Τεχνολογικό Εκπαιδευτικό Ίδρυμα Κρήτης



ΤΕΧΝΟΛΟΓΙΚΟ
ΕΚΠΑΙΔΕΥΤΙΚΟ
ΙΔΡΥΜΑ ΚΡΗΤΗΣ

Σχολή Τεχνολογικών Εφαρμογών
Τμήμα Εφαρμοσμένης Πληροφορικής & Πολυμέσων



ΤΜΗΜΑ
ΕΦΑΡΜΟΣΜΕΝΗΣ
ΠΛΗΡΟΦΟΡΙΚΗΣ
& ΠΟΛΥΜΕΣΩΝ

Πτυχιακή εργασία

ΤΙΤΛΟΣ ΠΤΥΧΙΑΚΗΣ : "ΜΕΛΕΤΗ ΚΑΙ ΣΥΓΚΡΙΣΗ ΜΕΘΟΔΩΝ ΕΚΤΙΜΗΣΗΣ ΥΠΟΚΕΙΜΕΝΙΚΗΣ ΠΟΙΟΤΗΤΑΣ ΤΗΛΕΟΠΤΙΚΗΣ ΕΙΚΟΝΑΣ"

ΜΑΡΙΑ ΣΙΣΜΑΝΙΔΟΥ Α.Μ 1725

Επιβλέπων καθηγητής: Κος Γαρδίκης Γεώργιος

Ευχαριστίες

Με την ολοκλήρωση της πτυχιακής μου εργασίας θα ήθελα να ευχαριστήσω τον επιβλέποντα καθηγητή μου κύριο Γαρδίκη Γεώργιο για την ευκαιρία που μου έδωσε να ασχοληθώ με το συγκεκριμένο θέμα καθώς και για την καθοδήγηση του.

Ένα μεγάλο ευχαριστώ στους γονείς μου, τόσο για την υλική, όσο και για τη ψυχική στήριξη που μου προσέφεραν όλο αυτό το διάστημα αλλά και για την υποστήριξη που μου δείχνουν σε όλη μου την πορεία μέχρι σήμερα.

Abstract

Digital television technology slowly replaces its predecessor (analog television) in every greek county . The aim of this transition is better and faster signal transmission . Yet again picture errors are still possible either from signal errors or compression distortion.

In this thesis is described and developed methods of objective and subjective evaluation of digital television pictures. At first there is a reference to an audio-visual content (video) characteristics and advantages of both formats , digital and analog . Then the structure and basic characteristics of the MPEG (digital signal compression) are described .In the last chapter there is a subjective quality assessment from a sample audience, but also there is a description of the algorithm(PSNR) which developed in Matlab programming environment and provides an objective image quality assessment .

Besides the development of the PSNR algorithm, the code which used for the deformation of the video used is provided (coded in maltab programming environment). In order to be more accurate , the rating of video quality based on what the viewer can see and not on the signal obtained from the receiver (TV) ,and this is because the perceived quality depends more on the video characteristics than its transmission parameters

Περίληψη

Η Ψηφιακή τηλεόραση αντικαθιστά την αναλογική σιγά σιγά και σε όλους τους Νομούς της Ελλάδας. Στόχος της είναι η καλύτερη και γρηγορότερη μετάδοση σήματος. Παρόλα αυτά όμως η παρουσία λαθών στην εικόνα είναι ακόμα εφικτή , είτε από λάθη του σήματος , είτε από παραμόρφωση είτε από συμπίεση.

Στην παρούσα πτυχιακή εργασία περιγράφονται και αναπτύσσονται τρόποι που μας δείχνουν την υποκειμενική και αντικειμενική αξιολόγηση της ψηφιακής τηλεοπτικής εικόνας. Αρχικά γίνεται αναφορά για ένα οπτικοακουστικό περιεχόμενο σε αναλογική και σε ψηφιακή μορφή αναφέροντας τα χαρακτηριστικά και τα πλεονεκτήματα και των δύο. Στην συνέχεια αναφέρεται ο τρόπος συμπίεσης ενός ψηφιακού οπτικοακουστικού σήματος με συμπιεστή MPEG , η ανατομία του και τα βασικά χαρακτηριστικά του. Στο τελευταίο κεφαλαίο έχουμε μια υποκειμενική εκτίμηση ποιότητας από κάποιο σύνολο θεατών , αλλά και περιγραφεί και αναπτύξει του αλγόριθμου PSNR σε περιβάλλον MATLAB όπου μας δίνει την αντικειμενική ποιότητα της ψηφιακής τηλεοπτικής εικόνας. Στα πλαίσια της πτυχιακής εργασίας αναπτύχτηκε εκτός από τον κώδικα του αλγόριθμου PSNR και κώδικας πάλι σε περιβάλλον MATLAB για την παραμόρφωση των video που χρησιμοποιήθηκαν.

Προκείμενου να είναι πιο ακριβής και πιο κοντά στην πραγματικότητα, πρέπει η βαθμολόγηση στην ποιότητα να γίνεται βάση του video που παρακολουθεί ο θεατής και όχι το σήμα που λαμβάνεται από τον δέκτη. Και αυτό επειδή η αντιληπτή ποιότητα εξαρτάται από τα χαρακτηριστικά του video και όχι τους παραμέτρους μετάδοσης του.

Περιεχόμενα

Ευχαριστίες....3

Abstract....4

Περίληψη....5

Ακρωνύμια....8

Εικόνες - Πίνακες....10

Κεφάλαιο 1 - Τηλεοπτική εικόνα....13

1.Εισαγωγή....13

1.1.Αναλογικό video....14

1.1.2.Βασικά μεγέθη αναλογικού video....16

1.1.3.Συστήματα αναλογικού video....16

1.1.4.Μειονεκτήματα/Πλεονεκτήματα αναλογικού video....17

1.2.Ψηφιακό video....17

1.2.1.Πρότυπα ψηφιακού video....18

1.2.2.Πλεονεκτήματα ψηφιακού video....19

Κεφάλαιο 2 - Συμπίεση....21

2.Συμπίεση....21

2.1.Λόγος συμπίεσης – Ρυθμός συμπίεσης....22

2.2.Συμπίεση Δεδομένων....22

2.2.1.Παράγοντες που σχετίζονται με την συμπίεση....23

2.3.Γενικές αρχές συμπίεσης video....25

2.3.1.Ποιοτική ανοχή....25

2.3.2.Πλεονασμός (Redundancy)....26

2.4.Δειγματοληψία....27

2.5.Κβαντοποίηση....27

2.6.Κωδικοποίηση....28

2.7.Μετασχηματισμός DCT Coding....30

2.8.Τμηματική Πρόβλεψη Κίνησης....32

2.9.Ακίνητες εικόνες καρέ....33

Κεφαλαίο 3 - Το πρότυπο συμπίεσης MPEG....37

3.Εισαγωγή....37

3.1.Ανατομία σήματος MPEG....38

3.2.Δύο βασικούς τρόπους συμπίεσης MPEG....40

3.2.1.Τα τρία είδη των frames στο MPEG 42
3.2.2.Κομμάτια Slices 47
3.3.Τα βήματα της συμπίεσης κατά το MPEG 49
3.4.Το πρότυπο κωδικοποίησης MPEG – 1 50
3.5.Το πρότυπο κωδικοποίησης MPEG-2 50
3.5.1.Οργάνωση προτύπου 51
3.5.2.Επίπεδα και προφίλ του προτύπου MPEG-2 52
3.5.3.Ατέλειες του MPEG-2 54
3.6.Το πρότυπο κωδικοποίησης MPEG-4 56
3.6.1.Αποδοχή του MPEG-4 59
3.6.2.Προφίλ και επίπεδα του MPEG-4 59
3.6.3.Τα βασικά μέρη του MPEG-4 60
3.6.4.Πλεονεκτήματα του MPEG-4 60
Κεφάλαιο 4 - Αλλοίωση video με χρήση Matlab 62
4.Εισαγωγή 62
4.1.Δομή του κώδικα 62
4.2.Παράδειγμα στο κωδικοποιημένο video 63
4.3.Το μέγεθος κλίμακας στην ποιότητα εικόνας 66
Κεφάλαιο 5 - Εκτίμηση της ποιότητας εικόνας video 69
5.Εισαγωγή 69
5.1.Υποκειμενικές Μέθοδοι 69
5.1.1.Πειραματικό μέρος υποκειμενικών μεθόδων 72
5.1.2.Επιλογή video 74
5.1.3.Διαμόρφωση ερωτηματολογίου 80
5.1.4.Διεξαγωγή/στατιστικά ερωτηματολογίου 83
5.2.Αντικειμενικές Μέθοδοι 92
5.2.1.Peak Signal to Noise Ratio (PSNR) 93
5.2.2.Πειραματικό μέρος 96
Κεφάλαιο - 6 Κώδικας 115
6.1.Κώδικας κεφαλαίο 4 115
6.2.Κώδικας κεφάλαιο 5 122
Λεξιλόγιο - Βιβλιογραφία

Ακρωνύμια

A

AE - Absolute Error
ASCII - American Standard Code for Information Interchange

C

CCIR - Consultative Committee for International Radio-communications
CD - Compact Disc
CIF - Common Interchange Format
CRT - Cathode ray tube

D

Db - decibel
DCT - Discrete cosine transform
DTV - Digital television
DV - Digital Video
DVB - Digital Video Broadcasting
DVD - Digital Video Disc ή Digital Versatile Disc

E

EBU - Euro pan broadcasting union

F

Fps - Frame per second
Fpp - Frame per pixel

G

GOP - Group of picture

H

HDTV - High Definition Television
Hz - Hertz

I

IDCT - Inverse Discrete Cosine Transform
ISDN -Integrated Services Digital Network
ITU -International Telecommunication Union
ITU-TS - International Telecommunication Union
ITU-RS - International Telecommunication Union

J

JPEG - Joint Photographic Experts Group

L

LSB - Least Significant Bit
LZW - Lempel, Ziv & Welsh

M

Mpeg - moving picture expect group

N

NTSC - National Television System Committee

P

PAL - Phase Alternating Line
PNG - Portable Network Graphics

Q

QCIF - Quarter CIF

R

RLE - Run Length Encoding

S

SEC AM - Sequential Couleur avec Memoire
SIF - Source Intermediate Format
SNR - Signal-to-noise ratio
SSCQE - Single Stimulus Continue Quality Evaluation

T

TIFF - Tagged Image File Format

V

VCD - Compact Disc digital video
VCR - Video Cassette Recording
VHS - Video Home System

Εικόνες - Πίνακες

Κεφάλαιο 1

Εικόνα 1.1-Διαπλεκόμενη σάρωση 15
Εικόνα 1.2-Συνεχής σάρωση 15

Κεφάλαιο 2

Εικόνα 2.1-Motion Compensation με χρήση του Trhree Step Search 33

Κεφάλαιο 3

Εικόνα 3.1-Το σύνολο των τυποποιήσεων MPEG 37
Εικόνα 3.2 -Σχηματικό διάγραμμα MPEG 38
Εικόνα 3.3 -Ανατομία του MPEG 39
Εικόνα 3.4 -intra-frame coding 41
Εικόνα 3.5 -inter-frame coding 42
Εικόνα 3.6 -Πίνακες που σχηματίζουν μια 4:4:4 και μια 4:2:2 εικόνα 43
Εικόνα 3.7 -Προς τα εμπρός πρόβλεψη στα P-frames του MPEG 45
Εικόνα 3.8 -Πρόβλεψη διπλής κατεύθυνσης στα B-frames του MPEG 46
Εικόνα 3.9 -Πίνακας δειγματοληψίας 48
Εικόνα 3.10 - Επίπεδα που ορίζουν την ανάλυση μιας εικόνας στο πρότυπο MPEG-2 ... 53
Εικόνα 3.11 -Δομή του MPEG-2 bit stream 54

Κεφάλαιο 4

Εικόνα 4.1-Frame 1 original – Compressed 63
Εικόνα 4.2-Frame 2 original – Compressed 63
Εικόνα 4.2-Frame 2 original – Compressed 64
Εικόνα 4.4-Frame 4 original – Compressed 64
Εικόνα 4.5-Frame 5 original – Compressed 64
Εικόνα 4.6-Frame 6 original – Compressed 65
Εικόνα 4.7-Frame 7 original – Compressed 65
Εικόνα 4.8-Frame 8 original – Compressed 65
Εικόνα 4.9-Frame 9 original – Compressed 66
Εικόνα 4.10-Frame 10 original – Compressed 66
Εικόνα 4.11-Αρχικό frame 67
Εικόνα 4.12-Κωδικοποίηση σε πολλά επίπεδα ποιότητας 68

Κεφαλαίο 5

Εικόνα 5.1 -Φωτεινότητα 72
Εικόνα 5.2 -Blockiness 73
Εικόνα 5.3 -Οξύτητα 73
Εικόνα 5.4 – video 1 αρχική ποιότητα και πληροφορίες 74
Εικόνα 5.5 – video 1 αλλοιωμένο και πληροφορίες 75
Εικόνα 5.6 – video 2 αρχική ποιότητα και πληροφορίες 75
Εικόνα 5.7 – video 2 αλλοιωμένο και πληροφορίες 76
Εικόνα 5.8 – video 3 αρχική ποιότητα και πληροφορίες 76
Εικόνα 5.9 – video 3 αλλοιωμένο και πληροφορίες 77
Εικόνα 5.10– video 4 αρχική ποιότητα και πληροφορίες 77
Εικόνα 5.11 – video 4 αλλοιωμένο και πληροφορίες 78
Εικόνα 5.12 – video 5 αρχική ποιότητα και πληροφορίες..... 78
Εικόνα 5.13 – video 5 αλλοιωμένο και πληροφορίες 79
Εικόνα 5.14 – Στοιχεία συμμετέχοντα 81
Εικόνα 5.15 – Απαντήσεις συμμετέχοντα 82
Εικόνα 5.16 – Φύλλο συμμετέχοντα 83
Εικόνα 5.17– Ηλικία συμμετεχόντων 84
Εικόνα 5.18– Συμμετοχή με γυαλιά ή μη 84
Εικόνα 5.19 – Γνώση συμμετεχόντων 84
Εικόνα 5.20 - Frame 102 αλλοιωμένο video 1 85
Εικόνα 5.21 - Frame 105 αρχικό video 1 85
Εικόνα 5.22 – Διάγραμμα 1ο video 1η στήλη γυναίκες 2η άντρες 86
Εικόνα 5.23 - Frame 108 αρχικό video 2 86
Εικόνα 5.24 - Frame 108 αλλοιωμένο video 2 87
Εικόνα 5.25 – Διάγραμμα 2ο video 1η στήλη γυναίκες 2η άντρες 87
Εικόνα 5.26 - Frame 50 αρχικό video 3 88
Εικόνα 5.27 - Frame 50 αλλοιωμένο video 3 88
Εικόνα 5.28 – Διάγραμμα 3ου video 1η στήλη γυναίκες 2η άντρες 89
Εικόνα 5.29 - Frame 250 αρχικό video 4 89
Εικόνα 5.30 - Frame 250 αλλοιωμένο video 4 90
Εικόνα 5.31 – Διάγραμμα 4ου video 1η στήλη γυναίκες 2η άντρες 90
Εικόνα 5.32 - Frame 120 αρχικό video 5 91
Εικόνα 5.33 - Frame 122 αλλοιωμένο video 5 91
Εικόνα 5.34 – Διάγραμμα 5ου video 1η στήλη γυναίκες 2η άντρες 92
Εικόνα 5.35 – psnr 45.53 95

Εικόνα 5.36 – psnr 36.81 95
Εικόνα 5.37 – psnr 31.45 96
Εικόνα 5.38 – PSNR και MSE σε διάφορες κλίμακες για το Frame 50 του video 3 98
Εικόνα 5.39 – (α)scale 1 και (β)scale 16 99
Εικόνα 5.40 - Οι τιμές είναι για 2 έως 10 frame του video 3 οι τιμές SNR είναι σε db..... 100
Εικόνα 5.41 - Οι τιμές είναι για 2 έως 10 frame του video 3 οι τιμές PSNR είναι σε db.... 101
Εικόνα 5.42 – Ακολουθία frame από 273-279 για το video 1 103
Εικόνα 5.43 – Ακολουθία frame από 282-286 για το video 1 104
Εικόνα 5.44 – Ακολουθία frame από 121-125 για το video 4 107
Εικόνα 5.45 – Ακολουθία frame από 271- 276 για το video 4 108
Εικόνα 5.46 – Ακολουθία frame από 101 – 108 για το video 5 110
Εικόνα 5.47 – Ακολουθία frame από 351 – 359 για το video 2 110
Εικόνα 5.48 – Ακολουθία frame από 592– 599 για το video 2 111
Εικόνα 5.49 –psnr 40.2561 frame 632 111
Εικόνα 5.50 –psnr 18.2671 frame 632 112

Κεφάλαιο 1 - Τηλεοπτική εικόνα

1. Εισαγωγή

Ο ανθρωπινός εγκέφαλος διατηρεί την αίσθηση μιας εικόνας για ένα ελάχιστο κλάσμα του δευτερόλεπτου αφού αυτήν χαθεί από το οπτικό του πεδίο. Η φυσιολογική αυτή ιδιαιτερότητα είναι γνωστή ως «*μετείκασμα*» και σε αυτήν αποδίδεται η αίσθηση της συνεχόμενης κίνησης που βλέπουν οι άνθρωποι ,όταν ακίνητες εικόνες που διαφέρουν ελάχιστα μεταξύ τους προβάλλονται γρήγορα η μια μετά την άλλη μπροστά στα μάτια τους.

Δηλαδή μετείκασμα είναι η διατήρηση της εικόνας στην συνείδηση για λίγο χρόνο μετά την απομάκρυνση του εξωτερικού ερεθίσματος. Στο μετείκασμα αποδίδεται η υποκειμενική αίσθηση της κίνησης κατά την γρήγορη προβολή ακίνητων εικόνων.

Δεν το αποδέχονται όλοι όμως , ορισμένοι ερευνητές θεωρούν το μετείκασμα ως το φυσιολογικό αίτιο της αίσθησης τις κινούμενης εικόνας υποστηρίζοντας την άποψη ότι είναι εξαιρετικά ασθενές φαινόμενο και η συνεχεία της όρασης είναι μια ψευδαίσθηση που δημιουργείται από τον θεατή περισσότερο με βάση ψυχολογικά αίτια και όχι την φυσιολογία του μετεικάσματος .

