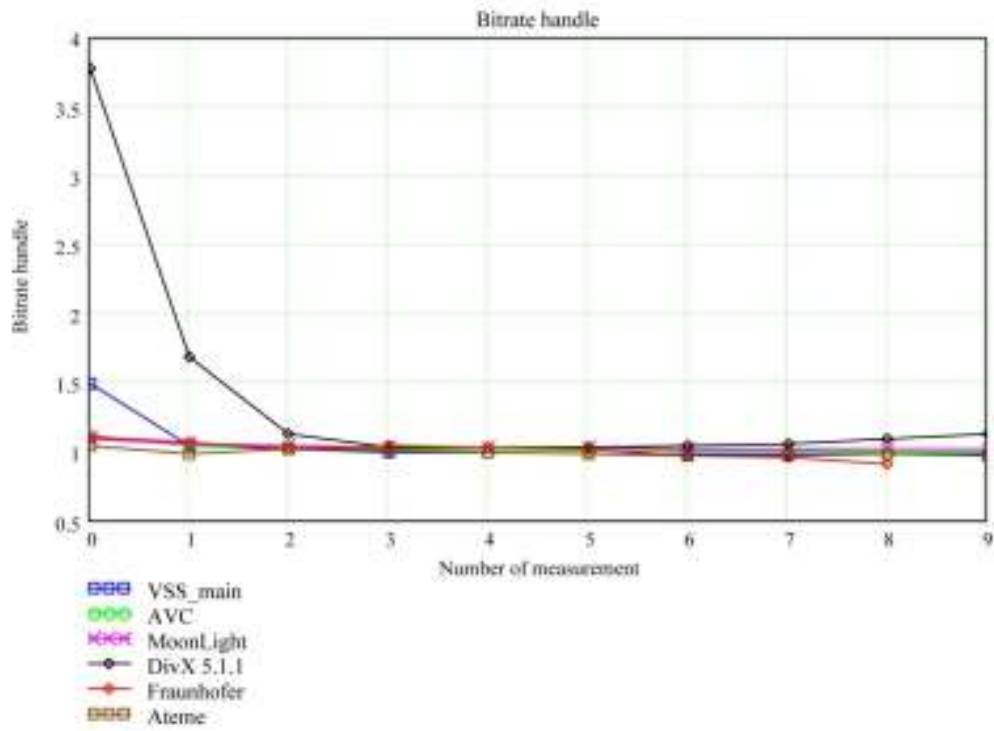
Εικόνα 117 - Διάγραμμα μεταβολής φωτεινότητας για το video Susidi

Όλα τα codecs που χρησιμοποιούν το H.264 αυξάνουν τη μέση φωτεινότητα. Η αλλαγή είναι περίπου σταθερή για όλους τους ρυθμούς bit.

4.3.4 Διαγράμματα διαχείρισης ρυθμού bit

Τα διαγράμματα διαχείρισης ρυθμού bit (Bit rate handling diagrams) δείχνουν πως τα codecs διαχειρίζονται διαφορετικούς ρυθμούς bit. Κάθε γραμμή αποτελείται από δέκα σημεία που αντιστοιχούν σε διαφορετικούς ρυθμούς bit. Η τιμή 1 σημαίνει ότι το codec διατηρεί τον επιθυμητό ρυθμό bit, ενώ τιμή μεγαλύτερη από 1 σημαίνει ότι το codec συμπιέζει τα δεδομένα με μεγαλύτερο ρυθμό bit από τον επιθυμητό.

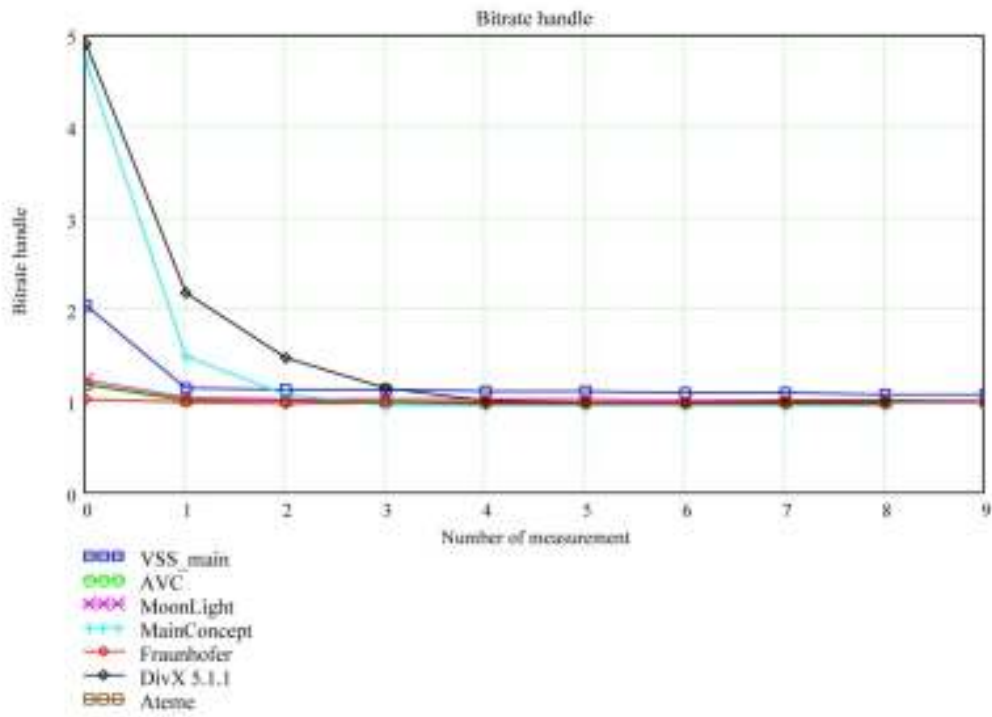
4.3.4.1 Video Bankomatdi



Εικόνα 118 - Διάγραμμα διαχείρισης ρυθμού bit για το video Bankomatdi

Όπως βλέπουμε τα codecs Videosoft και DivX αυξάνουν τους χαμηλούς ρυθμούς bit.

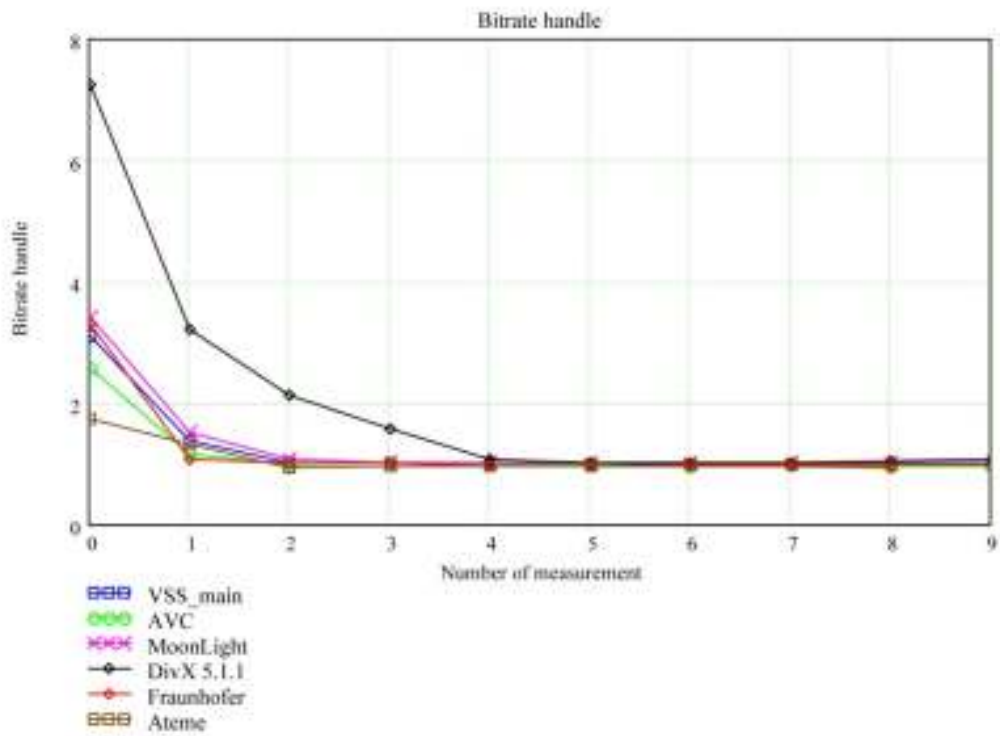
4.3.4.2 Video Battle



Εικόνα 119 - Διάγραμμα διαχείρισης ρυθμού bit για το video Battle

Το Videosoft αυξάνει ελαφρώς τους χαμηλούς ρυθμούς bit, ενώ τα codecs DivX και MainConcept τους αυξάνουν αρκετά.

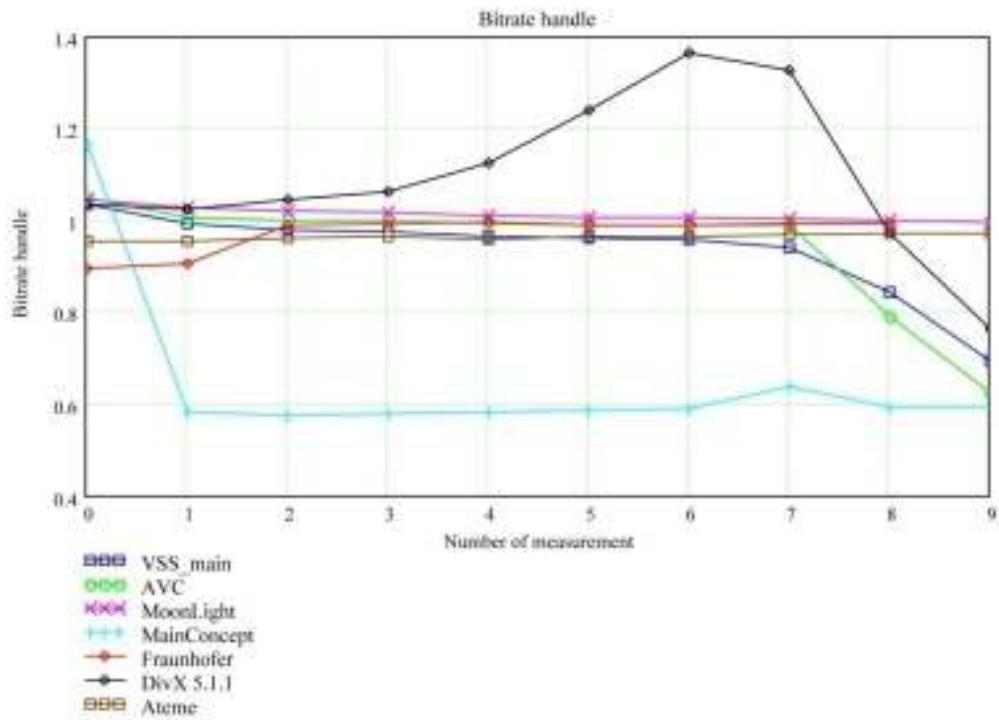
4.3.4.3 Video Bbc3di



Εικόνα 120 - Διάγραμμα διαχείρισης ρυθμού bit για το video Bbc3di

Όλα τα codecs που χρησιμοποιούν το πρότυπο H.264 αυξάνουν τους χαμηλούς ρυθμούς bit, ενώ το DivX τους αυξάνουν αρκετά.

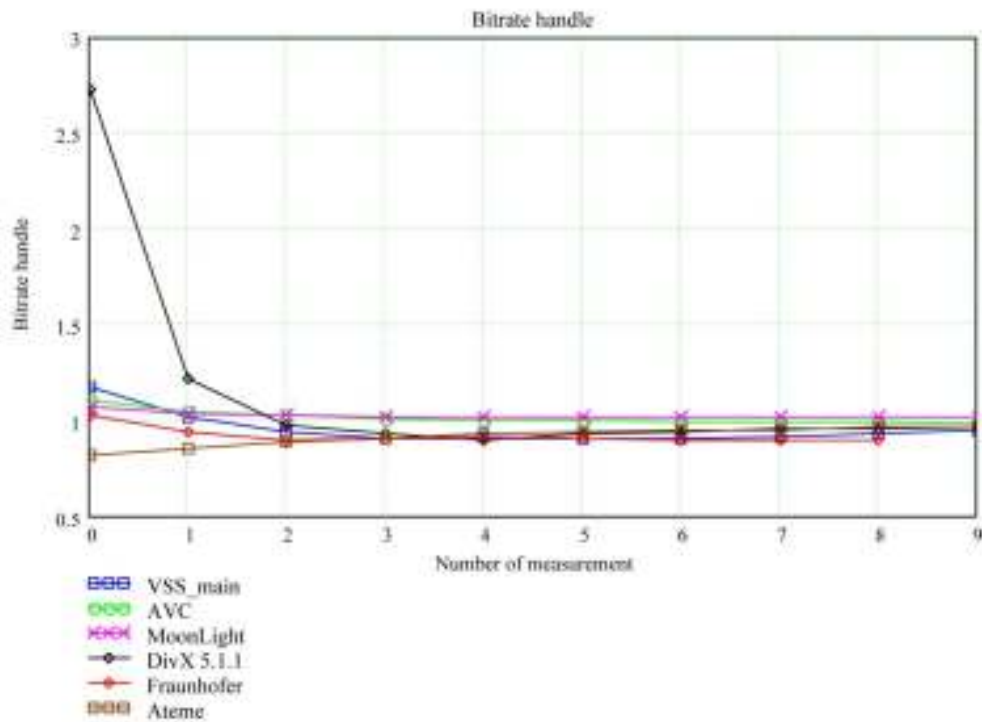
4.3.4.4 Video Foreman



Εικόνα 121 - Διάγραμμα διαχείρισης ρυθμού bit για το video Foreman

Το codec MainConcept μειώνει αρκετά τους επιθυμητούς ρυθμούς bit ενώ αντίθετα το DivX τους αυξάνει.

4.3.4.5 Video Susidi

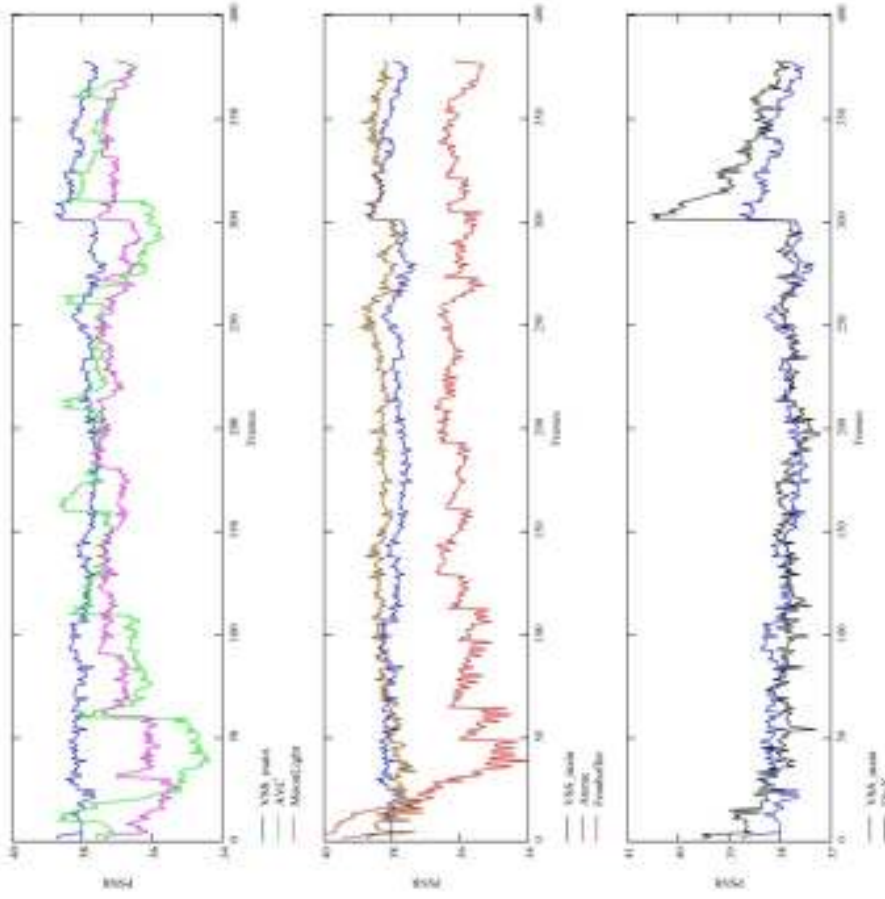


Εικόνα 122 - Διάγραμμα διαχείρισης ρυθμού bit για το video Susidi

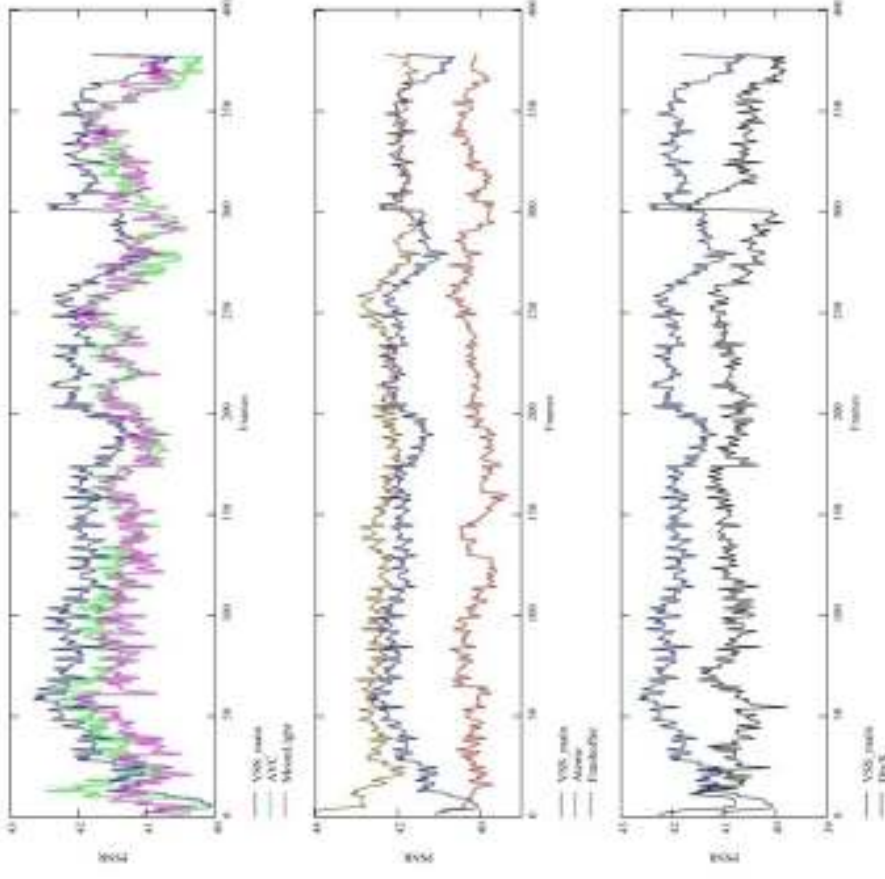
Όπως παρατηρούμε το DivX αυξάνει τους επιθυμητούς ρυθμούς bit.

4.3.5 Σύγκριση ανά πλαίσιο

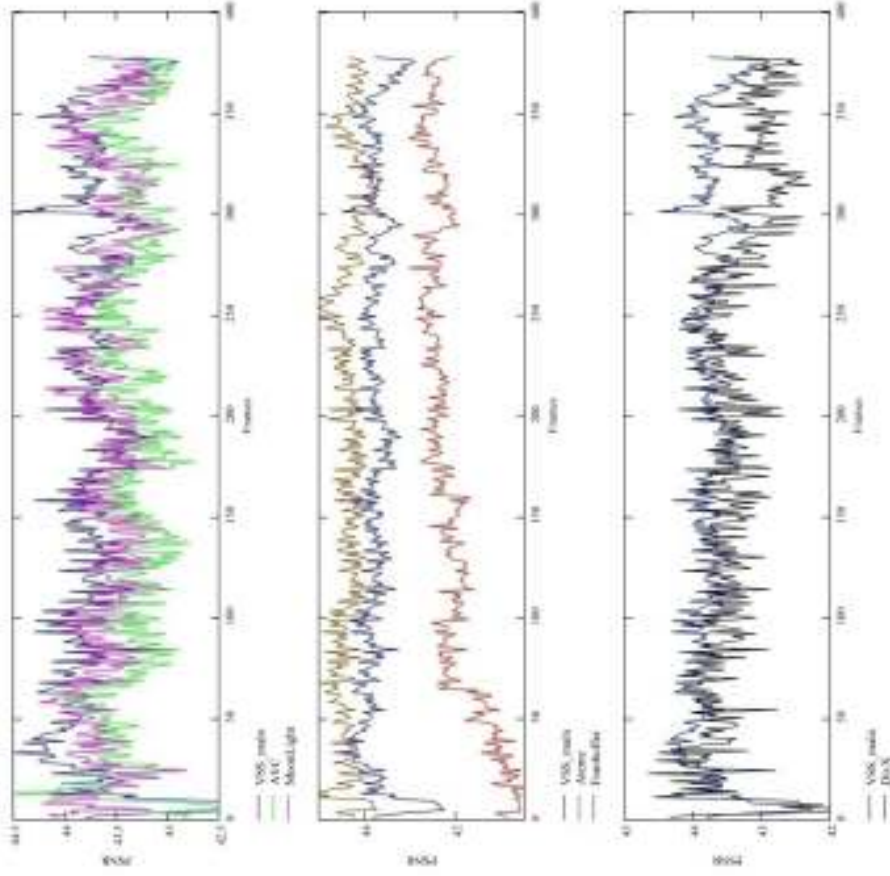
Τα διαγράμματα σύγκρισης ανά πλαίσιο (Per frame sequences comparison) δείχνουν πως η ποιότητα συμπίεσης αλλάζει από πλαίσιο σε πλαίσιο. Ο x-άξονας αντιπροσωπεύει τον αριθμό των πλαισίων και ο γ-άξονας την τιμή PSNR για το συγκεκριμένο πλαίσιο. Η συνεχής διαφορά στο video Susidi οφείλεται στον θόρυβο σε κάθε δεύτερο πλαίσιο.



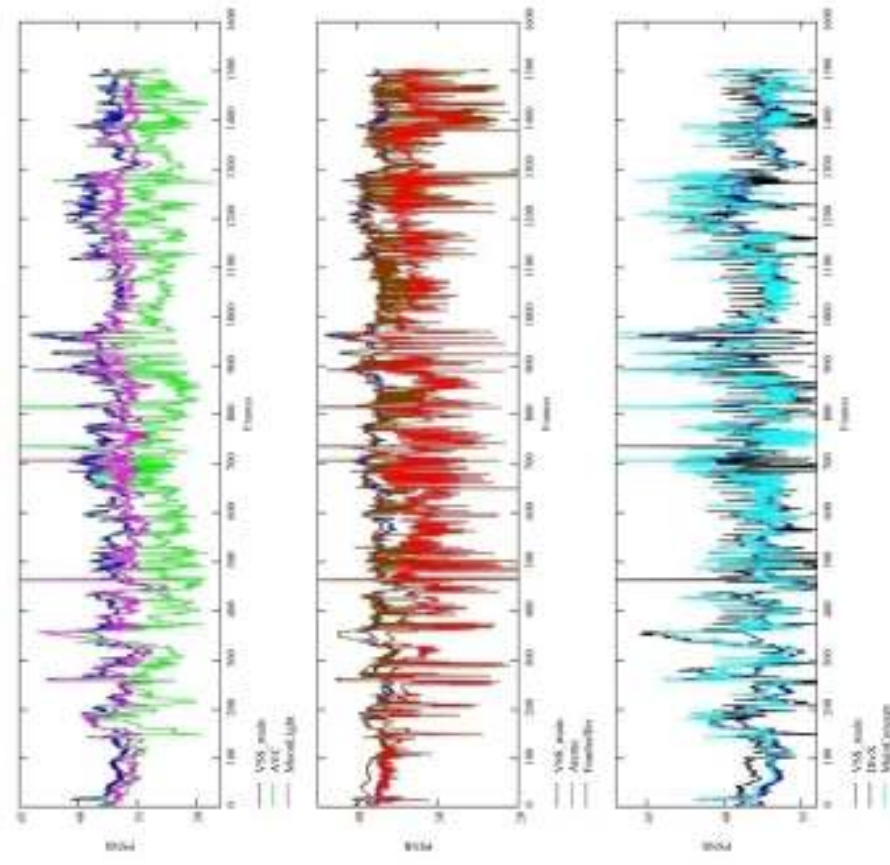
Εικόνα 123 - PSNR ανά πλαίσιο για το video BankomatIdi για ρυθμό bit 100kbps



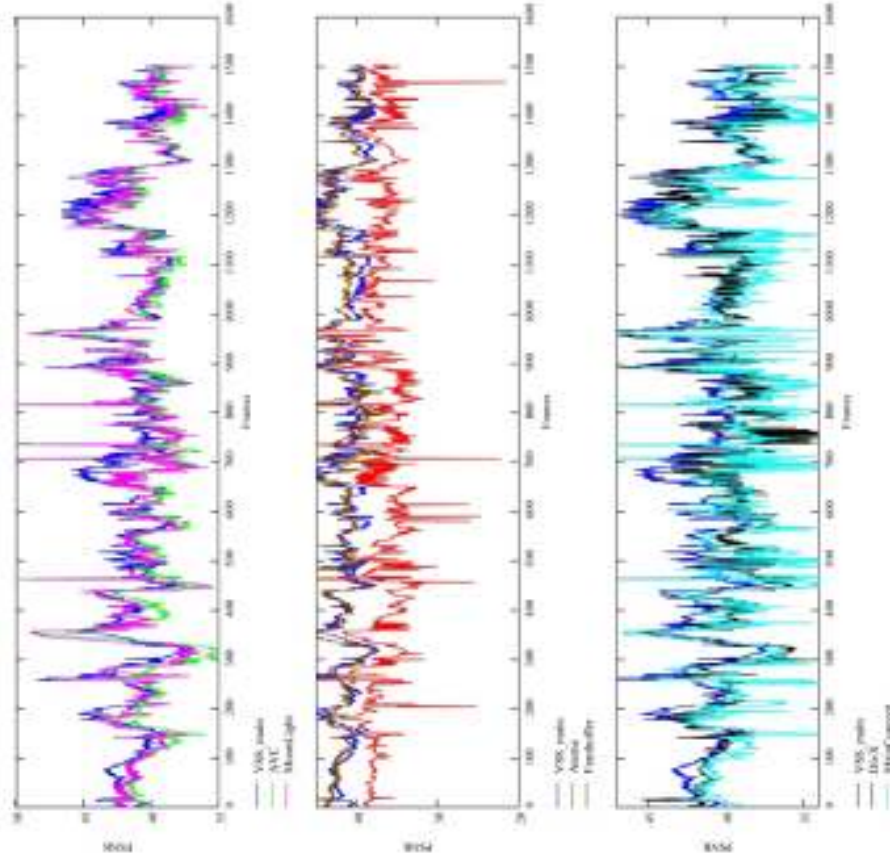
Εικόνα 124 - PSNR ανά πλαίσιο για το video BankomatIdi για ρυθμό bit 700kbps



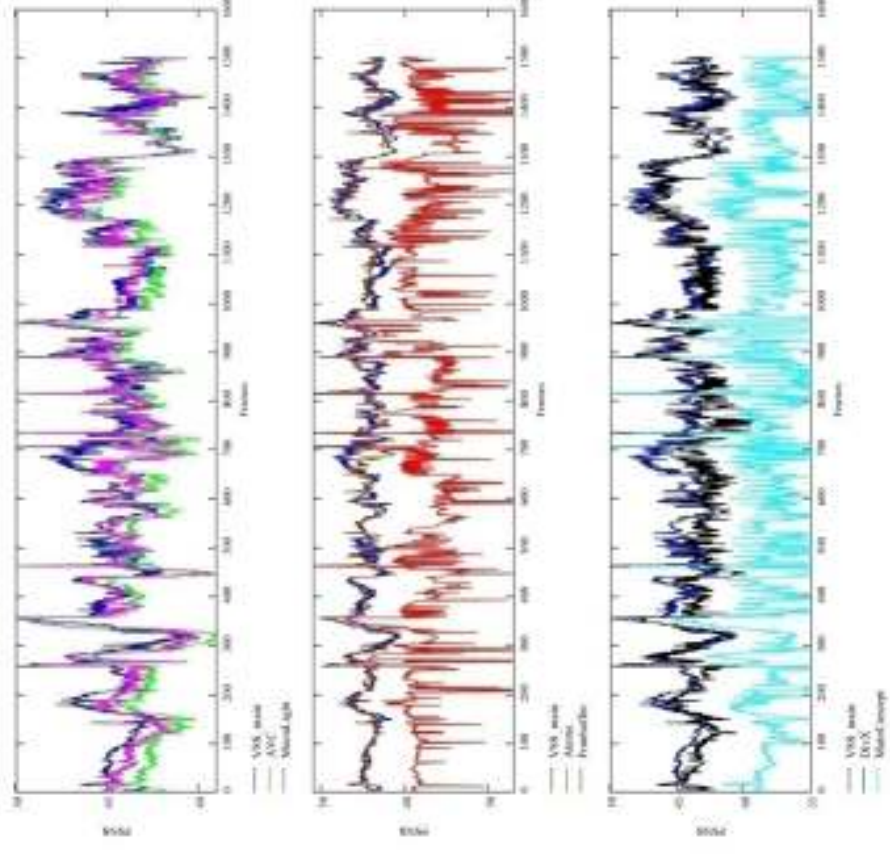
Εικόνα 125 - PSNR ανά πλάισιο για το video Bankomatdi για ρυθμό bit 2340kbps



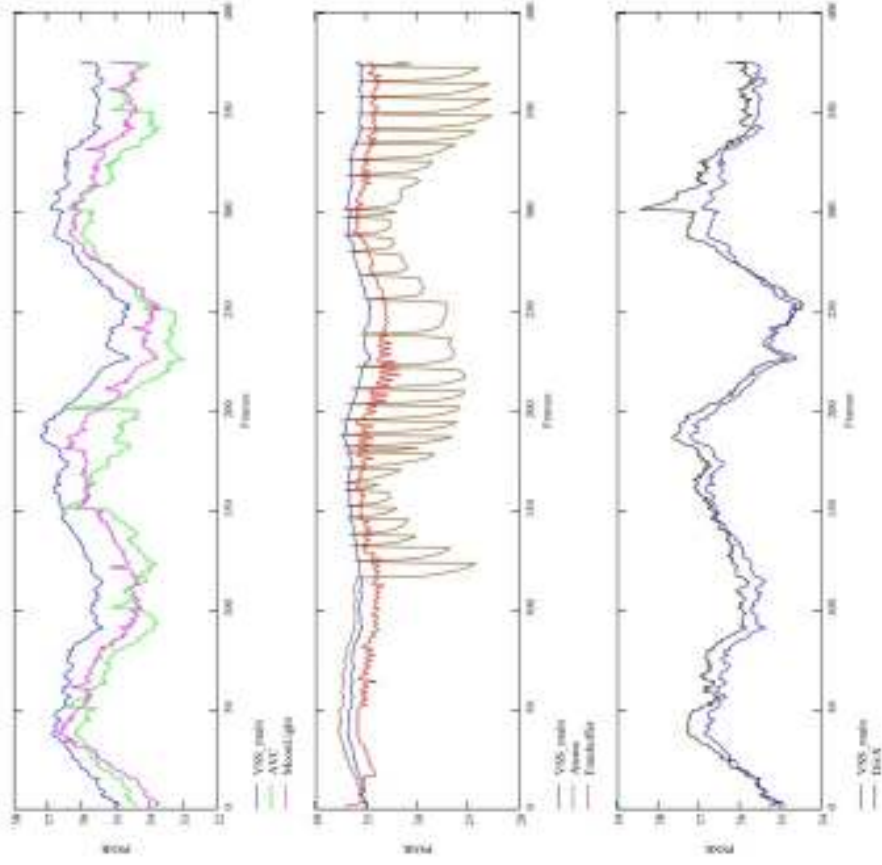
Εικόνα 126 - PSNR ανά πλάισιο για το video Battle για ρυθμό bit 100kbps



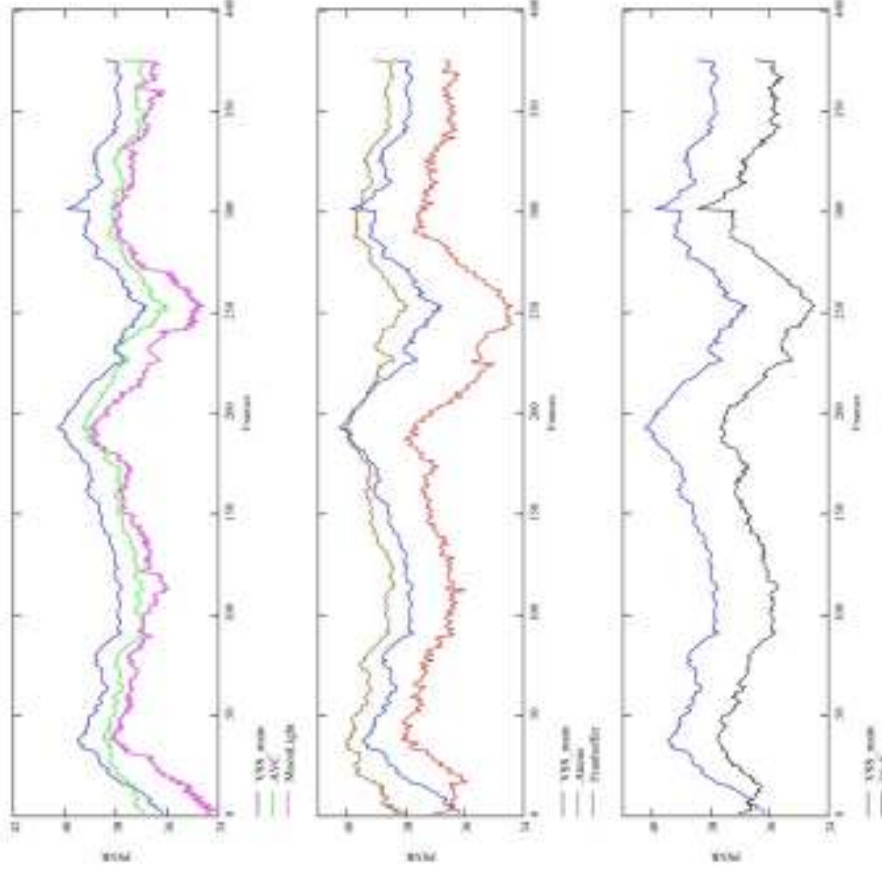
Εικόνα 127 - PSNR ανά πλαίσιο για το video Battle για ρυθμό bit 700kbps



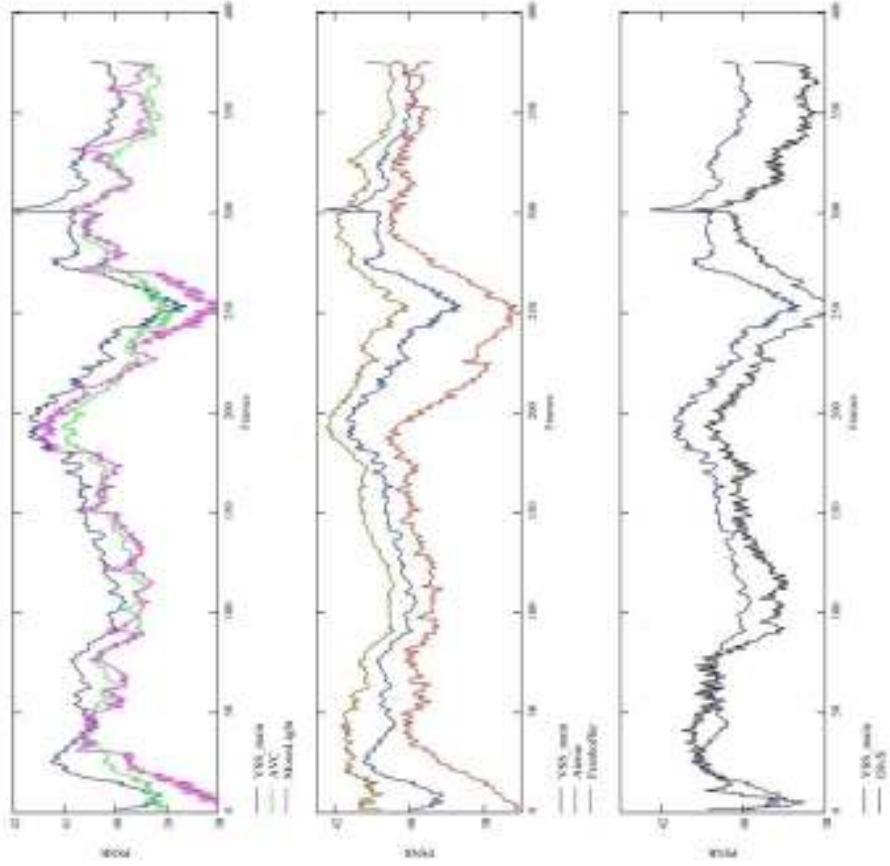
Εικόνα 128 - PSNR ανά πλαίσιο για το video Battle για ρυθμό bit 2340kbps



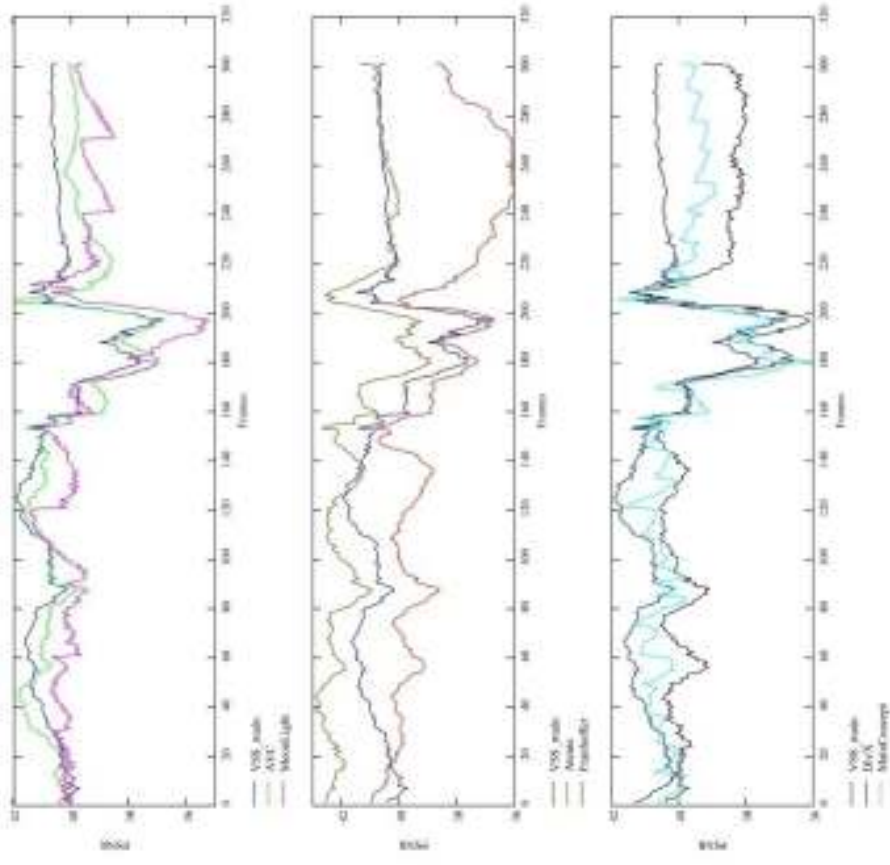
Εικόνα 129 - PSNR ανά πλαίσιο για το video Bbc3di για ρυθμό bit 100kbps



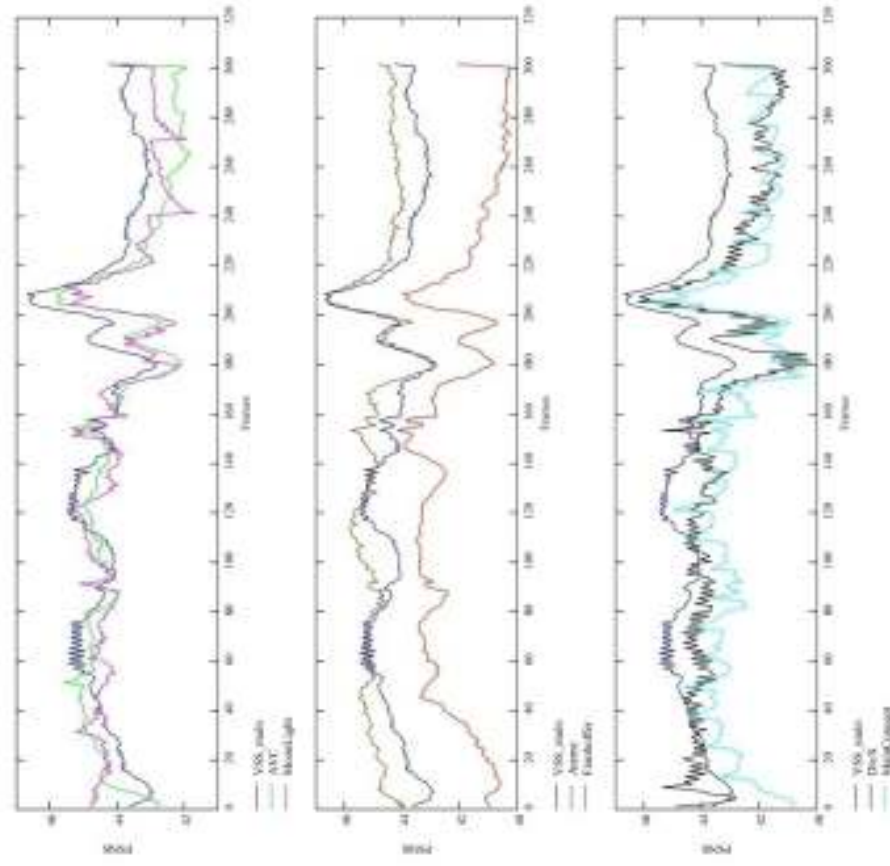
Εικόνα 130 - PSNR ανά πλαίσιο για το video Bbc3di για ρυθμό bit 700kbps