Όποια και να είναι η αιτία της ψευδαίσθησης της κινούμενης εικόνας δηλαδή ενός video για πολλές δεκαετίες αποτέλεσε ένα σημαντικό πολυμέσο για τις επικοινωνίες και για την ψυχαγωγία, αρχικά γραφόταν και μεταδιδόταν σε αναλογική μορφή για να φτάσει σήμερα στην ψηφιακή. Η εμφάνιση των ψηφιακών ολοκληρωμένων κυκλωμάτων και των υπολογιστών οδήγησε στην ψηφιοποίηση του video, και το ψηφιακό video υποκίνησε μια επανάσταση στη συμπίεση και τη μετάδοσή του. Η συμπίεση video έγινε έτσι , ένας σημαντικός τομέας έρευνας προς το τέλος της δεκαετίας του '80 και κατά τη δεκαετία του '90 και δημιούργησε ποικίλες εφαρμογές συμπεριλαμβανομένων της αποθήκευσης video σε DVD και CD, της ευρείας τηλεοπτικής ανοικτής μετάδοσης μέσω ψηφιακού καλωδίου, της δορυφορικής και επίγειας (από αέρος) ψηφιακής τηλεόρασης (DTV), και της video τηλεδιάσκεψης. Για το χρήστη, η ουσία της όλης διαδικασίας εστιάζει στην προσδοκία ότι το υλικό που θα δει θα αναπαραχθεί άμεσα , χωρίς διακοπές και φυσικά με την καλύτερη ποιότητα. Αυτός είναι και ο σημαντικότερος στόχος για τον οποίο αναπτύχθηκε η τεχνολογία αυτή.

Ποιοι είναι όμως, οι παράγοντες που διαμορφώνουν την τελική ποιότητα και ποια είναι τα μέσα που θα μας δώσουν μια υποκειμενική και αντικειμενική εκτίμηση της ποιότητας των video;

Οι παράγοντες αυτοί ποικίλλουν και επηρεάζουν με διαφορετική κάθε φορά βαρύτητα το τελικό αποτέλεσμα. Το πρώτο βήμα είναι η καταγραφή ή η δημιουργία του υλικού, είτε σε απευθείας ψηφιακή μορφή είτε σε αναλογική και κατόπιν η ψηφιοποίηση της. Η διαδικασία αυτή είναι καθοριστικής σημασίας, αφού η ποιότητα του αρχικού υλικού όπως μιας ταινίας θα πρέπει να είναι όσο το δυνατόν καλύτερη. Το επόμενο στάδιο αποτελείται από τη συμπίεση του υλικού, χρησιμοποιώντας τους κατάλληλους κωδικοποιητές / αποκωδικοποιητές (codecs) που έχουν αναπτυχθεί και φυσικά την μετάδοση τους με την καλύτερη αξιοπιστία.

1.1. Αναλογικό video

Με τον όρο «*αναλογικό video*» εννοούμε την ηλεκτρονική τεχνολογία κωδικοποίησης , μετάδοσης και αναπαραγωγής οπτικοακουστικής πληροφορίας (video), όπου η μορφή των χρησιμοποιούμενων σημάτων είναι αναλογικού τύπου.

Το πρώτο σήμα video που χρησιμοποιήθηκε ήταν το αναλογικό «ασπρόμαυρο» , δηλαδή σήμα αναλογικής μορφής που αναπαράγει την εικόνα στο δέκτη (συσκευή τηλεόρασης) ως αποχρώσεις του γκρίζου, μεταξύ των δύο ακραίων τιμών του απόλυτου λευκού και του απόλυτου μαύρου.

Τέτοια τεχνολογία χρησιμοποιεί η συμβατική «ασπρόμαυρη» τηλεόραση, όπου η αναπαραγωγή της εικόνας στην οθόνη τύπου CRT ,γίνεται καθώς η δέσμη ηλεκτρονίων που δημιουργείται από τον καθοδικό σωλήνα της συσκευής σαρώνει την οθόνη σε πολλές διαδοχικές οριζόντιες γραμμές. Οι διαδοχικές αυτές φωτεινές γραμμές δημιουργούν συνολικά την εικόνα που εμφανίζεται στην οθόνη. Ένα ειδικό σήμα «αμαύρωσης» στο τέλος κάθε γραμμής σάρωσης (horizontal blanking interval) καθοδηγεί τον μηχανισμό σάρωσης να επιστρέψει στην αρχή της οθόνης και να εκκινήσει τη σάρωση μιας νέας γραμμής.

Με αυτό τον τρόπο δηλαδή γραμμή προς γραμμή σαρώνεται ολόκληρη η οθόνη και δημιουργείται έτσι μια εικόνα που ονομάζεται frame(πλαίσιο). Στο τέλος σε κάθε frame υπάρχει ένα ακόμη τυποποιημένο σήμα αμαύρωσης (vertical blanking interval) που σηματοδοτεί τον τερματισμό του frame και μετακινεί το σήμα στην πάνω αριστερή γωνία της οθόνης ώστε να εκκινήσει τη σάρωση για τη δημιουργία ενός νέου frame. Η διαδικασία αυτή γίνεται με τέτοιο ρυθμό, ώστε τα frame να εναλλάσσονται στην οθόνη όσο γρήγορα απαιτείται ώστε να δημιουργείται η ψευδαίσθηση της κινούμενης εικόνας.

1.1.1. Διαπλεκόμενη και συνεχής σάρωση

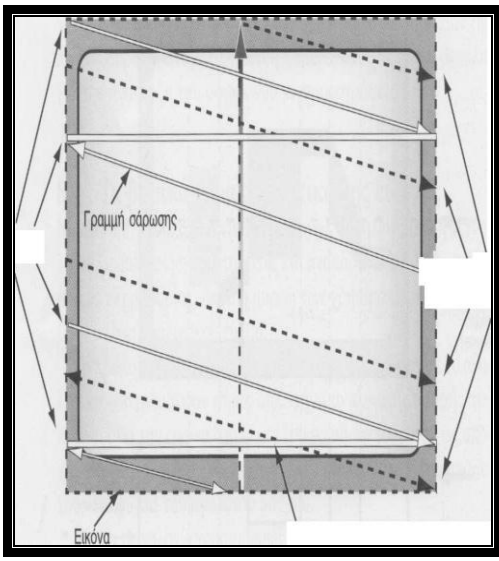
Η σάρωση της οθόνης μπορεί να γίνει με δύο τρόπους:

1. Διαπλεκόμενη (interlaced) σάρωση
2. Συνεχής (μη-διαπλεκόμενη) (non-interlaced ή progressive) σάρωση

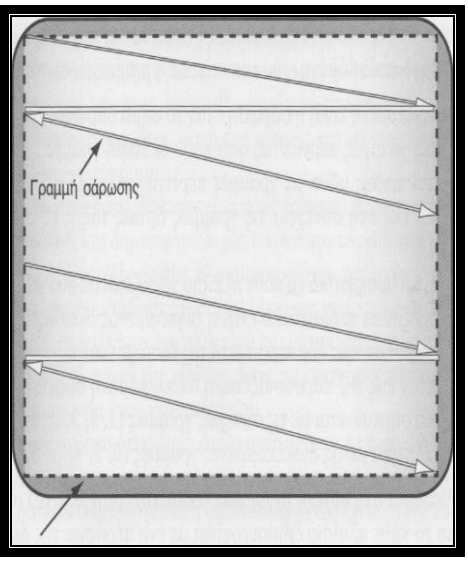
Διαπλεκόμενη είναι η σάρωση όταν το σήμα σαρώνει αρχικά τις μισές γραμμές περνώντας από κάθε δεύτερη γραμμή. Έτσι ουσιαστικά το κάθε frame χωρίζεται σε δύο μέρη που ονομάζονται πεδία (fields) και η σάρωση ενός ολοκληρωμένου frame απαιτεί δύο περάσματα της δέσμης ένα για κάθε πεδίο. *(Εικόνα 1.1)*

Στη ***συνεχή*** (ή μη διαπλεκόμενη) (progressive) σάρωση το κάθε frame δημιουργείται με ένα πέρασμα της δέσμης που σαρώνει με συνεχόμενο τρόπο τη μία γραμμή μετά την άλλη . *(Εικόνα 1.2)*

Διαπλεκόμενη είναι η σάρωση οθόνης στη τηλεόραση και το video ενώ συνεχής στους υπολογιστές.



Εικόνα 1.1 - *Διαπλεκόμενη σάρωση*



Εικόνα 1.2 - *Συνεχής σάρωση*

1.1.2. Βασικά μεγέθη αναλογικού video

Κάθετη Ανάλυση (vertical resolution) Είναι ο αριθμός των διακριτών οριζόντιων γραμμών στις οποίες αναλύεται ένα frame. Χαρακτηριστικές τιμές ανάλυσης είναι οι 525 γραμμές στο αμερικανικό σύστημα NTSC και 625 γραμμές στο ευρωπαϊκό σύστημα PAL.

Λόγος εικόνας (Aspect ratio) Είναι ο λόγος του πλάτους της εικόνας προς το ύψος της. Η συμβατική τηλεοπτική εικόνα έχει λόγο εικόνας 4:3 ενώ η εικόνα στην τηλεόραση υψηλής ευκρίνειας (High Definition TV ή HDTV) έχει 16:9.

Ρυθμός ανανέωσης πλαισίου (Frame rate, refresh rate) είναι το μέγεθος που εκφράζει το πόσο συχνά ανανεώνεται το πλαίσιο εικόνας στην οθόνη του δέκτη. Μετριέται σε μονάδες συχνότητας Hz ή ισοδύναμα σε fps (frames per second). Ο ρυθμός αυτός πρέπει να είναι αρκετά υψηλός ώστε να δημιουργείται με ικανοποιητικό τρόπο η ψευδαίσθηση της κινούμενης εικόνας. Σε αντίθετη περίπτωση ο θεατής αισθάνεται ότι η εικόνα τρεμοπαίζει ή τρεμοσβήνει, ένα φαινόμενο γνωστό με την ονομασία «*flickering*». Ο ελάχιστος απαιτούμενος ρυθμός ανανέωσης frame δεν είναι ο ίδιος στις διάφορες τεχνολογίες καθώς οι συνθήκες στο περιβάλλον προβολής της εικόνας επηρεάζουν την ποιότητα θέασης.

1.1.3. Συστήματα αναλογικού video

Υπάρχουν σήμερα δύο βασικά διαφορετικά όσο και ασύμβατα μεταξύ τους συστήματα μετάδοσης και αναπαραγωγής αναλογικού video και είναι γνωστά με τις ονομασίες: PAL και NTSC.

Τυπικά τα συστήματα αναλογικού σήματος video αναφέρονται με τρεις παραμέτρους στη σειρά:

ανάλυση/ρυθμός ανανέωσης πεδίου/διαπλοκή

Έτσι για τα παραπάνω συστήματα έχουμε τις κωδικοποιημένες ονομασίες:

- 625/50/2:1 Γνωστό κι ως PAL , χρησιμοποιείται σε Ευρώπη και Ασία με εύρος ζώνης 6.5 MHz. Εκπέμπει διαπλεκόμενη εικόνα με 625 συνολικά οριζόντιες γραμμές (312,5 γραμμές σε κάθε fields) και 25 frame ανά δευτερόλεπτο (δηλ. συχνότητα ανανέωσης πεδίου τα 50 Hz). Ο λόγος εικόνας είναι 4:3 και χρησιμοποιεί το μοντέλο χρώματος YUV .

- 525/59.94/2:1 Γνωστό κι ως NTSC , χρησιμοποιείται στη Βόρεια Αμερική και την Ιαπωνία και εκπέμπει με εύρος ζώνης 5.5 MHz. Εκπέμπει διαπλεκόμενη εικόνα με 525 συνολικά οριζόντιες γραμμές (262.5 γραμμές σε κάθε fields) και με ρυθμό ανανέωσης πεδίου τα 60 Hz (περίπου) (δηλ. 30 frame ανά δευτερόλεπτο). Ο λόγος εικόνας είναι 4:3 και χρησιμοποιεί το μοντέλο χρώματος YIQ.

Ένα ακόμη σύστημα αναλογικού video σε χρήση είναι το SEC AM χρησιμοποιείται σε Γαλλία, Ανατολική Ευρώπη (πχ. Ρωσία) & περιοχές της Αφρικής. Έχει ίδιο αριθμό γραμμών και ρυθμό ανανέωσης με το PAL αλλά διαφοροποιείται στον τρόπο κωδικοποίηση της πληροφορίας χρώματος. Υπάρχει τέλος και το σύστημα 1125/60/2:1 που χρησιμοποιείται στη τηλεόραση.

1.1.4. Μειονεκτήματα/Πλεονεκτήματα αναλογικού video

Τα κυριότερα μειονεκτήματα που προκύπτουν από τη χρήση video σε αναλογική μορφή είναι τα ακόλουθα:

1. Εξαιτίας της φύσης του αναλογικού video προσφέρεται στο χρήστη εξαιρετικά περιορισμένη αλληλεπίδραση δηλαδή οι επιλογές του είναι ελάχιστες.

2. Η ύπαρξη διαφορετικών προτύπων αναλογικού video format (NTSC ,PAL, SECAM δημιουργεί προβλήματα όταν απαιτείται η διαχείριση σήματος video από συσκευή κατά την περίπτωση όπου το σήμα έχει εγγραφεί σε διαφορετικό format από αυτό που υποστηρίζει η συσκευή.

3. Η αναζήτηση σε μια βάση (video database) από κασέτες video για την εύρεση συγκεκριμένων δεδομένων ή πληροφοριών είναι μια ιδιαίτερα επίπονη διαδικασία καθώς απαιτείται οπτική "σάρωση" του συνόλου των δεδομένων της βάσης.

4. Η αντιγραφή video σήματος από ένα αναλογικό μέσο σε άλλο έχει ως αποτέλεσμα μείωση της ποιότητας του σήματος η οποία γίνεται ιδιαίτερα σημαντική σε περίπτωση επαναλήψεων της αντιγραφής.

Τα πλεονεκτήματα του αναλογικού video είναι η εύκολη και οικονομική εκπομπή καθώς και η προσιτή τιμή του απαιτούμενου προς αναπαραγωγή εξοπλισμού.

1.2. Ψηφιακό video

Παλιότερα λέγοντας «*ψηφιακό video*» εννοούσαμε τα αρχεία ψηφιακού video που προέκυπταν από την ψηφιοποίηση σήματος αναλογικού video καθώς και τα εργαλεία

(λογισμικό) επεξεργασίας τους. Σήμερα ο όρος «ψηφιακό video» ή «τεχνολογία DV» (Digital Video ή DV) αναφέρεται γενικά σε ένα σύνολο ψηφιακών τεχνολογιών με τις οποίες είναι δυνατή η εξ αρχής παραγωγή, επεξεργασία, αποθήκευση και διαμοίραση αρχείων οπτικοακουστικής πληροφορίας σε ψηφιακή μορφή.

Υπάρχουν δύο τρόποι για να δημιουργήσει κανείς ψηφιακά αρχεία video:

1. Με ψηφιοποίηση αναλογικού video. Η τεχνική αυτή είναι γνωστή ως «σύλληψη αναλογικού video» (analog video capturing).

2. Με χρήση τεχνολογίας DV, δηλαδή να χρησιμοποιήσει κανείς εξ αρχής ψηφιακή τεχνολογία (πχ. ψηφιακή βιντεοκάμερα) και δημιουργεί απ' ευθείας ψηφιακό σήμα video.

Το ψηφιακό video μπορεί να περιγραφεί ως ένα σύνολο στατιστικών εικόνων που εναλλάσσονται γρήγορα πολλές φορές το δευτερόλεπτο. Άρα ένα video «κληρονομεί» τα χαρακτηριστικά των εικόνων που το αποτελούν (χρωματική και χωρική ανάλυση) και έχει επίσης το χαρακτηριστικό της συχνότητας εναλλαγής τους . Τα πλεονεκτήματα του ψηφιακού video είναι στην ανάλυση της εικόνας τους , στην απόδοση χρώματος τους και στον λόγο σήματος προς θόρυβο(SNR). Το ψηφιακό video επιτρέπει την επεξεργασία τυχαίας προσπέλασης σε αντίθεση με την γραμμική επεξεργασία του αναλογικού video, είναι πιο εξελίξιμο.

1.2.1. Πρότυπα ψηφιακού video

Πρότυπα:

1. **CCIR-601** - Διαδεδομένο format για τηλεοπτικά στούντιο με δειγματοληψία 4:2:2, Non-interlaced.

2. **DVB -** Διαδεδομένο format για τηλεοπτική μετάδοση δειγματοληψία 4:2:2, interlaced.

3. **HDTV -** Διαδεδομένο format για τηλεοπτική μετάδοση υψηλής ευκρίνειας ,δειγματοληψία 4:2:2 , interlaced.

4. **SIF -** Διαδεδομένο για ψηφιοποίηση και αποθήκευση αναλογικών κασετών VHS DV format, δειγματοληψία 4:1:1, Non-interlaced.

5. **CIF –** Διαδεδομένο για εφαρμογές τηλεδιάσκεψης , δειγματοληψία 4:1:1, Non-Interlaced.

6. **QCIF-** Διαδεδομένο για εφαρμογές τηλεδιάσκεψης με χρήση τηλεπικοινωνιακού καναλιού χαμηλού εύρους ζώνης ISDN.

Το χρωματικό μοντέλο YCr Cb χρησιμοποιείται σε όλα τα παραπάνω πρότυπα.

1.2.2. Πλεονεκτήματα ψηφιακού video

Λέγοντας ψηφιακό video αναφερόμαστε στο σύνολο των τεχνικών που υπάρχουν για την παράγωγη και την επεξεργασία οπτικοακουστικής πληροφορίας αποκλειστικά σε ψηφιακή μορφή. Το ψηφιακό video προσφέρει καλύτερη ποιότητα εικόνας σε σχέση με το αναλογικό.

Η καλύτερη εικόνα κρίνεται από τρεις βασικούς παράγοντες

1. Ανάλυση εικόνας , η κάθετη ανάλυση της ψηφιακής εικόνας είναι τουλάχιστον δύο φορές μεγαλύτερη από αυτήν που μας προσφέρει η αναλογική τεχνολογία. Προσφέρει καλύτερη ανάλυση κατά 25%.

2. Απόδοση χρώματος , η τεχνολογία DV χρησιμοποιεί κωδικοποίηση component που αποδίδει πολύ καλύτερη χρωματική πληροφορία από το Y/C η το composite. Έτσι δεν έχουμε διάχυση του χρώματος (color smear) ή θολή απόδοση (color blur).

3. Λόγος σήματος προς θόρυβο (signal to noise ratio) , είναι το κλάσμα με αριθμητή το πλάτος η την ισχύ του σήματος που μεταφέρει πληροφορία και παρανομαστή το επίπεδο θορύβου. Μετριέται σε decibel (dB) και όσο πιο μεγάλη είναι τόσο καλύτερη ποιότητα μετάδοσης και αναπαραγωγής.

Πέρα από την καλύτερη ποιότητα υπάρχουν και αλλά σημαντικά πλεονεκτήματα σε ένα ψηφιακό video. Επιτρέπει την επεξεργασία τυχαίας προσπέλασης , δεν υποβαθμίζεται η ποιότητα κάθε επόμενου αντίγραφου αφού κάθε νέο ψηφιακό αντίγραφο είναι πιστή αναπαραγωγή της πηγής. Δεν εμφανίζει τα προβλήματα ποιότητας που προκύπτουν κατά την ψηφιοποίηση αναλογικής πληροφορίας και πολλά άλλα.

Ρυθμός μετάδοσης ψηφιακού video

Για να υπολογιστεί ο ρυθμός μετάδοσης δεδομένων του ψηφιακού video χρησιμοποιούνται οι παρακάτω παράμετροι..

f = frame size (width * height – σε pixels)
r = frame rate (frames per second)

c = color depth (in bits/pixel)

t = time in seconds

Όταν σε ένα ψηφιακό video , γίνεται επαγγελματική επεξεργασία και σχεδόν καθόλου συμπίεση τότε ο ρυθμός δεδομένων είναι υψηλός. Αντίθετα, ο χαμηλός αριθμός δεδομένων παρουσιάζεται σε εφαρμογές με διάδραση ή και μετάδοση σε μεγάλες αποστάσεις, και επίσης όταν χρησιμοποιούνται τεχνικές συμπίεσης.

Όπως και να δέχεται ένα σήμα η τηλεόραση είτε αναλογικά είτε ψηφιακά το σήμα τις είναι ένα σύστημα τηλεπικοινωνίας που χρησιμεύει στη μετάδοση και λήψη κινούμενων εικόνων και ήχου εξ αποστάσεως. Αποτελεί το κυριότερο και δημοφιλέστερο Μέσο Μαζικής Επικοινωνίας και η χρήση της είναι ιδιαίτερα διαδεδομένη σε όλο τον κόσμο.

Λέγοντας "τηλεόραση" εννοούμε τη συσκευή, δηλαδή τον δέκτη, ο οποίος λαμβάνει το (τηλεοπτικό) σήμα που εκπέμπουν οι τηλεοπτικοί σταθμοί σε συγκεκριμένες συχνότητες (ή αλλιώς κανάλια) με την οθόνη που απεικονίζει το αποτέλεσμα της εκπομπής (μετατροπή του σήματος σε εικόνα και ήχο). Η λέξη προέρχεται από το αρχαίο ελληνικό πρόθεμα "τηλε-" ("μακριά") και τη λέξη "όραση".



Κεφάλαιο 2 – Συμπίεση

2. Συμπίεση

Η ανάγκη για ψηφιακή συμπίεση

Με δεδομένη την ολοένα και αυξανόμενη χρήση της ψηφιακής τεχνολογίας, η οποία αντικαθιστά την αναλογική που παραδοσιακά χρησιμοποιείται εδώ και δεκαετίες στις τηλεπικοινωνίες, την τηλεόραση, την ηχογράφηση και αναπαραγωγή της μουσικής αλλά και σε άλλους τομείς, έχει αρχίσει παράλληλα να γίνεται έρευνα και προς την κατεύθυνση της συμπίεσης της ψηφιακής πληροφορίας με στόχο την οικονομία εύρους φάσματος (bandwidth).

Οι λόγοι για τους οποίους είναι όχι απλά χρήσιμη αλλά απαραίτητη η ανάπτυξη αυτών των τεχνικών συμπίεσης δεν είναι ίσως τόσο προφανείς. Είναι φανερό ότι η ποιοτική ψηφιοποίηση κάποιου μέσου που απαιτεί υψηλή συχνότητα δειγματοληψίας και μεγάλο μέγεθος δείγματος θα δημιουργούσε ένα εξίσου μεγάλο αρχείο.

Το πρόβλημα λοιπόν των αρχείων με το μεγάλο μέγεθος είναι :

- ✓ Πρέπει να βρεθεί κατάλληλος αποθηκευτικός χώρος.
- ✓ Πρέπει η πληροφορία να μεταφέρεται αρκετά γρήγορα για αναπαραγωγή στα αντίστοιχα συστήματα.
- ✓ Πρέπει να μεταφέρεται ικανοποιητικά στο διαδίκτυο(internet).

Έτσι για μια εικόνα 1024 * 768 pixel με βάθος χρώματος 24 bit δημιουργεί αρχείο

➔ 1024*768*24=18874368 bit = 18 Mb = 2,25 MB

Ένα ασυμπίεστο video με 24 frame το δευτερόλεπτο χρειάζεται αποθηκευτικό χώρο

➔ 2,25 MB * 24 (frame) = 54 MB

Και για μια ταινία μιας ώρας απαιτείται χώρος

➔ 54MB *3600(sec) = 190 GB (στο περίπου)

Τέτοιου μεγέθους αρχεία δημιουργούν πρόβλημα και πρέπει με κάποιο τρόπο να συμπιεστούν ώστε να είναι εύκολα αποθηκεύσιμα και να μεταφέρονται και να αναπαράγονται εύκολα και γρήγορα.

2.1. Λόγος συμπίεσης – Ρυθμός συμπίεσης

Στόχος της συμπίεσης μπορεί να είναι είτε ο περιορισμός του μεγέθους ενός αρχείου είτε ο περιορισμός του ποσού πληροφορίας που πρέπει να μεταδοθεί. Αν ενδιαφερόμαστε για τον περιορισμό του αποθηκευτικού χώρου αναφερόμαστε στον λόγο συμπίεσης και κάνουμε λόγο για ένα μέγεθος που εκφράζει την σχέση του χώρου που καταλαμβάνουν τα συμπιεσμένα δεδομένα. Ο λόγος συμπίεσης είναι καθαρός αριθμός και δίνεται από την σχέση :

Λόγος συμπίεσης = αρχικό μέγεθος δεδομένων / μέγεθος συμπιεσμένων δεδομένων

Λόγος συμπίεσης 3 : 1 (3 προς 1) σημαίνει ότι μετά την συμπίεση το μέγεθος του αρχείου είναι το 1/3 του αρχικού.

Αν μας ενδιαφέρει η οικονομία που κάνουμε στην χρήση καναλιού τότε αναφερόμαστε στον ρυθμό συμπίεσης που εκφράζει τον ρυθμό μετάδοσης των συμπιεσμένων δεδομένων και μετριέται συνήθως σε bits per second , bits per sample , bits per pixel .

2.2. Συμπίεση Δεδομένων

Ένα σημαντικό ερώτημα που τίθεται είναι αν μπορεί η πληροφορία να κωδικοποιηθεί πιο αποδοτικά δηλαδή να βρεθεί ένας τρόπος να μεταδοθεί η πληροφορία με λιγότερα bit αλλά παράλληλα με την ίδια ποιότητα. Άρα συμπίεση είναι μια διαδικασία που επανακωδικοποιεί την εκπεμπόμενη πληροφορία με στόχο να μεταδώσει το ίδιο περιεχόμενο αλλά με λιγότερα bits. Η συμπίεση εκμεταλλεύεται κανόνες και περιορισμούς , προσπαθεί να εκμεταλλευτεί κάθε πλεονασμό που υπάρχει στην αρχική πληροφορία.

Δεν πρέπει να γίνεται σύγχυση μεταξύ σε ένα αλγόριθμο συμπίεσης (compession algorithm) που έχει στόχο την επίτευξη συμπίεσης των δεδομένων εισόδου εφαρμόζοντας μια απλή και συγκεκριμένη μεθοδολογία επεξεργασίας των δεδομένων έτσι ώστε να προκύψει μια περισσότερο συμπιεσμένη μορφή τους. Με έναν συμπιεστή (codec ή compressor) όπου είναι μια διαδικασία συμπίεσης και αποσυμπιέσεις και περιλαμβάνει περισσότερους από έναν αλγορίθμους στην σειρά δηλαδή τα δεδομένα εξόδου του ενός αποτελούν είσοδο για τον επόμενο. Γνωστοί τέτοιοι συμπιεστές που

εφαρμόζονται σε πληροφορία εικόνας και video είναι οι jpeg και mpeg. Και οι δυο αυτοί συμπιεστές βασίζονται στον " Διακριτό μετασχηματισμό συνημίτονου " (Discrete Cosine Transform).

Ειδικότερα στην πληροφορία πολυμέσων εκμεταλλευόμαστε το γεγονός ότι απευθύνεται σε ανθρώπους έτσι μπορούμε να εκμεταλλευτούμε τα φυσιολογικά χαρακτηριστικά και περιορισμούς των ανθρωπινών αισθητήριων οργάνων και να συμπιέσουμε την πληροφορία χωρίς να μειώσουμε αισθητά την ποιότητα της. Για παράδειγμα το ανθρώπινο μάτι μπορεί να δει κάποιες αποχρώσεις και το ανθρώπινο αυτί να αντιλήφθη κάποιες συχνότητες έτσι μπορούμε να ξεφορτωθούμε τέτοιες πληροφορίες χωρίς να μειώσουμε την ποιότητα. Η συμπίεση μπορεί να μειώσει τον όγκο των δεδομένων που στέλνονται ή αποθηκεύονται, με την ελαχιστοποίηση του ενυπάρχοντος πλεονασμού. Ο πλεονασμός παρουσιάζεται κατά τη δημιουργία των δεδομένων. Με τη διαδικασία της συμπίεσης η μεταφορά και η αποθήκευση γίνονται με πιο αποδοτικό τρόπο, ενώ παράλληλα διατηρείται η ακεραιότητα των δεδομένων.

2.2.1. Παράγοντες που σχετίζονται με την συμπίεση

Στόχος της συμπίεσης video είναι να μειώσει ένα ποσοστό της απαιτούμενης πληροφορίας, ώστε να μπορέσει να αποθηκευτεί το ψηφιακό σήμα, διατηρώντας βέβαια όσο είναι δυνατόν την ποιότητα του αρχικού video. Σύμφωνα λοιπόν με τα παραπάνω, υπάρχουν σημαντικοί παράγοντες που πρέπει να ληφθούν υπόψη κατά την αναφορά μας στην ψηφιακή συμπίεση video.

Πραγματικός χρόνος / Μη πραγματικός χρόνος

Μερικά συστήματα συμπίεσης συλλαμβάνουν, συμπιέζουν, αποσυμπιέζουν, και αναπαράγουν το video στα 30 fps, και όλα αυτά σε πραγματικό χρόνο, χωρίς καμία καθυστέρηση. Άλλα πάλι συστήματα είναι μόνο σε θέση να κάνουν σύλληψη μερικά από τα 30 fps και ή είναι σε θέση να αναπαράγουν μόνο μερικά από αυτά. Η ανεπαρκής συχνότητα ανανέωσης είναι ένα από τα πιο αξιοσημείωτα ελαττώματα του video. Κάτω των 24 πλαισίων το δευτερόλεπτο (fps), το video θα είναι αισθητά απότομο. Επιπλέον τα πλαίσια που παραλείπονται περιέχουν υπερβολικά σημαντικές πληροφορίες συγχρονισμού των δεδομένων. Με άλλα λόγια, εάν η κίνηση των χειλιών ενός ατόμου έχει χαθεί εξαιτίας των παραλειπόμενων πλαισίων κατά την διάρκεια της σύλληψης ή της αναπαραγωγής θα είναι αδύνατον να συγχρονιστεί ο ήχος σωστά με το

video. Άρα καταλαβαίνουμε ότι η έννοια του πραγματικού χρόνου έχει να κάνει με την συχνότητα ανανέωσης των πλαισίων.

Συμμετρική / Ασύμμετρη

Συμμετρική συμπίεση σημαίνει ότι εάν μπορεί κανείς να αναπαράγει ένα video μεγέθους 640 x 480 και με frame rate 30 fps, τότε μπορεί επίσης να κάνει σύλληψη και να το αποθηκεύσει με τις ίδιες τιμές. Ασύμμετρη συμπίεση σημαίνει ακριβώς το αντίθετο. Ο βαθμός ασυμμετρίας συνήθως εκφράζεται ως μια αναλογία. Η ασύμμετρη συμπίεση μπορεί να είναι πιο περίπλοκη και πιο αποδοτική όσον αφορά την ποιότητα και την ταχύτητα κατά την αναπαραγωγή για τον λόγο ότι χρησιμοποιείται πολύ περισσότερος χρόνος όταν συμπιέζεται το video. Η συμπίεση αυτή απαιτεί πάρα πολύ χρόνο.

Αναλογίες συμπίεσης

Η αναλογία της συμπίεσης αντιπροσωπεύει αριθμητικά τα αρχικό video σε σύγκριση με το συμπιεσμένο. Όσο πιο πολύ συμπιέζεται το video , τόσο πιο υψηλή είναι η αναλογία συμπίεσης ή αλλιώς η αριθμητική διαφορά ανάμεσα στις δύο τιμές. Γενικά, όσο πιο υψηλή είναι η αναλογία συμπίεσης τόσο πιο χαμηλή θα είναι η ποιότητα της εικόνας. Οι αναλογίες συμπίεσης διαφέρουν όχι μόνο από μια μέθοδο συμπίεσης σε μια άλλη, αλλά και από το hardware και το software που χρησιμοποιείται από έναν υπολογιστή για να γίνει η συμπίεση.

Απωλεστική συμπίεση/ μη απωλεστική συμπίεση

Οι τεχνικές συμπίεσης κατηγοριοποιούνται σε απωλεστικής (lossy) και μη απωλεστικής (lossless) συμπίεσης. Η διαφορά είναι απλή, στην συμπίεση χωρίς απώλειες η πληροφορία μετά την αποσυμπίεση είναι ίδια με την αρχική ενώ στην απωλεστική συμπίεση η πληροφορία μετά την αποσυμπίεση είναι λιγότερη σε σχέση με την αρχική κάποια πληροφορία έχει χαθεί. Στην μη απωλεστική συμπίεση (lossless compression) διατηρείται η ακεραιότητα των δεδομένων. Τα αρχικά δεδομένα και τα δεδομένα μετά τη συμπίεση και την αποσυμπίεση είναι ακριβώς τα ίδια, επειδή σε αυτές τις μεθόδους ο αλγόριθμος συμπίεσης και ο αλγόριθμος αποσυμπίεσης είναι ακριβώς αντίστροφοι. Κατά τη διαδικασία δε χάνεται κανένα μέρος των δεδομένων. Τα πλεονάζοντα δεδομένα κωδικοποιούνται κατά τη συμπίεση και αποκωδικοποιούνται κατά την αποσυμπίεση. Αυτοί οι μέθοδοι χρησιμοποιούνται όταν δεν πρέπει να χαθεί ούτε ένα bit δεδομένων όπως στην περίπτωση ενός αρχείου κειμένου ή ενός προγράμματος. Τα γνωστά προγράμματα 7z , bz , zip και rar χρησιμοποιούν μη απωλεστική συμπίεση. Ευρέως διαδεδομένοι μη απωλεστική αλγόριθμοι είναι ο Run Length Encoding (RLE) , Huffman , Delta , LZW .

Στην απωλεστική συμπίεση έχουμε απώλεια δεδομένων όπου μπορεί να μην είναι αποδεκτή σε αρχεία κειμένου ή σε προγράμματα είναι όμως αποδεκτή σε εικόνες και ταινίες. Ο λόγος είναι ότι τα μάτια μας και τα αφτιά μας δεν μπορούν να διακρίνουν πολύ μικρές αλλαγές. Για τέτοιες περιπτώσεις είναι κατάλληλες οι απωλέστηκες μέθοδοι συμπίεσης (lossy data compression). Οι μέθοδοι αυτές είναι οικονομικότερες και απαιτούν λιγότερο χρόνο και χώρο όταν πρέπει να σταλούν εκατομμύρια bit εικόνων και video το δευτερόλεπτο. Χαρακτηριστικό παράδειγμα απωλεστικής συμπίεσης εικόνας είναι η μέθοδος JPEG για video η μέθοδος και για ήχο το πρότυπο MP3.

2.3. Γενικές αρχές συμπίεσης video

2.3.1. Ποιοτική ανοχή

Σε αντίθεση με την αντίληψη ότι η ψηφιακή μετάδοση θα πρέπει να έχει το εύρος ζώνης (bandwith) που χρειάζεται για να μεταδώσει και το πιο απαιτητικό πλαίσιο (frame) μιας σειράς κινούμενων εικόνων, ακόμα και αν αυτό παρουσιάζεται με συχνότητα 1%, τώρα πια κάτω από την εμπορική πίεση που υπάρχει (λόγο κόστους) θεωρείται λογικό να δεχόμαστε κάποιο ποσοστό παραμόρφωσης σε τέτοιες σπάνιες σκηνές, με αντάλλαγμα να μπορούμε να χρησιμοποιήσουμε τα bit που εξοικονομούμε για την αναβάθμιση του μέσου όρου ανάλυσης του συνόλου των σκηνών. Έτσι σε αντίθεση με την ηχογράφηση ήχου σε CD, που γίνεται χωρίς καμία συμπίεση, στη συμπίεση video αναζητείται μία χρυσή τομή (sweet spot) ανάμεσα στην ποιότητα και το bandwith που πολλές φορές βασίζεται σε υποκειμενικές μετρήσεις που γίνονται σε άτομα που θεωρούνται ικανά “δείγματα” πάνω στην εκτίμηση της εικόνα και του ήχου λόγω επαγγελματικής εμπειρίας όπως θα αναφέρουμε μερικά παρακάτω. Αν η εμπειρία τους δεν τους επιτρέπει να αντιληφθούν σημαντικές διαφορές στην ποιότητα του κωδικοποιημένου ήχου από τον αρχικό και συγχρόνως θεωρούν ανεκτές τις ατέλειες (artifacts) της εικόνας μετά από κάποιο ποσοστό συμπίεσης, η μέθοδος και το ποσοστό αυτό θεωρούνται αποδεκτά και τυποποιούνται. Να σημειωθεί εδώ ότι ενώ ο ήχος προσφέρεται για μεγάλα ποσοστά συμπίεσης χωρίς να μπορεί να γίνει αισθητή υποβάθμιση στην ποιότητά του, στον τομέα της εικόνας τα πράγματα είναι λίγο πιο δύσκολα και οι όποιες ατέλειες είναι ορατές, απλά γίνεται προσπάθεια να μην γίνονται ενοχλητικές για το θεατή.