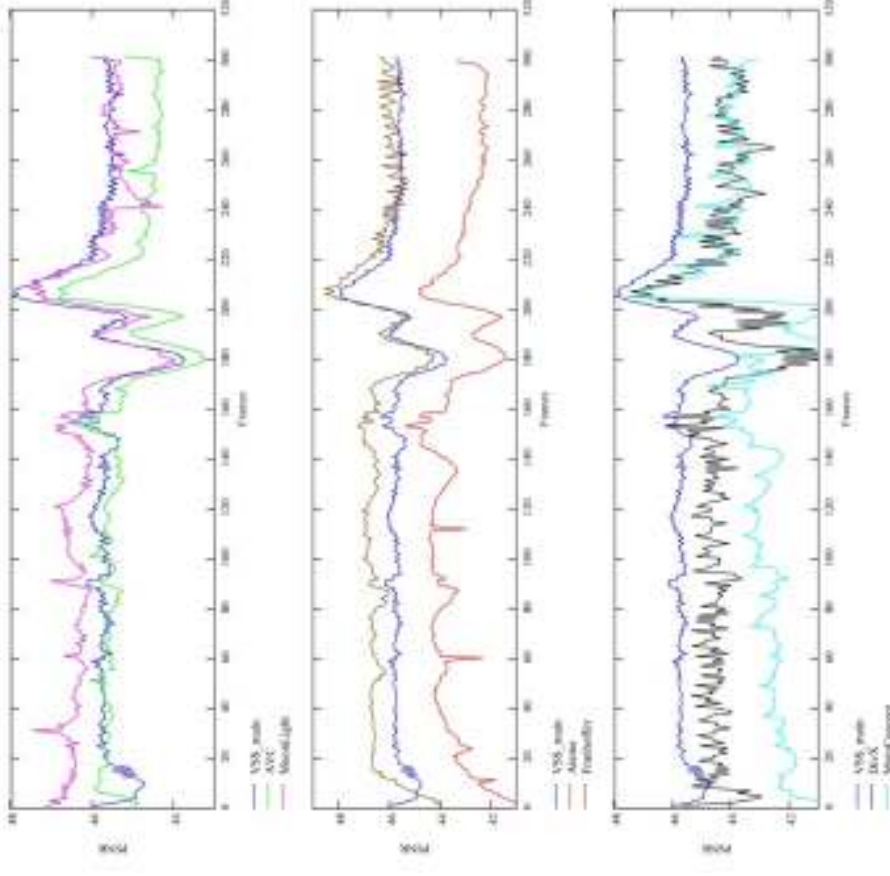
Εικόνα 131 - PSNR ανά πλαίσιο για το video Bbc3di για ρυθμό bit 2340kbps



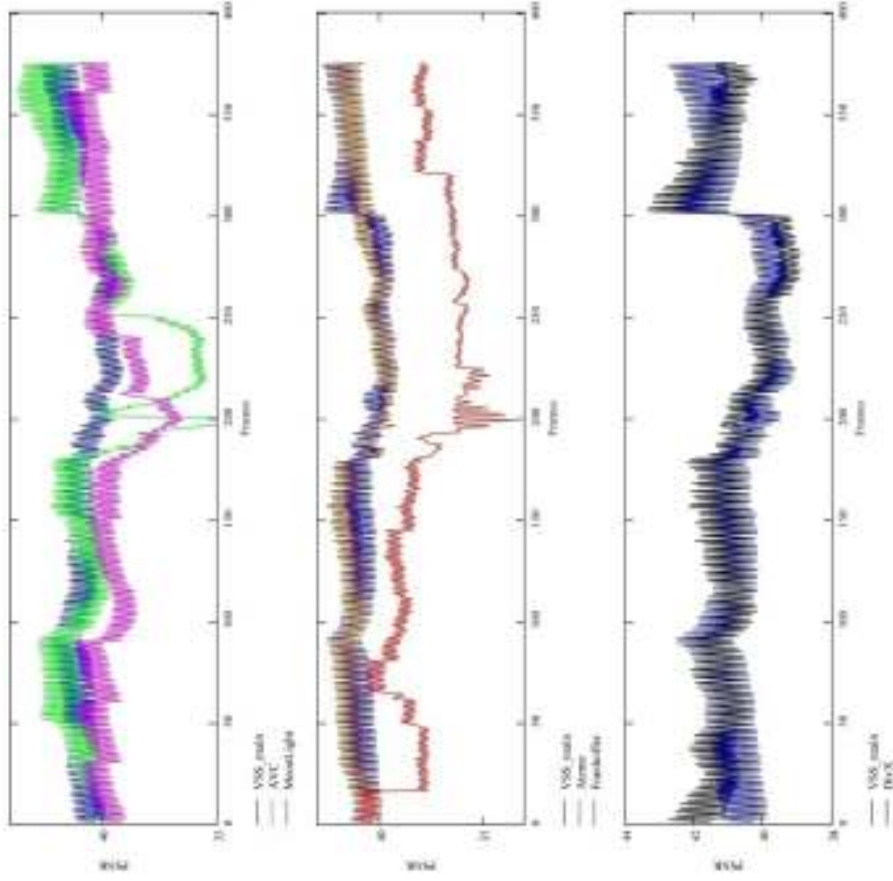
Εικόνα 132 - PSNR ανά πλαίσιο για το video Foreman για ρυθμό bit 100kbps



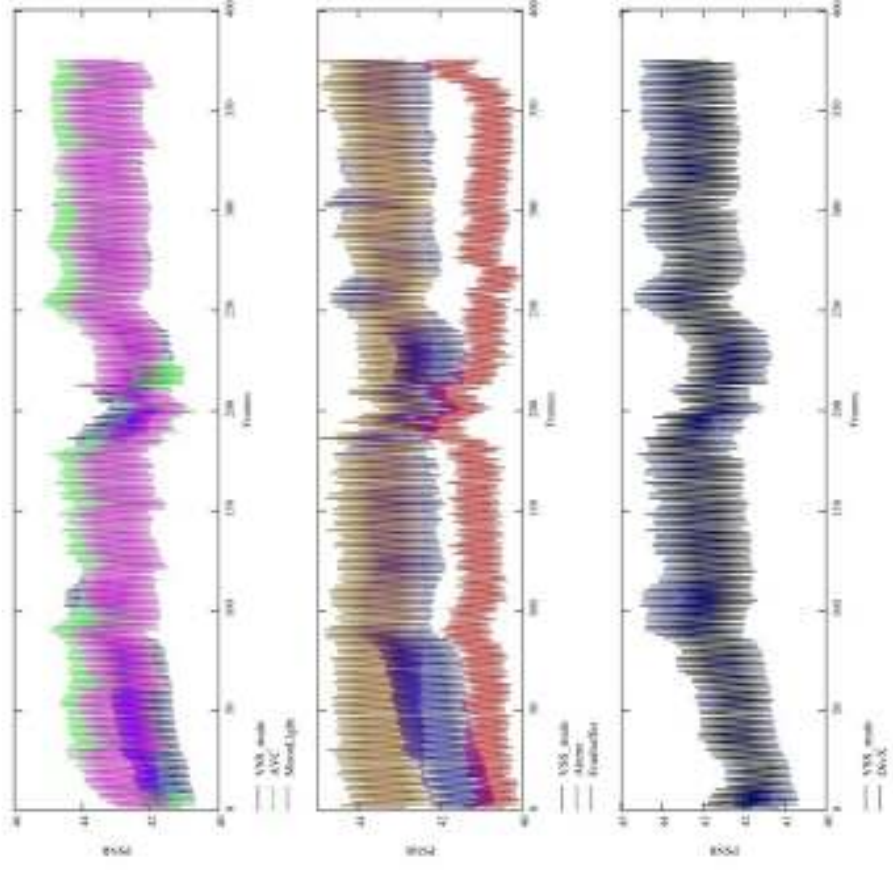
Εικόνα 133 - PSNR ανά πλαίσιο για το video Foreman για ρυθμό bit 700kbps



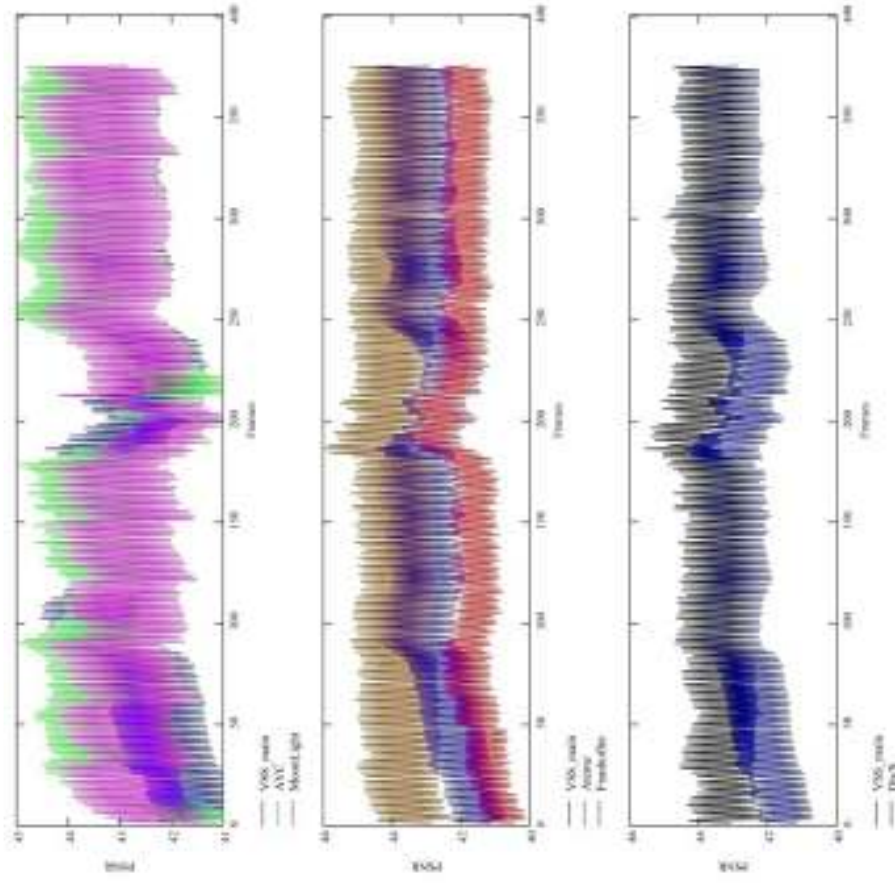
Εικόνα 134 - PSNR ανά πλαίσιο για το video Foreman για ρυθμό bit 2340kbps



Εικόνα 135 - PSNR ανά πλαίσιο για το video Susidi για ρυθμό bit 100kbps



Εικόνα 136 - PSNR ανά πλαίσιο για το video Susidi για ρυθμό bit 700kbps



Εικόνα 137 - PSNR ανά πλαίσιο για το video Susidi για ρυθμό bit 2340kbps

4.3.6 Οπτική σύγκριση των H.264 codecs και των DivX codecs

Η σύγκριση έγινε μεταξύ των codecs Ateve και DivX με ρυθμό bit 700kbps. Για τη σύγκριση χρησιμοποιήθηκαν τα video Bbc2di και Foreman.

4.3.6.1 Video Bbc3di, frame 280

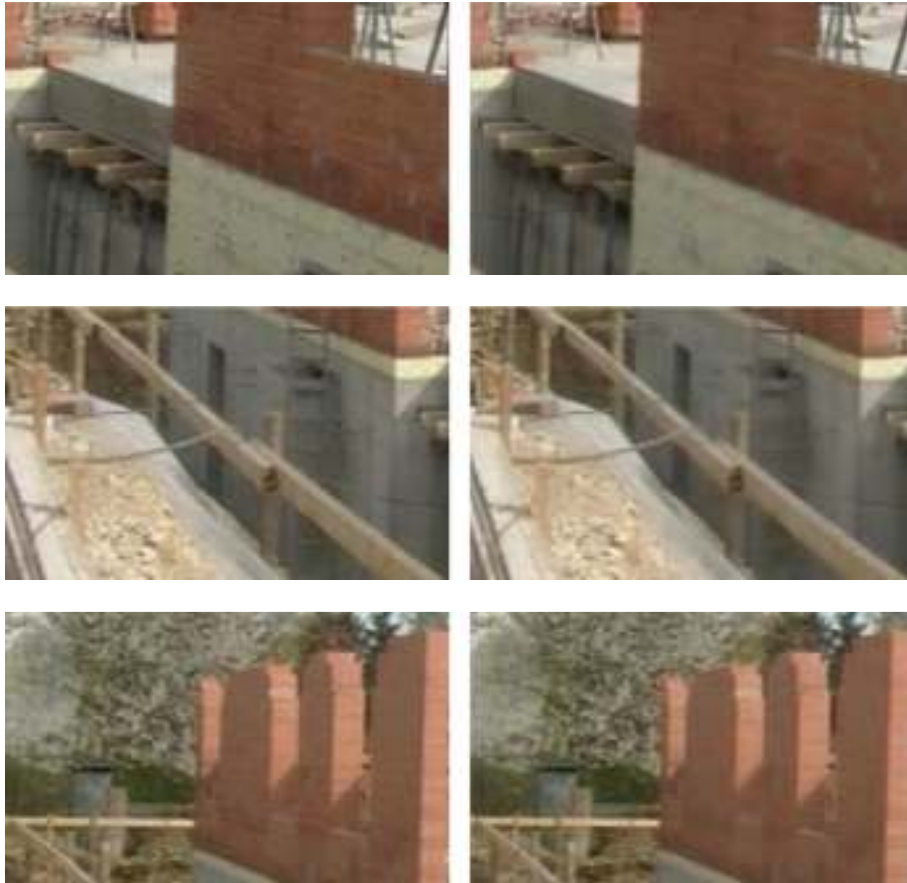
Οι παρακάτω εικόνες δείχνουν καλύτερη ποιότητα για το codec που χρησιμοποιεί το πρότυπο H.264.



Εικόνα 138 - Video Bbc3di πλαίσιο 280 συμπιεσμένο με το codec Ateve (αριστερά) και DivX (δεξιά)

4.3.6.2 Video Foreman, frame 282 και 9

Οι παρακάτω εικόνες δείχνουν καλύτερη ποιότητα για το codec που χρησιμοποιεί το πρότυπο H.264.

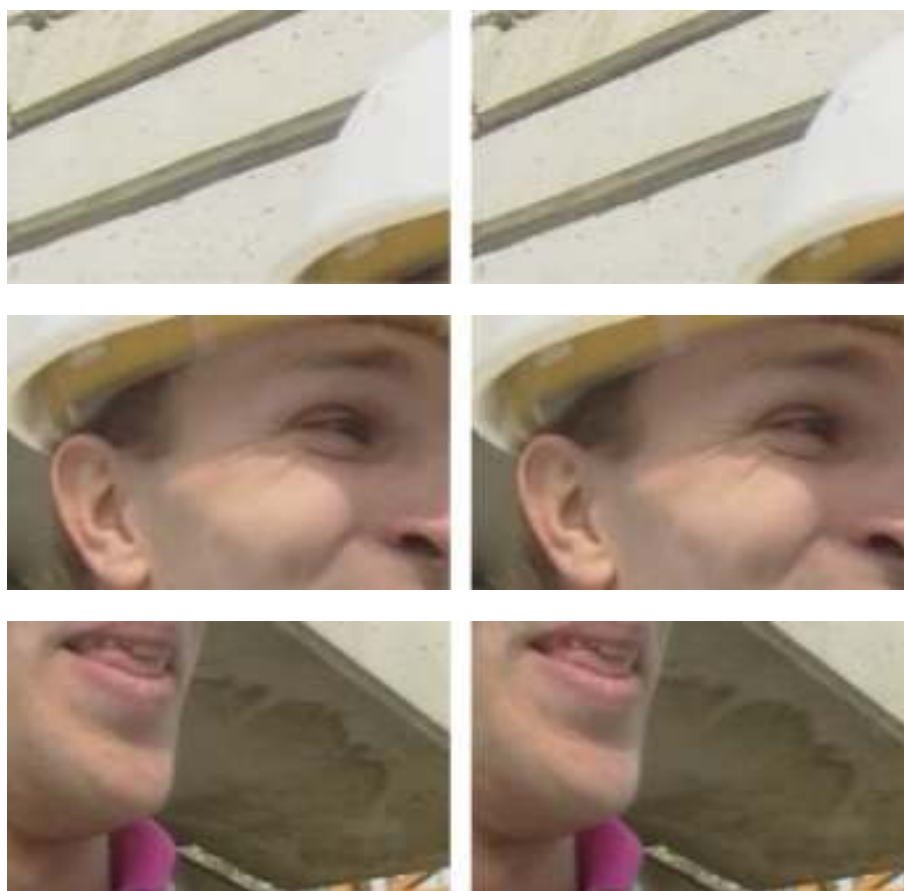


Εικόνα 139 - Video Foreman πλαίσιο 282 συμπιεσμένο με το codec AteMe (αριστερά) και DivX (δεξιά)

4.3.7 Συγκεντρωτικά συμπεράσματα

Για ίδιο PSNR τα codecs που χρησιμοποιούν το πρότυπο H.264 φαίνεται να αποδίδουν αρκετά καλύτερα από τα προγενέστερα codecs. Τα περισσότερα από τα codecs που ελέγχθηκαν είναι βελτιστοποιημένα για μέγιστη ταχύτητα για τα μηχανήματα της συγκεκριμένης χρονικής στιγμής και δεν χρησιμοποιήθηκαν όλες οι δυνατότητες του προτύπου H.264.

Οι ακόλουθες εικόνες δείχνουν με μεγαλύτερη λεπτομέρεια την ποιότητα των H.264 και DivX.



Εικόνα 140 - Video Foreman πλαίσιο 9 συμπιεσμένο με το codec Ateme (αριστερά) και DivX (δεξιά)

4.3.8 Ένας ανεπίσημος τρόπος σύγκρισης

Όπως παρατηρείται από τα παραπάνω διαγράμματα κάθε codec έχει διαφορετικά αποτελέσματα για κάθε εξεταζόμενο video. Αυτό συμβαίνει λόγω των διαφορετικών χαρακτηριστικών του κάθε video (κίνηση, θόρυβος κτλ). Αυτό μας επιτρέπει να ορίσουμε ποια codecs αποδίδουν καλά για κάθε είδος video και ποια έχουν καλά αποτελέσματα μόνο για κάποια συγκεκριμένα είδη video. Ωστόσο θα είχε ενδιαφέρον να δούμε κάποια γενικά χαρακτηριστικά κάθε ενός codec στο σύνολο του πειράματος.

Ο μέσος όρος από όλες τις τιμές του πειράματος για κάθε codec δεν θα ήταν αντιπροσωπευτικό δείγμα και έτσι εφαρμόστηκε μία ανεπίσημη εκτίμηση (informal estimation). Κάθε codec έπαιρνε ένα συγκεκριμένο βαθμό για κάθε αποτέλεσμα που είχε.

Αν κάποιο codec ήταν πιο σταθερό από τα άλλα σε περισσότερα από ένα σημεία έπαιρνε βαθμό 3 ανεξάρτητα από τα υπόλοιπα αποτελέσματα.

Αν κάποιο codec ήταν χειρότερο από όλα τα άλλα σε κάποιο σημείο τότε έπαιρνε βαθμό 1.

Αλλιώς ο βαθμός που έπαιρνε ήταν 2.

Με αυτή τη μέθοδο υπολογίστηκαν τα μεγέθη Y-PSNR, U-PSNR, V-PSNR και διαφορά Y (Y-difference).

	Bankomatdi	Battle	Bbc3di	Foreman	Susidi	Σύνολο	Θέση
Ateme	3	3	3	2	3	14	1
DivX	2	2	2	2	3	11	2
Fraunhofer	1	2	2	2	1	8	4,5
MainCocnept	-	1	-	1	-	2	-
Moonlight	1	2	1	2	2	8	4,5
Mpegable AVC	2	2	2	2	2	10	3
Videosoft	2	2	2	3	1	10	3

Πίνακας 45 - Αποτελέσματα της ανεπίσημης μεθόδου για το Y-PSNR

	Bankomatdi		Battle		Bbc3di		Foreman		Susidi		Σύνολο	Θέση
	U	V	U	V	U	V	U	V	U	V		
Ateme	3	3	3	3	3	3	3	3	3	3	30	1
DivX	1	1	2	2	1	1	2	1	2	1	14	5, 6
Fraunhofer	1	2	1	1	2	2	1	2	1	1	14	5, 6
MainCocnept			2	2			2	2			8	-
Moonlight	2	2	2	2	2	2	2	2	2	2	20	3
Mpegable AVC	2	1	2	2	2	2	2	1	2	1	17	4
Videosoft	2	2	2	2	2	2	2	3	2	2	21	2

Πίνακας 46 - Αποτελέσματα της ανεπίσημης μεθόδου για το U-PSNR και V-PSNR

	Bankomatdi	Battle	Bbc3di	Foreman	Susidi	Y-diff	Υ	UV/2	Σύνολο	Θέση
	Y-diff	Y-diff	Y-diff	Y-diff	Y-diff	Y-diff				
Ateme	3	2	2	1	2	10	14	15	39	1
DivX	3	3	3	3	3	15	11	7	33	2
Fraunhofer	1	2	2	2	1	8	8	7	23	6
MainCocnept	-	3	-	2	-	5	2	4	11	-
Moonlight	2	2	2	2	2	10	8	10	28	5
Mpegable AVC	2	2	3	2	1	10	10	8.5	28.5	4
Videosoft	2	1	2	2	2	9	10	10.5	29.5	3

Πίνακας 47 – Γενικά αποτελέσματα της ανεπίσημης μεθόδου

4.4 Video που χρησιμοποιήθηκαν

4.4.1 Videos τηλεδιάσκεψης

4.4.1.1 Salesman

Τίτλος video	Salesman
Ανάλυση	176x144
Αριθμός πλαισίων	449
Χρώμα	ΥV12
Πλαίσια ανά δευτερόλεπτο	30
Πηγή	Ασυμπιεστο (πρότυπο video), σταδιακό (progressive)

Πίνακας 48 – Χαρακτηριστικά του video Salesman



Εικόνα 141 - Πλαίσιο 1 (δεξιά) και 100 (αριστερά) του video Salesman

Στο συγκεκριμένο video η κάμερα είναι σταθερή και δεν υπάρχει αρκετή κίνηση. Έτσι το video αυτό μπορεί να χρησιμοποιηθεί για τον έλεγχο της συμπεριφοράς των codecs στις σταθερές σκηνές με μικρή χωρική ανάλυση.

4.4.1.2 Foreman

Τίτλος video	Foreman
Ανάλυση	352x288
Αριθμός πλαισίων	300
Χρώμα	ΥV12
Πλαίσια ανά δευτερόλεπτο	30
Πηγή	Ασυμπιεστο (πρότυπο video), σταδιακό (progressive)

Πίνακας 49 - Χαρακτηριστικά του video Foreman



Εικόνα 142 - Πλαίσιο 77 (αριστερά) και 258 (δεξιά) του video Foreman

Το video αυτό περιλαμβάνει ένα πρόσωπο σε μία ιδιαίτερη έκφραση. Δεν υπάρχει μεγάλο ποσοστό κίνησης αλλά η κίνηση που υπάρχει είναι ασυνάρτητη και δεν έχει κάποια γραμμικά χαρακτηριστικά. Η περιπλοκότητα του προσώπου και η κίνησή του δημιουργεί προβλήματα στη διαδικασία απόδοσης της κίνησης. Επίσης η κάμερα τρέμει και έτσι δημιουργούνται επιπλέον προβλήματα. Στο τέλος της ακολουθίας η κάμερα στρέφεται ξαφνικά προς το κτίριο και ακολουθεί μία σκηνή με σχεδόν καθόλου κίνηση. Έτσι η ακολουθία αυτή μπορεί να χρησιμοποιηθεί για τον έλεγχο της συμπεριφοράς των codecs για video που συνδυάζουν στατικές σκηνές και σκηνές με αυξημένη κίνηση.

4.4.1.3 News

Τίτλος video	News
Ανάλυση	352x288
Αριθμός πλαισίων	300
Χρώμα	ΥV12
Πλαίσια ανά δευτερόλεπτο	30
Πηγή	Ασυμπιεστο (πρότυπο video), σταδιακό (progressive)

Πίνακας 50 - χαρακτηριστικά του video News



Εικόνα 143 - Πλαίσιο 1 (αριστερά) και 100 (δεξιά) του video News

Το video αυτό παρουσιάζει δύο παρουσιαστές μπροστά από ένα στατικό φόντο. Το φόντο όμως έχει μία τηλεόραση με κινούμενες εικόνες. Η κάμερα είναι σταθερή. Η κίνηση των παρουσιαστών δεν είναι έντονη αλλά η κίνηση στην τηλεόραση στο φόντο είναι έντονη. Έτσι η ακολουθία αυτή μπορεί να χρησιμοποιηθεί για τον έλεγχο της συμπεριφοράς των codecs σε video όπου έχουμε στατικές σκηνές αλλά με κάποια σημεία έντονης κίνησης.

4.4.1.4 Akiyo

Τίτλος video	Akiyo
Ανάλυση	352x288
Αριθμός πλαισίων	300
Χρώμα	ΥV12
Πλαίσια ανά δευτερόλεπτο	25
Πηγή	Πρότυπο video, σταδιακό (progressive)

Πίνακας 51 - Χαρακτηριστικά του video Akiyo



Εικόνα 144 - Πλαίσιο 1 του video Akiyo

Το video “Akiyo” είναι ένα κλασικό video τηλεδιάσκεψης. Έχει σταθερό φόντο και έναν ομιλητή στο προσκήνιο με χαμηλή κίνηση, χωρίς εναλλαγή σκηνών.

4.4.1.5 Carphone

Τίτλος video	Carphone
Ανάλυση	176x144
Αριθμός πλαισίων	382
Χρώμα	ΥV12
Πλαίσια ανά δευτερόλεπτο	25
Πηγή	Πρότυπο video, σταδιακό (progressive)

Πίνακας 52 - Χαρακτηριστική του video Carphone



Εικόνα 145 - Πλαίσιο 319 του video Carphone

Πρόκειται για ένα video τηλεδιάσκεψης με αργή αλλαγή του φόντου, ένα τυπικό κούνημα της κάμερας και ένα άντρα που μιλάει στο προσκήνιο. Η κίνηση του ομιλητή είναι έντονη.

4.4.2 Videos ταινιών

4.4.2.1 Battle

Τίτλος video	Battle
Ανάλυση	704x288
Αριθμός πλαισίων	1599
Χρώμα	ΥV12
Πλαίσια ανά δευτερόλεπτο	24
Πηγή	MPEG-2 (DVD), απόπλεξη (deinterlace) FlaskMPEG

Πίνακας 53 - Χαρακτηριστικά του video Battle



Εικόνα 146 - Πλαίσιο 839 του video Battle

Η εικόνα αυτή είναι από την ταινία "Terminator 2". Σε σχέση με τα άλλα video είναι το πιο δύσκολο για συμπίεση λόγω της συνεχούς εναλλαγής του φωτισμού (εκρήξεις και laser), της γρήγορης κίνησης και της συχνής εναλλαγής σκηνών. Αυτά τα χαρακτηριστικά οδηγούν συνήθως τα codecs να συμπιέζουν τα πλαίσια σαν πλαίσια τύπου-I (I-frames).

4.4.2.2 Smith

Τίτλος video	Smith
Ανάλυση	720x432
Αριθμός πλαισίων	772
Χρώμα	ΥV12
Πλαίσια ανά δευτερόλεπτο	24
Πηγή	MPEG-2 (DVD)

Πίνακας 54 - Χαρακτηριστικά του video Smith



Εικόνα 147 - Πλαίσιο 650 του video Smith

Το video αυτό είναι από την ταινία “Mr and Mrs Smith” και η δυσκολία του στη συμπίεση είναι λόγω της κίνησης της πανοραμικής κάμερας. Δεν υπάρχουν σχεδόν καθόλου στατικές σκηνές.

4.4.2.3 Iceage

Τίτλος video	Iceage
Ανάλυση	720x576
Αριθμός πλαισίων	491
Χρώμα	ΥV12
Πλαίσια ανά δευτερόλεπτο	24
Πηγή	MPEG-2 (DVD)

Πίνακας 55 - Χαρακτηριστικά του video Iceage



Εικόνα 148 - Πλαίσιο 160 του video Iceage

Το video αυτό είναι από την ταινία "Iceage" και αποτελείται από δύο μέρη. Το πρώτο περιλαμβάνει πάρα πολύ κίνηση ενώ το δεύτερο περιλαμβάνει ένα σταθερό φόντο με αρκετή κίνηση στο προσκήνιο. Η δυσκολία στη συμπίεση προκύπτει από τη δυσκολία υπολογισμού και απόδοσης της κίνησης.

4.4.2.4 Lord of the rings

Τίτλος video	Lord of the Rings
Ανάλυση	720x416
Αριθμός πλαισίων	292
Χρώμα	YV12
Πλαίσια ανά δευτερόλεπτο	24
Πηγή	MPEG-2 (DVD)

Πίνακας 56 - Χαρακτηριστικά του video Lord of the Rings



Εικόνα 149 - Πλαίσιο 100 του video Lord of the Rings

Το video αυτό είναι από την ταινία “Lord of the rings” και περιλαμβάνει δύο μέρη. Το πρώτο περιλαμβάνει ένα σταθερό φόντο με αργή κίνηση στο προσκήνιο αλλά με πολλές λεπτομέρειες. Το δεύτερο δείχνει πρόσωπα τα οποία κινούνται συνεχώς. Η ακολουθία αυτή δεν είναι δύσκολη όσο αναφορά τη συμπίεση, αλλά λόγω των δύο διαφορετικών τμημάτων κάποια codecs ίσως να έχουνε δυσκολίες αν δεν χρησιμοποιούν σωστές παραμέτρους εσωτερικής κωδικοποίησης.

4.4.2.5 Rancho

Τίτλος video	Rancho
Ανάλυση	704x288
Αριθμός πλαισίων	1237
Χρώμα	ΥV12
Πλαίσια ανά δευτερόλεπτο	24
Πηγή	MPEG-2 (DVD), απόπλεξη (deinterlace) FlaskMPEG

Πίνακας 57 - Χαρακτηριστικά του video Rancho



Εικόνα 150 - Πλαίσιο 570 του video Rancho

Πρόκειται για μία σκηνή από την ταινία “Terminator 2”. Η κίνηση σε αυτό το video είναι ιδιαίτερα ομαλή αλλά υπάρχουν απότομες εναλλαγές σκηνών.

4.4.2.6 Futurama

Τίτλος video	Futurama
Ανάλυση	720x576
Αριθμός πλαισίων	292
Χρώμα	ΥV12
Πλαίσια ανά δευτερόλεπτο	25
Πηγή	MPEG-2 (DVD), σταδιακό (progressive)

Πίνακας 58 - Χαρακτηριστικά του video Futurama



Εικόνα 151 - Πλαίσιο 262 του video Futurama

Πρόκειται για ένα video από την ταινία κινουμένων σχεδίων “Futurama”. Το video περιέχει αρκετές μονόχρωμες περιοχές διαχωρισμένες με έντονα όρια.

4.4.2.7 Matrix

Τίτλος video	Matrix
Ανάλυση	720x416
Αριθμός πλαισίων	239
Χρώμα	YV12
Πλαίσια ανά δευτερόλεπτο	25
Πηγή	MPEG-2 (DVD), Έξυπνη απόπλεξη (Smart deinterlace)

Πίνακας 59 - Χαρακτηριστικά του video Matrix (ομάδα «ταινίες»)



Εικόνα 152 - Πλαίσιο 226 του video Matrix (ομάδα «ταινίες»)

Το video αυτό είναι από την ταινία “Matrix”. Λόγω της σχετικά απλής κίνησης, των αρκετά σκοτεινών χρωμάτων και της χαμηλής ανάλυσης το codec μπορεί να κωδικοποιήσει αρκετά εύκολα αυτό το video.

4.4.2.8 Bankomatdi

Τίτλος video	Bankomatdi
Ανάλυση	704x352
Αριθμός πλαισίων	376
Χρώμα	RGB
Πλαίσια ανά δευτερόλεπτο	30
Πηγή	MPEG-2 (DVD), Έξυπνη απόπλεξη (Smart deinterlace)

Πίνακας 60 - Χαρακτηριστικά του video Bankomatdi



Εικόνα 153 - Πλαίσιο 168 του video Bankomatdi

Το video αυτό είναι από την ταινία “Terminator 2” και είναι μια σκηνή σε ένα μηχάνημα ανάληψης. Η σκηνή έχει χαμηλή κίνηση, πολύ μικρές αλλαγές στο φόντο (η κάμερα στο δεύτερο μέρος πάει αργά προς τα δεξιά) και αρκετά υψηλή ανάλυση.

4.4.2.9 Bbc3di

Τίτλος video	Bbc3di
Ανάλυση	704x576
Αριθμός πλαισίων	374
Χρώμα	RGB
Πλαίσια ανά δευτερόλεπτο	25
Πηγή	Πρότυπο video, Έξυπνη απόπλεξη (Smart deinterlace)

Πίνακας 61 - Χαρακτηριστικά του video Bbc3di



Εικόνα 154 - Πλαίσιο 185 (αριστερά) και 258 (δεξιά) του video Bbc3di

Το video αυτό χαρακτηρίζεται από μία έντονη περιστροφική κίνηση. Δείχνει ένα ριγυτό δίσκο με διάφορες εικόνες επάνω του. Η ποιότητα του συμπιεσμένου video μπορεί να υπολογιστεί από τις λεπτομέρειες των εικόνων.

4.4.2.9 Susidi

Τίτλος video	Susidi
Ανάλυση	704x576
Αριθμός πλαισίων	374
Χρώμα	RGB
Πλαίσια ανά δευτερόλεπτο	25
Πηγή	MPEG-2 (40bit), Έξυπνη απόπλεξη (Smart deinterlace)

Πίνακας 62 - Χαρακτηριστικά του video Susidi



Εικόνα 155 - Πλαίσιο 193 του video Susidi

Το video αυτό χαρακτηρίζεται από υψηλό θόρυβο και αργή κίνηση. Στο πρώτο μέρος είναι σχεδόν στατική (μόνο η γυναίκα ανοιγοκλείνει τα μάτια της), στη συνέχεια η γυναίκα κινεί απότομα το κεφάλι της και μετά η σκηνή γίνεται και πάλι σχεδόν στατική. Ο θόρυβος αποκρύπτεται σε κάθε δεύτερο πλαίσιο λόγω της επιλογής πλαισίων τύπου-B (B-frames) του codec MPEG-2.

4.4.3 Video HDTV

4.4.3.1 Troy

Τίτλος video	Troy
Ανάλυση	1920x1072
Αριθμός πλαισίων	300
Χρώμα	ΥV12
Πλαίσια ανά δευτερόλεπτο	24
Πηγή	MPEG-2 (DVD)

Πίνακας 63 - Χαρακτηριστικά του video Troy



Εικόνα 156 - Πλαίσιο 1 του video Troy

Το video αυτό είναι από την ταινία “Troy” και περιλαμβάνει τρία τμήματα με έντονη εναλλαγή σκηνών. Το video περιέχει μέτρια κίνηση και μικρή κίνηση της κάμερας. Η συμπίεση του συγκεκριμένου video είναι δύσκολη λόγω των πολλών λεπτομερειών.

4.4.3.2 Matrix

Τίτλος video	Matrix
Ανάλυση	1920x1072
Αριθμός πλαισίων	250
Χρώμα	YV12
Πλαίσια ανά δευτερόλεπτο	30
Πηγή	MPEG-2 (DVD)

Πίνακας 64 - Χαρακτηριστικά του video Matrix (ομάδα «HDTV»)



Εικόνα 157 - Πλαίσιο 1 του video Matrix (ομάδα «HDTV»)

Το video αυτό προέρχεται από την ταινία “Matrix” και περιέχει γρήγορη εναλλαγή σκηνών με μέτρια κίνηση της κάμερας. Η κωδικοποίηση είναι δύσκολη λόγω του υψηλού ποσοστού κίνησης και των πολλών λεπτομερειών.

4.4.3.3 Concert

Τίτλος video	Concert
Ανάλυση	1664x1088
Αριθμός πλαισίων	390
Χρώμα	ΥV12
Πλαίσια ανά δευτερόλεπτο	25
Πηγή	MPEG-2 (HDTV broadcast), Έξυπνη απόπλεξη (Smart deinterlace)

Πίνακας 65 - Χαρακτηριστικά του video Concert



Εικόνα 158 - Πλαίσιο 128 του video Concert

Το video αυτό είναι κομμάτι μιας υψηλής ευκρίνειας μετάδοση (HDTV broadcast) μιας συμφωνικής ορχήστρας. Η χωρική ανάλυση είναι πολύ υψηλή, ενώ η κίνηση είναι αρκετά απλή (μερικές στιγμές σχεδόν ανύπαρκτη). Στο video υπάρχουν δύο αλλαγές σκηνών.

5. Αλγόριθμοι κωδικοποίησης ήχου

Σε αυτό το κεφάλαιο παρουσιάζονται ορισμένα βασικά χαρακτηριστικά των κυριότερων προτύπων κωδικοποίησης ήχου καθώς και κάποια βασικά στοιχεία θεωρίας σχετικά με τα δίκτυα VoIP (Voice over IP). Κάθε πρότυπο, όπως θα δούμε στη συνέχεια, χρησιμοποιεί διαφορετικό αλγόριθμο για την κωδικοποίηση, αποστολή και αποκωδικοποίηση των πακέτων πληροφορίας και έτσι αποδίδει καλύτερα κάτω από συγκεκριμένες συνθήκες. Ωστόσο όπως και στους αλγορίθμους κωδικοποίησης video έτσι και εδώ δεν είναι πάντα εφικτός ο σαφής διαχωρισμός για το εάν ένα πρότυπο αποδίδει καλύτερα από ένα άλλο υπό ορισμένες συνθήκες καθώς η απόδοσή του εξαρτάται από πολλούς, ανεξάρτητους πολλές φορές, παράγοντες.

5.1 Δίκτυα VoIP

Το Voice over IP ή VoIP ή τηλεφωνία μέσω διαδικτύου χαρακτηρίζει μια ομάδα πρωτοκόλλων – τεχνολογιών (πχ. SIP), η οποία προσφέρει φωνητική συνομιλία σε πραγματικό χρόνο. Οι συνομιλίες αυτές παραδοσιακά γίνονταν αποκλειστικά μέσω PC που ήταν συνδεδεμένο με το διαδίκτυο (Internet) και διέθετε μικρόφωνο, ακουστικά και το κατάλληλο λογισμικό. Η κλήση κατέληγε σε ένα άλλο, ανάλογα εξοπλισμένο, PC χωρίς να υπάρχει κάποια επιπλέον χρέωση, εκτός από αυτή της πρόσβασης στο διαδίκτυο, αφού στη συγκεκριμένη επικοινωνία δεν μεσολαβεί κάποιος παραδοσιακός φορέας τηλεπικοινωνιών παρά μόνο το διαδίκτυο. Πλέον υπάρχει η δυνατότητα προώθησης των κλήσεων VoIP σε σταθερά δίκτυα τηλεπικοινωνιών, αλλά δεν υπάρχει η αντίστροφη δυνατότητα.

Στα δίκτυα VoIP υπάρχει ένα κατώφλι ποιότητας υπηρεσιών (minimum Quality of Service) το οποίο εξαρτάται από το εύρος ζώνης, τις καθυστερήσεις, το ρυθμό απώλειας πακέτων και το τρέμουλο (jitter). Τα βήματα που απαιτούνται για το VoIP είναι: δειγματοληψία, ψηφιοποίηση και κωδικοποίηση, ενθυλάκωση (encapsulation), μετάδοση, αποκωδικοποίηση, προσωρινή αποθήκευση (buffering) και αναπαραγωγή.

Καθώς το ανθρώπινο αφτί είναι πιο ευαίσθητο στην απώλεια ποιότητας στον ήχο απ' ότι το μάτι στην εικόνα αντίστοιχα, είναι σημαντικό να διατηρούμε σταθερή την ποιότητα του ήχου σε μία ηχητική μετάδοση. Έτσι πρέπει να μελετήσουμε πως οι διάφοροι παράγοντες ενός δικτύου όπως το εύρος, ο ρυθμός απωλειών πακέτων και η καθυστέρηση επηρεάζουν την ποιότητα του ήχου. Καθώς οι αλγόριθμοι διαφέρουν αρκετά μεταξύ τους στο τρόπο με τον οποίο συμπεριφέρονται τον ήχο μπορούμε να υποθέσουμε ότι «αντιδρούν» και διαφορετικά στις αλλαγές των παραπάνω παραγόντων.