2.3.2. Πλεονασμός (Redundancy)

Αν σκεφτούμε ότι για κάθε ένα δευτερόλεπτο μιας ταινίας video χρειάζονται περίπου 30 εικόνες, καθεμία από τις οποίες απαιτεί ίσως και 1 MB μνήμης, είναι προφανές ότι ένα αρχείο video χρειάζεται τεράστιο όγκο δεδομένων όπως είδαμε και στο πιο πάνω παράδειγμα. Η ποσότητα της πληροφορίας ενός αρχείου video εξαρτάται από την ποιότητα της εικόνας και τον αριθμό των καρέ ανά δευτερόλεπτο που απαιτούνται για την σωστή αναπαραγωγή κίνησης. Η ανάγκη που υπάρχει για την συμπίεση των δεδομένων ενός αρχείου είναι επιτακτική όχι μόνο για τον όγκο του αρχείου όπως αναφέραμε αλλά και για τους λόγους διακίνησης του σε τακτά χρονικά διαστήματα. Η βασική αρχή πάνω στην οποία στηρίζονται όλες οι μέθοδοι ψηφιακής συμπίεσης είναι το γεγονός ότι το σήμα περιέχει ένα ποσοστό πλεονασμού (Redundancy). Δηλαδή πληροφορία που είτε μπορεί να παραληφθεί, είτε να κωδικοποιηθεί με λιγότερη ακρίβεια, χωρίς αυτό να έχει αξιοσημείωτη επίδραση στο τελικό αποτέλεσμα. Υπάρχουν δύο είδη πλεονασμού:

- ✓ **Στατικός πλεονασμός (Spatial Redundancy):**

Το επίπεδο του σήματος μπορεί σε κάθε δεδομένη χρονική στιγμή να προϋπολογιστεί από την προηγούμενη τιμή του, γιατί οι τιμές δειγμάτων της εικόνας σε μεγάλο ποσοστό σχετίζονται μεταξύ τους. Αυτό μπορούμε να το επαληθεύσουμε και εποπτικά π.χ. σε μία εικόνα του δελτίου ειδήσεων ένα μεγάλο κομμάτι της εικόνας (φόντο) παραμένει αμετάβλητο και μόνο το κομμάτι της εικόνας που καταλαμβάνει ο παρουσιαστής μεταβάλλεται ελαφρά (κινήσεις των χειλιών, των βλεφάρων κτλ). Έτσι μπορούμε να υπολογίσουμε ένα τμήμα της εικόνας από μία προηγούμενη και να προσθέσουμε απλά τις διαφορές που έχουν προκύψει χωρίς να χρειάζεται να κωδικοποιούμε σε κάθε frame την πλεονάζουσα πληροφορία.

- ✓ **Υποκειμενικός Πλεονασμός (Temporal Redundancy):**

Ανάλογα με το περιεχόμενο της εικόνας το ανθρώπινο μάτι μπορεί να ανεχτεί ένα ποσοστό παραμόρφωσης ή αλλοίωσης ορισμένων παραμέτρων της εικόνας χωρίς αυτό να γίνει αντιληπτό. Π.χ. είναι γνωστό ότι η ανθρώπινη όραση είναι γενικά πολύ πιο ευαίσθητη στη φωτεινότητα της εικόνας παρά τα χρώματα. Αντίστοιχες ιδιότητες έχει και η ακοή. Άρα μπορούμε να αφιερώσουμε λιγότερο από το διαθέσιμο bandwidth στην περιγραφή των χρωμάτων και γενικά της πλεονάζουσας πληροφορίας και αυτό να περάσει απαρατήρητο από το θεατή. Ο υποκειμενικός πλεονασμός και οι ιδιαιτερότητες της ανθρώπινης όρασης έχουν αξιοποιηθεί εδώ και δεκαετίες στην

αναλογική τεχνολογία της τηλεόρασης, αλλά τώρα βρίσκουν εφαρμογή και στις ψηφιακές τεχνικές μετάδοσης.

2.4. Δειγματοληψία

Για να ψηφιοποιηθεί ένα σήμα από αναλογικό σε ψηφιακό, χρειάζονται τρία βήματα. Γίνεται η δειγματοληψία του σήματος δηλαδή από το πλήθος τιμών του συνεχούς σήματος κρατάμε μόνο ένα σύνολο διακριτών τιμών. Στην συνέχεια έχουμε κβαντοποίηση και κωδικοποίηση του σήματος .

Βασικό κανόνα για την δειγματοληψία οποιουδήποτε αναλογικού σήματος αποτελεί το θεώρημα Nyquist-Shannon όπου σύμφωνα με αυτό το θεώρημα της δειγματοληψίας, για να πάρουμε κάποια δείγματα από το αναλογικό video χρησιμοποιούμε μια προκαθορισμένη συχνότητα όπου θα πρέπει να ικανοποιεί το θεώρημα Nyquist. Το θεώρημα αυτό βασίζεται στον κανόνα ότι για την μετατροπή ενός αναλογικού σήματος σε ψηφιακό του οποίου η υψηλότερη συχνότητα είναι x Hertz, τότε για την συχνότητα δειγματοληψίας απαιτεί τουλάχιστον 2 x Hertz. Σε περίπτωση που δεν συμβεί αυτός ο διπλασιασμός της συχνότητας, τότε θα παρατηρηθεί το φαινόμενο aliasing. Όπως αναφέραμε στο κεφάλαιο 1 ένα αναλογικό video εκτείνεται προς τρεις διαστάσεις οριζόντια, κάθετα και χρονικά. Η συχνότητα δειγματοληψίας πρέπει να ορίζεται με τρόπο ώστε να πλήρη βασικές απαιτήσεις. Όπως τα δείγματα πρέπει να βρίσκονται σε απόλυτη στοίχιση στην κάθετη διάσταση. Η απόσταση μεταξύ των δειγμάτων πρέπει να είναι ίδια και στην οριζόντια αλλά και στην κάθετη διάσταση. Επίσης, η συχνότητα δειγματοληψίας πρέπει να παραμένει ίδια στα διάφορα συστήματα αναλογικού video. Υπάρχουν τρία βασικά μεγέθη που μπορούν και επηρεάζουν τον όγκο των ψηφιακών δεδομένων που θα παραχθούν. Αυτά είναι η ανάλυση της εικόνας, το βάθος χρώματος και ο αριθμός των πλαισίων.

Στόχος της δειγματοληψίας είναι να μειώσει τη διάσταση του video στην οριζόντια ή κάθετη διεύθυνση και ουσιαστικά τον αριθμό των pixels που θα κωδικοποιηθούν πριν την διαδικασία της κωδικοποίησης.

2.5. Κβαντοποίηση

Η μέθοδος που μας βοηθάει να απαλλαγούμε από σημαντικό μέρος της πληροφορίας είναι η κβαντοποίηση. Με τον όρο κβαντοποίηση γενικά εννοούμε τη μετατροπή ενός σήματος άπειρων (η πάρα πολλών) τιμών σε ένα σήμα ορισμένων

διακριτών τιμών π.χ. η κβαντοποίηση μιας εικόνας που περιέχει εκατομμύρια χρώματα οδηγεί σε μία εικόνα που έχει 256 διαφορετικές τιμές για το χρώμα (πρότυπο JPEG). Με άλλα λόγια κβαντοποίηση είναι ο περιορισμός των bits με τα οποία περιγράφουμε τα δείγματα του σήματος (προφανώς το 256 έχει πολύ λιγότερα bits από τους τεράστιους αριθμούς με τους οποίους έπρεπε να περιγράψουμε τα δείγματά μας αν δεν γινόταν κβαντοποίηση).

Ένα παράδειγμα είναι το παρακάτω που δείχνει ταυτόχρονα με τη μεγάλη οικονομία που γίνεται και την εισαγωγή σημαντικών σφαλμάτων (για τον περιορισμό των οποίων επιστρατεύονται άλλες μέθοδοι) :

Ο αριθμός 45 είναι 101101 έχει δηλαδή 6 bits.
Με 4 bits γίνεται 1011 = 11
Με 3 bits γίνεται 101 = 5 κτλ.

Είναι προφανές ότι η κβαντοποίηση εισάγει σφάλμα ανάλογο με τον αριθμό τον bits που απορρίπτονται και κατά συνέπεια ευθύνεται στο μεγαλύτερο βαθμό για την απώλεια πληροφορίας κατά τη συμπίεση (lossy compression) σε αντίθεση με το μετασχηματισμό DCT που είναι μία, σε μεγάλο βαθμό αντιστρεπτή διαδικασία.

Για να περιγράψουμε όσο περισσότερες από τις τιμές του σήματος γίνεται με δοσμένο αριθμό bits, διαιρούμε τις τιμές των δειγμάτων είτε με σταθερές τιμές (uniform Quantization) είτε με πίνακες κβαντοποίησης (Quantization tables). Στη συγκεκριμένη περίπτωση, δηλαδή στο πρότυπο MPEG χρησιμοποιείται η δεύτερη μέθοδος και μάλιστα υπάρχει ένας πίνακας για τα πλαίσια που έχουν κωδικοποιηθεί με ενδοπλαισιακή (intra-frame coding) και ένας για αυτά με διαπλαισιακή (inter-frame coding) όπως θα δούμε παρακάτω . Εν ολίγοις η κβαντοποίηση είναι η μετατροπή ενός μεγέθους από συνεχές σε ψηφιακό, δηλαδή σε κάτι που μετριέται σε διακριτά βήματα και όχι με συνεχή τρόπο. Άρα η κβαντοποίηση είναι η διαδικασία που μετατρέπει ένα σήμα από το χώρο των συνεχών τιμών στο χώρο των διακριτών. Αυτό πραγματοποιείται με το να υποδιαιρεθεί το εύρος τιμών του συνεχούς σήματος σε συγκεκριμένα επίπεδα και σε κάθε επίπεδο να εκχωρείται ένας ψηφιακός κωδικός.

2.6. Κωδικοποίηση

Υλοποιώντας τις διαδικασίες δειγματοληψίας και κβαντοποίησης, ένα συνεχές σήμα Μετατρέπεται σε ένα διακριτό σύνολο τιμών. Στην συνέχεια χρησιμοποιείται η διαδικασία της κωδικοποίησης για τη μετατροπή του διακριτού συνόλου των τιμών

των δειγμάτων σε μια πιο κατάλληλη μορφή. Κάθε σχέδιο για την αναπαράσταση καθενός από αυτά τα διακριτά σύνολα 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).

Δύο είδη κωδικοποίησης που θα αναφερθούν παρακάτω είναι η ενδοπλαισιακή (intraframe) και διαπλαισιακή (interframe). Έτσι σε πολλές περιπτώσεις (πιο πολύπλοκες υλοποιήσεις) ο κωδικοποιητής μπορεί να επιλέξει το είδος κωδικοποίησης ανάλογα με της ανάγκες της εικόνας. Δηλαδή οι δύο τρόποι κωδικοποίησης χρησιμοποιούνται περιοδικά για να «φιλτράρουν» την εικόνα από αυξημένες παραμορφώσεις, ή στην αντίθετη περίπτωση να ελαττώσουν το ρυθμό μετάδοσης. Αυτός είναι και ο λόγος που στο MPEG-2 εμφανίζεται μεταβαλλόμενος ρυθμός μετάδοσης.

Οι μέθοδοι Run Length Encoding και Huffman

✓ RLE (Run Length Encoding)

Ο αλγόριθμος RLE (Run Length Encoding) αποτελεί μια από τις απλούστερες τεχνικές συμπίεσης. Σύμφωνα με τη μέθοδο αυτή, διατρέχεται η ακολουθία των bytes που αποτελούν τα δεδομένα προς συμπίεση και εντοπίζονται οι διαδοχικές επαναλήψεις του ίδιου χαρακτήρα. Στη συνέχεια αντικαθίστανται οι συνεχόμενες επαναλήψεις με το πλήθος τους, ακολουθούμενο από τον χαρακτήρα. Στην πράξη ο αλγόριθμος διατρέχει την ακολουθία των δεδομένων και αναζητά διαδοχικές επαναλήψεις του ίδιου χαρακτήρα. Όταν βρεθεί μεμονωμένος χαρακτήρας που δεν επαναλαμβάνεται, ή διπλή επανάληψη του, μένει ως έχει, καθώς δεν επωφελούμαστε από την κωδικοποίηση RLE που απαιτεί δύο χαρακτήρες. Όταν όμως βρεθούν διαδοχικές επαναλήψεις ενός χαρακτήρα, τις αντικαθιστά με ένα ζεύγος (αριθμός, χαρακτήρας), όπου «αριθμός» είναι το σύνολο των συνεχόμενων εμφανίσεων του χαρακτήρα.
Εν ολίγοις η ακολουθία πολλαπλών εμφανίσεων του ίδιου συμβόλου αντικαθίσταται από 2 σύμβολα (a) το σύμβολο που εμφανίζεται (b) πόσες φορές εμφανίζεται.

✓ Huffman

Αυτή η τεχνική συμπίεσης παρουσιάστηκε από τον D. Huffman το 1952 και επιτυγχάνει μεγάλα ποσοστά συμπίεσης, εκμεταλλευόμενη τη στατιστική ανάλυση της ακολουθίας δεδομένων. Η ανάλυση αυτή είναι πιο σύνθετη από την απλή ανίχνευση των συνεχόμενων επαναλήψεων που κάνει ο RLE. Σύμφωνα με την κωδικοποίηση του προτύπου ASCII, σε κάθε χαρακτήρα αντιστοιχεί ένα byte, που είναι ο κωδικός ASCII του χαρακτήρα αυτού. Αν η κωδικοποίηση γίνεται με τον νέο κώδικα Unicode, χρειάζονται 2

bytes για κάθε χαρακτήρα. Το συνολικό μέγεθος της ακολουθίας σε bytes είναι ίσο με το πλήθος των χαρακτήρων επί 1 ή 2 bytes ανά χαρακτήρα (ανάλογα με την κωδικοποίηση). Στην κωδικοποίηση Huffman, επιτυγχάνεται συμπίεση καθώς ο κώδικας που αντιστοιχεί σε κάθε χαρακτήρα δεν έχει σταθερό μήκος αλλά μεταβλητό. Η κωδικοποίηση για τους πιο συχνά εμφανιζόμενους χαρακτήρες είναι μικρότερη από ότι για τους λιγότερο συχνά εμφανιζόμενους, με αποτέλεσμα το συνολικό μέγεθος να είναι μικρότερο, καθώς για τους συχνότερους χρησιμοποιούνται μόνο 2-3 bits και όχι 8 ή 16.

Η κωδικοποίηση Huffman ανήκει στην γενικότερη κατηγορία της κωδικοποίησης εντροπίας και την υποκατηγορία στατικής κωδικοποίησης .Χρησιμοποιείται και στην συμπίεση ακίνητης και κινούμενης εικόνας συνήθως για κάθε εικόνα δημιουργείται ένα νέο λεξικό κωδικών ενώ στην περίπτωση τις κινούμενης δημιουργείται για κάθε πλαίσιο η σειρά πλαισίων .

Η παραπάνω αλγόριθμοι ανήκουν στην κατηγορία κωδικοποίησης εντροπίας είναι αλγόριθμοι που αντιμετωπίζουν την πληροφορία απλά σαν μια σειρά από bits και επιχαρούν να την συμπιέσουν εφαρμόζοντας κάποιες στατικές μεθόδους προσπαθώντας για παράδειγμα να μειώσουν τον πλεονασμό προσφέρουν κωδικοποίηση χωρίς απώλειες και διακρίνονται σε δυο βασικές κατηγορίες

1. Περιορισμός των επαναλαμβανόμενων ακολουθιών (suppression of repetitive sequences)
2. Στατική κωδικοποίηση (statistical encoding)

2.7. Μετασχηματισμός DCT Coding

Ο Διακριτός Συνημιτονικός Μετασχηματισμός (Discrete Cosine Transform) είναι μία μέθοδος που βρίσκει μεγάλη εφαρμογή στην ψηφιακή συμπίεση αλλά και στο MPEG ειδικότερα. Αποτελεί μια μαθηματική τεχνική προετοιμασίας των ψηφιακών δεδομένων για συμπίεση. Εκείνο που κάνει είναι να δίνει τέτοια δομή στα δεδομένα ώστε να είναι εύκολο να εντοπιστούν εκείνα που μπορούν να αγνοηθούν χωρίς αισθητή απώλεια στην ποιότητα της πληροφορίας.

Πως λειτούργει ο DCT ;

Στον μετασχηματισμό DCT μπορούμε να μεταφέρουμε την πληροφορία που περικλείει η εικόνα από το πεδίο του χώρου (time domain) στο πεδίο της συχνότητας

(frequency domain) όπου η περιγραφή της μπορεί να γίνει με σημαντικά μικρότερο πλήθος bits. Για μια ψηφιακή εικόνα μετασχηματίζει την πληροφορία των pixel της σε πληροφορία που αφόρα τις συχνότητες μεταβολής της εικόνας. Άρα μετά την εφαρμογή του μετασχηματισμού της εικόνας η πληροφορία τις πλέον αποτελείται από αριθμούς που δείχνουν την βαρύτητα που έχει η κάθε συχνότητα μεταβολής του σήματος της εικόνας στην σύνθεση της.

Ο μετασχηματισμός DCT ορίζεται ως εξής :

Για κάθε pixel (x,y) εφαρμόζοντας τον τύπο :

$$DCT(i,j) = \frac{1}{\sqrt{2N}} C(i)C(j) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} pixel(x,y) \cos\left[\frac{(2x+1)i\pi}{2N}\right] \cos\left[\frac{(2y+1)j\pi}{2N}\right]$$

όπου C(x) = 0.7071, x =0

1 , x >0

παίρνουμε την τιμή DCT(i,j) που είναι η τιμή του συντελεστή του μετασχηματισμού στο πεδίο της συχνότητας. Έτσι αντιστοιχίζουμε τις τιμές των pixels στις αντίστοιχες τιμές συντελεστών.

Οι συντελεστές αυτοί μεταφέρουν ο καθένας ένα κομμάτι της αρχικής πληροφορίας (αυτό που αντιστοιχεί στο κομμάτι του φάσματος που περιγράφει). Επειδή όμως έχει παρατηρηθεί ότι η ανθρώπινη όραση αντιλαμβάνεται πολύ περισσότερο τα φαινόμενα που σχετίζονται με χαμηλές συχνότητες όπως (π.χ. χρώματα με μικρότερα μήκη κύματος), ενώ δείχνει κάποια ανοσία σε υψίσυχνες περιοχές του σήματος (π.χ. ακμές της εικόνας), οι συντελεστές του μετασχηματισμού που αντιστοιχούν σε χαμηλές συχνότητες έχουν μεγαλύτερη βαρύτητα από αυτούς που περιγράφουν τις υψηλές συχνότητες και για το λόγο αυτό οι πρώτοι περιγράφονται με τη μεγαλύτερη δυνατή ακρίβεια.

Έτσι για μια εικόνα αν υποθέσουμε ότι αποτελείται από N X M στοιχειά , η εικόνα χωρίζεται σε block μεγέθους 8X8 pixel. Οι συντελεστές X(i,j) αυτού του πίνακα (i,j=0…7) είναι 8-bit τιμές (0-255) που παριστάνουν την φωτεινή ένταση των pixel και αποτελούν τα δεδομένα εισόδου του DCT.

Κατά την αναπαραγωγή γίνεται η αντίστροφη διαδικασία με τη βοήθεια του μετασχηματισμού IDCT (Inverse Discrete Cosine Transform - Αντίστροφος Διακριτός Μετασχηματισμός Συνημίτονων) , που περιγράφεται από τον τύπο:

$$pixel(x,y) = \frac{1}{\sqrt{2N}} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} C(i)C(j)DCT(i,j) \cos\left[\frac{(2x+1)i\pi}{2N}\right] \cos\left[\frac{(2y+1)j\pi}{2N}\right]$$

Το αποτέλεσμα είναι να πάρουμε πίσω σχεδόν ανέπαφη την αρχική πληροφορία (εκτός από κάποια αναπόφευκτα σφάλματα στρογγυλοποίησης).

2.8. Τμηματική Πρόβλεψη Κίνησης

Ένας τρόπος με τον οποίο μπορούμε να περιγράψουμε πιο αποτελεσματικά εικόνες με κίνηση είναι η τμηματική πρόβλεψη κίνησης. Με τη βοήθεια αυτής της μεθόδου μπορούμε να εκτελέσουμε τη διαπλαισιακή (inter-frame) κωδικοποίηση και να περιγράψουμε την αλληλουχία εικόνων ως σειρά ομοιοτήτων και διαφορών.

Εναλλακτικά με το να κωδικοποιήσουμε ανεξάρτητα τα δύο frame μπορούμε να περιγράψουμε τη δεύτερη εικόνα με το να τη χωρίσουμε σε ίσα τμήματα και να φτιάξουμε ένα πίνακα που να περιέχει τμήματα που έχουν μείνει ίδια και διανύσματα που να δείχνουν τη νέα θέση των τμημάτων που άλλαξαν θέση. Έτσι αν έχουμε ήδη αποστείλει την πρώτη εικόνα μπορούμε να στείλουμε τη δεύτερη σαν ένα πίνακα 20 διανυσμάτων και ορισμένων σταθερών τμημάτων, που προφανώς έχει πολύ μικρότερο μέγεθος. Το παραπάνω είναι μία καλή προσέγγιση της μεθόδου αλλά στην πραγματικότητα οι πραγματικές εικόνες δεν θα είναι τόσο όμοιες μεταξύ τους . Θα έχουν κάποια κοινά τμήματα που αλλάζουν θέση από πλαίσιο σε πλαίσιο αλλά θα υπάρχουν και τμήματα που αλλάζουν θέση διατηρώντας το σχήμα τους αλλά μεταβάλλεται το χρώμα τους, καθώς και άλλα που δεν υπάρχουν σε προηγούμενο πλαίσιο αλλά εμφανίζονται σε κάποιο για πρώτη φορά. Για την κάλυψη αυτών των περιπτώσεων χρησιμοποιείται μία πιο βελτιωμένη εκδοχή της παραπάνω ιδέας (ή για την ακρίβεια διάφορες εκδοχές της παραπάνω ιδέας).

Η σύνταξη του MPEG καθορίζει πως θα αναπαρίσταται η πληροφορία για την κίνηση του κάθε macroblock, ότι θα γίνεται δηλαδή αυτή η αναπαράσταση με τη χρήση διανυσμάτων κίνησης, αλλά δεν καθορίζει πως τα διανύσματα αυτά θα υπολογίζονται και για το λόγο αυτό εμφανίζονται διάφορες υλοποιήσεις της μεθόδου εύρεσης των διανυσμάτων κίνησης οι οποίες στηρίζονται όλες στην ελαχιστοποίηση μίας συνάρτησης που υπολογίζει την ταύτιση του τρέχοντος με το macroblock αναφοράς.

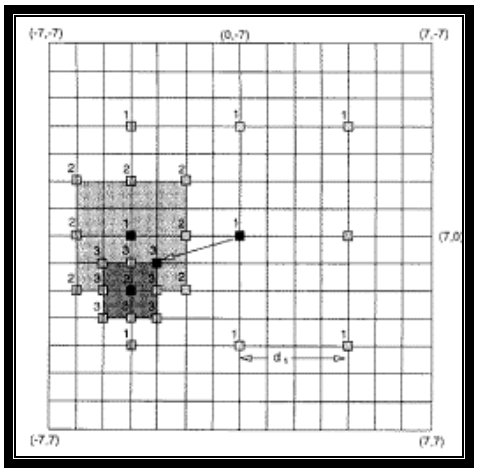
Αν και μπορεί να χρησιμοποιηθεί κάθε συνάρτηση σφάλματος που υπάρχει, η πιο συχνά χρησιμοποιούμενη συνάρτηση είναι η Απόλυτη Διαφορά (AE - Absolute Error) η οποία δίνεται από τον παρακάτω τύπο :

$$AE(d_x, d_y) = \sum_{i=0}^{15}\sum_{j=0}^{15}\left|f(i,j) - g(i-d_x, j-d_y)\right|$$

Στην παραπάνω εξίσωση το f(i,j) και g(i,j) αντιπροσωπεύουν τις συντεταγμένες των pixels στο τρέχον και το macroblock αναφοράς αντίστοιχα. Το macroblock αναφοράς που καθορίζεται από το διάνυσμα (dx,dy) αντιπροσωπεύει την περιοχή αναζήτησης. Το macroblock που παράγει το μικρότερο σφάλμα αντιστοιχεί στην τιμή του διανύσματος που ψάχνουμε.

Η πιο απλή διορατικά αλλά και η πιο πολύπλοκη από πλευράς υπολογιστικής πολυπλοκότητας είναι η πλήρης αναζήτηση (full search) η οποία καλύπτει κάθε pixel στην περιοχή αναζήτησης.

Για να μειωθεί λίγο η υπολογιστική πολυπλοκότητα έχει επινοηθεί η μέθοδος αναζήτησης τριών βημάτων (TSS - Three Step Search). Ο αλγόριθμος υπολογίζει την απόλυτη διαφορά (AE) στο κέντρο και σε οχτώ περιοχές της περιοχής αναζήτησης που είναι 32x32 pixels. Η περιοχή που θα έχει τη μικρότερη απόλυτη διαφορά γίνεται το κέντρο για την επόμενη αναζήτηση, η οποία έτσι έχει το μισό μέγεθος. Η διαδικασία αυτή επαναλαμβάνεται τρεις φορές, όπως φαίνεται στο παρακάτω σχήμα (η περιοχή του κάθε βήματος περικλείεται από σημεία που φέρουν την αντίστοιχη αρίθμηση).



Εικόνα 2.1 - *Motion Compensation με χρήση του Trhree Step Search*

2.9. Ακίνητες εικόνες καρέ

Μια ψηφιακή εικόνα είναι ένα ψηφιδωτό στοιχειωδών σημείων τα όποια διαθέτουν την απαραίτητη πληροφορία φωτεινότητας και χρώματος ώστε να δημιουργήσουν στον θεατή την αίσθηση της εικόνας. Θεωρητικά μια τέτοια εικόνα είναι οργανωμένη σε

σύνολα ορθογώνια διατεταγμένα που εκπέμπουν φως σε καθορισμένες εντάσεις αν μιλάμε για ασπρόμαυρες εικόνες ή αποχρώσεις είναι του γκρι αν μιλάμε για έγχρωμες . Τα στοιχειώδη αυτά σημεία ονομάζονται pixel .Η τιμή που εκχωρείται σε κάθε pixel είναι η μέση τιμή της φωτεινότητας στρογγυλοποιημένη προς τον πλησιέστερο ακέραιο. Οι εικόνες ή τα καρέ που εξάγονται από ακολουθίες εικόνων είναι σήματα δύο διαστάσεων I [M,N] αναπτύσσεται σε έναν διακριτό χώρο δυο διαστάσεων και παράγεται από την ψηφιοποίηση μιας αναλογικής εικόνας I[x,y] που αναπτύσσεται σε έναν συνεχή χώρο με διαστάσεις x και y. Η συνεχής εικόνα διαιρείται σε N σειρές και M στήλες , τα σημεία τομής των σειρών με των στηλών είναι τα pixel . Το μέγεθος που δείχνει από ποσά pixel αποτελείται είναι η ανάλυση εικόνας και μετριέται σε pixel per inch (όπου μια ίντσα είναι ίση με 2,54 εκατοστά). Κάθε pixel περιέχει την πληροφορία χρώματος που πρέπει να αναπαραχθέν για την παρουσίαση της εικόνας . Η πληροφορία αυτή αποθηκεύεται με χρήση κάποιων bit μέσα στο αρχείο της εικόνας . Αυτό είναι το βάθος χρώματος δηλαδή ο αριθμός που δηλώνει ποσά bit χρησιμοποιούνται για την αποθήκευση κάθε pixel.

Οι πιο συνηθισμένη βαθμοί χρώματος στις μέρες μας είναι

1. Χρώμα 8 bit - Χρησιμοποιούμε 8 bit δηλαδή 1 byte για κάθε pixel και επομένως μπορούμε να έχουμε 2^8 =256 χρώματα στην απεικόνιση μας.

2. Χρώμα 16 bit - Χρησιμοποιούμε 16 bit για κάθε pixel και επομένως μπορούμε να έχουμε 2^{16} =65536 χρώματα στην απεικόνιση μας .

3. Χρώμα 24 bit - Χρησιμοποιούμε 24 bit για κάθε pixel και επομένως μπορούμε να έχουμε 2^{24} =16777216 χρώματα στην απεικόνιση μας . Πραγματικό χρώμα (true color) και αναφέρεται έτσι επειδή αναπαριστά πιστά το ανθρώπινο μάτι . Στην πραγματικότητα αποδίδει πολύ περισσότερες αφού το ανθρώπινο μάτι μπορεί να ξεχωρίσει γύρω στις 300 με 350 χιλιάδες αποχρώσεις.

4. Χρώμα 32 bit – Αν και με το 24 bit είμαστε πλήρως καλυμμένη το επιπρόσθετο Byte περιέχει πληροφορία σχετικά με την διαφάνεια τις εικόνας και δεν αφορά αποχρώσεις. Χρησιμοποιείται καθαρά για επαγγελματικούς σκοπούς.

Άρα το μέγεθος ενός αρχείου εικόνας είναι

Μέγεθος αρχείου = αριθμός pixel * βάθος χρώματος

Όπου αριθμός pixel

Pixel = [pixel κατά πλάτος] *[pixel κατά μήκος] =
[ανάλυση]*[πλάτος]*[ανάλυση]*[μήκος]

Και όπως είναι φυσικό το πώς θα εμφανιστεί μια εικόνα στην οθόνη δεν εξαρτάται μόνο από την ανάλυση της εικόνας αλλά σε μεγάλο βαθμό και από την ανάλυση εξόδου του συστήματος. Η ανάλυση οθόνης είναι το μέγεθος που δείχνει από πόσα pixels παρουσιάζει σε κάθε διάσταση της εικόνας και εκφράζεται σε γινόμενο δυο αριθμών.

Άρα η κωδικοποίηση των ψηφιακών εικόνων είναι απαραίτητη ώστε να επιτυγχάνεται συμπίεση του όγκου των δεδομένων με σκοπό τη γρήγορη μετάδοση τους από διαύλους περιορισμένης χωρητικότητας ή τον περιορισμό των χώρων αποθήκευσης τους. Η συμπίεση μπορεί να επιτευχθεί αν γίνει εκμετάλλευση του πλεονασμού (Redundancy) που υπάρχει σε κάθε εικόνα. Οι μέθοδοι κωδικοποίησης διακρίνονται σε αυτές που δεν εισάγουν σφάλμα κατά την αποκωδικοποίηση (lossless) και σε εκείνες που έχουν απώλειες (lossy).

Μορφοποιήσεις (format) αρχείων εικόνων

Η πληροφορία σε κάθε αρχείο εικόνας χωρίζεται σε δύο τμήματα το πρώτο είναι η επικεφαλίδα αρχείου και περιλαμβάνει πληροφορίες σχετικά με τον τύπο της εικόνας τα χρώματα και τις διαστάσεις και το δεύτερο περιέχει συμπιεσμένη την πληροφορία της εικόνας

- ***GIF –*** *Το πρότυπο αυτό χρησιμοποιεί μέχρι 8 bits per pixel για αναπαράσταση έγχρωμης εικόνας και μέσα στο αρχείο αποθηκεύεται και η παλέτα (colormap) για την αντιστοιχία των χρωμάτων. Η κωδικοποίηση γίνεται με χρήση του αλγόριθμου LZW (Lempel, Ziv & Welsh) και στη συνέχεια με Run Length Encoding για περαιτέρω συμπίεση. Η κωδικοποίηση σύμφωνα με τις παραπάνω τεχνικές γίνεται χωρίς σφάλμα (lossless) και ο βαθμός συμπίεσης κυμαίνεται από 2:1 έως 5:1. Επειδή όμως υπάρχει δυνατότητα για απεικόνιση μόνο 256 χρωμάτων, η εικόνα έχει υποστεί κβαντισμό πριν από την κωδικοποίηση με συνέπεια την απώλεια πληροφορίας.*

- ***TIFF*** *- Η αναπαράσταση κάθε pixel γίνεται με 24 bits. Η κωδικοποίηση σε TIFF δίνει τη δυνατότητα χρησιμοποίησης διαφόρων αλγόριθμων συμπίεσης προσφέροντας επιλογή ανάλογα με το είδος της εικόνας. Πριν από την εφαρμογή του αλγόριθμου LZW οι εικόνες έχουν υποστεί αρχικά μια γραμμική διαδικασία οριζόντιας πρόβλεψης ενώ μεγαλύτερη συμπίεση επιτυγχάνεται εάν*

μηδενίσουμε το LSB (Least Significant Bit). Χωρίς μηδενισμό του LSB για εικόνα 24 bpp η συμπίεση είναι χαμηλή (1.1:1 έως 1.5:1) ενώ με μηδενισμό ο συντελεστής συμπίεσης αυξάνεται σε 1.8:1. Για έγχρωμες εικόνες με παλέτα ο μηδενισμός του LSB έχει ως αποτέλεσμα συντελεστή συμπίεσης μέχρι και 10:1.

- ***JPEG*** *Εξαιρετικά διαδεδομένη μορφή κωδικοποίησης που χρησιμοποιείται πολύ στο διαδίκτυο για μεταφορά εικόνων. Βασικό πλεονέκτημα είναι ότι διατηρεί όλοι την ποικιλία RGB χρωμάτων προσφέροντας ταυτόχρονα πολύ μικρό μέγεθος αρχείου . Η κωδικοποίηση κατά JPEG μπορεί να γίνεται χωρίς την εισαγωγή σφάλματος (lossless) αλλά και με σφάλμα (lossy). Οι δυνατοί τρόποι λειτουργίας του JPEG είναι τέσσερις: sequential, progressive, lossless και hierarchical, συνήθως όμως υλοποιείται η sequential τεχνική. Η lossless κωδικοποίηση χρησιμοποιεί γραμμικούς προβλεπτές και τεχνικές εντροπίας (Huffman) ενώ ο συντελεστής συμπίεσης είναι της τάξης του 2:1. Η lossy κωδικοποίηση βασίζεται στον διακριτό μετασχηματισμό συνημιτόνου, DCT ενώ στη συνέχεια οι συντελεστές που προκύπτουν κβαντίζονται και κωδικοποιούνται με την μέθοδο Huffman.*

- ***PNG*** *Η μορφοποίηση PNG είναι μια σχετικά πρόσφατη ιδέα και αποτελεί μια σημαντική βελτίωση της μορφοποίησης GIF. Πιο συγκεκριμένα, ένα αρχείο PNG μπορεί να περιέχει όλα τα χρώματα της αρχικής εικόνας και δεν περιορίζεται σε μια συγκεκριμένη παλέτα. .*

Άλλα format τα οποία χρησιμοποιούνται για την κωδικοποίηση ακίνητων εικόνων είναι τα ακόλουθα: PCX, Targa, PDF, BMP, Raw κλπ.

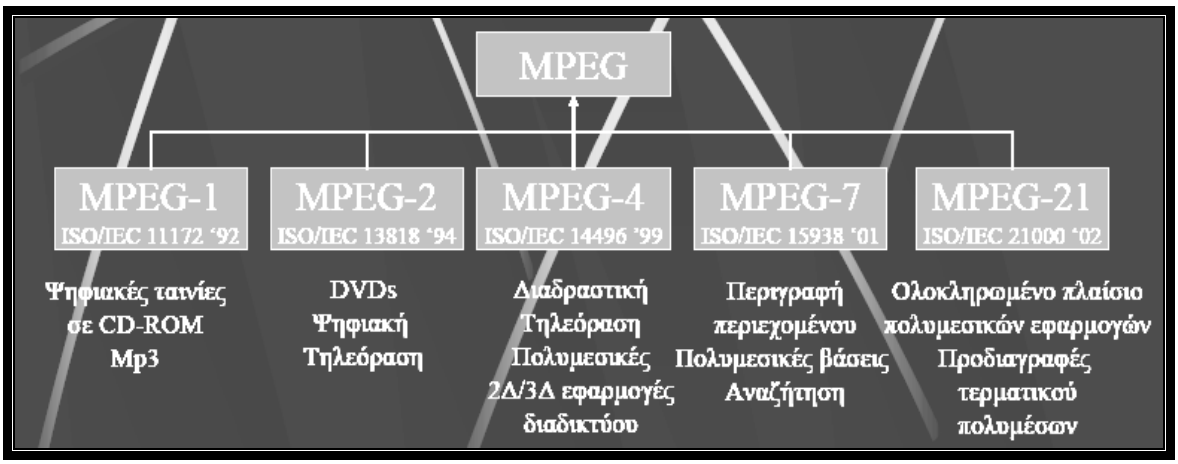
Κεφαλαίο 3 - Το πρότυπο συμπίεσης MPEG

3. Εισαγωγή

Τα αρχικά MPEG προέρχονται από τις λέξεις " moving picture expect group ". Πρόκειται για μια επιτροπή ειδικών που δραστηριοποιούνται στα πλαίσια του οργανισμού ISO με στόχο να αναπτύξει απωλεστικούς συμπιεστές (codes) για συμπίεση πληροφορίας video. Μέχρι σήμερα η επιτροπή MPEG έχει διατύπωση πολλούς συμπιεστές που φέρουν την ονομασία MPEG και συνδέονται από έναν χαρακτηριστικό αριθμό .