5.2 Παράγοντες που επηρεάζουν την ποιότητα της φωνής σε VoIP δίκτυα

Η ποιότητα και η αξιοπιστία είναι δύο πολύ σημαντικοί παράγοντες για τα VoIP δίκτυα. Υπάρχουν πολύ παράγοντες που επηρεάζουν την ποιότητα μιας VoIP σύνδεσης τους οποίους πρέπει να λαμβάνουμε υπόψη μας σε κάθε μέτρηση της ποιότητας φωνής που κάνουμε.

5.2.1 Εύρος ζώνης

Το εύρος ζώνης είναι ίσως ο πιο βασικός παράγοντας που επηρεάζει την ποιότητα της φωνής σε μία σύνδεση VoIP. Για παράδειγμα αν πραγματοποιήσετε μία VoIP σύνδεση πάνω από γραμμή dial-up τότε η ποιότητα της σύνδεσης είναι πολύ κακή σε αντίθεση με μία σύνδεση ευρείας ζώνης (broadband) (με την προϋπόθεση να μην διαμοιράζεται αυτή η σύνδεση σε πολλές άλλες εφαρμογές επικοινωνίας).

5.2.2 Εξοπλισμός

Ο εξοπλισμός που χρησιμοποιείται για την πραγματοποίηση της σύνδεσης VoIP είναι ένας πολύ σημαντικός παράγοντας που επηρεάζει την ποιότητα της σύνδεσης. Συνήθως ο φτηνός εξοπλισμός είναι αυτός που θα έχει και μειωμένη ποιότητα. Ωστόσο είναι σημαντικό να έχουμε όσο το δυνατό περισσότερες πληροφορίες σχετικά με το ATA, τον δρομολογητή (router) και το IP τηλέφωνο που χρησιμοποιούνται. Γενικά ο εξοπλισμός πρέπει να διαλέγεται ανάλογα με τις ανάγκες της κάθε σύνδεσης.

5.2.3 ATA / Δρομολογητής

Στην επιλογή ATA (Analog Telephone Adapter) / Δρομολογητή πρέπει να προσέξουμε τις τεχνολογίες συμπίεσης που υποστηρίζονται (codecs), το μηχανισμό ακύρωσης ηχούς (Echo cancellation) (έναν μηχανισμό ο οποίος μειώνει την ηχώ – echo) και την ασφάλειά του, όπως για παράδειγμα το τείχος προστασίας (firewall).

5.2.4 Συχνότητα του τηλεφώνου

Στη συχνότητα του τηλεφώνου μπορεί να παρεμβάλλονται σήματα από άλλες VoIP συσκευές και αυτό να μειώνει την ποιότητα της σύνδεσης. Έχουν παρατηρηθεί περιπτώσεις όπου παρουσιάζονταν προβλήματα στις συχνότητες 5.8Hz τα οποία λυνόντουσαν με τη χρήση τηλεφώνων σε χαμηλότερες συχνότητες (π.χ. 2.4Hz).

5.2.5 Καιρικές συνθήκες

Μερικές φορές η ποιότητα του ήχου επηρεάζεται από τον στατικό ηλεκτρισμό που δημιουργείται στα καλώδια της σύνδεσης από κεραυνούς, έντονη βροχόπτωση, δυνατούς ανέμους και αυξομειώσεις ρεύματος. Τα αποτελέσματα του στατικού ηλεκτρισμού δεν είναι τόσο εμφανή όταν σερφάρουμε στο διαδίκτυο ή κατεβάζουμε αρχεία, αλλά στην περίπτωση μετάδοσης φωνής τα αποτελέσματα του στη ποιότητα του μεταδιδόμενου ήχου είναι εμφανή. Για να απαλλαγούμε από τον στατικό ηλεκτρισμό μία λύση είναι να αποσυνδέσουμε τον εξοπλισμό, να περιμένουμε λίγο και μετά να τον ξανασυνδέσουμε. Γενικά η επίδραση των καιρικών φαινομένων στην ποιότητα της επικοινωνίας δεν μπορεί να ελεγχθεί, ωστόσο η σωστή συντήρηση των καλωδίων μπορεί να μειώσει τις αρνητικές επιδράσεις τους.

5.2.6 Το μέρος όπου βρίσκεται ο εξοπλισμός

Η παρεμβολές από άλλες συσκευές είναι ένας πολύ σημαντικός παράγοντας για την ποιότητα του ήχου στις VoIP συνδέσεις. Για παράδειγμα αν το ATA είναι πολύ κοντά στον δρομολογητή υπάρχει πιθανότητα να υπάρξει μείωση της ποιότητας, λόγω της ηλεκτρικής ανάδρασης (electrical feedback).

5.2.7 Το codec που χρησιμοποιείται

Στο VoIP μεταδίδονται συμπιεσμένα πακέτα, έτσι ώστε το φορτίο που μεταφέρεται να είναι μικρότερο. Κάθε codec είναι σχεδιασμένο για συγκεκριμένη χρήση. Αν ένα codec χρησιμοποιηθεί για άλλη χρήση από αυτήν για την οποία δημιουργήθηκε τότε το πιο πιθανό είναι να έχουμε μειωμένη ποιότητα στη σύνδεση.

5.3 Μέθοδοι υπολογισμού της ποιότητας

Η μεγαλύτερη δυσκολία σχετικά με την αξιολόγηση της ποιότητας του ήχου βρίσκεται στην επιλογή της μεθόδου που θα ακολουθεί. Πρέπει να χρησιμοποιηθεί μία μέθοδος η οποία βασίζεται στην υποκειμενική άποψη κάποιων ανθρώπων ή στην αναπαραγωγή αντικειμενικών υπολογισμών οι οποίοι είναι τελείως άσχετοι με τους προσωπικούς, υποκειμενικούς υπολογισμούς; Πρακτικά είναι αδύνατο να ορίσουμε ένα σύνολο αντικειμενικών μετρήσεων με σκοπό να χαρακτηρίσουμε την ποιότητα του ήχου. Και αυτό γιατί κάποιες μετρήσεις που είναι κατάλληλες για τον υπολογισμό μίας συγκεκριμένης παραμόρφωσης στον ήχο μπορεί να είναι ακατάλληλες για τον υπολογισμό μίας άλλης. Χρησιμοποιώντας κατάλληλο αριθμό ακροατών και ορισμένα τυποποιημένα tests, μπορούμε να μετριάσουμε το πρόβλημα αλλά έχουμε αύξηση του κόστους. Για αυτό το λόγο κατά τη διάρκεια ανάπτυξης ενός codec χρησιμοποιούνται κάποια tests μικρού κόστους σε συνδυασμό με την υποκειμενική άποψη ακροατών, πριν ακολουθήσει το πλήρες και ολοκληρωμένο test αξιολόγησής του.

5.3.1 Αντικειμενικές μέθοδοι υπολογισμού της ποιότητας

Οποιοδήποτε codec και να μελετάμε οι αντικειμενικές μέθοδοι υπολογισμού της ποιότητας (Objective Speech Quality Measures) είναι απαραίτητες για τον υπολογισμό της απόκλισης του τελικού σήματος από το αρχικό. Για τον υπολογισμό της ανομοιότητας του τελικού σήματος από το αρχικό μπορούμε να χρησιμοποιήσουμε μαθηματικές έννοιες όπως η συμμετρία (symmetry), η θετικά ορισμένη συνάρτηση (positive definitiveness) και η τριγωνική ανισότητα (triangle inequality), όπως ορίζονται παρακάτω:

- Συμμετρία: $d(x, y) = d(y, x)$
- Θετικά ορισμένη συνάρτηση: $d(x, x) = 0$ και $d(x, y) > 0$ για $x \neq y$
- Τριγωνική ανισότητα: $d(x, y) \leq d(x, z) + d(y, z)$

Η συμμετρία απαιτεί να μην υπάρχει διαφορά μεταξύ του σήματος αναφοράς και του σήματος που αξιολογείται σε όρους απόστασης. Ο θετικός ορισμός συνάρτησης δηλώνει ότι η απόσταση είναι μηδέν εάν το σήμα αναφοράς είναι πανομοιότυπο με το σήμα προς εξέταση. Η τριγωνική ανισότητα χρησιμοποιείται γιατί πρέπει η μέτρηση της απόστασης να έχει κάποια φυσική ερμηνεία, αλλά στην ουσία δεν έχει κάποια πρακτική εφαρμογή. Οι αντικειμενικές μέθοδοι υπολογισμού της ποιότητας του ήχου πρέπει να πληρούν και τα τρία παραπάνω κριτήρια.

Οι πιο διαδεδομένες μέθοδοι είναι ο σηματοθορυβικός λόγος (signal-to-noise ratio, SNR), ο συμβατικός SNR (conventional SNR), ο τμηματικός SNR (segmental SNR, SEGSNR) καθώς και ο SNR βάσει συχνότητας (frequency - weighted SNR).

5.3.2 Υποκειμενικές μέθοδοι υπολογισμού της ποιότητας

Τα αντικειμενικά tests ακολουθούνται από τα υποκειμενικά (Subjective Measures). Αναλόγως το ρυθμό του bit και την ποιότητα ενός codec απαιτούνται διαφορετικά υποκειμενικά tests για να αξιολογήσουμε την ποιότητα και το πόσο κατανοητό είναι το τελικό σήμα. Η ποιότητα συνήθως ελέγχεται με τη μέθοδο DAM (Diagnostic Acceptability Measure, DAM), ή την πιο διαδεδομένη μέθοδο MOS (Mean Opinion Score, MOS). Η καθαρότητα του σήματος, δηλαδή το πόσο κατανοητό είναι, αξιολογείται με τη μέθοδο CVC (Consonant-Vowel-Consonant, CVC) ή με τη μέθοδο DRT (Dynamic Rhythm Test, DRT).

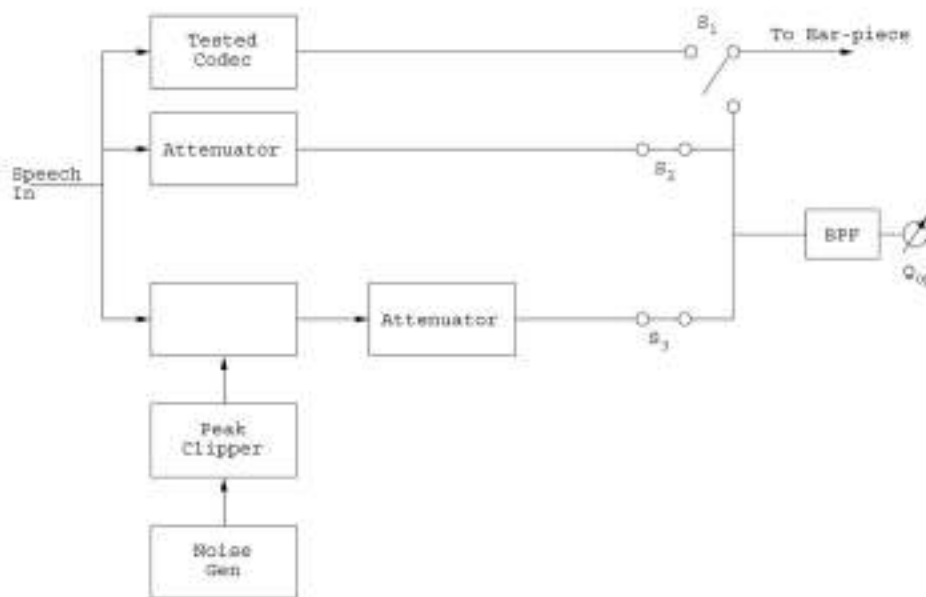
Στη μέθοδο DAM ζητείται από τον ακροατή να βαθμολογήσει το ηχητικό σήμα το οποίο αποτελείται από φωνητικά ισορροπημένες (phonetically balanced) προτάσεις, από την λίστα του Harvard (Harvard list), τόσο όσο αναφορά την ποιότητα της φωνής (speech quality) αλλά και την ποιότητα του ήχου στο φόντο (background quality). Η ποιότητα τόσο

της φωνής όσο και του φόντου είναι σε κλίμακα του 100, όπου ο κάθε ακροατής δίνει μία τιμή και στο τέλος προκύπτει ο μέσος όρος για κάθε κατηγορία. Πριν προκύψει ο μέσος όρος χρησιμοποιούνται και κάποιοι παράγοντες με διαφορετικό βάρος ο καθένας (weighting factors) για να δώσουμε έμφαση σε κάποιες συγκεκριμένες εφαρμογές του codec και να έχουμε καλύτερα αποτελέσματα.

Στις μεθόδους ελέγχου ανά ζεύγη (pair-wise preference tests) ο ακροατής συγκρίνει την ίδια πρόταση η οποία προκύπτει από διαφορετικά codecs. Για καλύτερα αποτελέσματα στα ηχητικά δείγματα περιλαμβάνεται και αυτούσιο το αρχικό σήμα (χωρίς να υποστεί συμπίεση). Τα αποτελέσματα συγκεντρώνονται σε ένα πίνακα και εάν δείχνουν μία καθαρή σειρά προτίμησης για όλα τα codecs αλλά και συγχρόνως μία απόκλιση περίπου 50% για το ίδιο codec, τότε τα αποτελέσματα γίνονται δεκτά. Αν ωστόσο δεν προκύψουν καθαρά αποτελέσματα πρέπει να πραγματοποιηθούν και άλλα tests.

5.3.3 Διαμορφωμένη μονάδα θορύβου αναφοράς (MNRU)

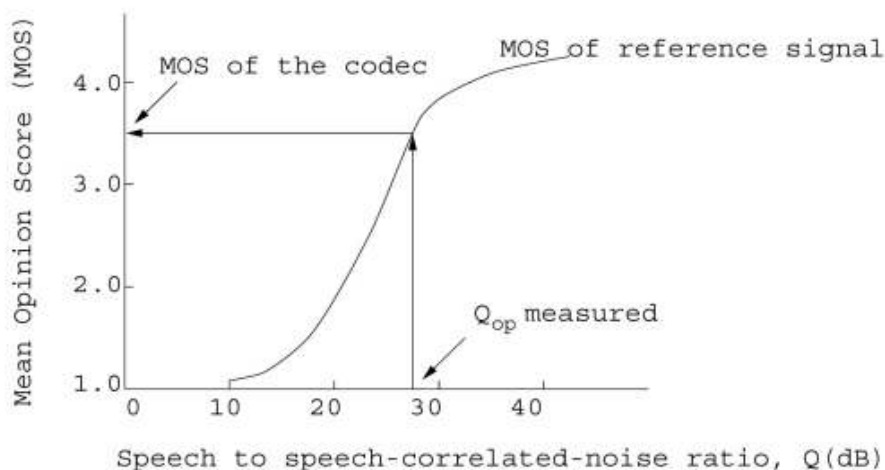
Η διαμορφωμένη μονάδα θορύβου αναφοράς (modulated noise reference unit, MNRU) προτάθηκε από τον Law και Seymour για να συσχετίσει την υποκειμενική ποιότητα με τις αντικειμενικές μετρήσεις και χρησιμοποιείται ευρέως από την CCITT. Το MNRU χρησιμοποιείται για την προσθήκη θορύβου, με πλάτος ανάλογο του ηχητικού σήματος, στο τελικό σήμα και έτσι προσδιορίζεται η σχέση θορύβου και σήματος.



Εικόνα 159 - Μπλοκ διάγραμμα MNRU

Ο σηματοθρομβικός λόγος (SNR) του σήματος αναφοράς (reference signal) μειώνεται σταδιακά από τον ακροατή (listener) χρησιμοποιώντας τους εξασθενητές (attenuators) όπως φαίνονται στην **Εικόνα 159**, έτσι ώστε να παρατηρήσουμε κατά πόσο το ηχητικό σήμα που ελέγχουμε ταιριάζει με το ηχητικό σήμα αναφοράς. Στη φάση αυτή οι διακόπτες S2 και S3 είναι κλειστοί και ο S1 εναλλάσσεται μεταξύ του σήματος αναφοράς και του ελεγχόμενου σήματος. Όταν τα δύο σήματα γίνουν πανομοιότυπα (υποκειμενικά από τον ακροατή) οι διακόπτες S2 και S3 χρησιμοποιούνται για τη μέτρηση της ισχύς του σήματος

αναφοράς και του σήματος θορύβου και έτσι έχουμε το Q σε dB (Opinion Equivalent Q, Q_{op}). Παρόλο που φαίνεται ότι το Q_{op} είναι μία αντικειμενική μέτρηση, εξαρτάται από την υποκειμενική κρίση του ακροατή και για αυτό το λόγο κατατάσσεται στις υποκειμενικές μεθόδους. Το Q_{op} μεταφράζεται εύκολα σε μονάδες MOS χρησιμοποιώντας τη γραφική παράσταση της **Εικόνας 160**.



Εικόνα 160 - Σχέση του Q_{op} με το MOS

5.3.4 Έλεγχοι κατανόησης

Στο έλεγχο κατανόησης ζητείται από τον ακροατή να αναγνωρίσει ανάμεσα από δύο ζευγάρια λέξεων ποιά είναι πλήρες. Οι δύο λέξεις διαφέρουν σε ένα σύμφωνο ή σε συνδυασμό σύμφωνο – φωνήεν – σύμφωνο (consonant – vowel – consonant, CVC). Σύμφωνα με το DRT (Diagnostic Rhythm Test, DRT), το οποίο αναπτύχθηκε από την Dynastat (Texas, USA), χρησιμοποιείται ένα σετ από 96 ζευγάρια λέξεων μερικές από τις οποίες είναι: meat-beat, pear-tear, saw-thaw, bond-rond κτλ. Τα ζευγάρια επιλέγονται κατάλληλα ώστε να ελέγχονται οι παρακάτω φωνητικές ιδιότητες: φώνηση (voicing), ρινισμός (nasality), σφύριγμα (sibilation), κρισιμότητα (graveness) και περιεκτικότητα (compactness). Για παράδειγμα εάν ο ακροατής αποτυγχάνει συνεχώς να διαχωρίσει τις λέξεις vast – fast, zoo – sue, goat – coat, συμπεραίνουμε ότι αδυνατεί να διαχωρίσει ήχους όπως ν, z, g κτλπ. Αυτό για τον σχεδιαστή σημαίνει ότι υπεύθυνος είναι ο μηχανισμός μακροπρόθεσμης πρόβλεψης (long-term predictor) ή ότι η φωνητική πληροφορία σε αυτό το φάσμα δεν λειτουργεί σωστά (voicing information in the spectrum does not work properly). Συλλέγοντας και μελετώντας τους τομείς όπου αποτυγχάνει ένα codec συλλέγουμε αρκετές πληροφορίες για τα μειονεκτήματα του codec. Συνηθισμένες τιμές του DRT είναι μεταξύ 75 και 95. Για εξαιρετική κατανόηση απαιτείται DRT 90.

5.4 Το πρότυπο G.711

Το G.711 είναι ένα πρότυπο για συμπίεση ήχου της ITU-T το οποίο χρησιμοποιήθηκε αρχικά στην τηλεφωνία. Το G.711 χρησιμοποιεί παλμοκωδική διαμόρφωση (pulse code modulation, PCM) για σήματα που έχουν δειγματοληφθεί με 8000 δείγματα ανά δευτερόλεπτο.

Στο πρότυπο προσδιορίζονται δύο αλγόριθμοι, ο μ-law ο οποίο χρησιμοποιείται στην Αμερική και στην Ιαπωνία και ο A-law ο οποίος χρησιμοποιείται στην Ευρώπη και στον υπόλοιπο κόσμο. Και οι δύο αλγόριθμοι κωδικοποιούν 14-bit και 13-bit PCM δείγματα σε 8-bit λογαριθμικά δείγματα. Ο κωδικοποιητής G.711 δημιουργεί μία ροή bit ταχύτητας 64kbits/sec για ένα σήμα που έχει δειγματοληφθεί στα 8kHz.

Ορισμένα από τα χαρακτηριστικά του G.711 είναι:

- Συχνότητα δειγματοληψίας 8kHz
- Ρυθμό bit 64kbps (8kHz συχνότητα δειγματοληψίας x 8bits ανά δείγμα)
- Τυπική καθυστέρηση αλγορίθμου 0.125ms
- Υποστηρίζεται η απόκρυψη χαμένων πακέτων (Packet Loss Concealment, PLC)
- Προσδιορίζεται αλγόριθμος ασυνεχούς μετάδοσης (discontinuous transmission, DTX) ο οποίος χρησιμοποιεί ανίχνευση φωνής (voice activity detection, VAD) και παραγωγή τεχνικού ήχου στο φόντο (comfort noise generation, CNG) για τη μείωση του εύρους ζώνης κατά τις περιόδους ησυχίας

Υπό ιδανικές συνθήκες η βαθμολογία MOS του G.711 είναι 4.45 για το μ-Law και 4.45 για το A-Law, ενώ υπό συνθήκες φόρτου είναι 4.13 για το μ-Law και 4.11 για το A-Law.

5.5 Το πρότυπο G.722

Το G.722 είναι ένα πρότυπο της ITU-T το οποίο λειτουργεί στα 48, 46 και 64 Kbps και στηρίζεται στη διαμόρφωση ADPCM (Adaptive Differential Pulse Code Modulation). Προσφέρει συμπίεση σε χαμηλούς ρυθμούς bit, δειγματοληπώντας τα ηχητικά δεδομένα στα 16kHz (διπλάσιο από αυτή της παραδοσιακής τηλεφωνίας).

Καθώς το πρότυπο G.722 υποστηρίζει διάφορους ρυθμούς bit το καθιστά κατάλληλο για χρήση σε VoIP εφαρμογές, όπου το εύρος ζώνης δεν είναι γνωστό και μπορεί να αλλάζει. Σε περιπτώσεις όπου το εύρος ζώνης είναι πιο σταθερό μπορούν να χρησιμοποιηθούν τα αποδοτικότερα codecs G.722.1 και G.722.2.

5.6 Το πρότυπο G.723

Το πρότυπο G.723 αποτελεί μία επέκταση του G.721 στα 24 και 40 Kbits/sec για εξοπλισμό ψηφιακών πολλαπλασιαστικών κυκλωμάτων (Digital circuit multiplication equipment, DCME) και στην πορεία επισκιάστηκε από την ανάπτυξη του προτύπου G.726.

5.7 Το πρότυπο G.723.1

Το πρότυπο G.723.1 κωδικοποιεί τον ήχο σε πλαίσια των 30ms (30ms frames) και μπορεί να λειτουργήσει σε δύο ρυθμούς bit, στα 6.3Kbps (χρησιμοποιώντας πλαίσια 24 bytes) και στα 5.3 Kbps (χρησιμοποιώντας πλαίσια των 20 bytes). Λόγω των χαμηλών του απαιτήσεων σε εύρος ζώνης χρησιμοποιείται κυρίως σε εφαρμογές VoIP.

Μερικά από τα χαρακτηριστικά του είναι:

- Συχνότητα δειγματοληψίας 8 KHz/16-bit (240 δείγματα για πλαίσια 30ms)
- Καθορισμένος ρυθμός bit (5.3Kbits/sec με 20 bytes σε πλαίσια 30ms και 6.3Kbits/sec με 24 bytes σε πλαίσια 30 ms)
- Καθορισμένο μέγεθος πλαισίου (frame size) για κάθε ρυθμό bit

- Καθυστέρηση αλγορίθμου 37.5ms ανά πλαίσιο
- Πρόκειται για ένα υβριδικό κωδικοποιητή όπου στους υψηλούς ρυθμούς bit χρησιμοποιεί MP-MLQ (Multi-Pulse Maximum Likelihood Quantization) ενώ στους χαμηλούς ACELP (Algebraic Code Excited Linear Prediction)
- Η πολυπλοκότητα του αλγορίθμου είναι 15, χρησιμοποιώντας μία σχετική κλίμακα όπου το G.711 είναι 1 και το G.729a 15

Η βαθμολογία MOS που πετυχαίνει υπό ιδανικές συνθήκες είναι 4.08 ενώ υπό συνθήκες φόρτου είναι 3.57.

5.8 Το πρότυπο G.726

Το G.726 είναι ένα πρότυπο συμπίεσης φωνής της ITU-T για μετάδοση σε ρυθμούς 16, 24, 32 και 40Kbps. Το G.726 είναι αυτό που εισήγαγε τα 16Kbits/sec. Ορισμένα βασικά χαρακτηριστικά του είναι:

- Συχνότητα δειγματοληψίας 8kHz
- Δυνατότητα 16, 24, 32 και 40 Kbps ρυθμών bit
- Δημιουργεί ροή bits και επομένως το μέγεθος πλαισίου καθορίζεται από τη διαδικασία δημιουργίας πακέτων (packetization) (συνήθως 80 δείγματα για μέγεθος πλαισίου 10 ms)
- Τυπική καθυστέρηση αλγορίθμου 0.125ms
- Χρησιμοποιεί τη διαμόρφωση ADPCM
- Η πολυπλοκότητα του αλγορίθμου βαθμολογείται με 10, βάσει μιας σχετικής κλίμακας όπου το G.711 έχει πολυπλοκότητα 1 και το G.72a έχει πολυπλοκότητα 15
- Στα 40Kbits/sec μπορεί να υποστηρίξει 12000bits/sec ενώ στα 32Kbits/sec μπορεί να υποστηρίξει 2400bits/sec

Η βαθμολογία MOS του G.726 υπό ιδανικές συνθήκες δικτύου είναι 4.30 ενώ υπό συνθήκες φόρτου είναι 3.79.

5.9 Το πρότυπο G.728

Το G.728 είναι ένα πρότυπο της ITU-T το οποίο λειτουργεί στα 16Kbps και χρησιμοποιεί την τεχνολογία LD-CELP (Low Delay Code Excited Linear Prediction, LD-CELP). Η καθυστέρηση του codec είναι μόλις 0.625ms (5 δείγματα).

5.10 Το πρότυπο G.729

Το G.729 είναι ένα πρότυπο το οποίο συμπιέζει τον ήχο σε κομμάτια των 10ms και χρησιμοποιείται κυρίως σε εφαρμογές VoIP και όχι τόσο για μετάδοση μουσικής, λόγω των χαμηλών του απαιτήσεων σε εύρος ζώνης. Λειτουργεί συνήθως στα 8Kbits/sec αλλά μπορεί να επεκταθεί και στα 6.4Kbits/sec και 11.8Kbits/sec.

Χρησιμοποιεί τεχνικές VAD και CNG για την καλύτερη εκμετάλλευση του διαθέσιμου εύρους ζώνης. Η ακουστική του συχνότητα επεκτείνεται από τα 50Hz μέχρι τα 7KHz.

5.11 Το πρότυπο G.729A

Το G.729A συμπιέζει τον ήχο σε κομμάτια των 10ms όπως το G.729 αλλά απαιτεί λιγότερους υπολογισμούς. Το πρότυπο αναπτύχθηκε από τη συνεργασία των France Telecom, Mitsubishi Electric Corporation, Nippon Telegraph and Telephone Corporation (NTT) και το Université de Sherbrooke.

Ορισμένα χαρακτηριστικά του είναι:

- Συχνότητα δειγματοληψίας 8KHz/16-bit (80 δείγματα για πλαίσια 10ms)
- Καθορισμένος ρυθμός bit (8Kbits/sec για πλαίσια 10ms)
- Καθορισμένο μέγεθος πλαισίων (10 bytes για πλαίσια 10ms)
- Καθυστέρηση αλγορίθμου 15ms ανά πλαίσιο
- Πρόκειται για ένα υβριδικό codec χρησιμοποιεί ACELP (Algebraic Code Excited Linear Prediction)
- Η πολυπλοκότητα του αλγορίθμου είναι 15, σε σχετική κλίμακα με το G.711 να έχει πολυπλοκότητα 1 και το G.723.1 να έχει 15

Η βαθμολογία του MOS υπό ιδανικές συνθήκες δικτύου είναι 4.04 ενώ υπό συνθήκες φόρτου δικτύου είναι 3.51.

6. Συγκρίσεις αλγορίθμων κωδικοποίησης ήχου

Σε αυτό το κεφάλαιο γίνεται μία προσπάθεια κατάταξης διαφόρων codecs ήχου, ανάλογα με την απόδοσή τους υπό ορισμένες συνθήκες. Όπως είδαμε και προηγουμένως η κατάταξη των codecs δεν είναι πάντα εφικτή, οι παράγοντες από τους οποίους εξαρτάται η ποιότητα του ήχου είναι πολλοί και ανεξάρτητοι.

Οι επιλογές που έχουμε για τη συμπίεση φωνής ποικίλουν από τεχνικές χαμηλής συμπίεσης, μέχρι τεχνικές πάρα υψηλής συμπίεσης. Ωστόσο για την σωστή επιλογή της μεθόδου που θα υλοποιηθεί, πριν αποφασίσουμε να «ανταλλάξουμε» την ποιότητα του ήχου με την ταχύτητα μετάδοσης, πρέπει να γνωρίζουμε τις επιπτώσεις της συγκεκριμένης τεχνικής μετάδοσης στην ποιότητα.

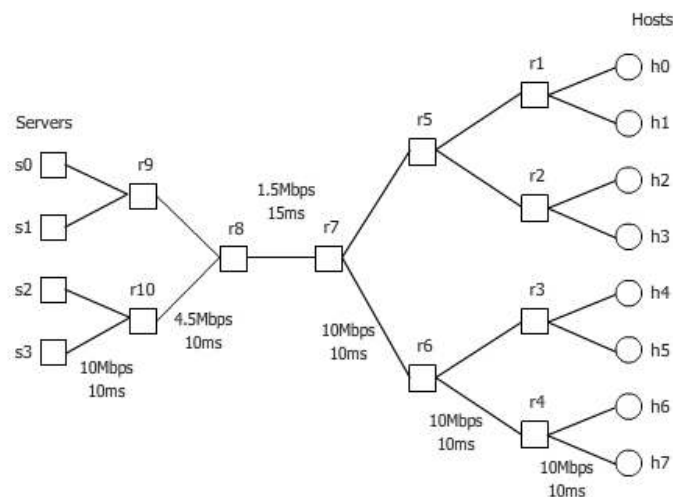
6.1 Σύγκριση των codecs G.711, G.723.1, G.726, G.729 και G.729A

Στο [18] μελετάται η συμπεριφορά των codecs G.711 και G.723 σε σχέση με τη τοπολογία του δικτύου, dumbbell ή γραμμική (linear) και το είδος της υπηρεσίας, καλύτερης προσπάθειας ή διαφοροποιημένης (best-effort / differentiated Services)³.

6.1.1 Πειραματική διαδικασία

Για την πραγματοποίηση των μετρήσεων χρησιμοποιήθηκαν δύο τοπολογίες δικτύων.

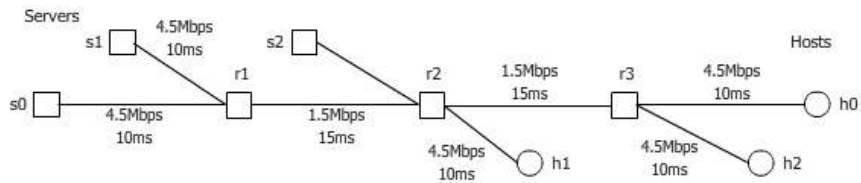
Η τοπολογία dumbbell, η οποία παρέχει διαδρομές με 6 ενδιάμεσους κόμβους για κάθε πηγή και προορισμό, με μία γραμμή 1.5Mbps στο κέντρο, όπως φαίνεται στην **Εικόνα 161**.



Εικόνα 161 - Τοπολογία δικτύου Dumbbell

Και η γραμμική τοπολογία η οποία αποτελείται από συνδέσεις 1.5 Mbps, όπως φαίνεται στην **Εικόνα 162**.

³ Βλ. Παράρτημα A5 και Παράρτημα A6



Εικόνα 162 – Γραμμική τοπολογία δικτύου

Αρχικά υποθέτουμε ομογενή ροή δεδομένων, όπου η κίνηση δημιουργείται από πηγές που χρησιμοποιούν το ίδιο codec. Έτσι μπορούμε να εξετάσουμε την κίνηση που δημιουργείται από διαφορετικά codecs. Επίσης υποθέτουμε ότι οι φάσεις των πηγών είναι τυχαίες. Τα ίδια χαρακτηριστικά ισχύουν και όταν μελετάμε τις υπηρεσίες καλύτερης προσπάθειας. Η συνολική κίνηση που δημιουργείται από τα ομαδοποιημένα δεδομένα (aggregates) περιορίζεται στο 30% του εύρους ζώνης (450kbps) και το ολικό εύρος ζώνης της γραμμής μοιράζεται μεταξύ της κίνησης των ομαδοποιημένων δεδομένων και της κίνησης του TCP (TCP traffic).

Στη συνέχεια θεωρούμε ετερογενή κίνηση, όπου η ροή των ομαδοποιημένων δεδομένων προκύπτει από πακέτα τα οποία δημιουργούνται από πηγές από τις οποίες μερικές χρησιμοποιούν το codec G.711 ενώ άλλες χρησιμοποιούν ένα από τα υπόλοιπα codecs. Τα πακέτα αυτά της ετερογενούς κίνησης μεταφέρονται ως ένα σύνολο. Η κίνηση δημιουργείται έτσι ώστε να εξυπηρετούνται πρώτα τα πακέτα του G.711 και μετά τα υπόλοιπα. Όταν μελετώνται οι υπηρεσίες καλύτερης προσπάθειας ισχύουν οι ίδιοι κανόνες για όλη την κίνηση ανεξαρτήτως codec.

Η από-άκρο-σε-άκρο καθυστέρηση για όλες τις ροές κατωφλιώνεται (έχει κάτω όριο) από την από-άκρο-σε-άκρο καθυστέρηση διάδοσης για τις δύο τοπολογίες. Στη τοπολογία dumbbell η καθυστέρηση είναι 65ms ενώ στη γραμμική είναι 50ms. Χρησιμοποιώντας επιταχυμένη προώθηση (Expedited Forwarding, EF)⁴ μπορούμε να ελέγξουμε μόνο την καθυστέρηση ουράς (queueing delay), δηλαδή το χρόνο που περιμένει ένα πακέτο στο δρομολογητή σε μία ουρά, σε όλους τους δρομολογητές.

6.1.2 Ρυθμίσεις των codecs

Στον Πίνακα 66 φαίνονται τα χαρακτηριστικά των codecs που χρησιμοποιήθηκαν κατά τη διάρκεια της πειραματικής διαδικασίας.

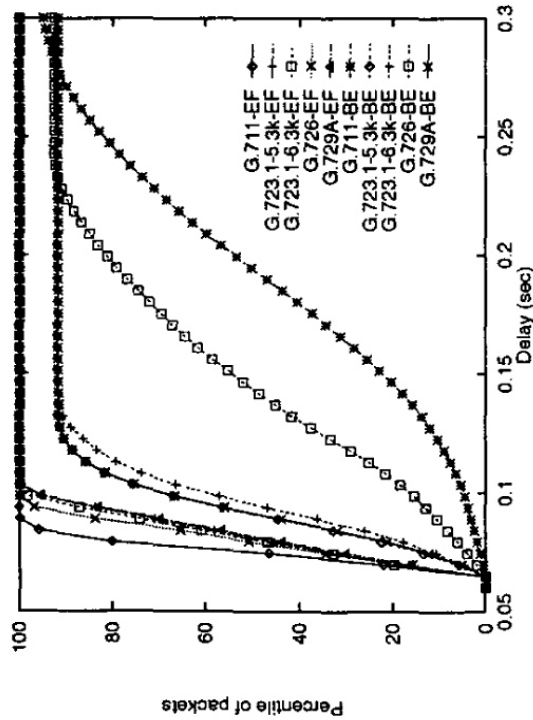
⁴ Βλ Παράρτημα Α7

Codec	G.711	G.723.1	G.726-32	G.729	G.729A
Coding speed (kbps)	64	5.3/6.3	32	8	8
Frame size (ms)	20	30	20	10	10
Processing Delay (ms)	20	30	20	10	10
Lookahead Delay (ms)	0	7.5	0	5	5
DSP MIPS	0.34	16	14	20	10.5
Payload (bytes)	160	20/24	80	20	20
Number of flows	7	84/71	14	56	56
Subscribed Rate Dacket time (ms)	20	30.2/30.5	20	20	20

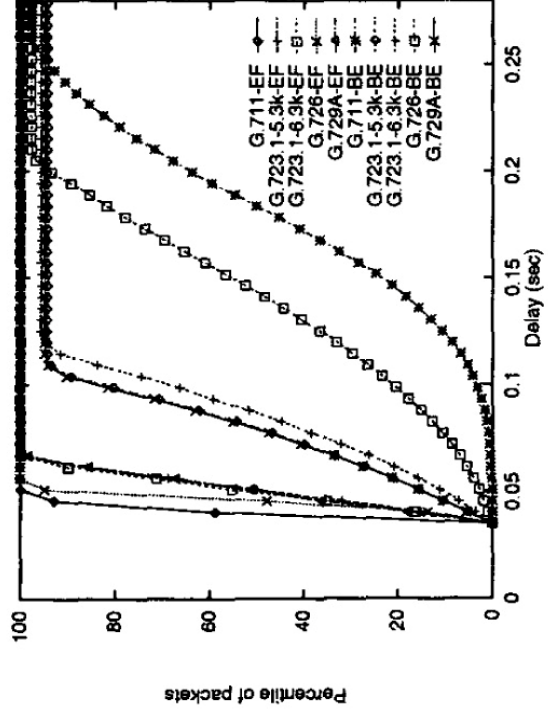
Πίνακας 66 - Χαρακτηριστικά των codecs που χρησιμοποιήθηκαν για τη σύγκριση των G.711, G.723.1, G.726, G.729 και G.729A

6.1.3 Αποτελέσματα και συμπεράσματα

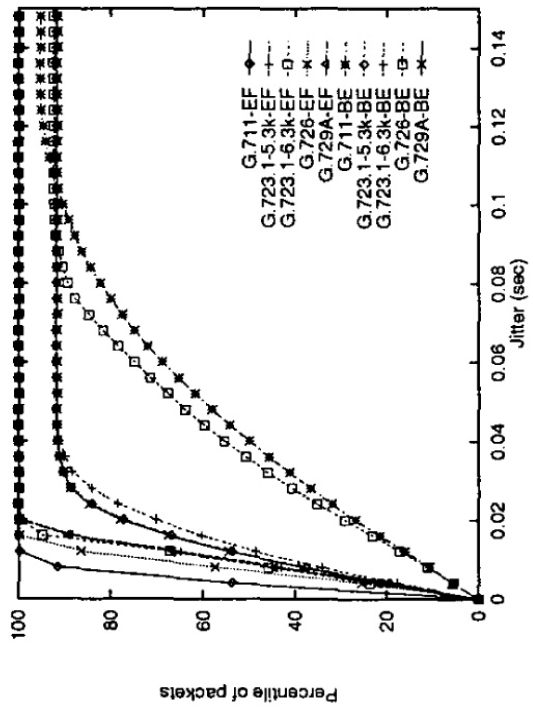
Τα αποτελέσματα σχετικά με την καθυστέρηση και το τρεμούλιασμα της φωνής που προέκυψαν για την ομογενή ροή δεδομένων φαίνονται στις **Εικόνες 163 μέχρι και 166**.



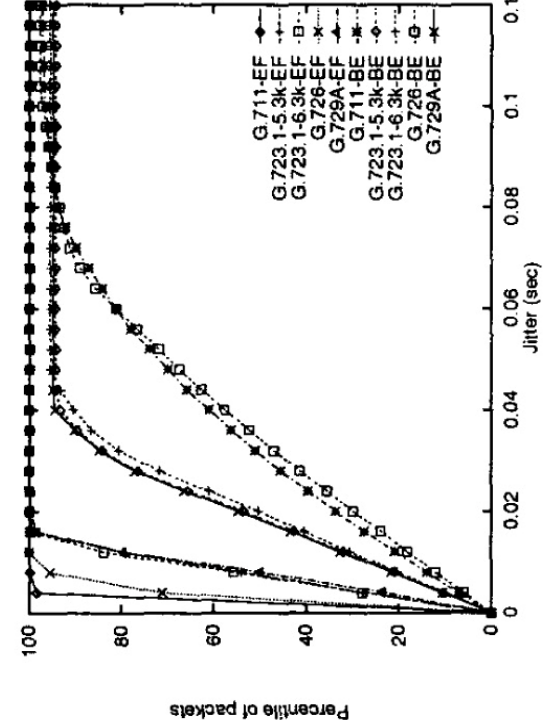
Εικόνα 163 – Διάγραμμα καθυστέρησης για συνθήκες ομογενούς ροής σε δίκτυο Dumbbell



Εικόνα 165 - Διάγραμμα καθυστέρησης για συνθήκες ομογενούς ροής σε δίκτυο Linear



Εικόνα 164 - Διάγραμμα jitter για συνθήκες ομογενούς ροής σε δίκτυο Dumbbell



Εικόνα 166 - Διάγραμμα jitter για συνθήκες ομογενούς ροής σε δίκτυο Linear

Από τα αποτελέσματα παρατηρούμε ότι η καθυστέρηση και το τρεμούλιασμα της φωνής βελτιώνονται αρκετά όταν αντί καλύτερη προσπάθεια χρησιμοποιήσουμε επιταχυμένη προώθηση (EF). Η βελτίωση είναι μεγαλύτερη για τα codecs με υψηλό ρυθμό κωδικοποίησης (coding rate), όπως το G.711 και επίσης παρατηρούμε ότι όταν χρησιμοποιούμε EF τα codecs με υψηλό ρυθμό κωδικοποίησης (πχ. G.711) έχουν καλύτερη απόδοση σε σχέση με τα codecs με χαμηλότερο ρυθμό κωδικοποίησης (πχ. G.723.1). Το ακριβώς αντίθετο συμβαίνει όταν αντί για EF χρησιμοποιηθούν υπηρεσίες καλύτερης προσπάθειας.

Λόγω του ότι το κάθε codec έχει διαφορετικό ρυθμό κωδικοποίησης χρειάζεται και διαφορετικός αριθμός ροών (number of flows) για να εκμεταλλευτεί όλο το διαθέσιμο εύρος ζώνης (450kbps). Έτσι ο αριθμός των ροών ποικίλει από 7 για το G.711 μέχρι 71/84 για το G.723.1. Επίσης το μέγεθος πακέτου διαφέρει από codec σε codec έτσι ο αριθμός των πακέτων που υπάρχει στο δίκτυο διαφέρει ανάλογα με το codec που χρησιμοποιείται. Όταν το μέγεθος των πακέτων είναι μικρό και ο αριθμός των ροών είναι μεγάλος, η αναμονή στην ουρά (queueing) των πακέτων συμβαίνει για δύο διαφορετικούς λόγους:

1. Υπάρχει αναμονή λόγω άλλων πακέτων που είναι ήδη στην ουρά EF
2. Αναμονή στη κορυφή της ουράς μέχρι να φύγει

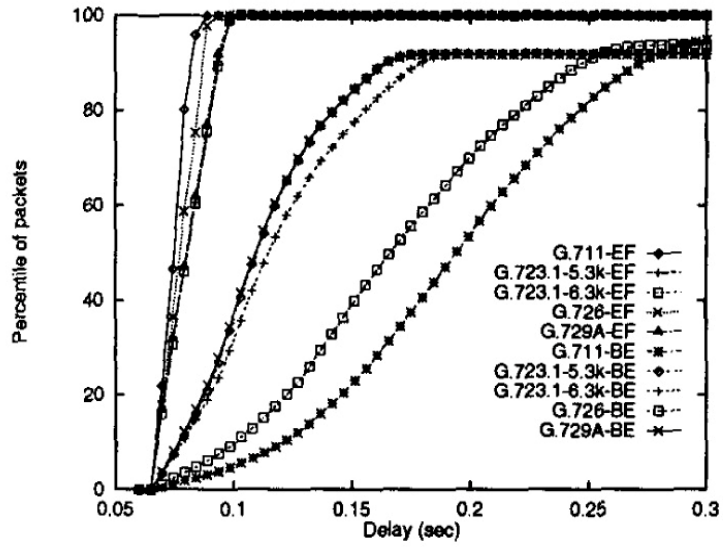
Για παράδειγμα το G.711 επειδή δημιουργεί μεγάλα πακέτα (160 bytes χωρίς τις επικεφαλίδες) και έχει μικρό αριθμό ροών, χρησιμοποιεί τη σύνδεση για μεγαλύτερο χρονικό διάστημα απ' όταν ξεκινήσει η μετάδοση (υπενθυμίζουμε ότι το εύρος ζώνης είναι σταθερό και μοιράζεται για όλα τα πακέτα φωνής). Έτσι η ανάμειξη TCP πακέτων είναι μικρότερη.

Από την άλλη το G.723.1 δημιουργεί μικρότερα πακέτα και έχει μεγαλύτερο αριθμό ροών και έτσι η ανάμειξη των πακέτων φωνής με τα πακέτα TCP είναι μεγαλύτερη.

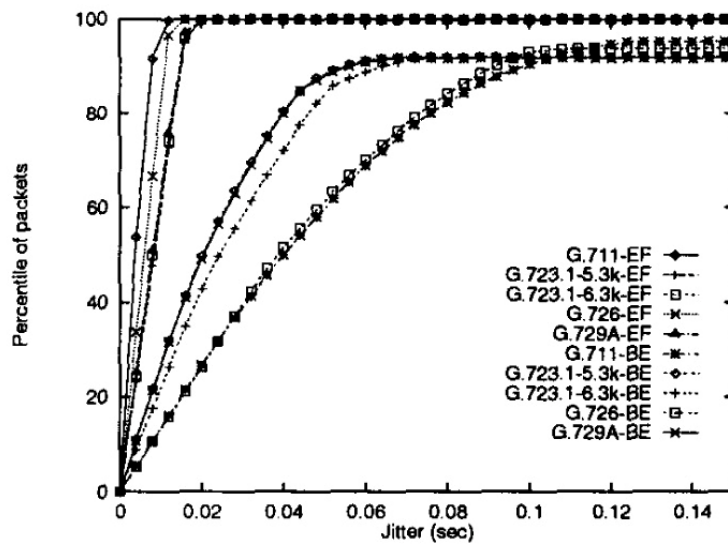
Άρα παρατηρούμε ότι το μέγεθος των πακέτων και ο αριθμός των ροών αλληλοεξαρτώνται. Με την αύξηση του αριθμού των ροών παρατηρείται μία χειροτέρευση της απόδοσης της κάθε ροής. Η μείωση του μεγέθους των πακέτων μαζί με τη μείωση του ρυθμού bit του codec δύναται να αντισταθμίσει την αύξηση αυτή των ροών κρατώντας το εύρος ζώνης σταθερό.

Επίσης παρατηρούμε ότι η καμπύλη για τις υπηρεσίες καλύτερης προσπάθειας δεν φτάνει το 100% αλλά περίπου στο 90%. Αυτό παρατηρείται διότι θεωρήσαμε την καθυστέρηση άπειρη για τα χαμένα πακέτα. Έτσι τα λαμβάναμε υπόψη μας αλλά με άπειρη καθυστέρηση (συμπερασματικά η γραφική παράσταση είναι μερικώς ελαττωματική). Η απώλεια πακέτων παρατηρείται όταν υπάρχει συμφόρηση στη γραμμή. Με το EF τα πακέτα φωνής εξυπηρετούνται από διαφορετική ουρά και έτσι δεν παρατηρούνται απώλειες πακέτων. Και στις δύο τοπολογίες παρατηρούμε περίπου ίδια συμπεριφορά παρ' όλο που οι τιμές είναι διαφορετικές, πράγμα που οφείλεται στο διαφορετικό αριθμό ενδιάμεσων σταθμών και καθυστερήσεων στις δύο τοπολογίες.

Για την ετερογενή ροή δεδομένων, όπου κάποιες πηγές χρησιμοποιούν το codec G.711 ενώ οι υπόλοιπες ένα από τα υπόλοιπα, προέκυψαν τα διαγράμματα των **Εικόνων 167 και 168** για την τοπολογία dumbbell.



Εικόνα 167 - Διάγραμμα καθυστέρησης για συνθήκες ετερογενούς ροής σε δίκτυο Dumbbell



Εικόνα 168 - Διάγραμμα jitter για συνθήκες ετερογενούς ροής σε δίκτυο Dumbbell

Παρουσιάζονται τα αποτελέσματα όπου η κίνηση του G.711 είναι περίπου 57% (256kpbs από τα 450kpbs διαθέσιμα). Και εδώ παρατηρούμε ότι αν χρησιμοποιήσουμε καλύτερη προσπάθεια τα codecs με υψηλό ρυθμό bit αποδίδουν χειρότερα από τα codecs με χαμηλό ρυθμό bit. Το αντίθετο συμβαίνει με χρήση του EF.

6.2 Σύγκριση των codecs G.711, G.721, G.722, G.726, G.729A, MPEG3 και GSM

Σκοπός στο [19] είναι η ανάλυση της απόδοσης των αλγορίθμων συμπίεσης ήχου G.711, G.722, G.726, G.721, G.729A, MPEG3 και GSM πάνω από IP δίκτυα, για μετάδοση πακέτων φωνής. Το να μεταδώσουμε τη φωνή σε πακέτα μας παρέχει καλύτερη χρησιμοποίηση του διαθέσιμου εύρους ζώνης και γενικά οι τεχνικές συμπίεσης παρέχουν καλύτερη μετάδοση σε σύγκριση με μία υλοποίηση χωρίς συμπίεση.

Στο πείραμα αυτό εξετάζουμε την απόδοση των κυριότερων αλγορίθμων συμπίεσης τόσο όσο αναφορά τη ταχύτητα μετάδοσης αλλά και όσο αναφορά την ποιότητα.

6.2.1 Πειραματική διαδικασία

Για την πραγματοποίηση του πειράματος χρησιμοποιήθηκε ένας προσομοιωτής μετάδοσης σε πραγματικό χρόνο (Real Time Transmission (RTP) Simulator) ο οποίος περιλαμβάνει ένα σύστημα με έναν εξυπηρετητή (server) και έναν πελάτη (client). Ο δέκτης (receiver) λαμβάνει την πληροφορία που υπάρχει στο δίκτυο, την αποκωδικοποιεί και την αναπαράγει. Από την άλλη ο εκπομπός (transmitter) κωδικοποιεί την πληροφορία (με το ανάλογο codec) και την μεταδίδει.

Αυτό που μας ενδιαφέρει είναι η ποιότητα του ήχου που φτάνει στο δέκτη κάτω από ειδικές συνθήκες όπως περιορισμένο εύρος ζώνης, καθυστερήσεις και απώλειες πακέτων. Κάθε αλγόριθμος αξιολογείται με βάση τα bits πληροφορίας που μεταφέρονται ανά μονάδα χρόνου (bit rate), την ποιότητα του αποκωδικοποιημένου ήχου στο δέκτη, την πολυπλοκότητά του, την καθυστέρηση που εισάγει αλλά και την ευαισθησία του στα λάθη και την απώλεια πακέτων.

Ο σηματοθορυβικός λόγος (Signal to Noise Ration, SNR) είναι το μέγεθος που χρησιμοποιείται για την αξιολόγηση της απόδοσης των αλγορίθμων συμπίεσης και υπολογίζεται από τη σχέση της **Εξίσωσης 2**.

$$SNR = 10 \log_{10} \left\{ \frac{\sum_{n=0}^{M-1} S(n)^2}{\sum_{n=0}^{M-1} (s(n) - \hat{s}(n))^2} \right\}$$

Εξίσωση 2 - Σηματοθορυβικός λόγος

Ο σηματοθορυβικός λόγος υπολογίζεται για όλη τη διάρκεια της μετάδοσης και για αυτό μπορεί να «αποκρύπτει» μικρές απώλειες και καθυστερήσεις σε μικρά χρονικά διαστήματα. Η καλύτερη αξιολόγηση για μικρά χρονικά διαστήματα γίνεται με τον υπολογισμό του τμηματικού σηματοθορυβικού λόγου (Segmental SNR) π.χ. για κάθε N διαστήματα και δίνεται από την **Εξίσωση 3**.

$$SEGSNR = \frac{10}{L} \sum_{i=0}^{L-1} \log_{10} \left\{ \frac{\sum_{n=0}^{N-1} S^2(iN + n)}{\sum_{n=0}^{N-1} (s(iN + n) - \hat{s}(iN + n))^2} \right\}$$

Εξίσωση 3 - Τμηματικός σηματοθορυβικός λόγος

Η αξιολόγηση της ποιότητας του ήχου βασίζεται στη βαθμολογία των ακροατών με ηλικίες από 18 μέχρι 40 ετών. Για τη μέτρηση της ποιότητας χρησιμοποιήθηκε η μέθοδος MOS (Mean Opinion Score, MOS).

Για την αξιολόγηση της απόδοσης των αλγορίθμων χρησιμοποιήθηκε ο αριθμός των bits πληροφορίας που μεταφέρονται ανά μονάδα.

Ο ρυθμός bit εξαρτάται από τη κωδικοποίηση που χρησιμοποιείται και παίζει σημαντικό ρόλο στην απόδοση του συστήματος.

Το πείραμα διεξήχθη χρησιμοποιώντας προτάσεις (ήχους) σε υψηλή ποιότητα, ηχογραφημένες τόσο από γυναίκες όσο και από άντρες. Στη συνέχεια οι ακροατές αξιολογούσαν τον αποσυμπιεσμένο ήχο χωρίς να γνωρίζουν την αρχική (πρωτότυπη) ηχογράφιση.

Η ψηφιοποίηση του αρχικού πρωτότυπου δείγματος έγινε με τη μέθοδο PCM (Pulse Code Modulation, Διαμόρφωση Παλμού) και τη ADPCM (Adaptive Differential PCM). Ωστόσο δεν συγκρίνονται αποτελέσματα που προέκυψαν από PCM και ADPCM μαζί, αλλά το καθένα ξεχωριστά.

Αρχικά σαν κατώφλι ποιότητας (toll-quality, ποιότητα ενός συνηθισμένου υπεραστικού τηλεφωνήματος) ορίζεται το MOS του G.726/32 ADPCM. Αν ένας άλλος αλγόριθμος πετύχει καλύτερο MOS τότε αυτός ο αλγόριθμος αυτός θεωρείται σαν κατώφλι.

6.2.2 Αποτελέσματα και συμπεράσματα

Στον **Πίνακα 67** παρουσιάζονται οι βαθμολογίες MOS που πέτυχε κάθε codec.

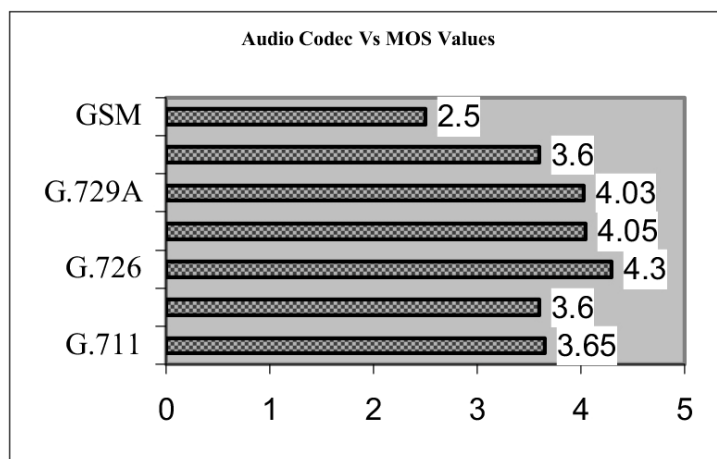
	G.711	G.722	G.726	G.721	G.729A	MPEG3	GSM
Ακροατής 1 (18 χρονών)	4	3	4	4	4	4	2.5
Ακροατής 2 (25 χρονών)	3.5	3.5	4.5	4.5	4.5	3.5	2
Ακροατής 3 (34 χρονών)	4	3.5	4	4	3.75	3.5	2.5
Ακροατής 4 (38 χρονών)	3.75	4	4.5	3.75	3.85	3	3
Ακροατής 5 (40 χρονών)	4	4	4.5	4	4	4	2.5
MOS	3.65	3.6	4.3	4.05	4.02	3.6	2.5

Πίνακας 67 - MOS τιμές των codecs **G.711, G.721, G.722, G.726, G.729A, MPEG3 και GSM**

Όπως παρατηρούμε το codec G.726 έχει αποσπάσει την μεγαλύτερη βαθμολογία (4.3) και έτσι μπορεί να θεωρηθεί ότι αποδίδει καλύτερα από όλα τα υπόλοιπα codecs, όσο αναφορά την ποιότητα του αναπαραγόμενου ήχου. Από την άλλη βλέπουμε ότι παρ' όλο που η διαφορά του G.726 από το δεύτερο σε βαθμολογία codec, το G.721 είναι αξιοσημείωτη (0.25 μονάδες MOS) προέρχεται από δύο μόνο μετρήσεις, αυτή του Ακροατή 4 (3.75) και του Ακροατή 5 (4). Όλοι οι υπόλοιποι ακροατές δώσανε την ίδια βαθμολογία στο G.726 και στο G.721. Αντίστοιχα η μικρή διαφορά στη βαθμολογία μεταξύ του G.721 και του G.729A οφείλεται και πάλι στη βαθμολογία δύο ακροατών, του Ακροατή 3 (3.75) και του Ακροατή 4 (3.85). Όλοι οι υπόλοιποι βαθμολόγησαν τα δύο codecs το ίδιο.

Συμπερασματικά και τα τρία codecs, G.726, G.721 και G.729A, έχουν αποσπάσει βαθμολογία MOS η οποία «μεταφράζεται» σε καλή ποιότητα ήχου, με λίγες αλλά όχι ενοχλητικές απώλειες.

Από τη άλλη τα codecs G.711, G.722, MPEG3 και GSM έχουν αποσπάσει αρκετά μικρότερη βαθμολογία σε σχέση με τα τρία πρώτα codecs (G.726, G.721 και G.729A) η οποία «μεταφράζεται» σε μέτρια μέχρι ανεπαρκή ποιότητα ήχου, με ενοχλητικές απώλειες.



Εικόνα 169 - MOS τιμές των codecs G.711, G.721, G.722, G.726, G.729A, MPEG3 και GSM

6.3 Σύγκριση του G.711 και G.728

Σε [20] εξετάζονται οι αλγόριθμοι συμπίεσης G.711 και G.728 κάτω από διάφορες συνθήκες απώλειας πακέτων και καθυστερήσεων.

6.3.1 Πειραματική διαδικασία

Το «σενάριο» του πειράματος για την εξαγωγή των αποτελεσμάτων έχει ως εξής: ηχητικά δείγματα συμπιέζονται με δύο διαφορετικούς αλγόριθμους και στέλνονται σε 14 εξυπηρετητές ηχούς (echo servers) σε όλο το κόσμο. Η απώλεια πακέτων, τα πακέτα εκτός σειράς (out of order), το τρέμουλο της φωνής και ο χρόνος κυκλικής (ολικής) διαδρομής (Round Trip Time, RTT) υπολογίζονται από τα δεδομένα που φτάνουν στους εξυπηρετητές. Στη συνέχεια μελετήθηκαν ορισμένα από τα ηχητικά δείγματα για να δούμε πως οι συνθήκες του δικτύου σε συνδυασμό με το codec επηρεάζουν την ποιότητα του ήχου.

Τα codecs που χρησιμοποιήθηκαν στο πείραμα είναι το G.711 PCM Mu-Law και το G.728 CELP. Η σύγκριση περιελάμβανε ολική απώλεια πακέτων, απώλεια μόνο ενός πακέτου και διαδοχική 2, 3, 4 και περισσότερων πακέτων.

Τα αρχικά ηχητικά δείγματα είχαν τέλεια ποιότητα και είχαν ετικέτα χρονικής στιγμής (time-stamped) και ήταν αριθμημένα ακολουθιακά (sequence-numbered) ώστε να μπορεί να υπολογιστεί η καθυστέρηση, η απώλεια πακέτων και τα πακέτα εκτός σειράς. Η ποιότητα των τελικών ηχητικών δειγμάτων μετρήθηκε με την μέθοδο MOS. Για την κατάσταση του δικτύου χρησιμοποιήθηκαν εντολές ping και traceroute. Η μετάδοση των ηχητικών δειγμάτων διαρκούσε περίπου 140 λεπτά και επαναλαμβανόταν περίπου κάθε 48 με 54 λεπτά. Έτσι συνολικά μεταδόθηκαν 2010 ηχητικά δείγματα, από τα οποία αξιολογήθηκαν τα 36.

Από τις εντολές ping και traceroute προέκυψε ο **Πίνακας 68** που παρουσιάζει τον ολικό χρόνο διαδρομής για τους διάφορους εξυπηρετητές.

Περιοχή / Χώρα	Round Trip Time (RTT) min/max (ms)
ΗΠΑ	
a.psc.edu, Pittsburgh, PA	41.16/61.09
ftp.cse.buffalo.edu, Buffalo, NY.	36.05/129.82
caltech.edu, Pasadena, CA.	101.70
ftp.gatech.edu, Atlanta, GA.	61.03/85.03
beach.utmb.edu, Galveston, TX.	169.57
www.udallas.edu, Dallas, TX.	46.72
www.ucsf.edu, San Francisco, CA.	70.26
International	
ftp.rhrz.uni-bonn.de, Bonn, Germany	139.83
ftp.contrib.de, Germany	115.35
psy.uq.edu.au, Brisbane, Australia	234.42/363.13
isc.asueng.eun.eg, Cairo, Egypt	241.44/329.42
ecnet.ec, Ecuador	613.91
sunsite.ust.hk, Hong Kong, Hong Kong	318.27
www.ust.hk, Hong Kong, Hong Kong	259.31

Πίνακας 68 - RTT για τους εξυπηρετητές

Το codec G.711 Mu-Law χρησιμοποιεί τη διαμόρφωση PCM με ρυθμό δειγματοληψίας 8000 δείγματα ανά δευτερόλεπτο, όπου χρησιμοποιούνται 8 bit για την κωδικοποίηση κάθε δείγματος με ρυθμό bit 64kbps.

6.3.2 Αποτελέσματα και συμπεράσματα

Το πείραμα έδειξε ότι τα δύο codecs κάτω από κοινές συνθήκες απωλειών πακέτων επηρεάζουν διαφορετικά την ποιότητα του ήχου. Έτσι επιβεβαιώνεται ότι η ποιότητα του ήχου δεν εξαρτάται μόνο από την κατάσταση του δικτύου αλλά και από το codec που χρησιμοποιείται για την κωδικοποίηση των δεδομένων. Έτσι μια βέλτιστη λύση θα ήταν 1) να υπάρχει η δυνατότητα εναλλαγής του codec που χρησιμοποιείται ανάλογα με την

κατάσταση του δικτύου και 2) δύο codecs μπορούν να συνδυαστούν ώστε να πετύχουμε την καλύτερη εκμετάλλευση του διαθέσιμου εύρους ζώνης.

Πιο συγκεκριμένα παρατηρήθηκε ότι:

1. Ο αριθμός των ενδιάμεσων κόμβων που πρέπει να διασχίσουμε για να φτάσουμε σε έναν ορισμένο κόμβο δεν είναι κριτήριο ποιότητας του δικτύου (δηλαδή δεν ισχύει: περισσότεροι ενδιάμεσοι κόμβοι = χειρότερη ποιότητα)
2. Ο RTT είναι ένα μέγεθος που τις περισσότερες φορές μας δείχνει και τον ρυθμό απώλειας πακέτων. Ωστόσο η συσχέτιση RTT και απώλειας πακέτων δεν είναι γραμμική.
3. Ο υπολογισμός του RTT από τον αρχικό κόμβο μέχρι τον κόμβο προορισμού (end-to-end RTT) δεν είναι αρκετός για το σωστό υπολογισμό του ρυθμού απώλειας πακέτων και της συμφόρησης του δικτύου. Πρέπει να υπολογιστεί ο RTT σε κάθε ενδιάμεσο κόμβο (per-hop RTT)
4. Ένας υψηλός RTT δεν σημαίνει πάντα συμφόρηση στο δίκτυο, όπως αντίστοιχα μόνο ο αριθμός των ενδιάμεσων κόμβων δεν είναι κριτήριο για την καταλληλότητα της γραμμής για VoIP.
5. Το RTT μίας εφαρμογής (application level RTT) και το RTT του δικτύου (network level RTT) είναι κατά κανόνα ίδια αλλά υπό ορισμένες συνθήκες μπορούν να διαφέρουν.

Ο Πίνακας 69 παρουσιάζει τις MOS τιμές για κάθε codec.

Αρχείο ήχου	Loss ratio %	Single loss ratio %	Loss burst ratio %	MOS
Mu-Law				
1	0	0	0	4.20
2	2.6	93.85	6.15	3.60
3	8.16	81.86	18.14	2.80
4	13.4	78.21	21.79	2.42
5	15.48	71.58	28.42	2.20
6	16.04	53.37	46.63	2.54
7	21.6	59.44	40.56	2.06
8	30.44	44.94	55.06	1.24
G.728				
1	0	0	0	4.00
2	2.88	91.67	8.33	3.74
3	8.04	67.16	32.84	3.30
4	10.92	87.91	12.09	2.92
5	11.28	61.7	38.3	2.52
6	15.25	62.2	37.8	2.16
7	21.49	57.54	42.46	1.72
8	28.93	52.7	47.3	1.04

Πίνακας 69 - MOS τιμές και ποσοστά απωλειών των codecs G.711 και G.728

6.3.3 Αναλυτικά συμπεράσματα

Κάτω από συνθήκες πολλών απωλειών πακέτων (ποσοστό απωλειών (loss ratio) 15% και πάνω) το G.711 Mu-Law αποδίδει καλύτερα σε σχέση με το G.728 όσο αναφορά την καθαρότητα του ήχου. Όσο όμως μειώνεται η απώλεια πακέτων (ποσοστό απωλειών 10% και κάτω) το G.728 αποδίδει καλύτερα. Ωστόσο και στα δύο codec το όριο στο οποίο είναι αποδεκτή η ποιότητα του ήχου είναι περίπου ίδιο (ποσοστό απωλειών 11% - 15%).

Την παραπάνω παρατήρηση μπορούμε να την εξηγήσουμε αν μελετήσουμε το μέγεθος των πακέτων του κάθε codec. Στο G.711 Mu-Law ένα πακέτο 960 byte περιέχει 120ms ηχητικής πληροφορίας, ενώ ένα πακέτο 720 bytes στο G.728 περιέχει 360ms. Δεδομένου ότι η πιο μικρή μονάδα (unit) φωνής που γίνεται κατανοητή είναι 40-80ms, η απώλεια ενός πακέτου για το G.728 σημαίνει απώλεια 6 μονάδων ενώ για το Mu-Law σημαίνει μόλις δύο. Έτσι υπό συνθήκες υψηλής απώλειας πακέτων το G.711 Mu-Law χάνει αρκετά πακέτα (741 πακέτα το Mu-Law σε αντίθεση με το G.728 που χάνει 241 σε ποσοστό απωλειών 30%) αλλά χάνει λιγότερη ηχητική πληροφορία και έτσι το ηχητικό αποτέλεσμα είναι πιο κατανοητό, ενώ το G.728 γίνεται σχεδόν ακατανόητο.

Σε συνθήκες χαμηλής απώλειας πακέτων το G.728 αποδίδει καλύτερα γιατί σε κάθε απώλεια χάνεται ένα μέρος της λέξης αλλά το υπόλοιπο παραμένει ανέπαφο, αλλά από την άλλη το G.711 Mu-Law χάνει μικρά κομμάτια σε κάθε απώλεια αλλά έχει περισσότερες απώλειες άρα πιο πολλές διακοπές.

Υποθέτοντας ότι το G.728 περιέχει τρεις φορές περισσότερη ηχητική πληροφορία από ότι το G.711 Mu-Law θα περιμέναμε ότι το G.728 θα έχει καλύτερα αποτελέσματα υπό συνθήκες απώλειας πακέτων και θα ήταν πιο κατανοητό. Ωστόσο προέκυψε ότι το όριο για να είναι αποδεκτό το codec (3,5 στην κλίμακα MOS) είναι περίπου ίδιο και για τα δύο codec.

Μπορούμε να κάνουμε εδώ τις παρατηρήσεις:

1. ότι θα ήταν χρήσιμος ένας μηχανισμός ο οποίος θα ενάλλαζε τα codecs ανάλογα με το ποσοστό απωλειών
2. υπό συνθήκες μεγάλης απώλειας πακέτων ένα codec υψηλού ρυθμού bit μπορεί να αντικατασταθεί από ένα codec χαμηλού ρυθμού bit κάνοντας έτσι καλύτερη χρήση του εύρους ζώνης, χωρίς όμως να έχουμε απώλεια στη ποιότητα της επικοινωνίας.

Μία έκρηξη απώλειας πακέτων έχει μεγαλύτερες επιπτώσεις στο G.728 από ότι στο G.711 Mu-Law.

Η απώλεια δύο ή τριών διαδοχικών πακέτων στο Mu-Law προκαλεί απώλεια 4 – 6 ηχητικών μονάδων, ποσότητα που επιτρέπει την αναγνώριση του ηχητικού αποτελέσματος από τον άνθρωπο. Αντίθετα απώλεια δύο ή τριών πακέτων στο G.728 σημαίνει απώλεια 12 – 18 ηχητικών μονάδων, πράγμα που σημαίνει μεγάλη μείωση στη ποιότητα του ηχητικού αποτελέσματος. Συνεπώς η MOS τιμή στο G.711 Mu-Law εξαρτάται από το ολικό ποσοστό απωλειών και όχι από τις εκρήξεις απωλειών. Αντίθετα η MOS τιμή του G.728 εξαρτάται σε μεγαλύτερο βαθμό από τις εκρήξεις απωλειών και την έκτασή τους.

Στον παραπάνω πίνακα βλέπουμε ότι το G.728 έχει πολύ μικρότερο MOS για το αρχείο 5 (MOS 2.52 για ποσοστό έκρηξης (burst ratio) 38.3%) σε σχέση με το αρχείο 4 (MOS 2.92 για ποσοστό έκρηξης 12.09%) παρ' όλο που το ποσοστό ολικών απωλειών (total loss ratio) δεν διαφέρει πολύ για τα δύο αρχεία, 11.28% και 10.92% αντίστοιχα. Αντίθετα στο G.711 Mu-Law παρατηρούμε ότι για το αρχείο 5 έχουμε μικρότερο MOS σε σχέση με το αρχείο 6, 2.20

έναντι 2.54 παρ' όλο που για το αρχείο 6 έχουμε υψηλότερο συνολικό ποσοστό ολικών απωλειών και ποσοστό απωλειών έκρηξης.

Σε περιπτώσεις υψηλών απωλειών πακέτων η διόρθωση του μεγαλύτερου ποσοστού με τη μέθοδο FEC⁵ καταναλώνει περισσότερο εύρος ζώνης στο G.711 Mu-Law σε σύγκριση με το G.728

Παρατηρήθηκε ότι κάτω από συνθήκες μεγάλων απωλειών και όταν υπάρχει συμφόρηση στο δίκτυο, η μέθοδος FEC σε συνδυασμό με τη χρήση του G.711 Mu-Law απαιτεί 4 φορές περισσότερο εύρος ζώνης σε σχέση με το G.728. Από την άλλη χρησιμοποιώντας τη μέθοδο FEC σε μία σύνδεση με μικρό ρυθμό απωλειών θα πρόσθετε απλώς κίνηση στο δίκτυο, χωρίς να βελτιώνει την ποιότητα της μετάδοσης.

Για παράδειγμα από τον παραπάνω πίνακα παρατηρούμε ότι για το G.728 αρχείο 2 για ποσοστό απωλειών 2.82% έχει MOS 3.74 η οποία είναι μία τιμή υψηλότερη από την αποδεκτή και για ποσοστό απωλειών 8.04% έχει MOS 3.30 τιμή κοντά στην αποδεκτή. Έτσι μπορούμε να πούμε ότι για το G.728 η χρήση FEC σε συνθήκες χαμηλών απωλειών (περίπου 3%) δεν θα προσφέρει ουσιαστική βελτίωση στη ποιότητα, αφού η ποιότητα του ήχου είναι ήδη σε αποδεκτό επίπεδο.

Έτσι για τη σωστή χρήση του FEC δεν πρέπει να λαμβάνεται υπόψη μόνο ο ρυθμός απωλειών πακέτων αλλά και το codec που χρησιμοποιείται και η διαθεσιμότητα του εύρους ζώνης, έτσι ώστε να έχουμε την καλύτερη δυνατή χρησιμοποίηση του ολικού εύρους ζώνης.

Σχεδόν για όλα τα sites το RTT που προκύπτει από το traceroute είναι μικρότερο από το άνω όριο καθυστέρησης που είναι αποδεκτό για τον ήχο. Σε μερικές περιπτώσεις όμως το RTT ανά πακέτο είναι μεγαλύτερο από αυτό το όριο. Επίσης το τρέμουλο ανά πακέτο σε μερικές περιπτώσεις είναι πιο πάνω από το αποδεκτό.

Η καθυστέρηση και το τρεμούλιασμα της φωνής είναι δύο πολύ σημαντικοί παράγοντες που επηρεάζουν την ποιότητα της μετάδοσης. Σε πραγματικό χρόνο αν κάποια πακέτα καθυστερήσουν αρκετά πρέπει να απορριφθούν, άρα θα έχουμε απώλεια πακέτων. Παρόμοια δεν είναι δυνατή η άμεση αναπαραγωγή του ήχου με τη λήψη του και έτσι έχουμε το τρέμουλο της φωνής, το οποίο επηρεάζει την ποιότητά της.

Σύμφωνα με την εισήγηση G.114 της ITU-T (ITU-T Recommendation G.114) η αποδεκτή καθυστέρηση για τις περισσότερες εφαρμογές πραγματικού χρόνου (real time) τηλεφωνίας είναι 150ms, με χρήση του codec G.711. Ο αποδεκτός χρόνος ο οποίος μεσολαβεί από τη στιγμή που μιλάμε μέχρι τη στιγμή που θα ακούσει ο ακροατής (Mouth to Ear) διαφέρει ανάλογα με το codec το ποίο χρησιμοποιείται. Έτσι για το G.711 είναι 400ms ενώ για το G.728 είναι 324ms. Πέρα από αυτούς τους χρόνους η καθυστέρηση δεν είναι αποδεκτή. Για το RTT έχουμε 800ms και 648ms αντίστοιχα. Όπως φαίνεται και στον πίνακα με τα RTT, παρ' όλο που οι τιμές του RTT διαφέρουν αρκετά από site σε site είναι όλες εντός των ορίων αυτών εκτός από ένα.

Αντίστοιχα το όριο για το τρέμουλο είναι περίπου 50ms έτσι ώστε ο ακροατής να μπορεί να ακούσει καθαρά. Παρατηρήθηκε ότι το όριο αυτό το ξεπερνάνε οι RTT ειδικά για τα διεθνή sites (international sites). Αυτό δείχνει ότι το τρέμουλο είναι ένα σημαντικό πρόβλημα του διαδικτύου και πρέπει να χρησιμοποιηθούν ειδικοί αλγόριθμοι για τη ρύθμιση της προσωρινής αποθήκευσης (buffering) και του τρέμουλου. Από την άλλη η καθυστέρηση

⁵ Βλ. Παράρτημα A8

φαίνεται να υπερβαίνει τα όρια μόνο σε ορισμένες περιπτώσεις. Αυτό σημαίνει ότι οι μηχανισμοί για τον έλεγχο της καθυστέρησης πρέπει να έχουν δύο μεθόδους:

Τη Συνηθισμένη (regular) – όπου θα εξαλείφονται φυσιολογικοί (normal) ρυθμοί απωλειών και θα έχουμε κανονική προσωρινή αποθήκευση

Την Εξάριση (exemption) – όπου θα διαχειρίζεται καλύτερα υψηλούς ρυθμούς απωλειών

6.4 Σύγκριση των προτύπων G.723.1, G.728, G.729 και GSM

Σε αυτό το κεφάλαιο θα εξεταστούν οι συμπεριφορές των G.723.1, G.728, G.729 και GSM όσο αφορά το ρυθμό bit, την πολυπλοκότητα, την καθυστέρηση και την ποιότητα βάσει των αποτελεσμάτων του [21].

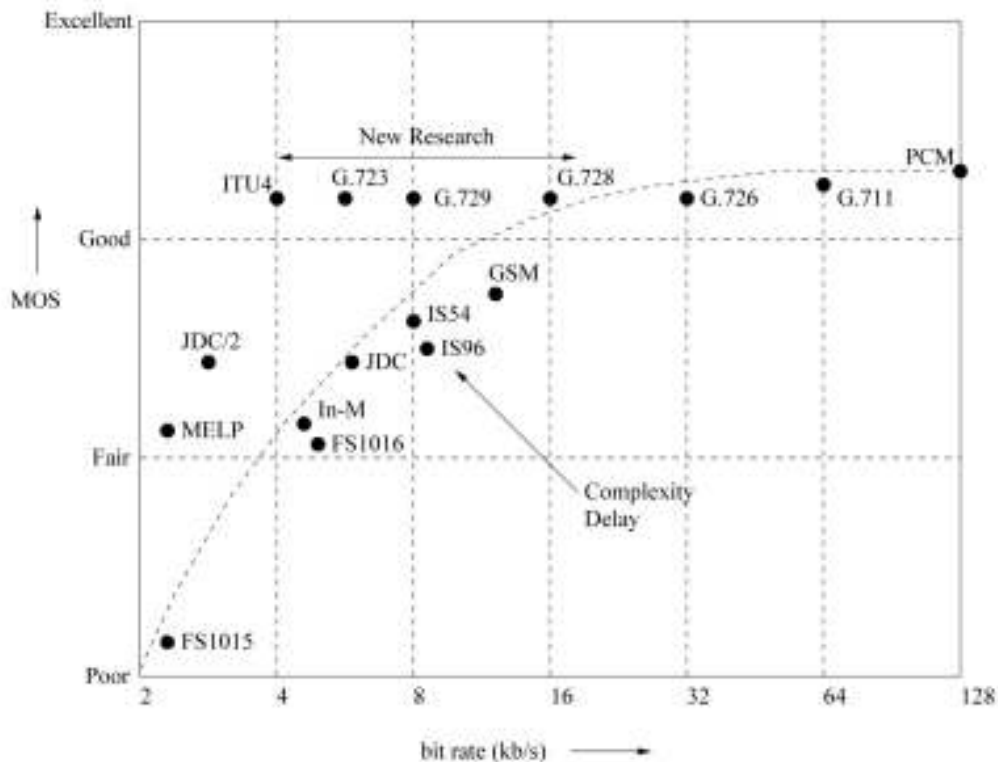
Η ITU (International Telecommunications Union, ITU) έχει προτυποποιήσει τα τρία codecs G.723.1, G.728 και G.729, τα οποία έχουν εφαρμογή σε εφαρμογές επικοινωνίας χαμηλού ρυθμού bit.

Η υπόδειξη για το G.729 8kb/s CS-ACLP της ITU έχει καθυστέρηση αλγορίθμου κωδικοποίησης 15ms για παροχή ποιότητας ήχου δικτύου (network quality speech). Στο πρότυπο G.729A μειώθηκε η πολυπλοκότητα του κωδικοποιητή CS-ACELP. Σχεδιάστηκε για ταυτόχρονη μετάδοση φωνής και δεδομένων σε επικοινωνίες όπου επικρατούν χαμηλοί ρυθμοί bit.

Η υπόδειξη για το G.723 6.3 με 5.3kb/s της ITU σχεδιάστηκε κυρίως για video-τηλέφωνα (videophones) χαμηλού ρυθμού bit. Λόγω του μεγέθους πλαισίου του 30ms και τον μονόδρομο αλγόριθμο κωδικοποίησης με καθυστέρηση 37.5ms (one-way algorithmic codec), επιτρέπει περαιτέρω μείωση του ρυθμού bit σε σχέση με το G.729. Σε εφαρμογές όπου η χαμηλή καθυστέρηση είναι σημαντική, η καθυστέρηση του G.723.1 ίσως είναι μεγάλη. Ωστόσο αν υπάρχει κάποια ανοχή στη καθυστέρηση τότε το G.723.1 είναι μία εναλλακτική του G.729 μικρότερης πολυπλοκότητας με κόστος την μικρή μείωση στην ποιότητα.

6.4.1 Υποκειμενική αξιολόγηση

Στην **Εικόνα 170** φαίνεται το διάγραμμα με τις βαθμολογίες MOS που έχουν πετύχει διάφορα codecs σύμφωνα με τις εργασίες [22] και [23].



Εικόνα 170 - Διάγραμμα MOS - bit rate των πιο διαδεδομένων codecs, IEEE 1996

Όπως παρατηρούμε τα τελευταία χρόνια έχουν αναπτυχθεί codecs τα οποία πετυχαίνουν ποιότητα όμοια με αυτή του 64kbps G.711 PCM codec, παρόλο το κόστος της αύξησης της καθυστέρησης κωδικοποίησης (coding delay) και της πολυπλοκότητας (implementation complexity). Το 8kbps G.729 codec είναι πρότυπο της ITU, το οποίο ξεπερνά κατά πολύ τα προηγούμενα πρότυπα σε όρους απόδοσης (robustness terms). Η απόδοση των 4kbps ITU codec (ITU4) συμπληρώνει το εντυπωσιακό αυτό σετ χαρακτηριστικών.

Υπό συνθήκες μετάδοσης ελεύθερων λαθών (error-free transmission) και χωρίς θόρυβο φόντου (no background-noise) η υποκειμενική ποιότητα του 16kbps G.728, του 8kbps G.729 και του 6.4kbps G.723.1 είναι περίπου 4 μονάδες MOS. Δηλαδή η ποιότητά τους είναι σχεδόν παρόμοια παρόλη τη διαφορά στο ρυθμό bit. Η όμοια ποιότητα για τόσο διαφορετικούς ρυθμούς bit είναι μία ένδειξη ότι αντιπροσωπεύουν διαφορετικές στιγμές στην εξέλιξη των codecs φωνής. Επίσης δείχνει ότι υπάρχει και διαφορά στην πολυπλοκότητα εφαρμογής.