Έτσι με την ονομασία MPEG αναφερόμαστε ουσιαστικά σε μια οικογένεια ισχυρών συμπιεστών οπτικοακουστικής πληροφορίας (video). Οι πιο γνωστοί από αυτούς είναι ο MPEG-1 , MPEG-2 , MPEG-4. Η γενική ιδέα πίσω από την κωδικοποίηση του video χρησιμοποιώντας MPEG είναι η καταγραφή των μεταβολών/διαφορών μεταξύ των frame του video, διορθώνοντας σε τακτά διαστήματα την μετάδοση με πλήρεις, αναλυτικές εικόνες. Ο στόχος είναι η ανάπτυξη μιας μεθόδου για αποθήκευση και μετάδοση μέσω δικτύου ψηφιακού video και audio.

Το MPEG standard σήμερα, αποτελείται από έξι φάσεις, MPEG-1, MPEG-2, MPEG-3, MPEG-4, MPEG-7 και MPEG-21 οι οποίες ορίζονται από τις ανάλογες συστάσεις της ISO/IEC. (Εικόνα 3.1)



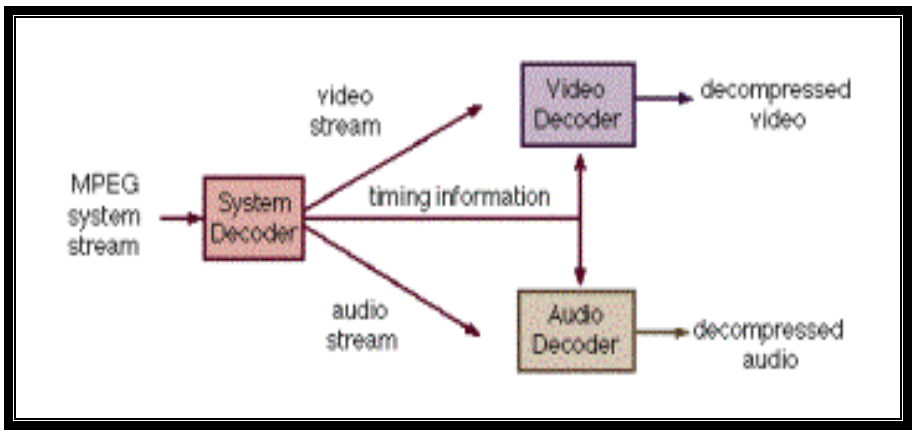
Εικόνα 3.1 - *Το σύνολο των τυποποιήσεων MPEG*

Ένα video αποτελείται από μια ακολουθία κινούμενων εικόνων, όπου οι διαδοχικές εικόνες που μεταδίδονται με σταθερό ρυθμό (συνήθως 25 εικόνες ανά sec)

παρουσιάζουν χωρική και χρονική συσχέτιση μεταξύ τους, γεγονός που εκμεταλλεύεται το MPEG για να συμπιέσει την πληροφορία. Τα frames του video ομαδοποιούνται ανά ομάδες GOP (Group of Pictures), τα οποία κωδικοποιούνται και αποκωδικοποιούνται ανεξάρτητα. Υπάρχουν τρία είδη κωδικοποιημένων πλαισίων που θα περιγράφουν παρακάτω.

3.1. Ανατομία σήματος MPEG

Ένα σήμα κωδικοποιημένο με MPEG αποτελείται από τρία επίπεδα: το επίπεδο συστήματος (system layer) , το τηλεοπτικό επίπεδο (video layer) και το ακουστικό επίπεδο (audio layer) όπως φαίνεται παρακάτω. (Εικόνα 3.2)



Εικόνα 3.2 - *Σχηματικό διάγραμμα MPEG*

Το επίπεδο system περιλαμβάνει πληροφορίες σχετικά με το συγχρονισμό , ελέγχει τη ροή του σήματος για να μην παρατηρείται έλλειψη ή πλεονασμός δεδομένων, παρέχει πληροφορίες για σημεία αναφοράς που διευκολύνουν την τυχαία προσπέλαση (random access) και τέλος περιέχει πληροφορίες για το διαχωρισμό του video από το audio και επίσης για την συγχρονισμένη απεικόνισή τους.

Τα επίπεδα video και audio περιέχουν κωδικοποιημένη την εικόνα και τον ήχο αντίστοιχα. Η κωδικοποίηση αυτών των επιπέδων μπορεί να έχει γίνει ταυτόχρονα ή ξεχωριστά. Σε κάθε περίπτωση, τα δεδομένα των τριών επιπέδων συνενώνονται σε ένα ενιαίο σήμα (bit-stream) μέσω μιας διαδικασίας που ονομάζεται πολυπλεξία (multiplexing ή muxing). Η αντίστροφη διαδικασία ονομάζεται demultiplexing (ή demuxing). Μερικά συστήματα κάνουν την πολυπλεξία σε πραγματικό χρόνο (real-time) και άλλα όχι.

Η διαδικασία της κωδικοποίησης και αποκωδικοποίησης του MPEG γίνεται με hardware ή με software. Η λύση του software είναι πιο φθηνή αλλά έχει μειονέκτημα στην ποιότητα της εικόνας και απαιτεί αρκετά ισχυρούς υπολογιστές για να λειτουργήσει. Η λύση του hardware είναι ακριβότερη, βασίζεται σε υλοποιήσεις των διαφόρων μεθόδων με ολοκληρωμένα VLSI και παράγει αυτόνομα συστήματα (κωδικοποιητές ή αποκωδικοποιητές) που δεν χρειάζονται υπολογιστή για να λειτουργήσουν, ενώ ταυτόχρονα έχουν πολύ καλή ποιότητα εικόνας.

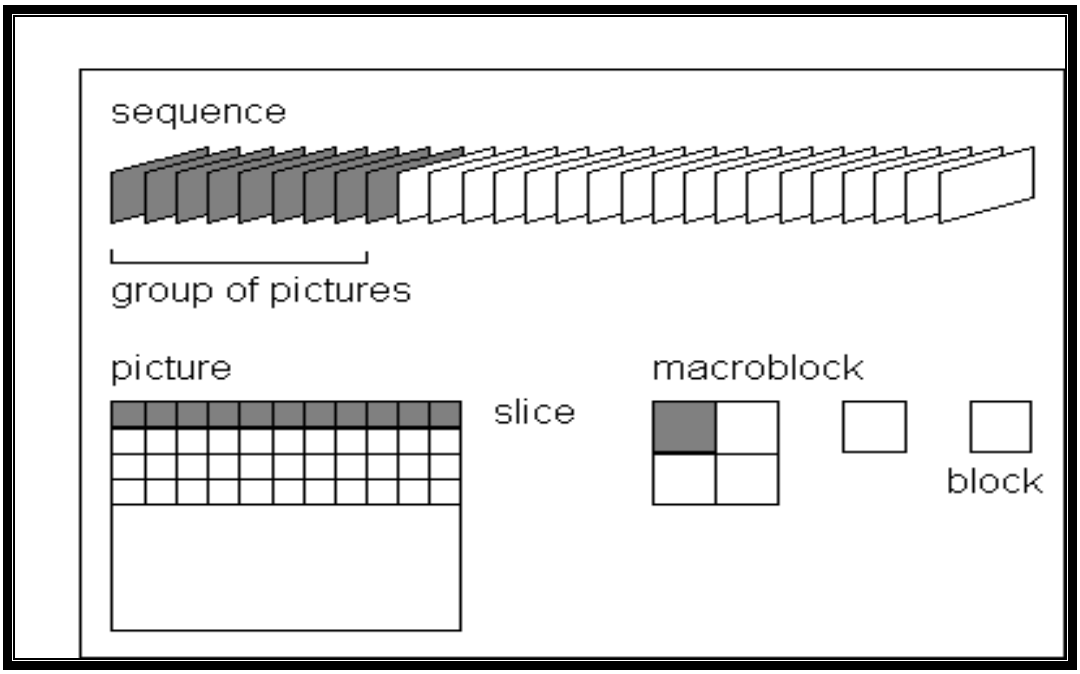
Είναι σημαντικό να σημειώσουμε ότι οι περισσότερες μέθοδοι κωδικοποίησης video όπως Cinepak, Indeo, Motion-JPEG, στηρίζονται στις μεθόδους συμπίεσης ακίνητης εικόνας JPEG αντιμετωπίζοντας την κινούμενη εικόνα σαν μια σειρά από διαδοχικές ακίνητες εικόνες. Η μέθοδος MPEG αν και δανείζεται σε μεγάλο βαθμό τις βασικές αρχές του JPEG επεκτείνει την προσέγγιση αυτή και περιγράφει την παράμετρο της κίνησης με ένα πιο λεπτομερή και αποτελεσματικό τρόπο από τις άλλες μεθόδους, με ευεργετικά αποτελέσματα από πλευράς ποσοστού συμπίεσης.

Προκειμένου να γίνει η επεξεργασία της εικόνας και η διαδικασία πρόβλεψης της κίνησης, το κάθε frame της εικόνας χωρίζεται σε τμήματα που ονομάζονται slices. Τα slices περιέχουν μία σειρά macroblocks, τα οποία όπως φανερώνει και το όνομά τους περιέχουν μία σειρά από blocks. Τα blocks είναι τμήματα της εικόνας διαστάσεων 8x8.

Ο ρόλος των slices είναι να περιορίσουν τη διάδοση των λαθών. Αν ένα slice ληφθεί με λάθος παραλείπεται χωρίς να χάνουμε όλη την εικόνα. Γενικά τα slices βοηθούν στην απόκρυψη των σφαλμάτων (error concealment). Το κάθε macroblock περιέχει ένα πίνακα με πληροφορίες φωτεινότητας (16x16) και δύο πίνακες του μισού μεγέθους (8x8) για το χρώμα. (Εικόνα 3.3)

Η ακολουθία αυτή είναι δομημένη ιεραρχικά:

ακολουθία (sequence) > καρέ (frame) > slice (σειρά από blocks) >block (τετράγωνο των 8x8 pixels).

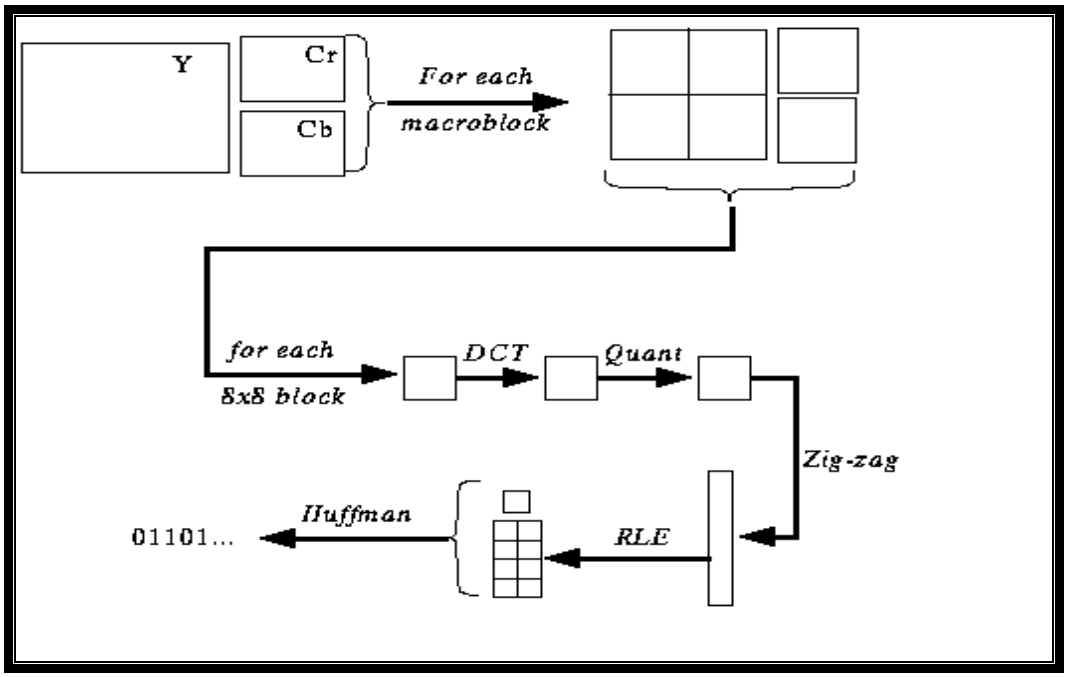


Εικόνα 3.3 - *Ανατομία του MPEG*

3.2. Δύο βασικούς τρόπους συμπίεσης MPEG

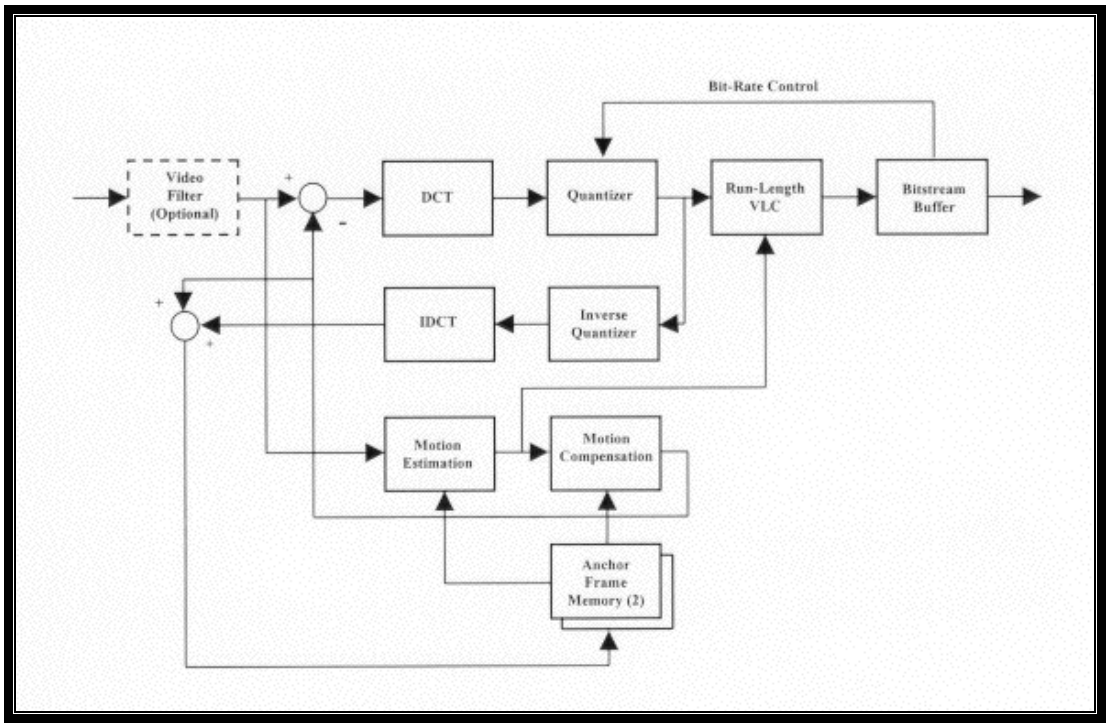
Δύο βασικές προσεγγίσεις στην κωδικοποίηση της κινούμενης εικόνας είναι οι παρακάτω. Η πρώτη είναι βασισμένη στη μέθοδο συμπίεσης ακίνητης εικόνας JPEG και η δεύτερη στην τμηματική πρόβλεψη κίνησης (block motion compensation).

Ενδοπλασιακά (intra-frame) Σύμφωνα με αυτή την προσέγγιση η κάθε εικόνα (frame) αντιμετωπίζεται σαν αυτόνομη μονάδα και κωδικοποιείται ανεξάρτητα από τις υπόλοιπες, οπότε το τελικό σήμα είναι μία σειρά από διακριτές ακίνητες εικόνες δηλαδή με τεχνικές συμπίεσης που εφαρμόζονται στην πληροφορία μόνο ενός frame. Το MPEG εφαρμόζει τον μετασχηματισμό DCT στην πληροφορία ενός frame και μέσω του κβαντισμού των συντελεστών DCT συμπιέζει την πληροφορία αυτήν. Τα frame που εφαρμόζονται κατά αυτόν τον τρόπο χαρακτηρίζονται ως I-frames και αποτελούν key –frames στην αλληλουχία των συμπιεστών πλαισίων που δημιουργεί το MPEG. (Εικόνα - 3.4)



Εικόνα 3.4 *- intra-frame coding*

<u>Διαπλαισιακά (inter-frame)</u> Λαμβάνονται κατά την κωδικοποίηση υπ' όψη οι πιθανές ομοιότητες μεταξύ των πλαισίων (frame) και κωδικοποιείται η διαφορά τους με χρήση της τμηματική πρόβλεψη κίνησης (block motion compensation). Ετσι στο τελικό σήμα υπάρχει μία εξάρτηση μεταξύ των πλαισίων αφού για να αποκωδικοποιηθεί κάποιο πιθανώς να πρέπει να ληφθούν πληροφορίες και από κάποιο άλλο (προηγούμενο ή επόμενο). Γίνεται έτσι πολύ μεγαλύτερη συμπίεση, αφού μόνο οι διαφορές μεταξύ των πλαισίων κωδικοποιούνται δηλαδή με τεχνικές συμπίεσης που εκμεταλλεύονται τον τρόπο που μεταβάλλεται η πληροφορία μεταξύ διαδοχικών πλαισίων. Μια τεχνική τις είναι η εκτίμηση κίνησης (motion estimation). Δύο διαδοχικά πλαίσια συνήθως δεν διαφέρουν πολύ. Έτσι η κωδικοποίηση κάθε επόμενου πλαισίου μπορεί να βασιστεί στην πληροφορία του προηγούμενου δηλαδή δεν χρειάζεται να αποθηκευτεί ολόκληρη η πληροφορία του επόμενου πλαισίου αλλά μόνο το τμήμα που διαφέρει από το προηγούμενο πλαίσιο που σίγουρα εμφανίζεται διαφοροποιημένο λόγο κάποιας κίνησης. (Εικόνα 3.5)



Εικόνα 3.5 - *inter-frame coding*

Για να εφαρμόσει το MPEG τη διαπλαισιακή συμπίεση δημιουργεί δύο νέα ειδή πλαισίων (frame)

- *Το πλαίσιο P (P-Frames, predicative frames)*
- *Το πλαίσιο B (B-Frames, bi-directional frames)*

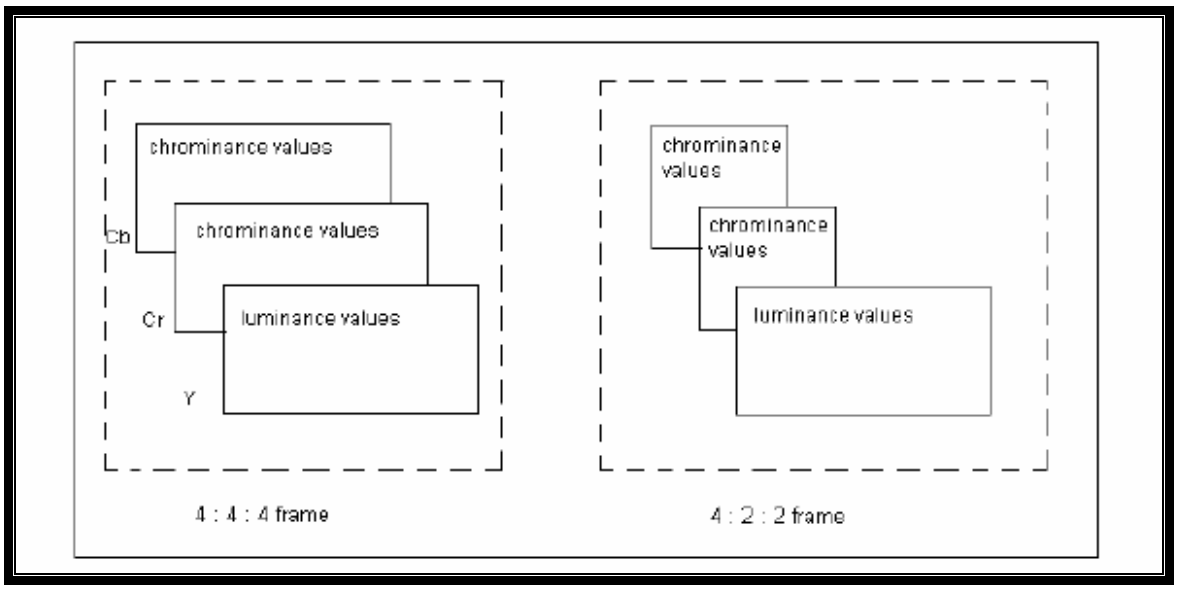
Τα όποια αναλύονται στην συνεχεία .