Το 13kbps GSM και το 5.3kbps G.723 είναι λίγο χαμηλότερα στην απόδοση με τιμή MOS περίπου 3.5. Εδώ πρέπει να σημειωθεί ότι υπάρχει το codec 13 kbps πλήρους ρυθμού (full rate) το οποίο έχει MOS περίπου 4.

Η οικογένεια των codecs τα οποία σχεδιάστηκαν για διάφορα κινητά ραδιοφωνικά συστήματα (mobile radio systems) όπως το 13kbps RPE GSM, 7.95kbps IS-54, IS-96, 6.7kbps JDC και 3.45kbps half-rate JDC (JDC/2) πετυχαίνουν ελαφρώς μικρότερες τιμές MOS σε σχέση με τα codecs της ITU.

Το 2.4kbps Federal Standard codec FS-1015 είναι ο μόνος κωδικοποιητής φωνής πηγής (source codec – vocoder)⁶ στο σύνολο των παραπάνω codecs και για αυτό είναι πιο πολύπλοκος ο υπολογισμός της ποιότητας του σε σχέση με την χαμηλή τιμή που φαίνεται στο διάγραμμα. Το 64kbps G.711 PCM codec και το G.726/G.727 ADPCM είναι codecs κυματομορφής (waveforms codecs)⁷. Πετυχαίνουν χαμηλή πολυπλοκότητα εφαρμογής σε συνδυασμό με μία μέτρια οικονομία στο ρυθμό bit. Τα υπόλοιπα codecs ανήκουν σε μία υβριδική οικογένεια (hybrid codec family) τα οποία πετυχαίνουν μεγάλη οικονομία σε ρυθμό bit αλλά με κόστος την αύξηση της πολυπλοκότητας και της καθυστέρησης.

Πιο συγκεκριμένα το προσαρμοστικό προς τα πίσω 16kbps G.728 (backward-adaptive scheme) (το οποίο προτυποποιήθηκε τις αρχές του 90') έχει ίδια ποιότητα στα 32 και στα 64 kbps με τα codecs κυματομορφής και συγχρόνως έχει πολύ χαμηλή καθυστέρηση, 2ms. Ένα codec παρόμοιας ποιότητας αλλά αρκετά πιο αποδοτικό είναι το 8kbps G.728 codec, το οποίο εγκρίθηκε τον Μάρτιο του 1996 από την ITU. Η ρύθμιση στα 6.4kbps διατηρεί ποιότητα όμοια με τα G.711, G.726, G.727 και G.728 ενώ για 5.3kbps mode πετυχαίνει ποιότητα παρόμοια με των κυψελωτών codecs (cellular codecs) στις αρχές του 80'.

Παράλληλα με τα πρότυπα της ITU για τα codecs φωνής προτάθηκαν και codecs για κυψελωτά κινητά συστήματα (cellular mobile systems). Το πρότυπο 13kbps RPE-LTP full-rate GSM (GSM-FR) codec το οποίο χρονολογείται στο δεύτερο μισό της δεκαετίας του 80', αποτελεί το πρώτο υβριδικό codec. Η πολυπλοκότητά του είναι αρκετά μικρότερη από αυτή των σχετικά πρόσφατων codecs CELP-based.

Παρατηρώντας το διάγραμμα βλέπουμε ότι το GSM-EFR (identical-rate enhanced full-rate GSM codec, GSM-EFR), πετυχαίνει ποιότητα παρόμοια με αυτή των G.729 και G.728. Το αρχικό GSM-FR codec ακολουθήθηκε λίγο αργότερα από το 8kbps VSELP IS-54 American κυψελωτό πρότυπο (cellular standard). Λόγω των πλεονεκτημάτων του το 7.95kbps IS-54 codec είχε παρόμοια υποκειμενική ποιότητα με το 13kbps GSM-FR scheme. Η ανάπτυξη του IS-54 codec ακολουθήθηκε από μία διαδικασία προτυποποίησης μισού ρυθμού (half-rate standardization process), η οποία οδήγησε στο 3.2kbps Pitch-Synchronous Innovation CELP (PSI-CELP). Το σύστημα IS-96 American CDMA έχει το δικό του προτυποποιημένο codec βασισμένο στο CELP (CELP-based speech codec), το οποίο λειτουργεί σε διάφορους ρυθμούς bit (variable-rate scheme) και επιτρέπει ρυθμούς bit από 1.2kbps μέχρι 14.4kbps, αναλόγως τη φωνητική δραστηριότητα που επικρατεί (prevalent voice activity). Οι παρατηρούμενες ποιότητες των αναφερθέντων κυψελωτών codecs, βρέθηκαν να είναι υποκειμενικά ίδιες μεταξύ τους υπό τέλειες συνθήκες διαύλου (perfect channel conditions). Τέλος, το 5.6kbps GSM codec (GSM-HR) μισού ρυθμού κατάφερε να πετύχει ποιότητα παρόμοια με του αρχικού 13kbps GSM-FR παρ' όλο το κόστος της τετραπλάσιας πολυπλοκότητας και της μεγαλύτερης καθυστέρησης (latency).

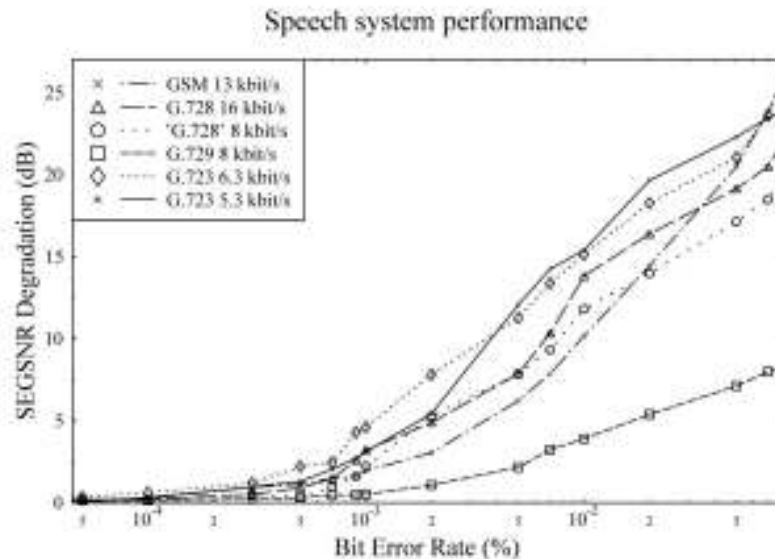
6.4.2 Σύγκριση της ευαισθησίας στα λάθη

Ένα στοιχείο αντικειμενικής μέτρησης της ποιότητας ενός codec είναι η σύγκριση της ευαισθησίας των bits (bit-sensitivity). Στο διάγραμμα της **Εικόνας 171** απεικονίζεται ο SEGSNR (τμηματικός σηματοθορυβικός λόγος) διαφόρων codecs για διάφορους ρυθμούς λαθών (bit error rates, BER) όταν εφαρμόζονται τυχαία λάθη. Το SEGSNR γενικά δεν είναι μία αξιόπιστη μέτρηση της ποιότητας, αλλά δείχνει ικανοποιητικά πόσο γρήγορα αυτή η αντικειμενική μέτρηση της ποιότητας μειώνεται για διάφορα codecs όταν εφαρμόζονται διαφορετικά BER. Όπως αναμενόταν το προσαρμοστικό προς τα πίσω codec (backwards-adaptive) G.728 και το προσαρμοστικό προς τα εμπρός (forward-adaptive) G.723.1, τα

⁶ Βλ. Παράρτημα A10

⁷ Βλ. Παράρτημα A9

οποία έχουν σχεδιαστεί βασικά για ήπιες καλωδιακές συνδέσεις (benign wireline connections), έχουν τη γρηγορότερη μείωση του SEGSNR καθώς αυξάνεται το BER. Η καλύτερη απόδοση πετυχαίνεται με μεγάλη διαφορά από το G.729, και ακολουθείται από το 13kbps GSM codec.



Εικόνα 171 - Διάγραμμα SEGSNR / BER των πιο διαδεδομένων codecs

6.4.3 Αντικειμενική αξιολόγηση

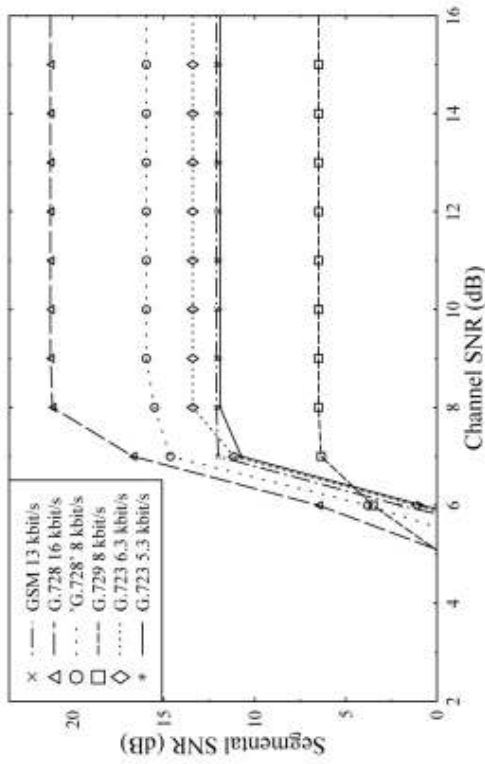
Σε αυτό το κεφάλαιο εξετάζεται η συμπεριφορά των codecs κάτω από όμοιες συνθήκες, όταν χρησιμοποιούνται όμοιοι transceivers (δηλαδή πομπός και δέκτης μαζί με χρήση του ίδιου κυκλώματος) πάνω από κανάλια Gaussian και Rayleigh.

Χρησιμοποιήθηκαν τέσσερα modems 1,2 και 4 bits ανά σύμβολο, διαμόρφωσης BPSK (Binary Phase Shift Keying, BPSK) και τετραγωνισμένη διαμόρφωση πλάτους τεσσάρων επιπέδων (4-level Quadrature Amplitude Modulation, 4QAM) και 16-QAM, σε συνδυασμό με έξι διαφορετικές ρυθμίσεις λειτουργίας των τεσσάρων codecs, τα οποία προστατεύονταν από το BCH (254,130,18) codec καναλιού (channel codec). Σημειωτέον ότι δεν χρησιμοποιήθηκε κανένας συνδυασμός ευαισθησίας πηγής (source-sensitivity) και καναλιού πολλών κλάσεων (multi-class channel) για να εξασφαλιστούν όμοιες συνθήκες για όλα τα codecs.

Το SEGSNR παρόλο που δεν είναι ένα πολύ αξιόπιστο μέγεθος μέτρησης, όταν συγκρίνονται codecs τα οποία λειτουργούν στη βάση διαφορετικών αλγορίθμων κωδικοποίησης, μπορεί να χρησιμοποιηθεί για τη μέτρηση της απόδοσης, παρουσιάζοντας μείωση για κακές συνθήκες διαύλου και έτσι επιτρέποντάς μας να βρούμε το ελάχιστο SNR του διαύλου που απαιτείται για τα διάφορα codecs και του transceivers.

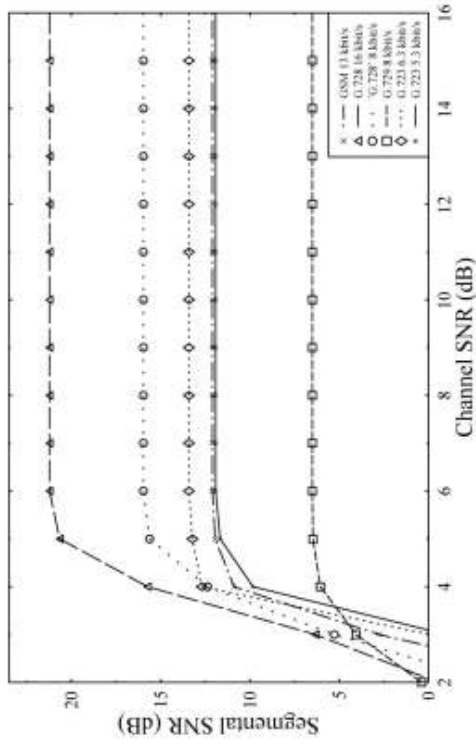
Στις Εικόνες 172 μέχρι και 177 παρουσιάζονται τα διαγράμματα του SEGSNR σε σχέση του SNR διαύλου.

Speech System Performance with 4QAM over Gaussian Channel



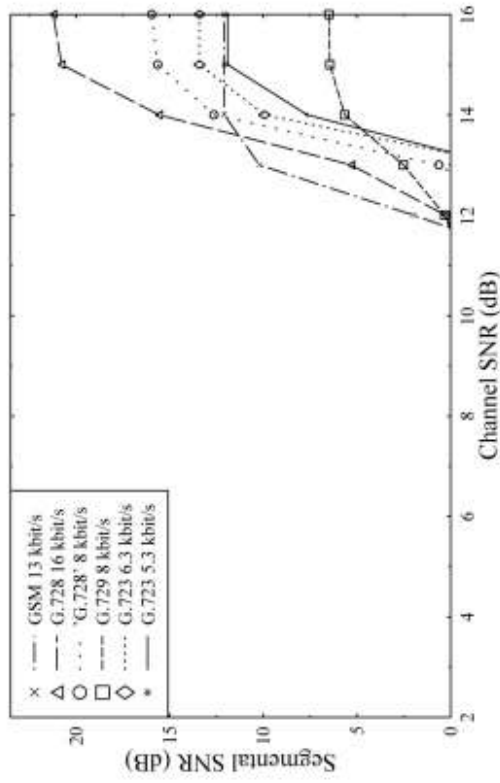
Εικόνα 173 - Διάγραμμα SEGSNR / SNR διαύλου με χρήση BCH(254,130,18) κώδικα και 4QAM πάνω από Gaussian κανάλι

Speech System Performance with BPSK over Gaussian Channel



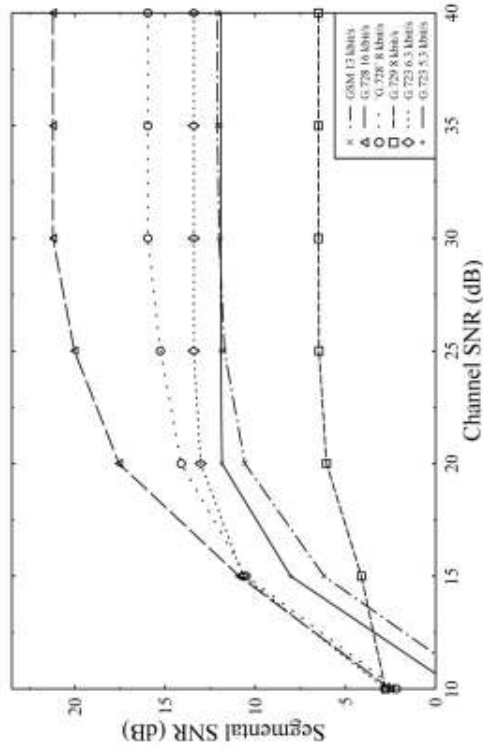
Εικόνα 172 – Διάγραμμα SEGSNR / SNR διαύλου με χρήση BCH(254,130,18) κώδικα και BPSK πάνω από Gaussian κανάλι

Speech System Performance with 16QAM over Gaussian Channel



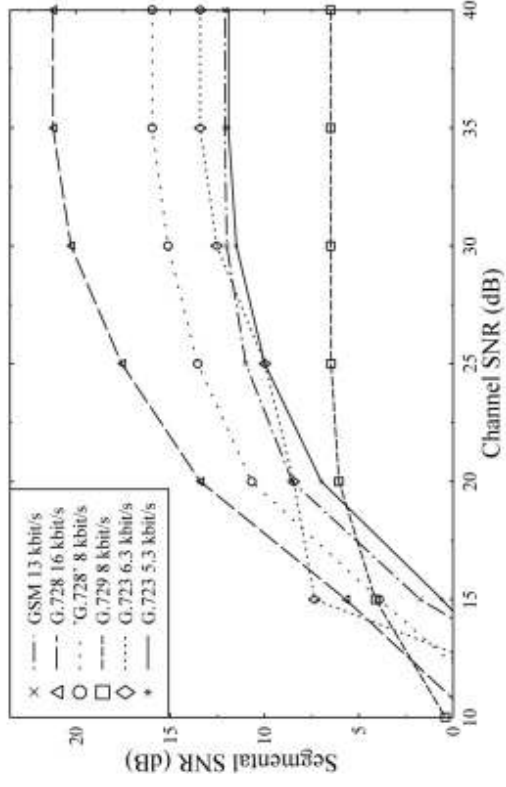
Εικόνα 174 - Διάγραμμα SEGSNR / SNR διαύλου με χρήση BCH(254,130,18) κώδικα και 16QAM πάνω από Gaussian κανάλι

Speech System Performance with BPSK over Rayleigh Channel



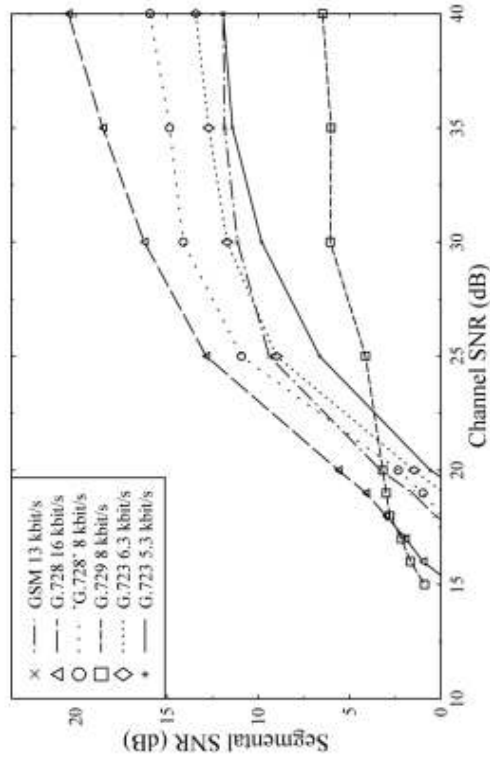
Εικόνα 175 - Διάγραμμα SEGSNR / SNR διαύλου με χρήση BCH(254,130,18) κώδικα και BPSK πάνω από Rayleigh κανάλι

Speech System Performance with 4QAM over Rayleigh Channel



Εικόνα 176 - Διάγραμμα SEGSR / SNR διαύλου με χρήση BCH(254,130,18) κώδικα και 4QAM πάνω από Rayleigh κανάλι

Speech System Performance with 16QAM over Rayleigh Channel



Εικόνα 177 - Διάγραμμα SEGSR / SNR διαύλου με χρήση BCH(254,130,18) κώδικα και 16QAM πάνω από Rayleigh κανάλι

Όπως βλέπουμε τα αποτελέσματα αυτά διαφέρουν από τα υποκειμενικά αποτελέσματα. Για παράδειγμα το 8kbps G.729 codec έχει μικρότερο SEGSNR παρόλο που έχει όμοια τιμή MOS με τα G.728 και 6.4kbps G.723.1. Αυτό οφείλεται στο υπερηχητικό φίλτρο στην είσοδο του καθώς επίσης και στη βελτιστοποίηση της ποιότητας των αντληπτών ήχων.

Ενδιαφέρον παρουσιάζει και η σύγκριση του 8kbps G.728 codec το οποίο δείχνει υψηλότερο SEGSNR από το G.729 όμοιου ρυθμού bit, αλλά ακούγεται πολύ χειρότερα σε σχέση με το G.729. Αυτές οι διαφορές γίνονται όλο και πιο φανερές όταν ελέγξουμε τα λάθη διαύλου στις περιοχές χαμηλού SNR των διαγραμμάτων.

Στον **Πίνακα 70** παρουσιάζονται οι ελάχιστες τιμές SNR του διαύλου για τους διάφορους transceivers πάνω από διαύλους Gaussian και Rayleigh.

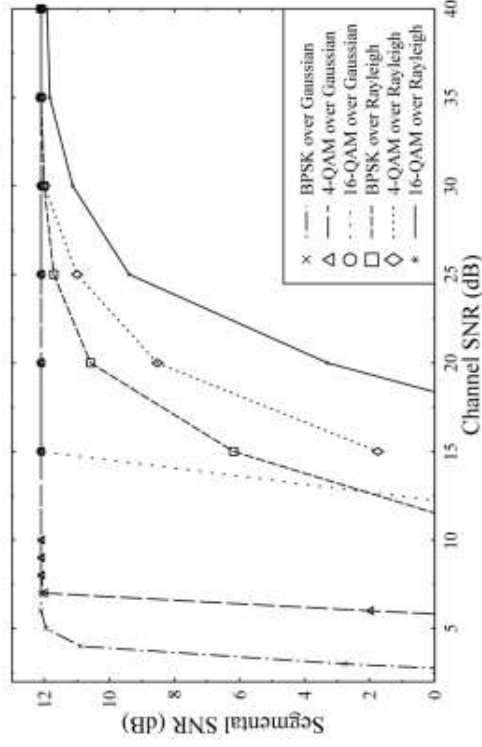
Codec	Ρυθμός bit (kbps)	BPSK		4-QAM		16-QAM	
		AWGN	Ray.	AWGN	Ray.	AWGN	Ray.
GSM	13	4	20	7	27	13	34
G.728	16	5	26	8	30	15	40
G,728	8	5	25	7	31	15	35
G.729	8	4	19	7	20	14	28
G.723.1	6.4	4	18	8	31	15	35
G.723.1	5.3	4	19	7	29	15	35

Πίνακας 70 – Ελάχιστο απαιτούμενο SNR διαύλου για τη διατήρηση της μείωσης του SEGSNR κάτω από 1dB με χρήση BCH(254,130,18) κώδικα και BPSK, 4QAM, 16QAM πάνω από Gaussian και Rayleigh κανάλια

Παρατηρούμε από τα διαγράμματα ότι σε δίαυλο Rayleigh τα προσαρμοστικά προς τα πίσω codecs (backward-adaptive codecs) έχουν μία μεγάλη μείωση της απόδοσης, ενώ το προσαρμοστικό προς τα εμπρός G.729 (forward-adaptive) ACELO έχει μία πιο αποδοτική συμπεριφορά.

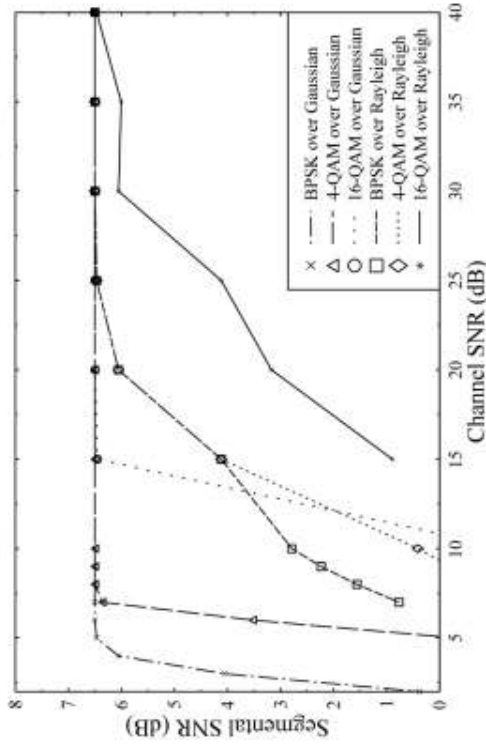
Στις **Εικόνες 178 μέχρι και 181** παρουσιάζονται όλα τα διαφορετικά SEGSNR σε σχέση με τις καμπύλες του SNR διαύλου, για ένα συγκεκριμένο codec. Έτσι μας δίνεται η δυνατότητα για απευθείας σύγκριση της αναμενόμενης απόδοσης των διαφόρων transceivers υπό ορισμένες συνθήκες διαύλου.

GSM 13 kbit/s performance over Gaussian and Rayleigh channels



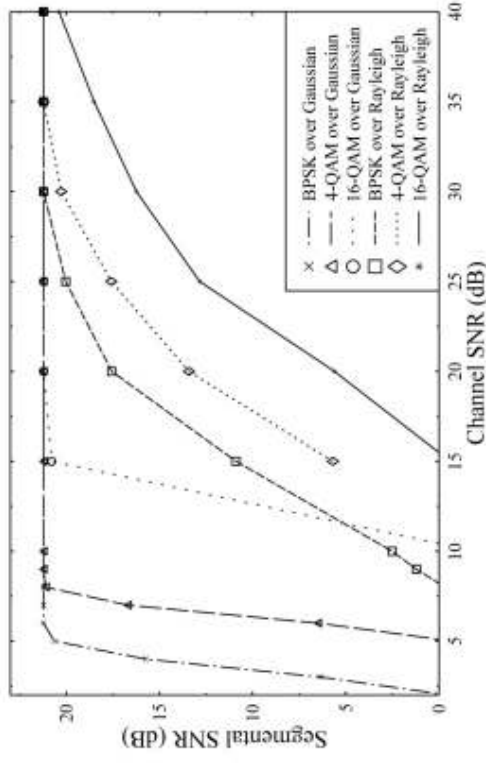
Εικόνα 178 – Διάγραμμα μείωσης SEGSNR / SNR διαύλου για το 13 kbps RPELTP GSM codec, με χρήση BCH(254,130,18) κώδικα και BPSK, 4QAM, 16QAM πάνω από Gaussian και Rayleigh κανάλια

G. 729 8 kbit/s performance over Gaussian and Rayleigh channels



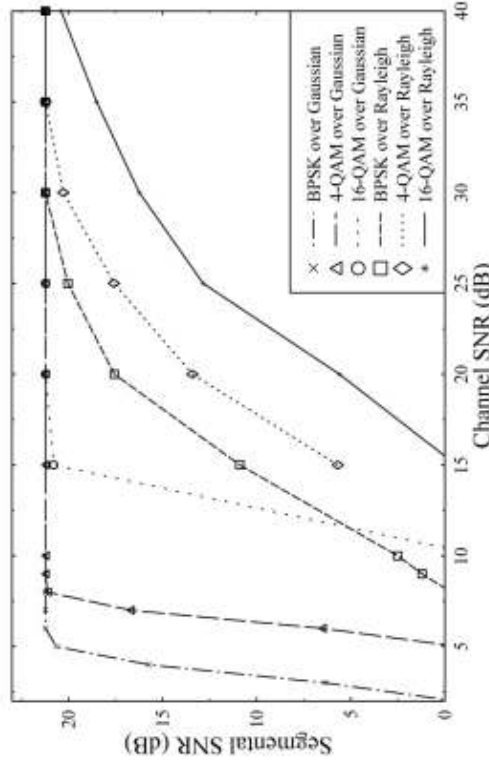
Εικόνα 180 - Διάγραμμα μείωσης SEGSNR / SNR διαύλου για το 8 kbps Forward-Adaptive G.729 codec, με χρήση BCH(254,130,18) κώδικα και BPSK, 4QAM, 16QAM πάνω από Gaussian και Rayleigh κανάλια

G.728 16 kbit/s performance over Gaussian and Rayleigh channels



Εικόνα 179 - Διάγραμμα μείωσης SEGSNR / SNR διαύλου για το 13 kbps Backward-Adaptive G.728 codec, με χρήση BCH(254,130,18) κώδικα και BPSK, 4QAM, 16QAM πάνω από Gaussian και Rayleigh κανάλια

G.728 16 kbit/s performance over Gaussian and Rayleigh channels



Εικόνα 181 - Διάγραμμα μείωσης SEGSNR / SNR διαύλου για το 5.3 kbps G.723.1 codec, με χρήση BCH(254,130,18) κώδικα και BPSK, 4QAM, 16QAM πάνω από Gaussian και Rayleigh κανάλια

6.5 Σύγκριση των G.711, G.729 και AMR

Στο [24] γίνεται μία ανάλυση της ποιότητας των αλγορίθμων συμπίεσης ήχου G.711, G.729 και AMP (Adaptive Multi-Rate). Μεγάλη βάση δίνεται στην απώλεια πακέτων, η οποία είναι βασικό πρόβλημα των VoIP εφαρμογών. Έχει σχεδιαστεί ένα σύνολο από ηχητικά δείγματα, το οποίο καλύπτει τις πιο συνηθισμένες συνθήκες απώλειας πακέτων. Για την αξιολόγηση των αλγορίθμων χρησιμοποιήθηκε η αντικειμενική μέθοδος MOS.

6.5.1 Πειραματική διαδικασία

Το σύνολο των ηχητικών δειγμάτων περιέχει ένα εύρος από συνθήκες απώλειας πακέτων όπως φαίνεται στον Πίνακα 71. Επιλέχθηκαν τα επικρατέστερα codecs τόσο στο ασύρματο VoIP όσο και στο ενσύρματο, όπως το G.711, G.729 και το Adaptive Multi-Rate (AMR). Το G.711 είναι ένα απλό codec κυματομορφής. Το G.729 και το AMR λειτουργούν σε πλαίσια και βασίζονται στις αρχές της γραμμικής πρόβλεψης (code excited linear prediction principles). Για το G.711, το ενσωματωμένο Σύστημα Απόκρυψης Χαμένων Πακέτων (Packet Loss Concealment, PLC) αναφέρεται σε ένα αλγόριθμο όπως παρουσιάζεται στο ITU-T Rec. G.711 Appendix I και η απουσία του PLC αναφέρεται στην αντικατάσταση των χαμένων πακέτων με σιγή. Στο σύνολο υπάρχουν 108 συνθήκες απώλειας πακέτων.

Codec	G.711, G.729, AMR (12.2kbps)
Μέγεθος πακέτου	10, 20, 30 ms για το G.711 και το G.729 και 20ms για το AMR
Ρυθμός απωλειών	1%, 2%, 4%, 7% και 10%
Είδος απωλειών	τυχαία, έκρηξη
Ζεύγη (tandeming)	G.729xG.729, AMRxAMR

Πίνακας 71 - Παράμετροι του πειράματος σύγκρισης των G.711, G.729 και AMR

Το σύνολο των ηχητικών δειγμάτων περιλαμβάνει επίσης και 16 συνθήκες αναφοράς (reference conditions) οι οποίες αναφέρονται στο φιλτράρισμα ή μη ένα τροποποιημένο ενδιάμεσο σύστημα αναφοράς (modified intermediate reference system, MIRS), το MNRU, τις κυματομορφές και τη γραμμική πρόβλεψη βάσει κώδικα (code excited linear prediction) codecs υπό συνθήκες καθαρού διαύλου. Αυτές οι συνθήκες αναφοράς βοηθάνε στη σύγκριση με άλλα σύνολα ηχητικών δειγμάτων.

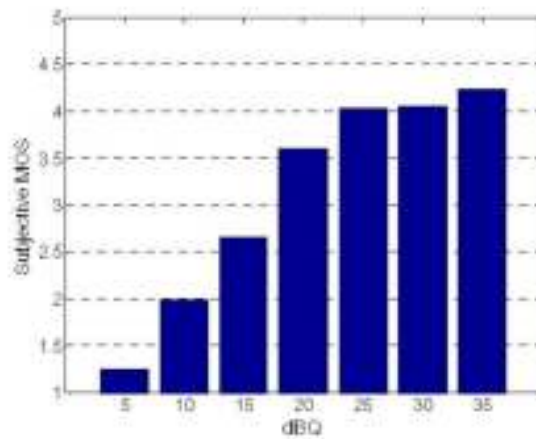
Τα δείγματα ηχογραφούνται σε υψηλή ποιότητα σε ένα ανηχοϊκό θάλαμο. Για τις ηχογραφήσεις χρησιμοποιήθηκαν τέσσερις Καναδοί (2 άντρες και 2 γυναίκες). Τα ηχητικά δείγματα είναι φιλτραρισμένα MIRS (εκτός από ένα αφιλτράριστο), στα -26dBov (dB(overload) – το πλάτος ενός σήματος, συνήθως ηχητικού σε σύγκριση με το μέγιστο πλάτος όπου μπορεί να διαχειριστεί η συσκευή χωρίς να υπάρξει απώλεια/αποκοπή) σωσμένο στα 8000Hz με διαμόρφωση στα 16bit (16bit format). Κάθε ηχητικό δείγμα (speech sample) περιέχει δύο προτάσεις χωρισμένες με ένα διάστημα σιγής 650ms.

Η βαθμολόγηση MOS έγινε στο Speech and Audio Processing Lab του πανεπιστημίου του Sherbrook το 2006. Το πείραμα ακολουθεί τις απαιτήσεις όπως αναφέρονται στο ITU-T Rec. P. 830. Στο πείραμα συμμετείχαν 60 ακροατές (38 άντρες και 22 γυναίκες). Το θόρυβος περιβάλλοντος μετριόταν καθημερινά στο ύψος του κεφαλιού του κάθε ακροατή (απουσία του ακροατή) και προέκυψε σε όλες τις περιπτώσεις περίπου 27-28dBA.

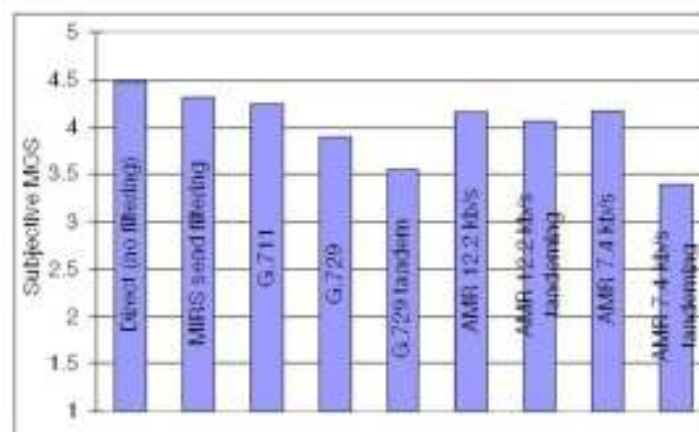
Για την αντικειμενική μέτρηση της ποιότητας χρησιμοποιήθηκαν το Mean Opinion Score - Listening Quality Objective (MOS-LQO, ITU Rec. P.862.1), το Perceptual Evaluation of Speech Quality (PESQ, ITU-T Rec. P.862) και το ITU-T Rec. P.563.

6.5.2 Αποτελέσματα και συμπεράσματα

Στην **Εικόνα 182** φαίνεται η γραφική παράσταση MNRU, όπου παρατηρείται ότι οι βαθμολογίες MOS αυξάνονται με την αύξηση του dBQ. Μετά τα 25dBQ ο ρυθμός αύξησης είναι μικρότερος όπως άλλωστε αναμενόταν. Οι βαθμολογίες MOS για το κάθε codec φαίνονται στην **Εικόνα 183**. Η απευθείας (χωρίς φίλτράρισμα) και η MIRS περιπτώσεις χρησιμοποιούν γραμμική κβαντοποίηση. Το G.711 παράγει μη-γραμμικά βήματα κβαντοποίησης και στις υπόλοιπες περιπτώσεις έχουμε συμπίεση με διάφορα codecs. Οι βαθμολογίες είναι οι αναμενόμενες: το απευθείας (αφιλτράριστο) έχει το μεγαλύτερο MOS ακολουθούμενο από το MIRS, το G.711 και μετά τα υπόλοιπα. Μία εξαίρεση είναι ότι η βαθμολογία του AMR 7.4kbps είναι λίγο καλύτερη από το κανονικό.

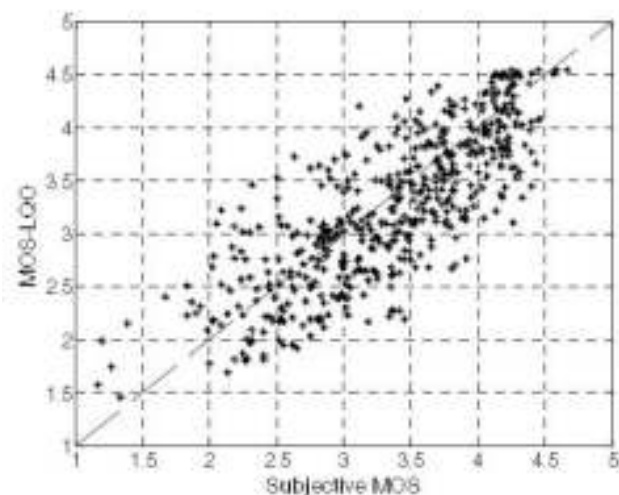


Εικόνα 182 - Διάγραμμα MNRU



Εικόνα 183 - MOS τιμές των codecs

Η **Εικόνα 184** δείχνει το υποκειμενικό και το αντικειμενικό MOS για κάθε δείγμα του συνόλου. Οι τιμές των υποκειμενικών και αντικειμενικών MOS, όπως παρατηρείται βρίσκονται ανάμεσα στο κατώτατο όριο 1.0 και το ανώτερο 4.55. Αυτό δείχνει ότι το MOS-LQO αποδίδει αρκετά καλά καθώς οι κουκίδες είναι κοντά και όμοια διατεταγμένες στην διαγώνιο.



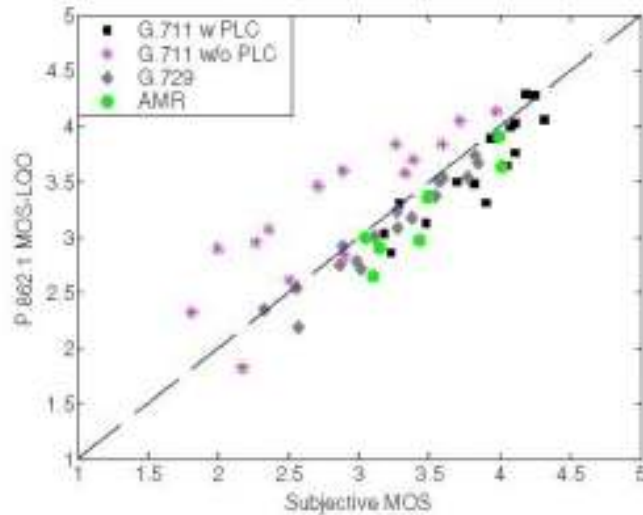
Εικόνα 184 - Διάγραμμα υποκειμενικού και αντικειμενικού MOS

Η απόδοση αυτών των αντικειμενικών αλγορίθμων χαρακτηρίζεται από τον συντελεστή συσχέτισης Pearson ρ και το μέσο τετραγωνικό σφάλμα σ (root mean square error, RMSE). Αρχικά υπολογίζεται η γραμμική σχέση μεταξύ της υποκειμενικής και αντικειμενικής βαθμολογίας MOS και έτσι προκύπτει πόσο κοντά είναι η αντικειμενική βαθμολογία στην υποκειμενική. Η ανάλυση γίνεται στο σύνολο των δειγμάτων και των αναφορών. Τα αποτελέσματα για το ρ και το σ φαίνονται στον **Πίνακα 72**. Παρατηρείται ότι το ρ είναι 0.89 για το PESQ και το MOS-LQO, ενώ για το P.563 είναι 0.56. Το σ για το PESQ και το MOS-LQO είναι μικρότερο από αυτό του P.563. Όπως ήταν αναμενόμενο οι αδιάκριτες (intrusive) μέθοδοι είναι πιο αποτελεσματικοί γιατί έχουμε πρόσβαση στην αναφορά. Το P.563 δεν είναι έγκυρη προσέγγιση και πρέπει να χρησιμοποιείται προσεκτικά. Παρόμοια απόδοση παρατηρείται όταν συγκρίνεται το MOS-LQO και το PESQ. Η διαφορά είναι πολύ μικρή για να καταλήξουμε σε ασφαλές συμπέρασμα.

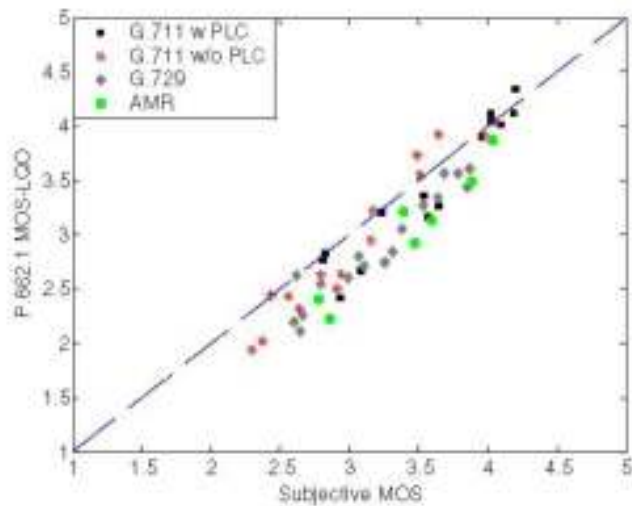
		Σύνολο	Αναφοράς	Τυχαίες απώλειες	Έκρηξη απωλειών
MOS-LQO	ρ	0.89	0.97	0.83	0.95
	σ	0.32	0.23	0.34	0.31
PESQ	ρ	0.89	0.95	0.83	0.94
	σ	0.30	0.32	0.36	0.21
P.563	ρ	0.56	0.87	0.67	0.32
	σ	0.53	0.67	0.47	0.24

Πίνακας 72 - Συντελεστής συσχέτισης (ρ) και μέσο τετραγωνικό σφάλμα (δ) των αντικειμενικών μεθόδων

Στην **Εικόνα 185** φαίνονται τα αποτελέσματα της απόδοσης των codecs για τυχαία απώλεια πακέτων, ενώ στην **Εικόνα 186** για έκρηξη απωλειών πακέτων. Για να είναι πιο καθαρά τα συμπεράσματα παρουσιάζεται η μέση βαθμολογία MOS για κάθε codec. Παρατηρείται μία γραμμική σχέση, ενώ το MOS-LQO είναι παρόμοιο και στις δύο εικόνες (για τυχαία απώλεια πακέτων και έκρηξη απώλειας πακέτων). Για το G.711 χωρίς PLC παρατηρούμαι ότι στην περίπτωση τυχαίας απώλειας πακέτων βρίσκεται πάνω από τη διαγώνιο ενώ στη περίπτωση έκρηξης απώλειας πακέτων βρίσκεται από κάτω. Αυτό σημαίνει ότι στη πρώτη περίπτωση αποδίδει καλύτερα, το ποίο είναι αναμενόμενο.



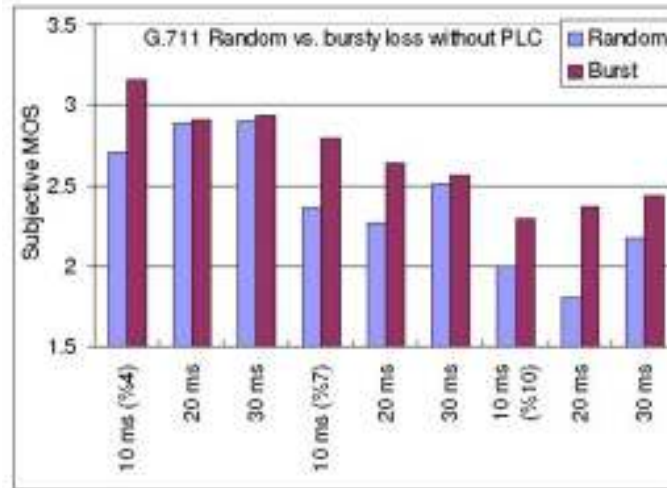
Εικόνα 185 - MOS τιμές για τυχαία απώλεια πακέτων



Εικόνα 186 - MOS τιμές για έκρηξη απώλειας πακέτων

Η απόδοση του G.711 χωρίς PLC εξετάζεται υπό το μοντέλο E (ITU-T Rec. G.107, The E-Model, a computational model for use in transmission planning, March 2005) με χρήση του MOS-LQO, περιμένοντας ότι θα είχαμε έκρηξη απώλειας πακέτων στη χαμηλή ποιότητα

ανάλογη της τυχαίας απώλειας πακέτων, υπό όμοιες συνθήκες, καθώς αποκόβονται μεγαλύτερα κομμάτια και το φαινόμενο αυτό αναμένεται να είναι πιο ενοχλητικό. Ωστόσο η **Εικόνα 187** δείχνει το ακριβώς αντίθετο. Αυτό μπορεί να εξηγηθεί καθώς όταν δεν χρησιμοποιείται το PLC το φαινόμενο της μειωμένης συχνότητας των clicks μπορεί να αντισταθμιστεί πιο εύκολα από την απώλεια πολλών πακέτων σε μία έκρηξη απώλειας πακέτων.



Εικόνα 187 - Η επίπτωση της έκρηξης απώλειας πακέτων για το G.711 χωρίς PLC

Συμπερασματικά το MOS-LQO και το PESQ είναι εφαρμόσιμα για την εκτίμηση των επιδράσεων της απώλειας πακέτων στο VoIP και έχουν παρόμοια απόδοση ωστόσο είναι περιοριστικά στο χειρισμό του G.711 χωρίς PLC. Το P.563 έχει μέτρια απόδοση ως προς την απώλεια πακέτων, καθώς είναι μία διακριτή (non-intrusive) μέτρηση.

7. Κυριότερα codecs για VoIP

Στο παρόν κεφάλαιο παρουσιάζονται τα πιο δημοφιλή codecs που χρησιμοποιούνται στο VoIP. Πρέπει να έχουμε στο μυαλό μας ότι όσο πιο μεγάλη συμπίεση χρησιμοποιείται τόσο περισσότερο χρόνο χρειάζεται για να πραγματοποιηθεί και άρα εισάγεται και μεγαλύτερη καθυστέρηση. Το ιδανικό codec είναι αυτό το οποίο κρατάει την ποιότητα σε ένα επίπεδο, αλλά συγχρόνως χρειάζεται τον λιγότερο δυνατό χρόνο για την κωδικοποίηση και αποκωδικοποίηση.

Codec	Συχνότητα δειγματοληψίας (kHz)	Εύρος ζώνης (kbps)	Ονομαστικό εύρος ζώνης (kbps)	Μέγεθος ωφέλιμου φορτίου (ms)	Άδεια	Σχόλια	Θετικά	Αρνητικά
DVI4	-	-	-	-	-	Πολύ σπάνιο codec	-	-
G.711	8	64	87.2	20	Ανοιχτού κώδικα	G.711 u/a συχνά αναφέρεται και ως u-law/a-law όπου το a-law είναι η Ευρωπαϊκή έκδοση και το u-law είναι η Αμερικανική / Ιαπωνική	Σχεδιάστηκε για να προσφέρει πολύ καλή μετάδοση φωνής. Πολύ μικρή επιβάρυνση επεξεργασίας.	Μαζί με την επιβάρυνση επεξεργασίας χρησιμοποιεί πάνω από 64kbps έτσι απαιτείται τουλάχιστον 128kbps (σε κάθε κατεύθυνση) δίκτυο.
G.722	16 16 16	48 56 64	- - -	30 30 30	Ανοιχτού κώδικα	Codec της ITU	-	-
G.723.1	8 8	5.3 6.3	20.8 21.9	30 30	Proprietary	Χρησιμοποιείται συχνά για Dialup VoIP για μέγιστη ποιότητα.	Πολύ υψηλή συμπίεση με πολύ υψηλή ποιότητα.	Χρειάζεται πολύ επεξεργαστική ισχύ.
G.726	8 8 8 8	16 24 32 40	- 47.2 55.2 -	- 20 20 -	Ανοιχτού κώδικα	Βελτιωμένη έκδοση του G.721 και G.723 (τελείως διαφορετικό από το G.723.1)	-	Η επιβάρυνση του CPU σε σχέση με τη συμπίεση που πετυχαίνουμε.
G.728	-	16	31.5	-	Ανοιχτού κώδικα	Codec της ITU	-	-
G.729	8	8	31.2	20	Patented	Codec της ITU	Πολύ καλή χρήση του εύρους ζώνης. Αποδίδει καλά κάτω από συνθήκες τυχαία λανθασμένων bit	Χρειάζεται άδεια για τη χρήση.

GSM	8	13	-		Proprietary	Το ίδιο codec με το GSM που χρησιμοποιείται στη κινητή τηλεφωνία, αν και τώρα χρησιμοποιούνται πιο εξελιγμένες εκδόσεις.	Σχετικά υψηλής συμπίεσης codec. Royalty free δηλαδή είναι διαθέσιμο σε πολλά hardware και software.	-
iLBC	- -	13.33 15	- -	30 20	Free to use	-	Αρκετά καλό σε συνθήκες απώλειας πακέτων.	-
Siren	-	-	-	-	-	Δεν υποστηρίζεται αρκετά το συγκεκριμένο codec.	-	-
Speex	8 16 32	- - -	- - -	- - -	Ανοιχτού κώδικα	-	Χρησιμοποιεί διάφορους ρυθμούς bit για να ελαχιστοποιήσει τη χρήση εύρους ζώνης	-

Πίνακας 73 - Χαρακτηριστικά δημοφιλέστερων codecs ήχου

7.1 Εύρος ζώνης / ονομαστικό εύρος ζώνης

Οι τιμές στη στήλη του εύρους ζώνης αντιπροσωπεύουν το ποσοστό των δεδομένων στο ωφέλιμο φορτίου του IP πακέτου (payload of the IP packets). Η τιμή αντιπροσωπεύει το εύρος ζώνης σε κάθε κατεύθυνση, όχι το άθροισμα εξερχόμενης (upstream) και εισερχόμενης (downstream) ροής δεδομένων. Οι τιμές αυτές υπολογίζονται για συνεχόμενη μετάδοση και στις δύο κατευθύνσεις, χωρίς διακοπές σιγής. Το ονομαστικό εύρος ζώνης είναι το εύρος ζώνης που αναμένεται να χρησιμοποιήσει το codec.

7.2 Ρυθμός δειγματοληψίας

Ο ρυθμός δειγματοληψίας (sampling rate) είναι ο ρυθμός με τον οποίο το αναλογικό σήμα δειγματοληπτείται και γίνεται ψηφιακό. Η συχνότητα δειγματοληψίας γνωρίζουμε ότι πρέπει να είναι τουλάχιστον διπλάσια από τη συχνότητα του σήματος ή αλλιώς η παρουσία δύο δειγμάτων ανά κύκλο είναι απαραίτητη για να μην χαθεί η πληροφορία που αφορά τη τιμή της φυσικής συχνότητας του σήματος. Συνεπώς όσο μεγαλύτερη η συχνότητα δειγματοληψίας τόσο μεγαλύτερο το εύρος των συχνοτήτων στις οποίες μπορούμε να κωδικοποιήσουμε. Το ανθρώπινο αυτί ακούει ένα εύρος ζώνης συχνοτήτων από 20Hz μέχρι 20kHz. Η συνηθισμένη ομιλία είναι περίπου στα 100Hz – 4kHz και έτσι απαιτείται συχνότητα δειγματοληψίας τουλάχιστον 8kHz για την κωδικοποίησή της. Μεγαλύτερες συχνότητες δειγματοληψίας θα κωδικοποιήσουν και υψηλότερες συχνότητες (το οποίο είναι χρήσιμο σε ορισμένες μόνο περιπτώσεις, όπως για παράδειγμα για κωδικοποίηση μουσικής) αλλά επίσης θα αυξηθεί και το απαιτούμενο εύρος ζώνης αφού θα έχουμε περισσότερα δείγματα για κωδικοποίηση και μετάδοση.

7.3 Μέγεθος ωφέλιμου φορτίου

Το μέγεθος ωφέλιμου φορτίου (payload size) κάθε κωδικοποιημένου πακέτου επηρεάζει δύο πράγματα, την καθυστέρηση (lag) και το εύρος ζώνης. Σε κάθε κωδικοποιημένο πακέτο που στέλνεται προστίθενται επιπλέον επικεφαλίδες. Έτσι πακέτα με μεγαλύτερο ωφέλιμο φορτίο περιέχουν αναλογικά λιγότερη πληροφορία στις επικεφαλίδες, έτσι μειώνεται το ονομαστικό εύρος ζώνης. Ωστόσο χρησιμοποιώντας περισσότερο ωφέλιμο φορτίο, περισσότερη ηχητική πληροφορία (για παράδειγμα μεγαλύτερη χρονική διάρκεια) δημιουργείται ένα μόνο πακέτο, το οποίο αυξάνει το χρόνο που χρειάζεται για να φτάσει στην άλλη άκρη της σύνδεσης καθώς επίσης και το χρόνο που χρειάζεται για να αποκωδικοποιηθεί και έτσι αυξάνεται η καθυστέρηση. Τα περισσότερα codecs έχουν μέγεθος ωφέλιμου φορτίου 10-40ms.

A. Παράρτημα

A1. Μετάδοση σε πραγματικό χρόνο

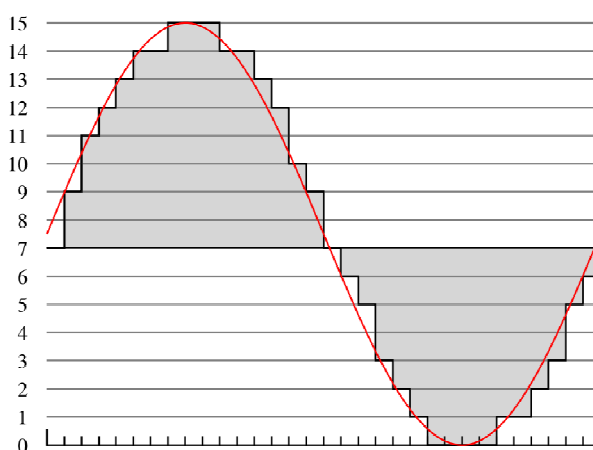
Η μετάδοση σε πραγματικό χρόνο (Real Time Transmission) βασίζεται σε IP δίκτυα και χρησιμοποιείται για την μετάδοση τόσο video όσο και ήχου. Παρέχονται υπηρεσίες όπως χρονική ανακατασκευή (time reconstruction), ανίχνευση λαθών, ασφάλεια και αναγνώριση περιεχομένου και αρχικά σχεδιάστηκε για μετάδοση σε πολλούς παραλήπτες συγχρόνως (multicast), ωστόσο μπορεί να χρησιμοποιηθεί και για μετάδοση σε ένα μόνο παραλήπτη (unicast). Μπορεί να χρησιμοποιηθεί για υπηρεσίες κατ' απαίτηση (video on demand) όπου η μετάδοση είναι μονόδρομη (one-way) αλλά και για διαδραστικές (interactive) υπηρεσίες όπως η τηλεφωνία VoIP.

Για τη μετάδοση σε πραγματικό χρόνο χρησιμοποιούνται δύο κανάλια. Ένα για την μετάδοση των δεδομένων πληροφορίας (data) και ένα για την μετάδοση των δεδομένων ελέγχου (control data).

A2. Παλμοκωδική διαμόρφωση (PCM)

Με τη παλμοκωδική διαμόρφωση (Pulse Code Modulation, PCM) ένα αναλογικό σήμα αναπαριστάται με ένα ψηφιακό, όπου το αναλογικό σήμα δειγματοληπτείται ανά συγκεκριμένο χρονικό διάστημα χρησιμοποιώντας ένα σήμα παλμών. Στη συνέχεια το σήμα ψηφιοποιείται ανάλογα με τις τιμές που έχουμε από το σήμα της δειγματοληψίας.

Στην **Εικόνα 188** βλέπουμε πως ένα ημιτονοειδές σήμα δειγματοληπτείται. Στη συνέχεια μπορούμε να το κωδικοποιήσουμε σε ψηφιακά δεδομένα. Στο σχεδιάγραμμα παρατηρούμε ότι οι τιμές του σήματος δειγματοληψίας είναι 7, 9, 11, 12, 13, 14, 15, 15, 15, 14 κτλ., οι τιμές αυτές μπορεί να κωδικοποιηθούν σε δυαδική μορφή 0111, 1001, 1011, 1100, 1101, 1110, 1110, 1111, 1111, 1111, 1110 κλπ.



Εικόνα 188 - Δειγματοληψία ημιτονοειδούς σήματος

A3. Διαφορική παλμοκωδική διαμόρφωση (DPCM)

Με την διαφορική παλμοκωδική διαμόρφωση (Differential Pulse Code Modulation, DPCM) κωδικοποιούνται μόνο οι διαφορές μεταξύ προηγούμενης και επόμενης τιμής. Με αυτό τον τρόπο καταφέρνουμε να μειώσουμε τα bits που χρειάζονται ανά δείγμα μέχρι και 25% σε

σύγκριση με την PCM. Κατά την αποκωδικοποίηση η τιμή του δείγματος υπολογίζεται βάσει των τελευταίων τιμών των αποκωδικοποιημένων δειγμάτων.

Μία συνηθισμένη τεχνική κωδικοποίησης φωνής θα προσπαθήσει να προβλέψει την τιμή του επόμενου δείγματος, βάσει του προηγούμενου. Αυτό είναι δυνατό λόγω του συσχετισμού που υπάρχει στα δείγματα φωνής (speech samples) λόγω της έκτασης της φωνής και των δονήσεων των φωνητικών χορδών. Αν οι προβλέψεις είναι αποδοτικές τότε το σήμα λάθους (error signal) μεταξύ των δειγμάτων που προβλέπονται και των τελικών δειγμάτων θα έχει μικρή διαφορά από τα αρχικά δείγματα. Έτσι θα αυτό το λάθος σήμα μπορεί να κβαντιστεί με λιγότερα bits από το αρχικό σήμα φωνής. Αυτή είναι και η βάση της DPCM διαμόρφωσης, όπου κβαντίζεται η διαφορά μεταξύ του αρχικού και του προβλεπόμενου σήματος.

A4. Προσαρμοσμένη διαφορική παλμοκωδική διαμόρφωση (ADPCM)

Στην προσαρμοσμένη διαφορική παλμοκωδική διαμόρφωση (Adaptive Differential Pulse Code Modulation, ADPCM) μπορούμε να ορίζουμε το βήμα κβαντοποίησης (quantization step) έτσι ώστε να πετύχουμε μείωση του απαιτούμενου εύρους ζώνης, για έναν συγκεκριμένο SNR (Signal to Noise Ratio, Λόγος Σήματος προς Θόρυβο).

Για καλύτερη απόδοση του codec ο μηχανισμός πρόβλεψης και ο κβαντιστής (quantizer) μπορούν να προσαρμόζονται στα χαρακτηριστικά της φωνής που κωδικοποιείται. Αυτή είναι η βάση της ADPCM διαμόρφωσης.

A5. Υπηρεσίες καλύτερης προσπάθειας

Στις Υπηρεσίες Καλύτερης Προσπάθειας (Best-effort) όλοι οι συνδρομητές που χρησιμοποιούν τη συγκεκριμένη υπηρεσία, ανταγωνίζονται για το διαθέσιμο εύρος ζώνης. Τυπικά χρησιμοποιείται σε εφαρμογές όπως το διαδίκτυο. [3]

A6. Διαφοροποιημένες υπηρεσίες

Στις διαφοροποιημένες υπηρεσίες (Differentiated Services, Diffserv) υπάρχει ένας μηχανισμός ο οποίος αναπτύσσεται στο πυρήνα του δικτύου. Η κίνηση η οποία μπαίνει σε ένα DS δίκτυο ταξινομείται, στη συνέχεια δίνεται μία τιμή στο DSCP (Diffserv Code Point, DSCP). Οι δρομολογητές (routers) κοιτάνε το DSCP και παρέχουν την ανάλογη συμπεριφορά ανά κόμβο (Per Hop Behavior, PHB) στη κίνηση. Το PHB είναι ένα μέσο για να δεσμεύονται πόροι στο δίκτυο για σύνολα με διαφορετικά DSCP.

Το EF PHB προϋποθέτει ότι ο ρυθμός αναχώρησης του συνόλου με ίδιο DSCP είναι ίδιος ή μεγαλύτερος από το ρυθμό που φτάνουν τα σύνολα με ίδιο DSCP.

A7. Επιταχυμένη προώθηση PHB

Η Επιταχυμένη Προώθηση PHB (Expedited Forwarding PHB, EF PHB) έχει μικρή καθυστέρηση, μικρή απώλεια πακέτων και μικρό τρεμούλιασμα φωνής (jitter). Στη κίνηση που έχει χαρακτηριστεί EF δίνεται προτεραιότητα σε σχέση με τις άλλες υπηρεσίες. Επειδή αυξημένη EF κίνηση θα οδηγήσει σε μεγάλες ουρές άρα και καθυστερήσεις σε όλο το δίκτυο, έχουν αναπτυχθεί διάφοροι μηχανισμοί ελέγχου και συνήθως τα τυπικά δίκτυα περιορίζουν την EF κίνηση στο 30% (μέγιστο) της συνολικής χωρητικότητας της σύνδεσης.

A8. Προοπτική διόρθωση λαθών

Η προοπτική διόρθωση λαθών (Forward Error Correction, FEC) είναι μία μέθοδος για τη διόρθωση λαθών στις τηλεπικοινωνίες, χωρίς να απαιτείται η αναμετάδοση του λανθασμένου πακέτου. Ο αποστολέας προσθέτει στο πακέτο κάποια επιπλέον δεδομένα (γνωστά ως κώδικας διόρθωσης λαθών – error correction code). Αυτό δίνει τη δυνατότητα στο λήπτη να εντοπίσει και να διορθώσει ορισμένα λάθη, χωρίς να απαιτείται η αποστολή επιπλέον δεδομένων από τον αποστολέα. Έτσι αποφεύγεται η μετάδοση δεδομένων για διόρθωση λαθών καθώς επίσης και αναμετάδοση όλου του πακέτου, κερδίζοντας έτσι σε εύρος ζώνης.

A9. Codecs κυματομορφής

Τα codecs κυματομορφής (waveform codecs) παράγουν ένα ανακατασκευασμένο σήμα του οποίου η κυματομορφή είναι όσο το δυνατό πιο όμοια με τη κυματομορφή του αρχικού σήματος, χωρίς να γνωρίζουμε πως έχει δημιουργηθεί το κωδικοποιημένο σήμα. Θεωρητικά αυτό σημαίνει ότι τα codecs κυματομορφής είναι ανεξάρτητα του σήματος και δουλεύουν εξίσου αποδοτικά και για σήματα που δεν είναι φωνητικά (non-speech signals). Γενικά είναι codecs χαμηλής πολυπλοκότητας, τα οποία παράγουν υψηλής ποιότητας φωνή για ρυθμούς bit πάνω από τα 16kbrs. Αν ο ρυθμός bit είναι κάτω από αυτό το όριο η ποιότητα μπορεί να μειωθεί πολύ γρήγορα. Η πιο απλή μέθοδος κωδικοποίησης κυματομορφής είναι η PCM.

A10. Codecs πηγής και vocoders

Τα codecs πηγής (source codecs) λειτουργούν χρησιμοποιώντας ένα μοντέλο για το πώς δημιουργήθηκε το αρχικό σήμα και προσπαθούν να εξάγουν από το σήμα που κωδικοποιείται τα χαρακτηριστικά του μοντέλου αυτού. Αυτά τα χαρακτηριστικά μεταδίδονται στον αποκωδικοποιητή. Τα codecs πηγής για τη φωνή λέγονται vocoders. Το εύρος φωνής απεικονίζεται σαν ένα φίλτρο διαφόρων χρόνων (time-varying filter), το οποίο διεγείρεται είτε από λευκό θόρυβο (white noise) (σε περιόδους όπου δεν υπάρχει φωνή) είτε από παλμούς οι οποίοι έχουν περίοδο την περίοδο της φωνής. Η πληροφορία που πρέπει να σταλεί στον αποκωδικοποιητή είναι τα χαρακτηριστικά του φίλτρου, όπως οι περίοδοι ύπαρξης και μη ύπαρξης φωνής (voiced/unvoiced flag), την απαραίτητη διαφορά του σήματος διέγερσης και η περίοδος των σημάτων φωνής. Αυτά ανανεώνονται κάθε 10-20ms λόγω της μη σταθερότητας της φωνής.

A11. Υβριδικά codecs ήχου

Τα υβριδικά codecs στοχεύουν να καλύψουν το κενό μεταξύ των codecs κυματομορφής και των codecs πηγής. Όπως αναφέρθηκε τα codecs κυματομορφής μπορούν να έχουν καλή ποιότητα για ρυθμούς bit πάνω από 16kbrs, αλλά κάτω από αυτό το ρυθμό bit η ποιότητα μειώνεται αισθητά. Από την άλλη, τα codecs πηγής μπορούν να αποδώσουν κατανοητή ομιλία στα 2.4kbrs και κάτω, αλλά δεν μπορούν να αποδώσουν ομιλία που να ακούγεται φυσικά σε οποιοδήποτε ρυθμό bit. Τα πιο διαδεδομένα υβριδικά codecs είναι τα Analysis-by-Synthesis (AbS) codecs. Χρησιμοποιούν την ίδια γραμμική πρόβλεψη που χρησιμοποιείται και στα codecs πηγής, αλλά αντί να εφαρμόζουν ένα μοντέλο δύο καταστάσεων (φωνή/όχι φωνή – voiced/unvoiced) για την εύρεση του σήματος, το σήμα διέγερσης επιλέγεται έτσι ώστε να ταιριάζει με την κυματομορφή του αρχικού σήματος.

A12. Υποκειμενική μέθοδος MOS

Μία μέθοδος για την αξιολόγηση της ποιότητας ενός codec είναι η μέθοδος MOS. Πρόκειται για μία αντιστοιχία αριθμών και ποιότητας του ήχου ή της εικόνας, μετά τη συμπίεση ή/και τη μεταφορά. Η κλίμακα των αριθμών είναι από το 1 μέχρι το 5, όπου το 1 υποδηλώνει τη χαμηλότερη ποιότητα και το 5 την υψηλότερη.

Ο συνολικός βαθμός MOS για ένα codec προκύπτει από το μέσο όρο των επιμέρους βαθμών που έχει δώσει ένα πλήθος ακροατών (αν πρόκειται για ήχο). Το ηχητικό δείγμα πρέπει να αποτελείται από προτάσεις τόσο από αντρικές φωνές όσο και από γυναικείες και ο ακροατής πρέπει να το αξιολογήσει με βάση τον παρακάτω πίνακα.

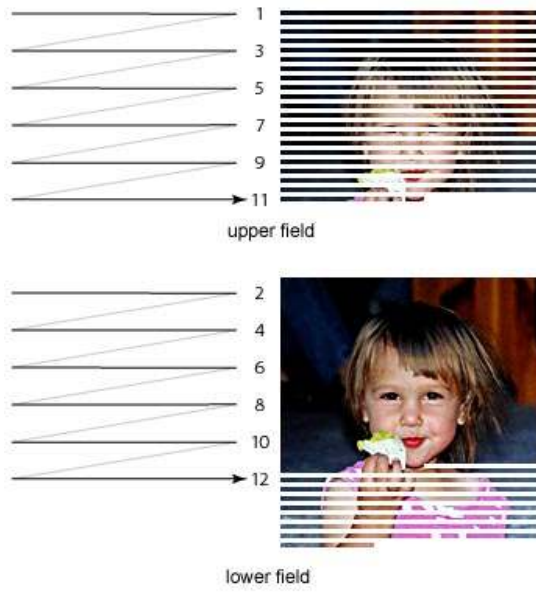
MOS	Ποιότητα	Απώλειες στη ποιότητα
5	Εξαιρετική	Ανεπαίσθητες
4	Καλή	Λίγες αλλά όχι ενοχλητικές
3	Μέτρια	Λίγο ενοχλητικές
2	Ανεπαρκής	Ενοχλητικές
1	Κακή	Πολύ ενοχλητικές

Πίνακας 74 - Η κλίμακα MOS

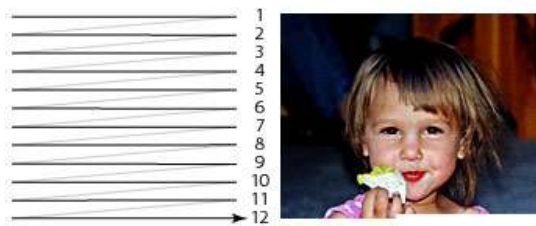
A13. Πεπλεγμένο και μη-πεπλεγμένο video (Interlaced/non-interlaced video)

Η μέθοδος του interlaced video χρησιμοποιείται από τις παραδοσιακές τηλεοράσεις. Όταν η τηλεόραση προβάλλει το αναλογικό σήμα του video, παρουσιάζει πρώτα τις περιττές γραμμές (πρώτο πεδίο – field) και εν συνεχεία τις άρτιες (δεύτερο πεδίο), όπου κάθε ζευγάρι σχηματίζει ένα καρτέ.

Οι οθόνες του υπολογιστή χρησιμοποιούν τη μέθοδο της σταδιακής σάρωσης (progressive scan). Σύμφωνα με αυτή τη μέθοδο, η οποία λέγεται και non-interlaced video, η οθόνη χωρίζεται σε πεδία και κάθε γραμμή εμφανίζεται με τη σειρά από πάνω προς τα κάτω.



Εικόνα 189 – Πεπλεγμένο video (Interlaced video)



Εικόνα 190 – Σταδιακή σάρωση (Progressive scan) (μη-πεπλεγμένο video)

Κεφάλαιο 3

ΕΙΚΟΝΑ 1 - ΔΙΑΓΡΑΜΜΑ ΣΧΕΣΗΣ ΚΑΤΑΝΟΜΗΣ ΤΩΝ BITS ΚΑΙ ΤΗΣ ΠΟΛΥΠΛΟΚΟΤΗΤΑΣ ΤΗΣ ΚΙΝΗΣΗΣ	31
ΕΙΚΟΝΑ 2 – ΣΧΕΣΗ ΝΕΑΣ ΜΕΘΟΔΟΥ ΥΠΟΛΟΓΙΣΜΟΥ ΤΗΣ ΠΟΛΥΠΛΟΚΟΤΗΤΑΣ ΤΗΣ ΚΙΝΗΣΗΣ ΚΑΙ ΚΑΤΑΝΟΜΗΣ ΤΩΝ BITS.....	32
ΕΙΚΟΝΑ 3 - ΚΑΜΠΥΛΗ ΤΟΥ ΠΟΣΟΣΤΟΥ ΠΑΡΑΜΟΡΦΩΣΗΣ ΚΑΙ ΕΠΙΛΟΓΗ ΤΗΣ ΜΕΘΟΔΟΥ ΚΩΔΙΚΟΠΟΙΗΣΗΣ.....	32
ΕΙΚΟΝΑ 4 - ΔΙΑΓΡΑΜΜΑ ΡΟΗΣ ΤΟΥ ΠΡΟΤΕΙΝΟΜΕΝΟΥ ΑΛΓΟΡΙΘΜΟΥ ΕΛΕΓΧΟΥ ΤΟΥ ΡΥΘΜΟΥ BIT	33
ΕΙΚΟΝΑ 5 - ΑΣΥΜΠΙΕΣΤΗ ΕΙΚΟΝΑ ΗΛΕΚΤΡΟΝΙΚΗΣ ΤΟΜΟΓΡΑΦΙΑΣ	37
ΕΙΚΟΝΑ 6 - ΣΥΜΠΙΕΣΜΕΝΗ ΕΙΚΟΝΑ ΗΛΕΚΤΡΟΝΙΚΗΣ ΤΟΜΟΓΡΑΦΙΑΣ ΜΕ MPEG-4	37
ΕΙΚΟΝΑ 7 - ΣΥΜΠΙΕΣΜΕΝΗ ΕΙΚΟΝΑ ΗΛΕΚΤΡΟΝΙΚΗΣ ΤΟΜΟΓΡΑΦΙΑΣ ΜΕ H.264	37
ΕΙΚΟΝΑ 8 - ΑΣΥΜΠΙΕΣΤΗ ΕΙΚΟΝΑ ΗΧΟΚΑΡΔΙΟΓΡΑΦΗΜΑΤΟΣ	37
ΕΙΚΟΝΑ 9 - ΣΥΜΠΙΕΣΜΕΝΗ ΕΙΚΟΝΑ ΗΧΟΚΑΡΔΙΟΓΡΑΦΗΜΑΤΟΣ ΜΕ MPEG-4.....	37
ΕΙΚΟΝΑ 10 - ΣΥΜΠΙΕΣΜΕΝΗ ΕΙΚΟΝΑ ΗΧΟΚΑΡΔΙΟΓΡΑΦΗΜΑΤΟΣ ΜΕ H.264	37
ΕΙΚΟΝΑ 11 - ΣΧΕΣΗ PSNR ΚΑΙ ΑΡΙΘΜΟΥ ΠΛΑΙΣΙΟΥ ΓΙΑ ΤΟ VIDEO ΗΧΟΚΑΡΔΙΟΓΡΑΦΗΜΑΤΟΣ, ΜΕ ΡΥΘΜΟ BIT 667.08KBPS 39	
ΕΙΚΟΝΑ 12 - ΣΧΕΣΗ PSNR ΚΑΙ ΑΡΙΘΜΟΥ ΠΛΑΙΣΙΟΥ ΓΙΑ ΤΟ VIDEO ΗΛΕΚΤΡΟΝΙΚΗΣ ΤΟΜΟΓΡΑΦΙΑΣ, ΜΕ ΡΥΘΜΟ BIT 285.03KBPS	40
ΕΙΚΟΝΑ 13 - ΔΙΑΓΡΑΜΜΑ ΡΥΘΜΟΥ BIT / PSNR ΓΙΑ ΤΟ VIDEO FOREMAN QCIF	43
ΕΙΚΟΝΑ 14 – ΔΙΑΓΡΑΜΜΑ ΡΥΘΜΟΥ BIT / PSNR ΓΙΑ ΤΟ VIDEO MISS AMERICA QCIF.....	43
ΕΙΚΟΝΑ 15 – ΔΙΑΓΡΑΜΜΑ ΡΥΘΜΟΥ BIT / PSNR ΓΙΑ ΤΟ VIDEO COAST GUARD QCIF	43
ΕΙΚΟΝΑ 16 - ΔΙΑΓΡΑΜΜΑ ΡΥΘΜΟΥ BIT / PSNR ΓΙΑ ΤΟ VIDEO CAR AND PHONE QCIF	43
ΕΙΚΟΝΑ 17 - ΔΙΑΓΡΑΜΜΑ ΡΥΘΜΟΥ BIT / PSNR ΓΙΑ ΤΟ VIDEO TENNIS QCIF	43
ΕΙΚΟΝΑ 18 - ΔΙΑΓΡΑΜΜΑ ΡΥΘΜΟΥ BIT / PSNR ΤΟΥ H.264 ΚΑΙ ΤΟΥ MPEG-2 ΓΙΑ ΤΟ VIDEO CAROUSEL	46
ΕΙΚΟΝΑ 19 - ΔΙΑΓΡΑΜΜΑ ΡΥΘΜΟΥ BIT / PSNR ΤΟΥ H.264 ΚΑΙ ΤΟΥ MPEG-2 ΓΙΑ ΤΟ VIDEO FLOWER.....	46
ΕΙΚΟΝΑ 20 – ΔΙΑΓΡΑΜΜΑ ΡΥΘΜΟΥ BIT / PSNR ΤΟΥ H.264, ΤΟΥ H.263 (ΒΑΣΙΚΟ ΠΡΟΦΙΛ) ΚΑΙ ΤΟΥ H.263 (ΠΡΟΦΙΛ ΥΨΗΛΗΣ ΣΥΜΠΙΕΣΗΣ) ΓΙΑ ΤΟ VIDEO PARIS CIF	49
ΕΙΚΟΝΑ 21 – ΔΙΑΓΡΑΜΜΑ ΡΥΘΜΟΥ BIT / PSNR ΤΟΥ H.264, ΤΟΥ H.263 (ΒΑΣΙΚΟ ΠΡΟΦΙΛ) ΚΑΙ ΤΟΥ H.263 (ΠΡΟΦΙΛ ΥΨΗΛΗΣ ΣΥΜΠΙΕΣΗΣ) ΓΙΑ ΤΟ VIDEO FOREMAN QCIF.....	49
ΕΙΚΟΝΑ 22 - ΔΙΑΓΡΑΜΜΑ ΡΥΘΜΟΥ BIT / PSNR ΤΩΝ H.264, MJPEG2000 ΚΑΙ PACC ΓΙΑ ΤΟ VIDEO PARIS	55
ΕΙΚΟΝΑ 23 - ΔΙΑΓΡΑΜΜΑ ΡΥΘΜΟΥ BIT / PSNR ΤΩΝ H.264, MJPEG2000 ΚΑΙ PACC ΓΙΑ ΤΟ VIDEO BUS.....	55
ΕΙΚΟΝΑ 24 - ΔΙΑΓΡΑΜΜΑ ΡΥΘΜΟΥ BIT / PSNR ΤΩΝ H.264, MJPEG2000 ΚΑΙ PACC ΓΙΑ ΤΟ VIDEO MOBILE	55
ΕΙΚΟΝΑ 25 - ΔΙΑΓΡΑΜΜΑ ΡΥΘΜΟΥ BIT / PSNR ΤΩΝ H.264, MJPEG2000 ΚΑΙ PACC ΓΙΑ ΤΟ VIDEO CANOE.....	55
ΕΙΚΟΝΑ 26 - ΔΙΑΓΡΑΜΜΑ ΡΥΘΜΟΥ BIT / PSNR ΤΩΝ H.264, MJPEG2000 ΚΑΙ PACC ΓΙΑ ΤΟ VIDEO CREW	57
ΕΙΚΟΝΑ 27 - ΔΙΑΓΡΑΜΜΑ ΡΥΘΜΟΥ BIT / PSNR ΤΩΝ H.264, MJPEG2000 ΚΑΙ PACC ΓΙΑ ΤΟ VIDEO HARBOUR.....	57
ΕΙΚΟΝΑ 28 - ΔΙΑΓΡΑΜΜΑ ΡΥΘΜΟΥ BIT / PSNR ΤΩΝ H.264, MJPEG2000 ΚΑΙ PACC ΓΙΑ ΤΟ VIDEO VINTAGE CAR.....	57
ΕΙΚΟΝΑ 29 - ΔΙΑΓΡΑΜΜΑ ΡΥΘΜΟΥ BIT / PSNR ΤΩΝ H.264, MJPEG2000 ΚΑΙ PACC ΓΙΑ ΤΟ VIDEO BOOK	57
ΕΙΚΟΝΑ 30 - ΜΠΛΟΚ ΔΙΑΓΡΑΜΜΑ ΛΕΙΤΟΥΡΓΙΑΣ ΕΝΟΣ ΚΩΔΙΚΟΠΟΙΗΤΗ ΕΝΟΣ ΣΥΣΤΗΜΑΤΟΣ ΠΟΥ ΧΡΗΣΙΜΟΠΟΙΕΙ ΕΝΑ ΠΛΑΙΣΙΟ ΧΑΜΗΛΗΣ ΑΝΑΛΥΣΗΣ MPEG-2 ΚΑΙ ΕΝΑ ΠΛΑΙΣΙΟ ΥΨΗΛΗΣ ΑΝΑΛΥΣΗΣ H.264	61
ΕΙΚΟΝΑ 31 - ΜΠΛΟΚ ΔΙΑΓΡΑΜΜΑ ΛΕΙΤΟΥΡΓΙΑΣ ΕΝΟΣ ΑΠΟΚΩΔΙΚΟΠΟΙΗΤΗ ΕΝΟΣ ΣΥΣΤΗΜΑΤΟΣ ΠΟΥ ΧΡΗΣΙΜΟΠΟΙΕΙ ΕΝΑ ΠΛΑΙΣΙΟ ΧΑΜΗΛΗΣ ΑΝΑΛΥΣΗΣ MPEG-2 ΚΑΙ ΕΝΑ ΠΛΑΙΣΙΟ ΥΨΗΛΗΣ ΑΝΑΛΥΣΗΣ H.264.....	61
ΕΙΚΟΝΑ 32 - ΔΙΑΓΡΑΜΜΑ ΡΥΘΜΟΥ BIT / PSNR ΤΟΥ ΠΡΟΤΕΙΝΟΜΕΝΟΥ ΣΥΝΔΥΑΣΜΟΥ ΠΛΑΙΣΙΩΝ H.264 ΚΑΙ MPEG-2 ΓΙΑ ΤΟ VIDEO FOREMAN	62
ΕΙΚΟΝΑ 33 – ΔΙΑΓΡΑΜΜΑ ΡΥΘΜΟΥ BIT / PSNR ΤΟΥ ΠΡΟΤΕΙΝΟΜΕΝΟΥ ΣΥΝΔΥΑΣΜΟΥ ΠΛΑΙΣΙΩΝ H.264 ΚΑΙ MPEG-2 ΓΙΑ ΤΟ VIDEO COAST GUARD	64
ΕΙΚΟΝΑ 34 - ΤΟ WMV9 ΕΞΑΛΕΙΨΕ ΤΗ ΜΑΥΡΗ ΛΩΡΙΔΑ ΣΤΗΝ ΚΟΡΥΦΗ ΤΟΥ VIDEO MOTHER AND DAUGHTER (A) ΑΡΧΙΚΗ (B) ΣΥΜΠΙΕΣΜΕΝΗ ΜΕ WMV9	71
ΕΙΚΟΝΑ 35 – ΕΞΟΙΚΟΝΟΜΗΣΗ ΡΥΘΜΟΥ BIT / PSNR ΜΕ ΧΡΗΣΗ ΤΟΥ H.264/AVC ΣΕ ΣΥΓΚΡΙΣΗ ΜΕ ΤΟ DIVX 5.1 ΓΙΑ ΟΛΑ ΤΑ VIDEOS.....	73