3.2.1. Τα τρία είδη των frames στο MPEG

Όλοι λίγο πολύ έχουμε ακούσει τον ορο frame rate. Ο όρος frame rate αναφέρεται στο ρυθμό ανανέωσης των αλλεπάλληλων εικόνων μίας οθόνης ανά δευτερόλεπτο. Αυτές οι εικόνες ονομάζονται frames (καρέ) και μετριούνται ανά δευτερόλεπτο (frames per second ή fps). Όσο πιο μεγάλος είναι αυτός ο αριθμός, τόσο πιο ομαλή είναι η κίνηση. Το ανθρώπινο μάτι αντιλαμβάνεται “σπάσιμο” στην κίνηση ενός μοντέλου, όταν ο ρυθμός ανανέωσης των καρέ της οθόνης ανά δευτερόλεπτο βρίσκεται κάτω των 30.

Ένα Frame περιέχει όλη την απαραίτητη πληροφορία για το χρώμα και την φωτεινότητα, έτσι ώστε να εμφανιστεί η εικόνα. Αυτή η πληροφορία οργανώνεται σε τρεις πίνακες που περιέχουν τις τιμές για την φωτεινοτητα “ luminance ” και χρωμα

"chrominance". Το μέγεθος αυτών των πινάκων διαφέρει ανάλογα με τον ρυθμό δειγματοληψίας χρώματος που εφαρμόζεται. Οι αντίστοιχοι πίνακες για ρυθμό 4:4:4 και 4:2:2 φαίνονται παρακάτω. (Εικόνα 3.6)



Εικόνα 3.6 - *Πίνακες που σχηματίζουν μια 4:4:4 και μια 4:2:2 εικόνα*

Πιο αναλυτικά: όπως είναι γνωστό για την κωδικοποίηση της εικόνας δε χρησιμοποιούνται τρία ανεξάρτητα σήματα για τις συνιστώσες R,G,B παρόλο που κάθε χρώμα μπορεί να περιγραφεί από αυτές και άλλο ένα για τη φωτεινότητα, αλλά στην πράξη χρησιμοποιείται μόνο αυτό της φωτεινότητας (Y) και δύο άλλα (Cr) και (Cb) που περιέχουν τις πληροφορίες για τα R,G,B αν συνδυαστούν με τη φωτεινότητα σύμφωνα με τους παρακάτω τύπους :

$$Y = 0,30R + O,59G + 0,14B$$
$$Cr = 0,74(R-Y) - 0,27(B-Y) = 0,60R + 0,28G + 0,32B$$
$$Cb = 0,48(R-Y) + 0,41(B-Y) = 0,21R + 0,52G + 0,31B \quad \text{(γιά το NTSC)}$$

$$Y = 0,30R + O,59G + 0,11B$$
$$Cr = 0,493(B-Y) = -0,15R - 0,29G + 0,44B$$
$$Cb = 0,877(R-Y) = 0,62R - 0,52G - 0,10B \quad \text{(για το PAL)}$$

Η ανθρώπινη όραση είναι λιγότερο ευαίσθητη στις χρωματικές μεταβολές παρά σε εκείνες της φωτεινότητας έτσι χρειάζονται λιγότερα δεδομένα χρωματικής πληροφορίας για να δημιουργηθεί η αίσθηση της ποιοτική εικόνας στον θεατή.

Έτσι κατά την δειγματοληψία μπορούν να ψηφιοποιηθούν λιγότερα δείγματα από τις συνιστώσες χρώματος (Cb και Cr) σε σχέση με το σήμα φωτεινότητας (luma σήμα Y). Είναι σημαντικός παράγοντας για την μείωση όγκου. Η αναλόγια δειγμάτων που παράγονται μεταξύ των σημάτων εκφράζεται ως ένας διπλός λόγος ακέραιων αριθμών όπου ο πρώτος είναι πάντοτε το 4. Αν δεν υπάρξει δειγματοληψίας ο λόγος αυτός είναι 4:4:4 . Οι πιο συνηθισμένες αναλογίες είναι 4:2:2 και 4:2:0. Κατά την αποκωδικοποίηση η πληροφορία χρώματος που λείπει αναπαράγεται με τον μέσο όρο των γειτονικών pixel.

Σε ένα σήμα MPEG υπάρχουν τριών ειδών πλαίσια (frame) :

- **I-frames (Intra frames)**

Τα I-frames είναι εικόνες που δεν χρησιμοποιούν εκτίμηση κίνησης , αλλά μεταδίδουν συνολικά την πληροφορία της εικόνας σε μορφή κωδικοποιημένη σύμφωνα με το πρότυπο JPEG.
Η διαδικασία που ακολουθείται είναι η εξής:

Η εικόνα χωρίζεται σε υποεικόνες (blocks) των 8x8 ή 16x16 pixels και εφαρμόζεται ο διακριτός μετασχηματισμός συνημιτόνου (Discrete Cosine Transformation - DCT) σε καθεμιά από αυτές. Στην μετασχηματισμένη πλέον εικόνα στο DCT πεδίο εφαρμόζεται η διαδικασία του κβαντισμού και της συμπίεσης-κωδικοποίησης (Huffman), οπότε τελικά κατασκευάζεται η εικόνα σε συμπιεσμένη μορφή, αλλά με όλη την πληροφορία και γι' αυτό χρησιμοποιείται και ως πλαίσιο αναφοράς για τα υπόλοιπα είδη πλαισίων. Άρα τα πλαίσια τύπου I είναι τα μόνα που είναι κωδικοποιημένα στο σύνολό τους και η αποκωδικοποίηση μπορεί να γίνει χωρίς αναφορά σε κάποιο άλλο. Εν ολίγοις η εικόνα χωρίζεται σε macroblocks και για κάθε block ξεχωριστά εφαρμόζεται DCT, Κβαντοποίηση, Zig-Zag Scanning,Run-Length-Encoding και Huffman Encoding .

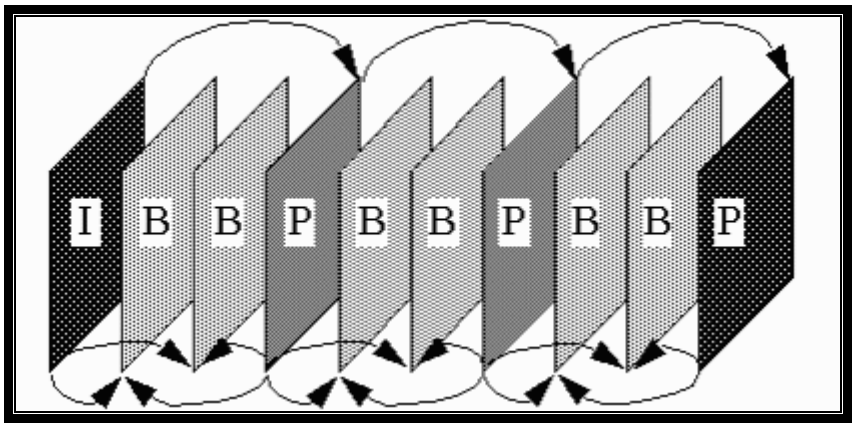
- **P-frames (Predicted frames)**

Στα P-frames χρησιμοποιείται η διαδικασία της εκτίμησης και αντιστάθμισης κίνησης (motion estimation and compensation), δηλαδή μεταδίδεται η πληροφορία της κίνησης των blocks από το προηγούμενο πλαίσιο αναφοράς (το προηγούμενο I ή P

frame) προς το τρέχον πλαίσιο P, καθώς και το σφάλμα που προκύπτει από αυτή τη διαδικασία. Με τη σειρά τους μπορούν να αποτελέσουν και αυτά σημείο αναφοράς για επόμενα πλαίσια και αυτός είναι και ο λόγος που συμβάλλουν στην εισαγωγή και διάδοση σφαλμάτων, αφού η διαδικασία της πρόβλεψης κίνησης δεν μπορεί να είναι 100% ακριβής. Η πληροφορία της κίνησης αναπαρίσταται μέσω των διανυσμάτων κίνησης, ενώ το σφάλμα μεταδίδεται με τη μορφή κωδικοποιημένης με τον DCT μετασχηματισμό εικόνας. Υπάρχει σημαντική ελευθερία στον τρόπο εξαγωγής της πληροφορίας κίνησης, γεγονός που επιτρέπει συνεχώς τη βελτίωση της ποιότητας της κωδικοποίησης και της εκμετάλλευσης των ιδιοτήτων του προτύπου. Δεν έχουν το μέγεθος των I πλαισίων γιατί δεν έχουν περιγραφεί με την ίδια ακρίβεια, δηλαδή παρουσιάζουν μεγαλύτερο ποσοστό συμπίεσης.

Η διαδικασία λοιπόν της κωδικοποίησης τους, είναι:

Σύγκριση macroblocks και δημιουργία ενός γραμμικού συνδυασμού αυτών που παρουσιάζουν σημαντική ομοιότητα, δημιουργία motion vector, μετασχηματισμός DCT σε κάθε block του νέου macroblock, Κβαντοποίηση, Run-Length-Encoding και το αποτέλεσμα κωδικοποιείται με κωδικοποίηση Huffman, όπως και στα I πλαίσια. Στην Εικόνα 3.7 φαίνεται ο τρόπος της αναφοράς για την προς τα εμπρός πρόβλεψη (forward prediction) των P-frames.

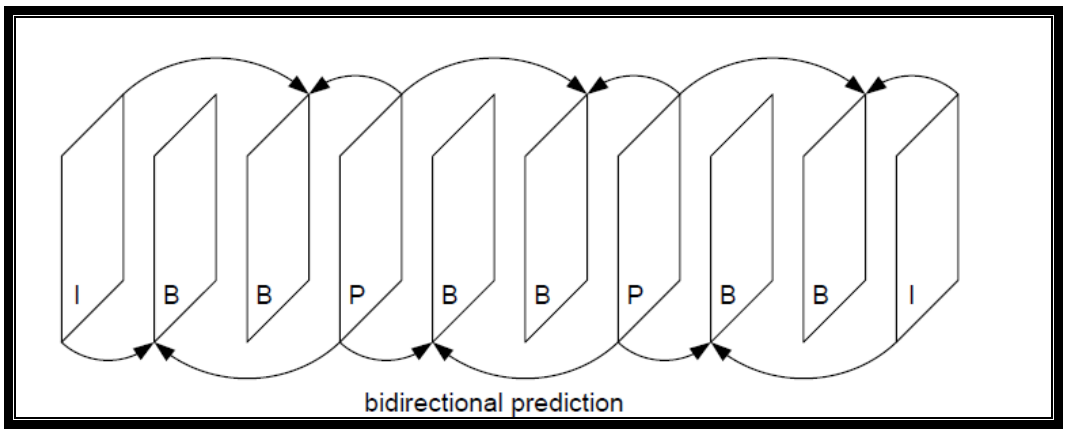


Εικόνα 3.7 - *Προς τα εμπρός πρόβλεψη στα P-frames του MPEG*

- **B-frames (Bidirectional Interpolated frames)**

Στα B-frames πραγματοποιείται και πάλι η διαδικασία εκτίμησης και αντιστάθμισης κίνησης, μόνο που εδώ χρησιμοποιείται το προηγούμενο πλαίσιο αναφοράς (το προηγούμενο I ή P frame) ή το επόμενο πλαίσιο αναφοράς (το επόμενο I ή P frame) ή

ακόμη και τα δύο ταυτόχρονα, ενώ το frame συντίθεται μέσω διαδικασίας χώρο-χρονικά ισοσταθμισμένης γραμμικής παρεμβολής. Δε συντελούν τόσο πολύ στη διάδοση των σφαλμάτων γιατί δεν χρησιμοποιούνται ως σημεία αναφοράς και επιπλέον μειώνουν σημαντικά το σφάλμα παίρνοντας το μέσο όρο από δύο πλαίσια. Μπορούμε να πούμε ότι ο 'κύκλος της ζωής' τους περιορίζεται μόνο σε αυτά και δεν επεκτείνεται με το να κληροδοτούν πληροφορίες σε άλλα πλαίσια, κάτι που πολλές φορές σε συνδυασμό και με την υπολογιστική πολυπλοκότητα που απαιτούν για την κωδικοποίηση και αποκωδικοποίηση τα κάνει μη επιθυμητά από τους κατασκευαστές. Η πληροφορία της κίνησης αναπαρίσταται μέσω διανυσμάτων κίνησης, που εδώ είναι δύο ειδών: διανύσματα προς τα εμπρός κίνησης (forward motion vectors) με αναφορά στο προηγούμενο πλαίσιο αναφοράς (forward prediction), και διανύσματα προς τα πίσω κίνησης (backward motion vectors) με αναφορά στο επόμενο πλαίσιο αναφοράς (backward prediction). Η διαδικασία της αναφοράς για την κίνηση προς τις δύο κατευθύνσεις (bidirectional prediction) για τα B-frames φαίνεται στο Εικόνα 3.8 .



Εικόνα 3.8 - *Πρόβλεψη διπλής κατεύθυνσης στα B-frames του MPEG*

Ο κύριος λόγος ύπαρξης των B-πλαισίων είναι η κάλυψη της περίπτωσης κάποιες πληροφορίες της εικόνας να υπάρχουν σε επόμενα πλαίσια και να μην υπάρχουν στα προηγούμενα. Συνεπώς η πρόβλεψή τους με τα P πλαίσια θα ήταν αδύνατη. Αφού τα πλαίσια P κατασκευάζονται με βάση τα I και τα B με βάση τα I και P είναι προφανές ότι τα I πρέπει να έχουν σταλεί πριν τα αντίστοιχα P. Επίσης και τα P αλλά και τα I πρέπει να έχουν σταλεί πριν από τα αντίστοιχα B, παρόλο που στη μετάδοση αυτά παρεμβάλλονται ανάμεσά τους.

Η μικρότερη μονάδα που μπορεί να αποκωδικοποιηθεί ανεξάρτητα ονομάζεται GOP (Group of Pictures) και περιέχει όλα τα I,P,B πλαίσια που χρειάζονται για την αποκωδικοποίηση, χωρίς να γίνονται αναφορές σε άλλο GOP.

Ομάδα πλαισίων (GOP)

Group of Pictures ονομάζεται το σύνολο των frame που βρίσκονται ανάμεσα στο key-frame (τα καρέ από τα οποία εξαρτώνται τα B frames) του τρέχοντος GOP, και του key-frame του προηγούμενου GOP. Ένα πλαίσιο I μαζί με τα πλαίσια P και B που το ακολουθούν αποτελεί μια ομάδα πλαισίων (Group of picture ,GOP). Το μέγεθος μιας τέτοιας ομάδας μπορεί να ρυθμιστεί σε 8 , 12 η 16 πλαίσια ανάλογα με το πιο μέγεθος οδηγεί σε καλύτερη κωδικοποίηση. Κάθε αρχείο video mpeg αποτελείται από μια τέτοια ακολουθία (GOPs) που με την σειρά αποτελούνται από πλαίσια (frames). Κάθε πλαίσιο αναλύεται σε macroblocks και κάθε macroblock μπορεί να αποτελείται από 6 απλά block (4 για το σήμα luma και 3 για το chroma υπό-δειγματοληψία 4:2:0) ή και λιγότερα. Η οργάνωση με βάση τις ομάδες (GOPs) επιτρέπει την τυχαία προσπέλαση στην πληροφορία του αρχείου καθώς μια ομάδα αποτελεί την ανεξάρτητα αποκωδικοποιημένη μονάδα πληροφορίας μέσα στο αρχείο.