Κεφάλαιο 4

ΕΙΚΟΝΑ 37 - ΜΕΣΟΣ ΠΟΣΟΣΤΟ ΡΥΘΜΟΥ ΒΙΤ ΓΙΑ ΣΥΓΚΕΚΡΙΜΕΝΗ ΠΟΙΟΤΗΤΑ ΓΙΑ ΤΑ ΒΙΔΕΟΣ ΒΙΔΕΟ-ΔΙΑΣΚΕΨΗΣ ΓΙΑ ΡΥΘΜΙΣΗ ΥΨΗΛΗΣ ΤΑΧΥΤΗΤΑΣ	78
ΕΙΚΟΝΑ 38 - ΜΕΣΟ ΠΟΣΟΣΤΟ ΡΥΘΜΟΥ ΒΙΤ ΓΙΑ ΣΥΓΚΕΚΡΙΜΕΝΗ ΠΟΙΟΤΗΤΑ ΓΙΑ ΤΑ ΒΙΔΕΟΣ ΒΙΔΕΟ-ΔΙΑΣΚΕΨΗΣ ΓΙΑ ΡΥΘΜΙΣΗ ΥΨΗΛΗΣ ΠΟΙΟΤΗΤΑΣ	78
ΕΙΚΟΝΑ 39 – ΔΙΑΓΡΑΜΜΑ ΡΥΘΜΟΥ ΒΙΤ / Y-PSNR ΓΙΑ ΤΟ ΒΙΔΕΟ BATTLE ΓΙΑ ΡΥΘΜΙΣΗ ΥΨΗΛΗΣ ΠΟΙΟΤΗΤΑΣ	80
ΕΙΚΟΝΑ 40 - ΔΙΑΓΡΑΜΜΑ ΡΥΘΜΟΥ ΒΙΤ / Y-PSNR ΓΙΑ ΤΟ ΒΙΔΕΟ LORD OF THE RINGS ΓΙΑ ΡΥΘΜΙΣΗ ΥΨΗΛΗΣ ΠΟΙΟΤΗΤΑΣ .	80
ΕΙΚΟΝΑ 41 – ΔΙΑΓΡΑΜΜΑ ΡΥΘΜΟΥ ΒΙΤ / Y-PSNR ΓΙΑ ΤΟ ΒΙΔΕΟ BATTLE ΓΙΑ ΡΥΘΜΙΣΗ ΥΨΗΛΗΣ ΤΑΧΥΤΗΤΑΣ.....	82
ΕΙΚΟΝΑ 42 - ΔΙΑΓΡΑΜΜΑ ΡΥΘΜΟΥ ΒΙΤ / Y-PSNR ΓΙΑ ΤΟ ΒΙΔΕΟ LORD OF THE RINGS ΓΙΑ ΡΥΘΜΙΣΗ ΥΨΗΛΗΣ ΤΑΧΥΤΗΤΑΣ..	82
ΕΙΚΟΝΑ 43 - ΤΑΧΥΤΗΤΑ ΚΩΔΙΚΟΠΟΙΗΣΗΣ ΓΙΑ ΤΟ ΒΙΔΕΟ BATTLE ΓΙΑ ΡΥΘΜΙΣΗ ΥΨΗΛΗΣ ΠΟΙΟΤΗΤΑΣ	84
ΕΙΚΟΝΑ 44 - ΤΑΧΥΤΗΤΑ ΚΩΔΙΚΟΠΟΙΗΣΗΣ ΓΙΑ ΤΟ ΒΙΔΕΟ BATTLE ΓΙΑ ΡΥΘΜΙΣΗ ΥΨΗΛΗΣ ΠΟΙΟΤΗΤΑΣ, ΕΚΤΟΣ ΤΟΥ CODEC AMD	84
ΕΙΚΟΝΑ 45 - ΤΑΧΥΤΗΤΑ ΚΩΔΙΚΟΠΟΙΗΣΗΣ ΓΙΑ ΤΟ ΒΙΔΕΟ LORD OF THE RINGS ΓΙΑ ΡΥΘΜΙΣΗ ΥΨΗΛΗΣ ΠΟΙΟΤΗΤΑΣ.....	84
ΕΙΚΟΝΑ 46 - ΤΑΧΥΤΗΤΑ ΚΩΔΙΚΟΠΟΙΗΣΗΣ ΓΙΑ ΤΟ ΒΙΔΕΟ LORD OF THE RINGS ΓΙΑ ΡΥΘΜΙΣΗ ΥΨΗΛΗΣ ΠΟΙΟΤΗΤΑΣ, ΕΚΤΟΣ ΤΟΥ CODEC AMD.....	84
ΕΙΚΟΝΑ 47 - ΤΑΧΥΤΗΤΑ ΚΩΔΙΚΟΠΟΙΗΣΗΣ ΓΙΑ ΤΟ ΒΙΔΕΟ BATTLE ΓΙΑ ΡΥΘΜΙΣΗ ΥΨΗΛΗΣ ΤΑΧΥΤΗΤΑΣ.....	85
ΕΙΚΟΝΑ 48 - ΤΑΧΥΤΗΤΑ ΚΩΔΙΚΟΠΟΙΗΣΗΣ ΓΙΑ ΤΟ ΒΙΔΕΟ BATTLE ΓΙΑ ΡΥΘΜΙΣΗ ΥΨΗΛΗΣ ΤΑΧΥΤΗΤΑΣ, ΕΚΤΟΣ ΤΟΥ CODEC AMD	85
ΕΙΚΟΝΑ 49 - ΤΑΧΥΤΗΤΑ ΚΩΔΙΚΟΠΟΙΗΣΗΣ ΓΙΑ ΤΟ ΒΙΔΕΟ LORD OF THE RINGS ΓΙΑ ΡΥΘΜΙΣΗ ΥΨΗΛΗΣ ΤΑΧΥΤΗΤΑΣ	85
ΕΙΚΟΝΑ 50 - ΤΑΧΥΤΗΤΑ ΚΩΔΙΚΟΠΟΙΗΣΗΣ ΓΙΑ ΤΟ ΒΙΔΕΟ LORD OF THE RINGS ΓΙΑ ΡΥΘΜΙΣΗ ΥΨΗΛΗΣ ΤΑΧΥΤΗΤΑΣ, ΕΚΤΟΣ ΤΟΥ CODEC AMD.....	85
ΕΙΚΟΝΑ 51 –ΔΙΑΓΡΑΜΜΑ ΤΑΧΥΤΗΤΑΣ ΚΩΔΙΚΟΠΟΙΗΣΗΣ / Y-PSNR ΓΙΑ ΤΟ ΒΙΔΕΟ BATTLE ΓΙΑ ΡΥΘΜΙΣΗ ΥΨΗΛΗΣ ΠΟΙΟΤΗΤΑΣ	87
ΕΙΚΟΝΑ 52 - ΔΙΑΓΡΑΜΜΑ ΤΑΧΥΤΗΤΑΣ ΚΩΔΙΚΟΠΟΙΗΣΗΣ / Y-PSNR ΓΙΑ ΤΟ ΒΙΔΕΟ LORD OF THE RINGS ΓΙΑ ΡΥΘΜΙΣΗ ΥΨΗΛΗΣ ΠΟΙΟΤΗΤΑΣ	88
ΕΙΚΟΝΑ 53 - ΔΙΑΓΡΑΜΜΑ ΤΑΧΥΤΗΤΑΣ ΚΩΔΙΚΟΠΟΙΗΣΗΣ / Y-PSNR ΓΙΑ ΟΛΑ ΤΑ ΒΙΔΕΟΣ ΤΗΣ ΟΜΑΔΑΣ ΤΑΙΝΙΕΣ ΓΙΑ ΡΥΘΜΙΣΗ ΥΨΗΛΗΣ ΠΟΙΟΤΗΤΑΣ.....	88
ΕΙΚΟΝΑ 54 - ΔΙΑΓΡΑΜΜΑ ΤΑΧΥΤΗΤΑΣ ΚΩΔΙΚΟΠΟΙΗΣΗΣ / Y-PSNR ΓΙΑ ΤΟ ΒΙΔΕΟ BATTLE ΓΙΑ ΡΥΘΜΙΣΗ ΥΨΗΛΗΣ ΤΑΧΥΤΗΤΑΣ	90
ΕΙΚΟΝΑ 55 - ΔΙΑΓΡΑΜΜΑ ΤΑΧΥΤΗΤΑΣ ΚΩΔΙΚΟΠΟΙΗΣΗΣ / Y-PSNR ΓΙΑ ΤΟ ΒΙΔΕΟ LORD OF THE RINGS ΓΙΑ ΡΥΘΜΙΣΗ ΥΨΗΛΗΣ ΤΑΧΥΤΗΤΑΣ.....	90
ΕΙΚΟΝΑ 56 - ΔΙΑΓΡΑΜΜΑ ΤΑΧΥΤΗΤΑΣ ΚΩΔΙΚΟΠΟΙΗΣΗΣ / Y-PSNR ΓΙΑ ΟΛΑ ΤΑ ΒΙΔΕΟΣ ΤΗΣ ΟΜΑΔΑΣ ΤΑΙΝΙΕΣ ΓΙΑ ΡΥΘΜΙΣΗ ΥΨΗΛΗΣ ΤΑΧΥΤΗΤΑΣ	91
ΕΙΚΟΝΑ 57 – ΔΙΑΧΕΙΡΙΣΗ ΡΥΘΜΟΥ ΒΙΤ ΓΙΑ ΤΟ ΒΙΔΕΟ BATTLE ΓΙΑ ΡΥΘΜΙΣΗ ΥΨΗΛΗΣ ΠΟΙΟΤΗΤΑΣ	92
ΕΙΚΟΝΑ 58 - ΔΙΑΧΕΙΡΙΣΗ ΡΥΘΜΟΥ ΒΙΤ ΓΙΑ ΤΟ ΒΙΔΕΟ LORD OF THE RINGS ΓΙΑ ΡΥΘΜΙΣΗ ΥΨΗΛΗΣ ΠΟΙΟΤΗΤΑΣ	92
ΕΙΚΟΝΑ 59 – ΔΙΑΧΕΙΡΙΣΗ ΡΥΘΜΟΥ ΒΙΤ ΓΙΑ ΤΟ ΒΙΔΕΟ BATTLE ΓΙΑ ΡΥΘΜΙΣΗ ΥΨΗΛΗΣ ΤΑΧΥΤΗΤΑΣ.....	92
ΕΙΚΟΝΑ 60 - ΔΙΑΧΕΙΡΙΣΗ ΡΥΘΜΟΥ ΒΙΤ ΓΙΑ ΤΟ ΒΙΔΕΟ LORD OF THE RINGS ΓΙΑ ΡΥΘΜΙΣΗ ΥΨΗΛΗΣ ΤΑΧΥΤΗΤΑΣ.....	92
ΕΙΚΟΝΑ 61 - ΜΕΣΟ ΠΟΣΟΣΤΟ ΡΥΘΜΟΥ ΒΙΤ ΓΙΑ ΣΤΑΘΕΡΗ ΠΟΙΟΤΗΤΑ ΓΙΑ ΟΛΑ ΤΑ ΒΙΔΕΟ ΤΗΣ ΟΜΑΔΑΣ ΤΑΙΝΙΕΣ ΓΙΑ ΡΥΘΜΙΣΗ ΥΨΗΛΗΣ ΤΑΧΥΤΗΤΑΣ	94
ΕΙΚΟΝΑ 62 - ΜΕΣΟ ΠΟΣΟΣΤΟ ΡΥΘΜΟΥ ΒΙΤ ΓΙΑ ΣΤΑΘΕΡΗ ΠΟΙΟΤΗΤΑ ΓΙΑ ΟΛΑ ΤΑ ΒΙΔΕΟΣ ΤΗΣ ΟΜΑΔΑΣ ΤΑΙΝΙΕΣ ΓΙΑ ΡΥΘΜΙΣΗ ΥΨΗΛΗΣ ΠΟΙΟΤΗΤΑΣ	94
ΕΙΚΟΝΑ 63 - ΜΕΣΟ ΠΟΣΟΣΤΟ ΡΥΘΜΟΥ ΒΙΤ ΓΙΑ ΣΤΑΘΕΡΗ ΠΟΙΟΤΗΤΑ ΓΙΑ ΟΛΑ ΤΑ ΒΙΔΕΟΣ ΤΗΣ ΟΜΑΔΑΣ HDTV ΓΙΑ ΡΥΘΜΙΣΗ ΥΨΗΛΗΣ ΤΑΧΥΤΗΤΑΣ	96
ΕΙΚΟΝΑ 64 - ΜΕΣΟ ΠΟΣΟΣΤΟ ΡΥΘΜΟΥ ΒΙΤ ΓΙΑ ΣΤΑΘΕΡΗ ΠΟΙΟΤΗΤΑ ΓΙΑ ΟΛΑ ΤΑ ΒΙΔΕΟΣ ΤΗΣ ΟΜΑΔΑΣ HDTV ΓΙΑ ΡΥΘΜΙΣΗ ΥΨΗΛΗΣ ΠΟΙΟΤΗΤΑΣ.....	96

ΕΙΚΟΝΑ 65 - ΜΕΣΟΣ ΣΧΕΤΙΚΟΣ ΡΥΘΜΟΣ BIT ΓΙΑ ΣΤΑΘΕΡΗ ΠΟΙΟΤΗΤΑ ΓΙΑ ΟΛΑ ΤΑ VIDEOS ΤΗΣ ΟΜΑΔΑΣ VIDEO-ΔΙΑΣΚΕΨΗ, ΓΙΑ ΟΛΕΣ ΤΙΣ ΡΥΘΜΙΣΕΙΣ - YSSIM.....	98
ΕΙΚΟΝΑ 66 - ΜΕΣΟΣ ΣΧΕΤΙΚΟΣ ΧΡΟΝΟΣ ΚΩΔΙΚΟΠΟΙΗΣΗΣ ΓΙΑ ΣΤΑΘΕΡΗ ΠΟΙΟΤΗΤΑ ΓΙΑ ΟΛΑ ΤΑ VIDEOS ΤΗΣ ΟΜΑΔΑΣ VIDEO-ΔΙΑΣΚΕΨΗ, ΓΙΑ ΟΛΕΣ ΤΙΣ ΡΥΘΜΙΣΕΙΣ	98
ΕΙΚΟΝΑ 67 - ΜΕΣΟΣ ΣΧΕΤΙΚΟΣ ΡΥΘΜΟΣ BIT ΓΙΑ ΣΤΑΘΕΡΗ ΠΟΙΟΤΗΤΑ ΓΙΑ ΟΛΑ ΤΑ VIDEOS ΤΗΣ ΟΜΑΔΑΣ ΤΑΙΝΙΕΣ, ΓΙΑ ΟΛΕΣ ΤΙΣ ΡΥΘΜΙΣΕΙΣ - YSSIM.....	100
ΕΙΚΟΝΑ 68 - ΜΕΣΟΣ ΣΧΕΤΙΚΟΣ ΧΡΟΝΟΣ ΚΩΔΙΚΟΠΟΙΗΣΗΣ ΓΙΑ ΣΤΑΘΕΡΗ ΠΟΙΟΤΗΤΑ ΓΙΑ ΟΛΑ ΤΑ VIDEOS ΤΗΣ ΟΜΑΔΑΣ ΤΑΙΝΙΕΣ, ΓΙΑ ΟΛΕΣ ΤΙΣ ΡΥΘΜΙΣΕΙΣ.....	100
ΕΙΚΟΝΑ 69 - ΜΕΣΟΣ ΣΧΕΤΙΚΟΣ ΡΥΘΜΟΣ BIT ΓΙΑ ΣΤΑΘΕΡΗ ΠΟΙΟΤΗΤΑ ΓΙΑ ΟΛΑ ΤΑ VIDEOS ΤΗΣ ΟΜΑΔΑΣ HDTV, ΓΙΑ ΟΛΕΣ ΤΙΣ ΡΥΘΜΙΣΕΙΣ - YSSIM.....	102
ΕΙΚΟΝΑ 70 - ΜΕΣΟΣ ΣΧΕΤΙΚΟΣ ΧΡΟΝΟΣ ΚΩΔΙΚΟΠΟΙΗΣΗΣ ΓΙΑ ΣΤΑΘΕΡΗ ΠΟΙΟΤΗΤΑ ΓΙΑ ΟΛΑ ΤΑ VIDEOS ΤΗΣ ΟΜΑΔΑΣ HDTV, ΓΙΑ ΟΛΕΣ ΤΙΣ ΡΥΘΜΙΣΕΙΣ.....	102
ΕΙΚΟΝΑ 71 - ΜΕΣΟΣ ΣΧΕΤΙΚΟΣ ΡΥΘΜΟΣ BIT ΓΙΑ ΣΤΑΘΕΡΗ ΠΟΙΟΤΗΤΑ ΓΙΑ ΟΛΑ ΤΑ VIDEOS ΟΛΩΝ ΤΩΝ ΟΜΑΔΩΝ, ΓΙΑ ΟΛΕΣ ΤΙΣ ΡΥΘΜΙΣΕΙΣ - YSSIM.....	103
ΕΙΚΟΝΑ 72 - ΜΕΣΟΣ ΣΧΕΤΙΚΟΣ ΧΡΟΝΟΣ ΚΩΔΙΚΟΠΟΙΗΣΗΣ ΓΙΑ ΣΤΑΘΕΡΗ ΠΟΙΟΤΗΤΑ ΓΙΑ ΟΛΑ ΤΑ VIDEOS ΟΛΩΝ ΤΩΝ ΟΜΑΔΩΝ, ΓΙΑ ΟΛΕΣ ΤΙΣ ΡΥΘΜΙΣΕΙΣ.....	103
ΕΙΚΟΝΑ 73 – ΔΙΑΓΡΑΜΜΑ ΡΥΘΜΟΥ BIT / Y-PSNR ΓΙΑ ΤΟ VIDEO SALESMAN, ΓΙΑ ΤΑ CODECS ARTEMIS H.264 ΚΑΙ x264 ΓΙΑ ΡΥΘΜΙΣΗ ΥΨΗΛΗΣ ΤΑΧΥΤΗΤΑΣ	105
ΕΙΚΟΝΑ 74 – ΔΙΑΓΡΑΜΜΑ ΡΥΘΜΟΥ BIT / Y-SSIM ΓΙΑ ΤΟ VIDEO SALESMAN, ΓΙΑ ΤΑ CODECS ARTEMIS H.264 ΚΑΙ x264 ΓΙΑ ΡΥΘΜΙΣΗ ΥΨΗΛΗΣ ΤΑΧΥΤΗΤΑΣ	105
ΕΙΚΟΝΑ 75 – ΔΙΑΓΡΑΜΜΑ ΡΥΘΜΟΥ BIT / Y-SSIM ΓΙΑ ΤΟ VIDEO FOREMAN ΓΙΑ ΡΥΘΜΙΣΗ ΥΨΗΛΗΣ ΠΟΙΟΤΗΤΑΣ	109
ΕΙΚΟΝΑ 76 – ΔΙΑΓΡΑΜΜΑ ΡΥΘΜΟΥ BIT / Y-SSIM ΓΙΑ ΤΟ VIDEO FOREMAN ΓΙΑ ΡΥΘΜΙΣΗ ΥΨΗΛΗΣ ΤΑΧΥΤΗΤΑΣ	109
ΕΙΚΟΝΑ 77 – ΣΧΕΤΙΚΟΣ ΡΥΘΜΟΣ BIT / ΣΧΕΤΙΚΗ ΤΑΧΥΤΗΤΑ ΚΩΔΙΚΟΠΟΙΗΣΗΣ ΓΙΑ ΟΛΑ ΤΑ VIDEOS ΤΗΣ ΟΜΑΔΑΣ VIDEO-ΔΙΑΣΚΕΨΗ ΓΙΑ ΡΥΘΜΙΣΗ ΥΨΗΛΗΣ ΠΟΙΟΤΗΤΑΣ	110
ΕΙΚΟΝΑ 78 – ΣΧΕΤΙΚΟΣ ΡΥΘΜΟΣ BIT / ΣΧΕΤΙΚΗ ΤΑΧΥΤΗΤΑ ΚΩΔΙΚΟΠΟΙΗΣΗΣ ΓΙΑ ΟΛΑ ΤΑ VIDEOS ΤΗΣ ΟΜΑΔΑΣ VIDEO-ΔΙΑΣΚΕΨΗ ΓΙΑ ΡΥΘΜΙΣΗ ΥΨΗΛΗΣ ΤΑΧΥΤΗΤΑΣ	111
ΕΙΚΟΝΑ 79 – ΔΙΑΓΡΑΜΜΑ ΡΥΘΜΟΥ BIT / Y-SSIM ΓΙΑ ΤΟ VIDEO RANCHO ΓΙΑ ΡΥΘΜΙΣΗ ΥΨΗΛΗΣ ΠΟΙΟΤΗΤΑΣ	113
ΕΙΚΟΝΑ 80 - ΔΙΑΓΡΑΜΜΑ ΡΥΘΜΟΥ BIT / Y-SSIM ΓΙΑ ΤΟ VIDEO RANCHO ΓΙΑ ΡΥΘΜΙΣΗ ΥΨΗΛΗΣ ΤΑΧΥΤΗΤΑΣ	114
ΕΙΚΟΝΑ 81 – ΔΙΑΓΡΑΜΜΑ ΔΙΑΧΕΙΡΙΣΗΣ ΡΥΘΜΟΥ BIT ΤΩΝ CODECS DIVX ΚΑΙ x264 ΓΙΑ ΤΟ VIDEO MATRIX.....	115
ΕΙΚΟΝΑ 82 - ΔΙΑΓΡΑΜΜΑ ΔΙΑΧΕΙΡΙΣΗΣ ΡΥΘΜΟΥ BIT ΤΩΝ CODECS DIVX ΚΑΙ x264 ΓΙΑ ΤΟ VIDEO FUTURAMA.....	116
ΕΙΚΟΝΑ 83 – ΔΙΑΓΡΑΜΜΑ ΣΧΕΤΙΚΟΥ ΡΥΘΜΟΥ BIT / ΣΧΕΤΙΚΗΣ ΤΑΧΥΤΗΤΑΣ ΚΩΔΙΚΟΠΟΙΗΣΗΣ ΓΙΑ ΤΟ VIDEO RANCHO ΓΙΑ ΡΥΘΜΙΣΗ ΥΨΗΛΗΣ ΠΟΙΟΤΗΤΑΣ	117
ΕΙΚΟΝΑ 84 - ΔΙΑΓΡΑΜΜΑ ΣΧΕΤΙΚΟΥ ΡΥΘΜΟΥ BIT / ΣΧΕΤΙΚΗΣ ΤΑΧΥΤΗΤΑΣ ΚΩΔΙΚΟΠΟΙΗΣΗΣ ΓΙΑ ΤΟ VIDEO RANCHO ΓΙΑ ΡΥΘΜΙΣΗ ΥΨΗΛΗΣ ΤΑΧΥΤΗΤΑΣ	118
ΕΙΚΟΝΑ 85 – ΔΙΑΓΡΑΜΜΑ ΡΥΘΜΟΥ BIT / ΠΟΙΟΤΗΤΑΣ ΓΙΑ ΤΟ VIDEO CONCERT ΓΙΑ ΡΥΘΜΙΣΗ ΥΨΗΛΗΣ ΠΟΙΟΤΗΤΑΣ.....	120
ΕΙΚΟΝΑ 86 - ΔΙΑΓΡΑΜΜΑ ΠΟΙΟΤΗΤΑΣ ΑΝΑ ΠΛΑΙΣΙΟ ΓΙΑ ΤΟ VIDEO CONCERT ΓΙΑ ΡΥΘΜΙΣΗ ΥΨΗΛΗΣ ΠΟΙΟΤΗΤΑΣ	121
ΕΙΚΟΝΑ 87 – ΔΙΑΓΡΑΜΜΑ ΣΧΕΤΙΚΟΥ ΡΥΘΜΟΥ BIT / ΣΧΕΤΙΚΗΣ ΤΑΧΥΤΗΤΑΣ ΚΩΔΙΚΟΠΟΙΗΣΗΣ ΓΙΑ ΤΟ VIDEO CONCERT ΓΙΑ ΡΥΘΜΙΣΗ ΥΨΗΛΗΣ ΠΟΙΟΤΗΤΑΣ	122
ΕΙΚΟΝΑ 88 – ΜΕΣΟ ΜΕΓΕΘΟΣ VIDEO ΓΙΑ ΣΤΟ ΣΥΝΟΛΟ ΤΟΥ ΠΕΙΡΑΜΑΤΟΣ	124
ΕΙΚΟΝΑ 89 – ΔΙΑΓΡΑΜΜΑ ΡΥΘΜΟΥ BIT / ΠΟΙΟΤΗΤΑΣ ΓΙΑ ΤΟ VIDEO FOREMAN ΓΙΑ ΡΥΘΜΙΣΗ ΥΨΗΛΗΣ ΠΟΙΟΤΗΤΑΣ.....	127
ΕΙΚΟΝΑ 90 – ΔΙΑΓΡΑΜΜΑ ΡΥΘΜΟΥ BIT / ΠΟΙΟΤΗΤΑΣ ΓΙΑ ΤΟ VIDEO ΑΚΙΥΟ ΓΙΑ ΡΥΘΜΙΣΗ ΥΨΗΛΗΣ ΠΟΙΟΤΗΤΑΣ	127
ΕΙΚΟΝΑ 91 – ΔΙΑΓΡΑΜΜΑ ΡΥΘΜΟΥ BIT / ΠΟΙΟΤΗΤΑΣ ΤΩΝ CODECS x264 ΚΑΙ APPLE ΓΙΑ ΤΟ VIDEO MATRIX ΓΙΑ ΡΥΘΜΙΣΗ ΥΨΗΛΗΣ ΠΟΙΟΤΗΤΑΣ.....	127
ΕΙΚΟΝΑ 92 – ΔΙΑΓΡΑΜΜΑ ΡΥΘΜΟΥ BIT / ΠΟΙΟΤΗΤΑΣ ΤΩΝ CODECS x264 ΚΑΙ APPLE ΓΙΑ ΤΟ VIDEO MATRIX ΓΙΑ ΡΥΘΜΙΣΗ ΥΨΗΛΗΣ ΠΟΙΟΤΗΤΑΣ, ΜΕ ΜΕΤΡΗΣΗ ΤΟΥ Y-BI-PSNR.....	127
ΕΙΚΟΝΑ 93 - ΔΙΑΓΡΑΜΜΑ ΡΥΘΜΟΥ BIT RATE / Y-PSNR ΓΙΑ ΤΟ VIDEO BANKOMATDI.....	130
ΕΙΚΟΝΑ 94 - ΔΙΑΓΡΑΜΜΑ ΡΥΘΜΟΥ BIT / DELTA Y-PSNR ΓΙΑ ΤΟ VIDEO BANKOMATDI.....	130
ΕΙΚΟΝΑ 95 – ΔΙΑΓΡΑΜΜΑ ΡΥΘΜΟΥ BIT / U-PSNR ΓΙΑ ΤΟ VIDEO BANKOMATDI.....	130
ΕΙΚΟΝΑ 96 – ΔΙΑΓΡΑΜΜΑ ΡΥΘΜΟΥ BIT / V-PSNR ΓΙΑ ΤΟ VIDEO BANKOMATDI	130

ΕΙΚΟΝΑ 97 – ΔΙΑΓΡΑΜΜΑ ΡΥΘΜΟΥ BIT / Y-PSNR ΓΙΑ ΤΟ VIDEO BATTLE	132
ΕΙΚΟΝΑ 98 - ΔΙΑΓΡΑΜΜΑ ΡΥΘΜΟΥ BIT / DELTA Y-PSNR ΓΙΑ ΤΟ VIDEO BATTLE.....	132
ΕΙΚΟΝΑ 99 - ΔΙΑΓΡΑΜΜΑ ΡΥΘΜΟΥ BIT / U-PSNR ΓΙΑ ΤΟ VIDEO BATTLE	132
ΕΙΚΟΝΑ 100 – ΔΙΑΓΡΑΜΜΑ ΡΥΘΜΟΥ BIT / V-PSNR ΓΙΑ ΤΟ VIDEO BATTLE	132
ΕΙΚΟΝΑ 101 – ΔΙΑΓΡΑΜΜΑ ΡΥΘΜΟΥ BIT / Y-PSNR ΓΙΑ ΤΟ VIDEO Bbc3DI	134
ΕΙΚΟΝΑ 102 – ΔΙΑΓΡΑΜΜΑ ΡΥΘΜΟΥ BIT / DELTA Y-PSNR ΓΙΑ ΤΟ VIDEO Bbc3DI	134
ΕΙΚΟΝΑ 103 – ΔΙΑΓΡΑΜΜΑ ΡΥΘΜΟΥ BIT / U-PSNR ΓΙΑ ΤΟ VIDEO Bbc3DI	134
ΕΙΚΟΝΑ 104 – ΔΙΑΓΡΑΜΜΑ ΡΥΘΜΟΥ BIT / V-PSNR ΓΙΑ ΤΟ VIDEO Bbc3DI.....	134
ΕΙΚΟΝΑ 105 – ΔΙΑΓΡΑΜΜΑ ΡΥΘΜΟΥ BIT / Y-PSNR ΓΙΑ ΤΟ VIDEO FOREMAN	136
ΕΙΚΟΝΑ 106 - ΔΙΑΓΡΑΜΜΑ ΡΥΘΜΟΥBIT / DELTA Y-PSNR ΓΙΑ ΤΟ VIDEO FOREMAN	136
ΕΙΚΟΝΑ 107 – ΔΙΑΓΡΑΜΜΑ ΡΥΘΜΟΥ BIT / U-PSNR ΓΙΑ ΤΟ VIDEO FOREMAN.....	136
ΕΙΚΟΝΑ 108 – ΔΙΑΓΡΑΜΜΑ ΡΥΘΜΟΥ BIT / V-PSNR ΓΙΑ ΤΟ VIDEO FOREMAN	136
ΕΙΚΟΝΑ 109 – ΔΙΑΓΡΑΜΜΑ ΡΥΘΜΟΥ BIT / Y-PSNR ΓΙΑ ΤΟ VIDEO SUSIDI	138
ΕΙΚΟΝΑ 110 – ΔΙΑΓΡΑΜΜΑ ΡΥΘΜΟΥ BIT / DELTA Y-PSNR ΓΙΑ ΤΟ VIDEO SUSIDI.....	138
ΕΙΚΟΝΑ 111 – ΔΙΑΓΡΑΜΜΑ ΡΥΘΜΟΥ BIT / U-PSNR ΓΙΑ ΤΟ VIDEO SUSIDI.....	138
ΕΙΚΟΝΑ 112 – ΔΙΑΓΡΑΜΜΑ ΡΥΘΜΟΥ BIT / V-PSNR ΓΙΑ ΤΟ VIDEO SUSIDI	138
ΕΙΚΟΝΑ 113 – ΔΙΑΓΡΑΜΜΑ ΜΕΤΑΒΟΛΗΣ ΦΩΤΕΙΝΟΤΗΤΑΣ ΓΙΑ ΤΟ VIDEO BANKOMATDI	139
ΕΙΚΟΝΑ 114 - ΔΙΑΓΡΑΜΜΑ ΜΕΤΑΒΟΛΗΣ ΦΩΤΕΙΝΟΤΗΤΑΣ ΓΙΑ ΤΟ VIDEO BATTLE.....	140
ΕΙΚΟΝΑ 115 - ΔΙΑΓΡΑΜΜΑ ΜΕΤΑΒΟΛΗΣ ΦΩΤΕΙΝΟΤΗΤΑΣ ΓΙΑ ΤΟ VIDEO Bbc3DI	141
ΕΙΚΟΝΑ 116 - ΔΙΑΓΡΑΜΜΑ ΜΕΤΑΒΟΛΗΣ ΦΩΤΕΙΝΟΤΗΤΑΣ ΓΙΑ ΤΟ VIDEO FOREMAN.....	142
ΕΙΚΟΝΑ 117 - ΔΙΑΓΡΑΜΜΑ ΜΕΤΑΒΟΛΗΣ ΦΩΤΕΙΝΟΤΗΤΑΣ ΓΙΑ ΤΟ VIDEO SUSIDI.....	143
ΕΙΚΟΝΑ 118 - ΔΙΑΓΡΑΜΜΑ ΔΙΑΧΕΙΡΙΣΗΣ ΡΥΘΜΟΥ BIT ΓΙΑ ΤΟ VIDEO BANKOMATDI.....	144
ΕΙΚΟΝΑ 119 - ΔΙΑΓΡΑΜΜΑ ΔΙΑΧΕΙΡΙΣΗΣ ΡΥΘΜΟΥ BIT ΓΙΑ ΤΟ VIDEO BATTLE	145
ΕΙΚΟΝΑ 120 - ΔΙΑΓΡΑΜΜΑ ΔΙΑΧΕΙΡΙΣΗΣ ΡΥΘΜΟΥ BIT ΓΙΑ ΤΟ VIDEO Bbc3DI.....	146
ΕΙΚΟΝΑ 121 - ΔΙΑΓΡΑΜΜΑ ΔΙΑΧΕΙΡΙΣΗΣ ΡΥΘΜΟΥ BIT ΓΙΑ ΤΟ VIDEO FOREMAN	147
ΕΙΚΟΝΑ 122 - ΔΙΑΓΡΑΜΜΑ ΔΙΑΧΕΙΡΙΣΗΣ ΡΥΘΜΟΥ BIT ΓΙΑ ΤΟ VIDEO SUSIDI	148
ΕΙΚΟΝΑ 123 - PSNR ΑΝΑ ΠΛΑΙΣΙΟ ΓΙΑ ΤΟ VIDEO BANKOMATDI ΓΙΑ ΡΥΘΜΟ BIT 100KBPS	149
ΕΙΚΟΝΑ 124 - PSNR ΑΝΑ ΠΛΑΙΣΙΟ ΓΙΑ ΤΟ VIDEO BANKOMATDI ΓΙΑ ΡΥΘΜΟ BIT 700KBPS	149
ΕΙΚΟΝΑ 125 - PSNR ΑΝΑ ΠΛΑΙΣΙΟ ΓΙΑ ΤΟ VIDEO BANKOMATDI ΓΙΑ ΡΥΘΜΟ BIT 2340KBPS	150
ΕΙΚΟΝΑ 126 - PSNR ΑΝΑ ΠΛΑΙΣΙΟ ΓΙΑ ΤΟ VIDEO BATTLE ΓΙΑ ΡΥΘΜΟ BIT 100KBPS.....	150
ΕΙΚΟΝΑ 127 - PSNR ΑΝΑ ΠΛΑΙΣΙΟ ΓΙΑ ΤΟ VIDEO BATTLE ΓΙΑ ΡΥΘΜΟ BIT 700KBPS.....	151
ΕΙΚΟΝΑ 128 - PSNR ΑΝΑ ΠΛΑΙΣΙΟ ΓΙΑ ΤΟ VIDEO BATTLE ΓΙΑ ΡΥΘΜΟ BIT 2340KBPS.....	151
ΕΙΚΟΝΑ 129 - PSNR ΑΝΑ ΠΛΑΙΣΙΟ ΓΙΑ ΤΟ VIDEO Bbc3DI ΓΙΑ ΡΥΘΜΟ BIT 100KBPS	152
ΕΙΚΟΝΑ 130 - PSNR ΑΝΑ ΠΛΑΙΣΙΟ ΓΙΑ ΤΟ VIDEO Bbc3DI ΓΙΑ ΡΥΘΜΟ BIT 700KBPS	152
ΕΙΚΟΝΑ 131 - PSNR ΑΝΑ ΠΛΑΙΣΙΟ ΓΙΑ ΤΟ VIDEO Bbc3DI ΓΙΑ ΡΥΘΜΟ BIT 2340KBPS	153
ΕΙΚΟΝΑ 132 - PSNR ΑΝΑ ΠΛΑΙΣΙΟ ΓΙΑ ΤΟ VIDEO FOREMAN ΓΙΑ ΡΥΘΜΟ BIT 100KBPS.....	153
ΕΙΚΟΝΑ 133 - PSNR ΑΝΑ ΠΛΑΙΣΙΟ ΓΙΑ ΤΟ VIDEO FOREMAN ΓΙΑ ΡΥΘΜΟ BIT 700KBPS.....	154
ΕΙΚΟΝΑ 134 - PSNR ΑΝΑ ΠΛΑΙΣΙΟ ΓΙΑ ΤΟ VIDEO FOREMAN ΓΙΑ ΡΥΘΜΟ BIT 2340KBPS.....	154
ΕΙΚΟΝΑ 135 - PSNR ΑΝΑ ΠΛΑΙΣΙΟ ΓΙΑ ΤΟ VIDEO SUSIDI ΓΙΑ ΡΥΘΜΟ BIT 100KBPS.....	155
ΕΙΚΟΝΑ 136 - PSNR ΑΝΑ ΠΛΑΙΣΙΟ ΓΙΑ ΤΟ VIDEO SUSIDI ΓΙΑ ΡΥΘΜΟ BIT 700KBPS.....	155
ΕΙΚΟΝΑ 137 - PSNR ΑΝΑ ΠΛΑΙΣΙΟ ΓΙΑ ΤΟ VIDEO SUSIDI ΓΙΑ ΡΥΘΜΟ BIT 2340KBPS.....	156
ΕΙΚΟΝΑ 138 - VIDEO Bbc3DI ΠΛΑΙΣΙΟ 280 ΣΥΜΠΙΕΣΜΕΝΟ ΜΕ ΤΟ CODEC ATEME (ΑΡΙΣΤΕΡΑ) ΚΑΙ DIVX (ΔΕΞΙΑ)	157
ΕΙΚΟΝΑ 139 - VIDEO FOREMAN ΠΛΑΙΣΙΟ 282 ΣΥΜΠΙΕΣΜΕΝΟ ΜΕ ΤΟ CODEC ATEME (ΑΡΙΣΤΕΡΑ) ΚΑΙ DIVX (ΔΕΞΙΑ)	158
ΕΙΚΟΝΑ 140 - VIDEO FOREMAN ΠΛΑΙΣΙΟ 9 ΣΥΜΠΙΕΣΜΕΝΟ ΜΕ ΤΟ CODEC ATEME (ΑΡΙΣΤΕΡΑ) ΚΑΙ DIVX (ΔΕΞΙΑ)	159
ΕΙΚΟΝΑ 141 - ΠΛΑΙΣΙΟ 1 (ΔΕΞΙΑ) ΚΑΙ 100 (ΑΡΙΣΤΕΡΑ) ΤΟΥ VIDEO SALESMAN	162
ΕΙΚΟΝΑ 142 - ΠΛΑΙΣΙΟ 77 (ΑΡΙΣΤΕΡΑ) ΚΑΙ 258 (ΔΕΞΙΑ) ΤΟΥ VIDEO FOREMAN	163
ΕΙΚΟΝΑ 143 - ΠΛΑΙΣΙΟ 1 (ΑΡΙΣΤΕΡΑ) ΚΑΙ 100 (ΔΕΞΙΑ) ΤΟΥ VIDEO NEWS.....	164
ΕΙΚΟΝΑ 144 - ΠΛΑΙΣΙΟ 1 ΤΟΥ VIDEO ΑΚΙΥΟ	165
ΕΙΚΟΝΑ 145 - ΠΛΑΙΣΙΟ 319 ΤΟΥ VIDEO CARPHONE.....	166

ΕΙΚΟΝΑ 146 - ΠΛΑΙΣΙΟ 839 ΤΟΥ VIDEO BATTLE	167
ΕΙΚΟΝΑ 147 - ΠΛΑΙΣΙΟ 650 ΤΟΥ VIDEO SMITH	168
ΕΙΚΟΝΑ 148 - ΠΛΑΙΣΙΟ 160 ΤΟΥ VIDEO ICEAGE	169
ΕΙΚΟΝΑ 149 - ΠΛΑΙΣΙΟ 100 ΤΟΥ VIDEO LORD OF THE RINGS	170
ΕΙΚΟΝΑ 150 - ΠΛΑΙΣΙΟ 570 ΤΟΥ VIDEO RANCHO	171
ΕΙΚΟΝΑ 151 - ΠΛΑΙΣΙΟ 262 ΤΟΥ VIDEO FUTURAMA	172
ΕΙΚΟΝΑ 152 - ΠΛΑΙΣΙΟ 226 ΤΟΥ VIDEO MATRIX (ΟΜΑΔΑ «ΤΑΙΝΙΕΣ»).....	173
ΕΙΚΟΝΑ 153 - ΠΛΑΙΣΙΟ 168 ΤΟΥ VIDEO BANKOMATDI	174
ΕΙΚΟΝΑ 154 - ΠΛΑΙΣΙΟ 185 (ΑΡΙΣΤΕΡΑ) ΚΑΙ 258 (ΔΕΞΙΑ) ΤΟΥ VIDEO Bbc3DI.....	175
ΕΙΚΟΝΑ 155 - ΠΛΑΙΣΙΟ 193 ΤΟΥ VIDEO SUSIDI	176
ΕΙΚΟΝΑ 156 - ΠΛΑΙΣΙΟ 1 ΤΟΥ VIDEO TROY.....	177
ΕΙΚΟΝΑ 157 - ΠΛΑΙΣΙΟ 1 ΤΟΥ VIDEO MATRIX (ΟΜΑΔΑ «HDTV»).....	178
ΕΙΚΟΝΑ 158 - ΠΛΑΙΣΙΟ 128 ΤΟΥ VIDEO CONCERT	179

Κεφάλαιο 5

ΕΙΚΟΝΑ 159 - ΜΠΛΟΚ ΔΙΑΓΡΑΜΜΑ MNRU	183
ΕΙΚΟΝΑ 160 - ΣΧΕΣΗ ΤΟΥ QOP ΜΕ ΤΟ MOS	184

Κεφάλαιο 6

ΕΙΚΟΝΑ 161 - ΤΟΠΟΛΟΓΙΑ ΔΙΚΤΥΟΥ DUMBELL	188
ΕΙΚΟΝΑ 162 – ΓΡΑΜΜΙΚΗ ΤΟΠΟΛΟΓΙΑ ΔΙΚΤΥΟΥ.....	189
ΕΙΚΟΝΑ 163 – ΔΙΑΓΡΑΜΜΑ ΚΑΘΥΣΤΕΡΗΣΗΣ ΓΙΑ ΣΥΝΘΗΚΕΣ ΟΜΟΓΕΝΟΥΣ ΡΟΗΣ ΣΕ ΔΙΚΤΥΟ DUMBELL.....	191
ΕΙΚΟΝΑ 164 - ΔΙΑΓΡΑΜΜΑ JITTER ΓΙΑ ΣΥΝΘΗΚΕΣ ΟΜΟΓΕΝΟΥΣ ΡΟΗΣ ΣΕ ΔΙΚΤΥΟ DUMBELL	191
ΕΙΚΟΝΑ 165 - ΔΙΑΓΡΑΜΜΑ ΚΑΘΥΣΤΕΡΗΣΗΣ ΓΙΑ ΣΥΝΘΗΚΕΣ ΟΜΟΓΕΝΟΥΣ ΡΟΗΣ ΣΕ ΔΙΚΤΥΟ LINEAR.....	191
ΕΙΚΟΝΑ 166 - ΔΙΑΓΡΑΜΜΑ JITTER ΓΙΑ ΣΥΝΘΗΚΕΣ ΟΜΟΓΕΝΟΥΣ ΡΟΗΣ ΣΕ ΔΙΚΤΥΟ LINEAR.....	191
ΕΙΚΟΝΑ 167 - ΔΙΑΓΡΑΜΜΑ ΚΑΘΥΣΤΕΡΗΣΗΣ ΓΙΑ ΣΥΝΘΗΚΕΣ ΕΤΕΡΟΓΕΝΟΥΣ ΡΟΗΣ ΣΕ ΔΙΚΤΥΟ DUMBELL	193
ΕΙΚΟΝΑ 168 - ΔΙΑΓΡΑΜΜΑ JITTER ΓΙΑ ΣΥΝΘΗΚΕΣ ΕΤΕΡΟΓΕΝΟΥΣ ΡΟΗΣ ΣΕ ΔΙΚΤΥΟ DUMBELL	193
ΕΙΚΟΝΑ 169 - MOS ΤΙΜΕΣ ΤΩΝ CODECS G.711, G.721, G.722, G.726, G.729A, MPEG3 ΚΑΙ GSM.....	196
ΕΙΚΟΝΑ 170 - ΔΙΑΓΡΑΜΜΑ MOS - BIT RATE ΤΩΝ ΠΙΟ ΔΙΑΔΕΔΟΜΕΝΩΝ CODECS, IEEE 1996	202
ΕΙΚΟΝΑ 171 - ΔΙΑΓΡΑΜΜΑ SEGSNR / BER ΤΩΝ ΠΙΟ ΔΙΑΔΕΔΟΜΕΝΩΝ CODECS.....	204
ΕΙΚΟΝΑ 172 – ΔΙΑΓΡΑΜΜΑ SEGSNR / SNR ΔΙΑΥΛΟΥ ΜΕ ΧΡΗΣΗ BCH(254,130,18) ΚΩΔΙΚΑ ΚΑΙ BPSK ΠΑΝΩ ΑΠΟ GAUSSIAN ΚΑΝΑΛΙ	206
ΕΙΚΟΝΑ 173 - ΔΙΑΓΡΑΜΜΑ SEGSNR / SNR ΔΙΑΥΛΟΥ ΜΕ ΧΡΗΣΗ BCH(254,130,18) ΚΩΔΙΚΑ ΚΑΙ 4QAM ΠΑΝΩ ΑΠΟ GAUSSIAN ΚΑΝΑΛΙ	206
ΕΙΚΟΝΑ 174 - ΔΙΑΓΡΑΜΜΑ SEGSNR / SNR ΔΙΑΥΛΟΥ ΜΕ ΧΡΗΣΗ BCH(254,130,18) ΚΩΔΙΚΑ ΚΑΙ 16QAM ΠΑΝΩ ΑΠΟ GAUSSIAN ΚΑΝΑΛΙ	206
ΕΙΚΟΝΑ 175 - ΔΙΑΓΡΑΜΜΑ SEGSNR / SNR ΔΙΑΥΛΟΥ ΜΕ ΧΡΗΣΗ BCH(254,130,18) ΚΩΔΙΚΑ ΚΑΙ BPSK ΠΑΝΩ ΑΠΟ RAYLEIGH ΚΑΝΑΛΙ.....	206
ΕΙΚΟΝΑ 176 - ΔΙΑΓΡΑΜΜΑ SEGSNR / SNR ΔΙΑΥΛΟΥ ΜΕ ΧΡΗΣΗ BCH(254,130,18) ΚΩΔΙΚΑ ΚΑΙ 4QAM ΠΑΝΩ ΑΠΟ RAYLEIGH ΚΑΝΑΛΙ.....	207
ΕΙΚΟΝΑ 177 - ΔΙΑΓΡΑΜΜΑ SEGSNR / SNR ΔΙΑΥΛΟΥ ΜΕ ΧΡΗΣΗ BCH(254,130,18) ΚΩΔΙΚΑ ΚΑΙ 16QAM ΠΑΝΩ ΑΠΟ RAYLEIGH ΚΑΝΑΛΙ.....	207
ΕΙΚΟΝΑ 178 – ΔΙΑΓΡΑΜΜΑ ΜΕΙΩΣΗΣ SEGSNR / SNR ΔΙΑΥΛΟΥ ΓΙΑ ΤΟ 13 KBPS RPELTP GSM CODEC, ΜΕ ΧΡΗΣΗ BCH(254,130,18) ΚΩΔΙΚΑ ΚΑΙ BPSK, 4QAM, 16QAM ΠΑΝΩ ΑΠΟ GAUSSIAN ΚΑΙ RAYLEIGH ΚΑΝΑΛΙΑ	209
ΕΙΚΟΝΑ 179 - ΔΙΑΓΡΑΜΜΑ ΜΕΙΩΣΗΣ SEGSNR / SNR ΔΙΑΥΛΟΥ ΓΙΑ ΤΟ 13 KBPS BACKWARD-ADAPTIVE G.728 CODEC, ΜΕ ΧΡΗΣΗ BCH(254,130,18) ΚΩΔΙΚΑ ΚΑΙ BPSK, 4QAM, 16QAM ΠΑΝΩ ΑΠΟ GAUSSIAN ΚΑΙ RAYLEIGH ΚΑΝΑΛΙΑ	209

ΕΙΚΟΝΑ 180 - ΔΙΑΓΡΑΜΜΑ ΜΕΙΩΣΗΣ SEGSNR / SNR ΔΙΑΥΛΟΥ ΓΙΑ ΤΟ 8 ΚΒΡS FORWARD-ADAPTIVE G.729 CODEC, ΜΕ ΧΡΗΣΗ BCH(254,130,18) ΚΩΔΙΚΑ ΚΑΙ BPSK, 4QAM, 16QAM ΠΑΝΩ ΑΠΟ GAUSSIAN ΚΑΙ RAYLEIGH ΚΑΝΑΛΙΑ	209
ΕΙΚΟΝΑ 181 - ΔΙΑΓΡΑΜΜΑ ΜΕΙΩΣΗΣ SEGSNR / SNR ΔΙΑΥΛΟΥ ΓΙΑ ΤΟ 5.3 ΚΒΡS G.723.1 CODEC, ΜΕ ΧΡΗΣΗ BCH(254,130,18) ΚΩΔΙΚΑ ΚΑΙ BPSK, 4QAM, 16QAM ΠΑΝΩ ΑΠΟ GAUSSIAN ΚΑΙ RAYLEIGH ΚΑΝΑΛΙΑ	209
ΕΙΚΟΝΑ 182 - ΔΙΑΓΡΑΜΜΑ MNRU	211
ΕΙΚΟΝΑ 183 - MOS ΤΙΜΕΣ ΤΩΝ CODECS	211
ΕΙΚΟΝΑ 184 - ΔΙΑΓΡΑΜΜΑ ΥΠΟΚΕΙΜΕΝΙΚΟΥ ΚΑΙ ΑΝΤΙΚΕΙΜΕΝΙΚΟΥ MOS	212
ΕΙΚΟΝΑ 185 - MOS ΤΙΜΕΣ ΓΙΑ ΤΥΧΑΙΑ ΑΠΩΛΕΙΑ ΠΑΚΕΤΩΝ	213
ΕΙΚΟΝΑ 186 - MOS ΤΙΜΕΣ ΓΙΑ ΕΚΡΗΣΗ ΑΠΩΛΕΙΑΣ ΠΑΚΕΤΩΝ	213
ΕΙΚΟΝΑ 187 - Η ΕΠΙΠΤΩΣΗ ΤΗΣ ΕΚΡΗΣΗΣ ΑΠΩΛΕΙΑΣ ΠΑΚΕΤΩΝ ΓΙΑ ΤΟ G.711 ΧΩΡΙΣ PLC	214

Παράρτημα Α

ΕΙΚΟΝΑ 188 - ΔΕΙΓΜΑΤΟΛΗΨΙΑ ΗΜΙΤΟΝΟΕΙΔΟΥΣ ΣΗΜΑΤΟΣ	219
ΕΙΚΟΝΑ 189 - ΠΕΠΛΕΓΜΕΝΟ VIDEO (INTERLACED VIDEO)	223
ΕΙΚΟΝΑ 190 - ΣΤΑΔΙΑΚΗ ΣΑΡΩΣΗ (PROGRESSIVE SCAN) (ΜΗ-ΠΕΠΛΕΓΜΕΝΟ VIDEO)	223

Ευρετήριο πινάκων

ΠΙΝΑΚΑΣ 1 - ΔΙΕΘΝΗΣ ΟΡΟΛΟΓΙΑ ΚΑΙ ΑΠΟΔΟΣΗ ΣΤΑ ΕΛΛΗΝΙΚΑ	7
---	---

Κεφάλαιο 2

ΠΙΝΑΚΑΣ 2 - ΠΡΟΦΙΛ ΤΟΥ ΠΡΟΤΥΠΟΥ H.264	24
ΠΙΝΑΚΑΣ 3 - ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΤΩΝ ΠΡΟΤΥΠΩΝ H.263, H.263+ ΚΑΙ H.263++	26
ΠΙΝΑΚΑΣ 4 - ΠΡΟΦΙΛ ΤΟΥ ΠΡΟΤΥΠΟΥ MPEG-2	27

Κεφάλαιο 3

ΠΙΝΑΚΑΣ 5 - ΒΙΔΕΟΣ ΠΟΥ ΧΡΗΣΙΜΟΠΟΙΗΘΗΚΑΝ ΓΙΑ ΤΗ ΣΥΓΚΡΙΣΗ ΤΟΥ MPEG-4 ΚΑΙ ΤΟΥ H.264	34
ΠΙΝΑΚΑΣ 6 - ΡΥΘΙΣΕΙΣ ΤΟΥ ΠΡΟΤΥΠΟΥ H.264 ΓΙΑ ΤΗΝ ΔΙΕΞΑΓΩΓΗ ΤΟΥ ΠΕΙΡΑΜΑΤΟΣ	34
ΠΙΝΑΚΑΣ 7 - PSNR ΚΑΙ ΡΥΘΜΟΙ BIT ΤΟΥ MPEG-4 ΚΑΙ ΤΟΥ H.264 ΓΙΑ ΤΑ ΒΙΔΕΟΣ ΗΛΕΚΤΡΟΝΙΚΗΣ ΤΟΜΟΓΡΑΦΙΑΣ ΚΑΙ ΗΛΕΚΤΡΟΚΑΡΔΙΟΓΡΑΦΗΜΑΤΟΣ	35
ΠΙΝΑΚΑΣ 8 - ΜΕΣΗ ΔΙΑΦΟΡΑ PSNR ΚΑΙ ΡΥΘΜΟΙ BIT ΜΕΤΑΞΥ MPEG-4 ΚΑΙ H.264	35
ΠΙΝΑΚΑΣ 9 - PSNR ΚΑΙ ΡΥΘΜΟΣ BIT ΤΩΝ ΜΕΘΟΔΩΝ H014 ΚΑΙ ΤΗΣ ΠΡΟΤΕΙΝΟΜΕΝΗΣ ΓΙΑ ΤΑ ΒΙΔΕΟΣ ΗΛΕΚΤΡΟΝΙΚΗΣ ΤΟΜΟΓΡΑΦΙΑΣ ΚΑΙ ΗΧΟΚΑΡΔΙΟΓΡΑΦΗΜΑΤΟΣ	38
ΠΙΝΑΚΑΣ 10 - ΜΕΣΗ ΔΙΑΦΟΡΑ PSNR ΚΑΙ ΡΥΘΜΟΙ BIT ΜΕΤΑΞΥ ΤΩΝ ΜΕΘΟΔΩΝ H014 ΚΑΙ ΤΗΣ ΠΡΟΤΕΙΝΟΜΕΝΗΣ	39
ΠΙΝΑΚΑΣ 11 - ΥΠΟΚΕΙΜΕΝΙΚΟ SCORE ΓΙΑ ΤΗ ΜΕΘΟΔΟ ΕΛΕΓΧΟΥ ΤΟΥ ΡΥΘΜΟΥ BIT H014 ΚΑΙ ΤΗΝ ΠΡΟΤΕΙΝΟΜΕΝΗ	40
ΠΙΝΑΚΑΣ 12 - ΜΕΣΟ PSNR ΦΩΤΕΙΝΟΤΗΤΑΣ ΤΟΥ H.264, H.263 ΚΑΙ H.263+	42
ΠΙΝΑΚΑΣ 13 - ΚΕΡΔΟΣ ΚΩΔΙΚΟΠΟΙΗΣΗΣ ΤΟΥ H.264 ΣΕ ΣΧΕΣΗ ΜΕ ΤΟ MPEG-2 ΓΙΑ ΒΙΔΕΟ CCIR	45
ΠΙΝΑΚΑΣ 14 - ΚΕΡΔΟΣ ΚΩΔΙΚΟΠΟΙΗΣΗΣ ΤΟΥ H.264 ΣΕ ΣΧΕΣΗ ΜΕ ΤΟ MPEG-2 ΓΙΑ ΒΙΔΕΟ CIF	45
ΠΙΝΑΚΑΣ 15 - ΚΕΡΔΟΣ ΚΩΔΙΚΟΠΟΙΗΣΗΣ ΤΟΥ H.264 ΣΕ ΣΧΕΣΗ ΜΕ ΤΟ H.263 (ΒΑΣΙΚΟ ΠΡΟΦΙΛ) ΓΙΑ ΒΙΔΕΟ CIF	47
ΠΙΝΑΚΑΣ 16 - ΚΕΡΔΟΣ ΚΩΔΙΚΟΠΟΙΗΣΗΣ ΤΟΥ H.264 ΣΕ ΣΧΕΣΗ ΜΕ ΤΟ H.263 (ΒΑΣΙΚΟ ΠΡΟΦΙΛ) ΓΙΑ ΒΙΔΕΟ QCIF	48
ΠΙΝΑΚΑΣ 17 - ΚΕΡΔΟΣ ΚΩΔΙΚΟΠΟΙΗΣΗΣ ΤΟΥ H.264 ΣΕ ΣΧΕΣΗ ΜΕ ΤΟ H.263 (ΠΡΟΦΙΛ ΥΨΗΛΗΣ ΣΥΜΠΙΕΣΗΣ) ΓΙΑ ΒΙΔΕΟ CIF	50
ΠΙΝΑΚΑΣ 18 - ΚΕΡΔΟΣ ΚΩΔΙΚΟΠΟΙΗΣΗΣ ΤΟΥ H.264 ΣΕ ΣΧΕΣΗ ΜΕ ΤΟ H.263 (ΠΡΟΦΙΛ ΥΨΗΛΗΣ ΣΥΜΠΙΕΣΗΣ) ΓΙΑ ΒΙΔΕΟ QCIF	51
ΠΙΝΑΚΑΣ 19 - ΒΙΔΕΟΣ ΠΟΥ ΧΡΗΣΙΜΟΠΟΙΗΘΗΚΑΝ ΓΙΑ ΤΟΝ ΕΛΕΓΧΟ ΤΟΥ ΡΥΘΜΟΥ ΠΑΡΑΜΟΡΦΩΣΗΣ (RD)	53
ΠΙΝΑΚΑΣ 20 - ΠΟΣΟΣΤΟ ΜΕΙΩΣΗΣ ΤΟΥ ΡΥΘΜΟΥ BIT ΤΟΥ ΠΡΟΤΕΙΝΟΜΕΝΟΥ ΣΥΝΔΥΑΣΜΟΥ ΠΛΑΙΣΙΩΝ H.264 ΚΑΙ MPEG-2 ΕΝΑΝΤΙ ΤΟΥ MPEG-2 ΓΙΑ ΤΟ ΒΙΔΕΟ FOREMAN	63
ΠΙΝΑΚΑΣ 21 - ΠΟΣΟΣΤΟ ΜΕΙΩΣΗΣ ΤΟΥ ΡΥΘΜΟΥ BIT ΤΟΥ ΠΡΟΤΕΙΝΟΜΕΝΟΥ ΣΥΝΔΥΑΣΜΟΥ ΠΛΑΙΣΙΩΝ H.264 ΚΑΙ PEG-2 ΕΝΑΝΤΙ ΤΟΥ MPEG-2 ΓΙΑ ΤΟ ΒΙΔΕΟ COAST GUARD	64
ΠΙΝΑΚΑΣ 22 - ΒΙΔΕΟΣ ΠΟΥ ΧΡΗΣΙΜΟΠΟΙΗΘΗΚΑΝ ΓΙΑ ΤΗ ΣΥΓΚΡΙΣΗ ΤΩΝ H.264, DIVX, WMV, XVID ΚΑΙ MC-EZBC	66
ΠΙΝΑΚΑΣ 23 - ΣΥΝΟΛΙΚΟ ΚΕΡΔΟΣ ΚΩΔΙΚΟΠΟΙΗΣΗΣ ΤΟΥ H.264/AVC ΕΝΑΝΤΙ ΤΩΝ ΥΠΟΛΟΙΠΩΝ CODECS	68
ΠΙΝΑΚΑΣ 24 - ΜΕΤΡΗΣΕΙΣ ΠΟΙΟΤΗΤΑΣ ΤΩΝ H.264/AVC, DIVX, WMV9, XVID ΚΑΙ MC-EZBC	69

Κεφάλαιο 4

ΠΙΝΑΚΑΣ 25 - CODECS ΠΟΥ ΧΡΗΣΙΜΟΠΟΙΗΘΗΚΑΝ ΓΙΑ ΤΗ ΣΥΓΚΡΙΣΗ XVID, MAINCONCEPT H.264, INTEL H.264, x264, AMD H.264 ΚΑΙ ARTEMIS H.264	75
ΠΙΝΑΚΑΣ 26 - ΒΙΔΕΟΣ ΠΟΥ ΧΡΗΣΙΜΟΠΟΙΗΘΗΚΑΝ ΓΙΑ ΤΗ ΣΥΓΚΡΙΣΗ XVID, MAINCONCEPT H.264, INTEL H.264, x264, AMD H.264 ΚΑΙ ARTEMIS H.264	75

ΠΙΝΑΚΑΣ 27 – ΜΕΣΟ ΠΟΣΟΣΤΟ ΡΥΘΜΟΥ ΒΙΤ ΓΙΑ ΣΥΓΚΕΚΡΙΜΕΝΗ ΠΟΙΟΤΗΤΑ ΓΙΑ ΤΑ VIDEOS VIDEO-ΔΙΑΣΚΕΨΗΣ ΓΙΑ ΡΥΘΜΙΣΗ ΥΨΗΛΗΣ ΤΑΧΥΤΗΤΑΣ	77
ΠΙΝΑΚΑΣ 28 - ΜΕΣΟ ΠΟΣΟΣΤΟ ΡΥΘΜΟΥ ΒΙΤ ΓΙΑ ΣΥΓΚΕΚΡΙΜΕΝΗ ΠΟΙΟΤΗΤΑ ΓΙΑ ΤΑ VIDEOS VIDEO-ΔΙΑΣΚΕΨΗΣ ΓΙΑ ΡΥΘΜΙΣΗ ΥΨΗΛΗΣ ΠΟΙΟΤΗΤΑΣ.....	77
ΠΙΝΑΚΑΣ 29 – ΜΕΣΟ ΠΟΣΟΣΤΟ ΡΥΘΜΟΥ ΒΙΤ ΓΙΑ ΣΤΑΘΕΡΗ ΠΟΙΟΤΗΤΑ ΓΙΑ ΟΛΑ ΤΑ VIDEOS ΤΗΣ ΟΜΑΔΑΣ ΤΑΝΙΕΣ ΓΙΑ ΡΥΘΜΙΣΗ ΥΨΗΛΗΣ ΤΑΧΥΤΗΤΑΣ	93
ΠΙΝΑΚΑΣ 30 - ΜΕΣΟ ΠΟΣΟΣΤΟ ΡΥΘΜΟΥ ΒΙΤ ΓΙΑ ΣΤΑΘΕΡΗ ΠΟΙΟΤΗΤΑ ΓΙΑ ΟΛΑ ΤΑ VIDEOS ΤΗΣ ΟΜΑΔΑΣ ΤΑΝΙΕΣ ΓΙΑ ΡΥΘΜΙΣΗ ΥΨΗΛΗΣ ΠΟΙΟΤΗΤΑΣ.....	93
ΠΙΝΑΚΑΣ 31 - ΜΕΣΟ ΠΟΣΟΣΤΟ ΡΥΘΜΟΥ ΒΙΤ ΓΙΑ ΣΤΑΘΕΡΗ ΠΟΙΟΤΗΤΑ ΓΙΑ ΟΛΑ ΤΑ VIDEOS ΤΗΣ ΟΜΑΔΑΣ HDTV ΓΙΑ ΡΥΘΜΙΣΗ ΥΨΗΛΗΣ ΤΑΧΥΤΗΤΑΣ	95
ΠΙΝΑΚΑΣ 32 - ΜΕΣΟ ΠΟΣΟΣΤΟ ΡΥΘΜΟΥ ΒΙΤ ΓΙΑ ΣΤΑΘΕΡΗ ΠΟΙΟΤΗΤΑ ΓΙΑ ΟΛΑ ΤΑ VIDEOS ΤΗΣ ΟΜΑΔΑΣ ΤΑΝΙΕΣ ΓΙΑ ΡΥΘΜΙΣΗ ΥΨΗΛΗΣ ΠΟΙΟΤΗΤΑΣ.....	95
ΠΙΝΑΚΑΣ 33 - CODECS ΠΟΥ ΧΡΗΣΙΜΟΠΟΙΗΘΗΚΑΝ ΓΙΑ ΤΗ ΣΥΓΚΡΙΣΗ DIVX 6.2.5, MAINCONCEPT H.264, INTEL H.264, VSS H.263, x264 ΚΑΙ APPLE H.264	106
ΠΙΝΑΚΑΣ 34 - VIDEOS ΠΟΥ ΧΡΗΣΙΜΟΠΟΙΗΘΗΚΑΝ ΓΙΑ ΤΗ ΣΥΓΚΡΙΣΗ DIVX 6.2.5, MAINCONCEPT H.264, INTEL H.264, VSS H.263, x264 ΚΑΙ APPLE H.264	106
ΠΙΝΑΚΑΣ 35 – ΜΕΣΟ ΜΕΓΕΘΟΣ ΑΡΧΕΙΟΥ ΓΙΑ ΣΤΑΘΕΡΗ ΠΟΙΟΤΗΤΑ ΓΙΑ ΟΛΑ ΤΑ VIDEOS ΤΗΣ ΟΜΑΔΑΣ VIDEO-ΔΙΑΣΚΕΨΗ, ΓΙΑ ΡΥΘΜΙΣΗ ΥΨΗΛΗΣ ΠΟΙΟΤΗΤΑΣ ΚΑΙ ΓΙΑ ΡΥΘΜΟΥΣ ΒΙΤ 100 – 300 ΚΒΡS	112
ΠΙΝΑΚΑΣ 36 - ΜΕΣΟ ΜΕΓΕΘΟΣ ΑΡΧΕΙΟΥ ΓΙΑ ΣΤΑΘΕΡΗ ΠΟΙΟΤΗΤΑ ΓΙΑ ΟΛΑ ΤΑ VIDEOS ΤΗΣ ΟΜΑΔΑΣ VIDEO-ΔΙΑΣΚΕΨΗ, ΓΙΑ ΡΥΘΜΙΣΗ ΥΨΗΛΗΣ ΤΑΧΥΤΗΤΑΣ ΚΑΙ ΓΙΑ ΡΥΘΜΟΥΣ ΒΙΤ 100 – 300 ΚΒΡS.....	112
ΠΙΝΑΚΑΣ 37 - ΜΕΣΟ ΜΕΓΕΘΟΣ ΑΡΧΕΙΟΥ ΓΙΑ ΣΤΑΘΕΡΗ ΠΟΙΟΤΗΤΑ ΓΙΑ ΟΛΑ ΤΑ VIDEOS ΤΗΣ ΟΜΑΔΑΣ ΤΑΝΙΕΣ, ΓΙΑ ΡΥΘΜΙΣΗ ΥΨΗΛΗΣ ΤΑΧΥΤΗΤΑΣ ΚΑΙ ΓΙΑ ΡΥΘΜΟΥΣ ΒΙΤ 600-1800ΚΒΡS	119
ΠΙΝΑΚΑΣ 38 - ΜΕΣΟ ΜΕΓΕΘΟΣ ΑΡΧΕΙΟΥ ΓΙΑ ΣΤΑΘΕΡΗ ΠΟΙΟΤΗΤΑ ΓΙΑ ΟΛΑ ΤΑ VIDEOS ΤΗΣ ΟΜΑΔΑΣ ΤΑΝΙΕΣ, ΓΙΑ ΡΥΘΜΙΣΗ ΥΨΗΛΗΣ ΠΟΙΟΤΗΤΑΣ ΚΑΙ ΓΙΑ ΡΥΘΜΟΥΣ ΒΙΤ 600-1800ΚΒΡS	119
ΠΙΝΑΚΑΣ 39 - ΜΕΣΟ ΜΕΓΕΘΟΣ ΑΡΧΕΙΟΥ ΓΙΑ ΣΤΑΘΕΡΗ ΠΟΙΟΤΗΤΑ ΓΙΑ ΟΛΑ ΤΑ VIDEOS ΤΗΣ ΟΜΑΔΑΣ ΤΑΝΙΕΣ, ΓΙΑ ΡΥΘΜΙΣΗ ΥΨΗΛΗΣ ΠΟΙΟΤΗΤΑΣ - 2 ΠΕΡΑΣΜΑΤΩΝ ΚΑΙ ΓΙΑ ΡΥΘΜΟΥΣ ΒΙΤ 600-1800ΚΒΡS	119
ΠΙΝΑΚΑΣ 40 - ΜΕΣΟ ΜΕΓΕΘΟΣ ΑΡΧΕΙΟΥ ΓΙΑ ΣΤΑΘΕΡΗ ΠΟΙΟΤΗΤΑ ΓΙΑ ΟΛΑ ΤΑ VIDEOS ΤΗΣ ΟΜΑΔΑΣ HDTV, ΓΙΑ ΡΥΘΜΟΥΣ ΒΙΤ 2-10ΜΒΡS	123
ΠΙΝΑΚΑΣ 41 – ΑΠΟΤΕΛΕΣΜΑΤΑ ΤΟΥ 2006 ΓΙΑ ΤΟ ΜΕΣΟ ΜΕΓΕΘΟΣ ΑΡΧΕΙΟΥ ΓΙΑ ΣΤΑΘΕΡΗ ΠΟΙΟΤΗΤΑ	124
ΠΙΝΑΚΑΣ 42 - ΑΠΟΤΕΛΕΣΜΑΤΑ ΤΟΥ 2006 ΓΙΑ ΤΗ ΜΕΣΗ ΤΑΧΥΤΗΤΑ ΚΩΔΙΚΟΠΟΙΗΣΗΣ.....	125
ΠΙΝΑΚΑΣ 43 - CODECS ΠΟΥ ΧΡΗΣΙΜΟΠΟΙΗΘΗΚΑΝ ΓΙΑ ΤΗ ΣΥΓΚΡΙΣΗ ΤΩΝ MPEGABLE AVC, MOONLIGHT H.264, MAINCONCEPT H.264, FRAUNHOFER IIS, ATEME MPEG-4 AVC / H.264, VIDEOSOFT H.264 ΚΑΙ DIVX PRO 5.1.1	128
ΠΙΝΑΚΑΣ 44 - VIDEOS ΠΟΥ ΧΡΗΣΙΜΟΠΟΙΗΘΗΚΑΝ ΓΙΑ ΤΗ ΣΥΓΚΡΙΣΗ ΤΩΝ MPEGABLE AVC, MOONLIGHT H.264, MAINCONCEPT H.264, FRAUNHOFER IIS, ATEME MPEG-4 AVC / H.264, VIDEOSOFT H.264 ΚΑΙ DIVX PRO 5.1.1	128
ΠΙΝΑΚΑΣ 45 - ΑΠΟΤΕΛΕΣΜΑΤΑ ΤΗΣ ΑΝΕΠΙΣΗΜΗΣ ΜΕΘΟΔΟΥ ΓΙΑ ΤΟ Y-PSNR	160
ΠΙΝΑΚΑΣ 46 - ΑΠΟΤΕΛΕΣΜΑΤΑ ΤΗΣ ΑΝΕΠΙΣΗΜΗΣ ΜΕΘΟΔΟΥ ΓΙΑ ΤΟ U-PSNR ΚΑΙ V-PSNR.....	160
ΠΙΝΑΚΑΣ 47 – ΓΕΝΙΚΑ ΑΠΟΤΕΛΕΣΜΑΤΑ ΤΗΣ ΑΝΕΠΙΣΗΜΗΣ ΜΕΘΟΔΟΥ	161
ΠΙΝΑΚΑΣ 48 – ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΤΟΥ VIDEO SALESMAN.....	162
ΠΙΝΑΚΑΣ 49 - ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΤΟΥ VIDEO FOREMAN	163
ΠΙΝΑΚΑΣ 50 - ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΤΟΥ VIDEO NEWS.....	164
ΠΙΝΑΚΑΣ 51 - ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΤΟΥ VIDEO AKIYO	165
ΠΙΝΑΚΑΣ 52 - ΧΑΡΑΚΤΗΡΙΣΤΙΚΗ ΤΟΥ VIDEO CARPHONE.....	166
ΠΙΝΑΚΑΣ 53 - ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΤΟΥ VIDEO BATTLE	167
ΠΙΝΑΚΑΣ 54 - ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΤΟΥ VIDEO SMITH	168
ΠΙΝΑΚΑΣ 55 - ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΤΟΥ VIDEO ICEAGE	169
ΠΙΝΑΚΑΣ 56 - ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΤΟΥ VIDEO LORD OF THE RINGS	170
ΠΙΝΑΚΑΣ 57 - ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΤΟΥ VIDEO RANCHO	171

ΠΙΝΑΚΑΣ 58 - ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΤΟΥ VIDEO FUTURAMA	172
ΠΙΝΑΚΑΣ 59 - ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΤΟΥ VIDEO MATRIX (ΟΜΑΔΑ «ΤΑΙΝΙΕΣ»)	173
ΠΙΝΑΚΑΣ 60 - ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΤΟΥ VIDEO BANKOMATDI	174
ΠΙΝΑΚΑΣ 61 - ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΤΟΥ VIDEO BBC3DI.....	175
ΠΙΝΑΚΑΣ 62 - ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΤΟΥ VIDEO SUSIDI	176
ΠΙΝΑΚΑΣ 63 - ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΤΟΥ VIDEO TROY.....	177
ΠΙΝΑΚΑΣ 64 - ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΤΟΥ VIDEO MATRIX (ΟΜΑΔΑ «HDTV»)	178
ΠΙΝΑΚΑΣ 65 - ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΤΟΥ VIDEO CONCERT	179

Κεφάλαιο 6

ΠΙΝΑΚΑΣ 66 - ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΤΩΝ CODECS ΠΟΥ ΧΡΗΣΙΜΟΠΟΙΗΘΗΚΑΝ ΓΙΑ ΤΗ ΣΥΓΚΡΙΣΗ ΤΩΝ G.711, G.723.1, G.726, G.729 ΚΑΙ G.729A	190
ΠΙΝΑΚΑΣ 67 - MOS ΤΙΜΕΣ ΤΩΝ CODECS G.711, G.721, G.722, G.726, G.729A, MPEG3 ΚΑΙ GSM	195
ΠΙΝΑΚΑΣ 68 - RTT ΓΙΑ ΤΟΥΣ ΕΞΥΠΗΡΕΤΗΤΕΣ	197
ΠΙΝΑΚΑΣ 69 - MOS ΤΙΜΕΣ ΚΑΙ ΠΟΣΟΣΤΑ ΑΠΩΛΕΙΩΝ ΤΩΝ CODECS G.711 ΚΑΙ G.728	198
ΠΙΝΑΚΑΣ 70 – ΕΛΑΧΙΣΤΟ ΑΠΑΙΤΟΥΜΕΝΟ SNR ΔΙΑΛΟΥ ΓΙΑ ΤΗ ΔΙΑΤΗΡΗΣΗ ΤΗΣ ΜΕΙΩΣΗΣ ΤΟΥ SEGSNR ΚΑΤΩ ΑΠΟ 1dB ΜΕ ΧΡΗΣΗ BCH(254,130,18) ΚΩΔΙΚΑ ΚΑΙ BPSK, 4QAM, 16QAM ΠΑΝΩ ΑΠΟ GAUSSIAN ΚΑΙ RAYLEIGH ΚΑΝΑΛΙΑ	208
ΠΙΝΑΚΑΣ 71 - ΠΑΡΑΜΕΤΡΟΙ ΤΟΥ ΠΕΙΡΑΜΑΤΟΣ ΣΥΓΚΡΙΣΗΣ ΤΩΝ G.711, G.729 ΚΑΙ AMR	210
ΠΙΝΑΚΑΣ 72 - ΣΥΝΤΕΛΕΣΤΗΣ ΣΥΣΧΕΤΙΣΗΣ (ρ) ΚΑΙ ΜΕΣΟ ΤΕΤΡΑΓΩΝΙΚΟ ΣΦΑΛΜΑ (Δ) ΤΩΝ ΑΝΤΙΚΕΙΜΕΝΙΚΩΝ ΜΕΘΟΔΩΝ ...	212

Κεφάλαιο 7

ΠΙΝΑΚΑΣ 73 - ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΔΗΜΟΦΙΛΕΣΤΕΡΩΝ CODECS ΗΧΟΥ	217
--	-----

Παράρτημα Α

ΠΙΝΑΚΑΣ 74 - Η ΚΛΙΜΑΚΑ MOS.....	222
---------------------------------	-----

Βιβλιογραφία

- [1] Δ. Κουτσούρης, Χ. Παυλόπουλος, Α. Πρέντζα «Εισαγωγή στη βιοϊατρική τεχνολογία και ανάλυση ιατρικών σημάτων», εκδόσεις Τζιόλα, 2003
- [2] Simon Haykin, Ε.Δ. Συκάς, Μ.Ε. Θεολόγου «Συστήματα Επικοινωνίας», εκδόσεις Παπασωτηρίου, 1995
- [3] Ιάκβος Στ. Βενιέρης «Δίκτυα Ευρείας Ζώνης», 2^η Έκδοση, εκδόσεις Τζιόλα, 2007
- [4] Fred Halsall, «Multimedia Communications, Applications, Networks, Protocols and Standards», Addison - Wesley 2001
- [5] Hongtao Yu, Zhiping Lin, Feng Pan “Applications and Improvement of H.264 in Medical Video Compression” (IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS—I: REGULAR PAPERS, VOL. 52, NO. 12, DECEMBER 2005)
- [6] M. Q. Jiang “M. Q. Jiang, X. Q. Yi, and N. Ling, “Improved frame-layer rate control for H.264 using MAD ratio,” in Proc. 2004 IEEE Int. Symp. Circuits Syst., vol. III, Vancouver, BC, Canada, May 2004, pp. 813–816”
- [7] C. J. Tsai, C. W. Tang, C. H. Chen, and Y. H. Yu, “Adaptive rate-distortion optimization using perceptual hints,” in Proc. 2004 IEEE Int. Conf. Multimedia and Expo, Taipei, China, Jun. 2004, pp. 667–670
- [8] M. M. Ghandi and M. Ghanbari, “A Lagrangian optimization rate control algorithm for the H.264/AVC encoder,” in Proc. 2004 IEEE Int. Conf. Image Processing, Singapore, Oct. 2004, pp. 123–126
- [9] Gulistan RAJA, Muhammad J. MIRZA, Department of Electrical Engineering, University of Engineering & Technology, Tanila, Pakistan «Performance Comparison of Advanced Video Coding H.264 Standard with Baseline H.263 and H.263+ Standards», International Symposium on Communications and Information Technologic, 2004
- [10] Nejat Kamaci, Yucel Altunbasak «Performance Comparison of the Emerging H.264 video coding standard with the existing standards», Center for Signal and Image Processing Georgia Institute of Technology, Atlanta, GA, USA, IEEE2003
- [11] Detlev Marpea, Valeri Georgeb, Hans L. Cyconb, Kai U. Barthelb «Performance evaluation of Motion-JPEG2000 in comparison with H.264/AVC operated in pure intra coding mode», Fraunhofer-Institute for Telecommunications, Heinrich-Hertz-Institute (HHI), Fachhochschule fur Technik und Wirtschaft (FHTW), University of Applied Sciences
- [12] D. Storey, P. Nasiopoulos «Combining H.264 and MPEG2 for multi-resolution video applications, The University of British Columbia, Department of Electrical and Computer Engineering», IEEE 2004
- [13] R. Dugad and N. Ahuja “A scheme for spatial scalability using non-scalable encoders,”, IEEE Transactions on Circuits and Systems for Video Technology, vol. 13, no. 10, pp. 993-99, Oct 2003
- [14] Peter Lambert, Wesley De Neve, Philippe De Neve, Ingrid Moerman, Piet Demeester, Rik Van de Walle «Rate-Distortion Performance of H.264/AVC Compared to State-of-the-Art Video Codecs», IEEE Transactions on circuits and systems for video technology, Vol. 16, No. 1, January 2006
- [15] «MPEG-4 AVC/H.264 Video Codecs Comparison» Short version of report, Project head: Dr. Dmitriy Vatolin, Measurements, analysis: Dmitriy Kulikov, Alexander Parshin, December 2007, CS MSU Graphics&Media Lab, Video Group, URL: <http://www.compression.ru/video/>
- [16] «MPEG-4 AVC/H.264 Video Codecs Comparison» Short version of report, Project head: Dr. Dmitriy Vatolin, Measurements, analysis: Dmitriy Kulikov, Alexander Parshin, Translation: Artem Titarenko, Verification: Maxim Smirnov, November 2006, CS MSU Graphics&Media Lab, Video Group, URL:<http://www.compression.ru/video/>
- [17] «MPEG-4 AVC/H.264 Video Codecs Comparison» Short version of report, Project

- head: Dr. Dmitriy Vatolin, Measuring, refinement: Dmitriy Kulikov, Alexander Parshin, Translation: Daria Kalinkina, Verification: Stass Soldatov, January 2005, CS MSU Graphics&Media Lab, Video Group, URL:<http://www.compression.ru/video/>
- [18] Jogesh. K. Muppala, Terdsak Banchedvanich, Anurag Tyagi, « VoIP Performance on Differentiated Services Enabled Network», Department of Computer Science, The Hong Kong University of Science and Technology, IEEE 2000
 - [19] J Light, A.Bhuvaneshwari, «Performance Analysis of Audio Codecs over Real-Time Transmission Protocol (RTP) for Voice Services over Internet Protocol», IEEE 2004
 - [20] Lopamudra Roychoudhuri, Ehab Al-Shaer, Hazem Hamed, Gregory B. Brewster, « Audio Transmission over the Internet: Experiments and Observations», School of Computer Science, Telecommunications and Information Systems DePaul University, IEEE 2003
 - [21] L. Hanzo, «Voice Compression and Communications: Principles and Applications for Fixed and Wireless Channels», Department of Electronics and Computer Science, University of Southampton, UK, 1999
 - [22] R. V. Cox and P. Kroon, "Low bit-rate speech coders for multimedia communications," IEEE Comms. Mag., pp. 34{41, December 1996
 - [23] R. V. Cox, "Speech coding and synthesis," in Speech coding standards (W. Kleijn and K. Paliwal, eds.), ch. 2, pp. 49{78, Elsevier, 1995
 - [24] Lijing Ding , Ayman Radwan, Mohamed Samy El-Hennawey, Rafik A. Goubrant, « Performance Study of Objective Voice Quality Measures in VoIP», IEEE 2007
 - [25] T. Wiegand et al., "Rate-Constrained Coder Control and Comparison of Video Coding Standards", IEEE Trans. on Circ. and Sys. for Video Technology, Vol. 13, No. 7, pp. 688–703, July 2003)
 - [26] D. Marpe, G. Blattermann, J. Ricke, and P. Maaß, "A Two-Layered Wavelet-Based Algorithm for Efficient Lossless and Lossy Image Compression", IEEE Trans. on Circ. and Sys. for Video Technology, Vol. 10, No. 7, pp. 1094–1102, Oct. 2000
 - [27] OzVoIP.com - The VoIP Specialists! URL: <http://ozvoip.com> (September 2008)
 - [28] Wikipedia URL: <http://www.wikipedia.org> (September 2008)
 - [29] Institute of Electrical and Electronics Engineers (IEEE) URL: <http://www.ieee.org> (September 2008)
 - [30] International Telecommunication Union (ITU) URL: <http://www.itu.int> (September 2008)
 - [31] Association for Computing Machinery URL: <http://www.acm.org> (September 2008)