GOP (Group of Pictures):

- ✓ Η μικρότερη μονάδα που αποκωδικοποιείται ανεξάρτητα.
- ✓ Περιέχει όλα τα I,P,B πλαίσια που χρειάζονται για αποκωδικοποίηση (χωρίς αναφορά σε άλλα GOP)

3.2.2. Κομμάτια Slices

Ένα MPEG Frame χωρίζεται σε κομμάτια (slices). Ένα slice αποτελείται από μια σειρά από macroblocks ακολουθώντας μια σάρωση του frame από αριστερά προς τα δεξιά και από πάνω προς τα κάτω. Στην κωδικοποιημένη ροή των bits κάθε slice ξεκινάει με μια επικεφαλίδα η οποία δεν έχει κωδικοποιηθεί όπως γίνεται σε όλα τα δεδομένα που υπάρχουν στην ροή μετάδοσης bits και συνεπώς είναι το χαμηλότερο επίπεδο ενοτήτων στη κωδικοποιημένη ροή που μπορεί να έχει κάποιος πρόσβαση χωρίς να χρειαστεί αποκωδικοποίησης. Η παρουσία τμημάτων δεν συμβάλλει στη συμπίεση της εικόνας ,παρά μόνο στην ύπαρξη πληροφορίας συγχρονισμού μέσα σε μία εικόνα σε ότι αφορά την απεικόνισή της. Σε περίπτωση ενός λάθους στη μετάδοση, και την απώλεια πληροφορίας σχετικά με την εικόνα, η διεργασία δεν απορρίπτει ολόκληρο το frame , παρά χρησιμοποιεί την πληροφορία στο τμήμα και συνεχίζει με την αποκωδικοποίηση του επόμενου block. Συμπεραίνουμε ότι έχοντας πολλά blocks σε μία εικόνα, έχουμε μεγάλη ανεκτικότητα σε λάθη. Παράλληλα όμως, αυτό

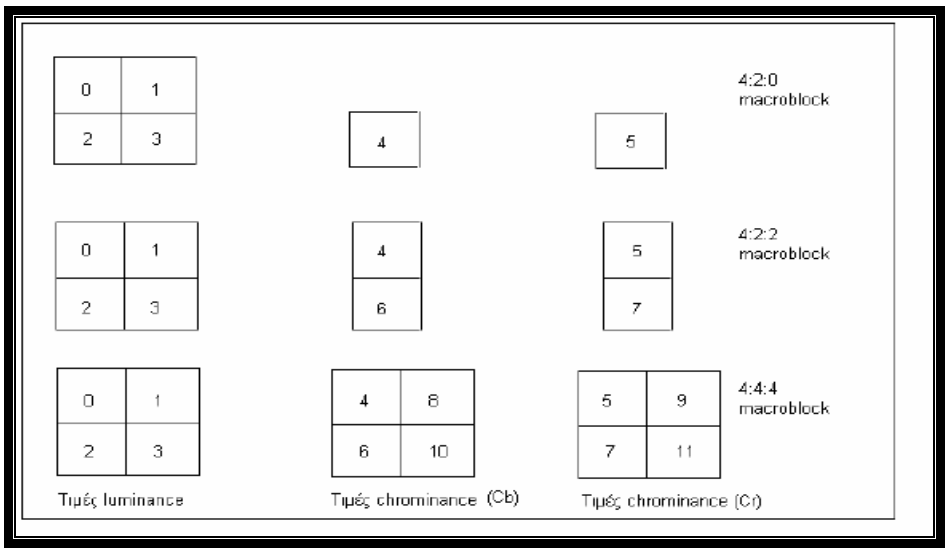
συνεπάγεται ένα μεγάλο overhead για κάθε εικόνα, οπότε πρέπει να βρούμε την χρυσή τομή ανάμεσα σε ποιότητα, παρουσία λαθών και απαιτούμενο εύρος ζώνης(bandwidth).

3.2.3. Τμήματα - Μακροτμήματα Blocks – Macroblocks

Τα blocks και τα macroblocks συνθέτουν μία ιεραρχία στη δομή του MPEG : κάθε frame διαιρείται σε έναν αριθμό blocks, τα οποία ομαδοποιούνται σε macroblocks και αυτά με την σειρά τους σε slices. Κάθε block αποτελείται από οχτώ γραμμές και κάθε γραμμή περιέχει οχτώ δείγματα τιμών luminance και chrominance από ένα εικονοπλαίσιο(pixel). Έτσι συνολικά, έχουμε 64 τιμές chrominance ή luminance σε ένα block, κάθε block συνεπώς αντιπροσωπεύει 64 εικονοστοιχεία(pixels). Κάθε macroblock περιέχει σταθερά τέσσερα blocks με τιμές luminance και έναν αριθμό από blocks με τιμές chrominance. Ο αριθμός των blocks chrominance σε ένα macroblock εξαρτάται από το ρυθμό δειγματοληψίας χρώματος

Έτσι υπάρχουν :

• macroblocks ρυθμού 4:2:0 : Περιέχουν δύο blocks με πληροφορία chrominance.
• macroblocks ρυθμού 4:2:2 : Περιέχουν τέσσερα blocks με πληροφορία chrominance.
• macroblocks ρυθμού 4:4:4 : Περιέχουν οχτώ blocks με πληροφορία chrominance.
Η παρακάτω εικόνα αποσαφηνίζει τη δομή ενός macroblock για τους τρεις ρυθμούς δειγματοληψίας χρώματος.



Εικόνα 3.9 - *Πίνακας δειγματοληψίας*

Στη Εικόνα 3.9 εφαρμόζεται ο μετασχηματισμός DCT και οι υπόλοιπες διαδικασίες συμπίεσης καθώς και η διαδικασία της πρόβλεψης της κίνησης.

3.3. Τα βήματα της συμπίεσης κατά το MPEG

1→Αναλύεται το αρχείο video για να καθορίσει το ποια πλαίσια θα κωδικοποιηθούν ως πλαίσια I ,P και B. Επίσης καθορίζεται και το μέγεθος των ομάδων πλαισίων(GOPs) και ο μέγιστος και ελάχιστος ρυθμός δεδομένων (bit rate).

2→Τα πλαίσια διαιρούνται σε macro-blocks (16X16pixels).

3→Το σήμα μετατρέπεται από RGB σε YUV ενώ ταυτόχρονα γίνεται και υποδειγματοληψία χρώματος. Το MPEG εφαρμόζει το σχήμα 4:2:0. Έτσι από κάθε macro-block 16X16 προκύπτουν τελικά τέσσερα 8X8 block με πληροφορία φωτεινότητας (Y) και δύο με πληροφορία χρώματος (ένα U και ένα V).

4→Εφαρμόζει την τεχνική αντιστάθμιση κίνησης σε κάθε πλαίσιο P και B. Δηλαδή υπολογίζει τα διανύσματα κίνησης και τούς συντελεστές πιθανού σφάλματος για κάθε macro-block.

5→Σε κάθε πλαίσιο I, P και B τα macro-blocks συμπιέζονται κατά JPEG δηλαδή

A) Εφαρμογή μετασχηματισμού DCT

B) Κβαντισμος ασυντέλεστων DCT

Γ) Συμπίεση RLE των συντελεστών DCT κάθε macro-block

6→Στην μορφή πληροφορίας που προκύπτει εφαρμόζεται και ο αλγόριθμος συμπίεσης Huffman.

Από το παραπάνω βλέπουμε ότι οι βασικές φάσεις συμπίεσης στην κωδικοποίηση κατά MPEG είναι :

1→Υποδειγματοληψία χρώματος (chroma sub sampling)

2→Η κβάντωση των συντελεστών DCT

3→Η αντιστάθμιση κίνησης

Επίσης ο συμπιεστής MPEG είναι ασύμμετρος συμπιεστής δηλαδή η κωδικοποίηση χρειάζεται περισσότερο χρόνο από την αποκωδικοποίηση .

3.4. Το πρότυπο κωδικοποίησης MPEG – 1

Το αποτέλεσμα των εργασιών των πρώτων συνεδριάσεων της επιτροπής MPEG (1988), ήταν η δημιουργία του πρώτου προτύπου MPEG-1 (ISO/IEC 11172 '92). Με το MPEG-1 παράγεται μια ανάλυση εικόνας 352x240 pixels με τα καρέ να έχουν μια ταχύτητα της τάξεως των 30 καρέ ανά δευτερόλεπτο (30fps). Το πρότυπο αυτό αφορά εγγραφή και αναπαραγωγή κινητού video , που χρησιμοποιείται σε εφαρμογές ατομικών υπολογιστών. Η ποιότητα της εικόνας video που παράγει, είναι λίγο κατώτερη των συμβατικών VCR (video). Οι εφαρμογές αυτού του προτύπου περιορίζονται σε μορφές οπτικού σήματος, που δεν έχει αλληλοδιαδοχή των πεδίων και υποστηρίζει κωδικοποιήσεις οπτικών σημάτων μέχρι 1,5 Mbit/s. Η πιο δημοφιλής χρήση του MPEG-1 είναι στα δισκάκια VCD και όπου αλλού χρειάζεται μικρό σε σχέση με το MPEG-2 bandwidth.

3.5. Το πρότυπο κωδικοποίησης MPEG-2

Το πρωτόκολλο MPEG – 2 παρουσιάστηκε ως επίσημο πρωτόκολλο αρχικά το Νοέμβριο του 1994 στη Σιγκαπούρη. Η πρώτη αυτή έκδοση περιελάμβανε μόνο τα τρία πρώτα κεφάλαια του πρωτοκόλλου. Τα υπόλοιπα οκτώ κεφάλαια παρουσιάστηκαν ένα χρόνο αργότερα. Η ανάπτυξη αυτού του νέου πρωτοκόλλου είχε ως κύριο στόχο την έλλειψη κάποιων δυνατοτήτων από το MPEG-1, οι οποίες άρχισαν να ζητούνται από τη βιομηχανία ψηφιακής εικόνας. Το βασικότερο πρόβλημα που ανέκυπτε στο MPEG – 1 ήταν ότι η ποιότητα του προτύπου VHS που απέδιδε ήταν χαμηλή σε σχέση με την ποιότητα των πρότυπων της τηλεόρασης PAL και του HDTV που χρησιμοποιούνταν εκείνη την εποχή από τη τηλεόραση. Επίσης, οι μελέτες μέχρι τότε έδειχναν ότι ο αλγόριθμος που χρησιμοποιούσε το MPEG-1 ήταν αρκετά αναποτελεσματικός σε σχέση με τους αλγορίθμους που είχαν αναπτυχθεί εκείνη την εποχή για τη συμπίεση video , ενώ παράλληλα ο νέος αλγόριθμος συμπίεσης του MPEG–2 θα έπρεπε να είναι συμβατός με αυτόν του MPEG–1 ώστε να υπάρχει μια συνέχεια μεταξύ των δύο πρωτοκόλλων. Παρόμοιο πρόβλημα αντιμετώπιζε η ομάδα ανάπτυξης και με τον ήχο. Για να σιγουρευτεί μια αρμονική λύση, που θα κάλυπτε μια ευρεία γκάμα εφαρμογών, η ομάδα εργασίας ανέπτυξε το MPEG–2 σε στενή συνεργασία με τις ομάδες μελέτης ITU-TS Study Group και Expert Group for ATM Video Coding. Εκτός των δύο αυτών

ομάδων, άλλοι αντιπρόσωποι των σχεδιαστών των ITU-TS, EBU, ITU-RS, SMPTE και North America HDTV εργάστηκαν επίσης στο MPEG–2.

Με λίγα λόγια η επιτυχία του προτύπου MPEG-1 οδήγησε στη δημιουργία του επόμενου προτύπου MPEG-2 το οποίο προσπάθησε να φέρει μια πιο γενική περιγραφή των μεθόδων κωδικοποίησης εικόνας και ήχου, με περισσότερες επιλογές διαβάθμισης της ποιότητας, το οποίο θα στόχευε κυρίως σε εφαρμογές μεσαίας και υψηλής ποιότητας. Άρα μιλάμε για ένα πρότυπο κωδικοποίησης video και audio για καλύτερη ποιότητα μετάδοσης στις τηλεοράσεις. Χρησιμοποιείται στις περισσότερες ταινίες DVD.

3.5.1. Οργάνωση προτύπου

Για να κατανοήσουμε την οργάνωση του αρχικά θα πρέπει να δεχτούμε ότι το βασικό στοιχείο της κωδικοποιήσεως του είναι η εικόνα, η οποία υφίσταται την ενδοπλαισιακή συμπίεση. Για να εφαρμοσθεί λοιπόν η συμπίεση , οι εικόνες οργανώνονται όπως ήδη αναφέραμε σε ομάδες εικόνων (GOP), συνήθως 12 στον αριθμό, όμως μόνο μία στις δώδεκα, που ονομάζεται και πλαίσιο I, (από το intra-frame), προέρχεται από όλα τα στοιχεία της εικόνας δεν χάνει δηλαδή καμία πληροφορία. Οι υπόλοιπες σχηματίζονται από την εκπομπή στοιχείων που περιέχουν τις διαφορές με τις προηγούμενες και ονομάζονται προβλεπόμενα πλαίσια (Predicted frames). Τα intra-frame τα οποία περιέχουν πολλές πληροφορίες εκπέμπονται με πολύ μικρή συχνότητα, ενώ τα υπόλοιπα πλαίσια που περιέχουν πιο λίγες εκπέμπονται με μεγαλύτερη συχνότητα. Με τον τρόπο αυτό επιτυγχάνεται μεγάλη συμπίεση (διαπλαισιακή). Δηλαδή η συμπίεση στο MPEG-2 επιτυγχάνεται με την ενδοπλαισιακή και την διαπλαισιακή συμπίεση. Πιο αναλυτικά όμως .

Στην ενδοπλαισιακή συμπίεση MPEG-2 η εικόνα διαιρείται και επεξεργάζεται σε δύο διαστάσεις με ομάδες των 8 pixel οριζοντίως και 8 pixel καθέτως. Δηλαδή αν η εικόνα των 625 γραμμών περιέχει 720 στοιχεία οριζοντίως και 576 καθέτως, η εικόνα θα περιέχει συνολικά να αποτελείται από 90 ομάδες οριζοντίως και 72 κατακορύφως, δηλαδή η εικόνα θα περιέχει 90*72=6480 ομάδες των 8*8=64 pixel. Στο σύστημα MPEG χρησιμοποιείται ο διακριτός μετασχηματισμός του συνημιτόνου (DCT). Με το μετασχηματισμό αυτό μετασχηματίζονται και απεικονίζονται τα στοιχεία της ομάδας pixel από το διάστημα του χώρου σ’ ένα διάστημα συχνότητας στο χώρο. Με το μετασχηματισμό αυτό δεν μειώνουμε τον αριθμό των δυφίων, που απαιτούνται για να παραστήσουμε την ομάδα. Στην πράξη για ομάδα 8*8 pixel των 8 δυφίων , ο μετασχηματισμός παράγει ομάδα 8*8 συντελεστών, στην οποία ο κάθε συντελεστής είναι κωδικοποιημένος με 11 δυφία για να έχουμε αναστρέψιμη διαδικασία. Η μείωση στον αριθμό των δυφίων προκύπτει γιατί οι ομάδες φυσικών και συνήθων εικόνων η

κατανομή των ομοίων pixel δεν είναι ομοιόμορφη. Με το μετασχηματισμό η ενέργεια συγκεντρώνεται στους συντελεστές των χαμηλών συχνοτήτων, ενώ οι άλλοι συντελεστές έχουν σχεδόν μηδενική τιμή. Η μείωση στη ροή των δυφίων εξασφαλίζεται επίσης και με τη κβαντοποίηση και με τη μη εκπομπή των συντελεστών, που έχουν πολύ μικρή τιμή, όπως επίσης και με την κωδικοποίηση των υπολοίπων συντελεστών. Θα πρέπει να τονίσουμε ότι στο μετασχηματισμό του συνημιτόνου η έννοια της συχνότητας δεν αναφέρεται στο χρόνο αλλά στο μήκος. Η οριζόντια συχνότητα αυξάνει από αριστερά προς τα δεξιά και η κατακόρυφη συχνότητα από πάνω προς τα κάτω. Όταν έχουμε μεγάλο πλεονασμό στο χώρο, τότε οι περισσότεροι συντελεστές θα είναι μηδενικοί.

Αν όλα τα στοιχεία της ομάδας είναι ομοιόμορφα, θα υπάρχει ένας μόνο συντελεστής στην επάνω αριστερή γωνία και θα παριστά τη συνεχή συνιστώσα του οπτικού σήματος. Μετά το μετασχηματισμό οι συντελεστές κβαντοποιούνται, ώστε να μειώσουμε τον αριθμό δυφίων που απαιτούνται για την εκπομπή. Για κάθε συντελεστή εφαρμόζεται διαφορετική κβαντοποίηση. Η κβαντοποίηση εισάγει σφάλμα και παραμορφώσεις στην εικόνα. Εν γένει είναι ανεκτό μεγαλύτερο σφάλμα κβαντοποίησης για τις υψηλές συχνότητες, όπως επίσης στα σήματα χρωματισμού. Μετά την κβαντοποίηση οι συντελεστές σαρώνονται σε σχήμα zing–zang , και αυτοί που παράγονται από την σάρωση, κωδικοποιούνται με μέτρηση των συντελεστών που έχουν μηδενική τιμή και προηγούνται κάποιου που δεν έχει μηδενική τιμή. Η κωδικοποίηση αυτή καλείται κωδικοποίηση μέτρησης μήκους (run-length code). Επειδή πολλές από τις τιμές των συντελεστών είναι μηδενικές, η χρήση του κώδικα αυτού μειώνει το μέγεθος του απαιτούμενου κώδικα για την περιγραφή των μηδενικών. Συνεπώς ο κώδικας μέτρησης μήκους αποτελείται από δύο αριθμούς. Ο πρώτος δίνει τον αριθμό των μηδενικών που και ο δεύτερος το μέγεθος του αμέσως επόμενου συντελεστή που δεν έχει μηδενική τιμή. Η κωδικοποίηση μέτρησης μήκους υφίσταται περαιτέρω κωδικοποίηση με έναν κώδικα που είναι γνωστός με το όνομα κώδικας μεταβλητού μήκους (variable length code). Ο συγκεκριμένος κώδικας καθορίζει άλλους που έχουν μεταβλητό μήκος, το οποίο εξαρτάται από τη συχνότητα εμφάνισης του κάθε συνδυασμού (αριθμού προηγουμένων μηδενικών και μεγέθους επομένου συντελεστή). Οι συχνά εμφανιζόμενοι συνδυασμοί έχουν μικρό μήκος ενώ οι σπανιότερα εμφανιζόμενοι μεγαλύτερο μήκος.

3.5.2. Επίπεδα και προφίλ του προτύπου MPEG-2

Το πρότυπο MPEG-2 υποστηρίζει έναν αριθμό από προφίλ που στοχεύουν σε μια συγκεκριμένη περιοχή εφαρμογών και καθορίζουν το σύνολο των εργαλείων συμπίεσης

που χρησιμοποιούνται. Υποστηρίζονται τέσσερα προφίλ, εκ των οποίων το κύριο είναι για χρήση γενικού σκοπού, το απλό προφίλ είναι παρόμοιο με το κύριο, ενώ τα άλλα διαπραγματεύονται την επεκτασιμότητα (scalability) και την τηλεόραση υψηλής ευκρίνειας (HDTV). Τα προφίλ διαφέρουν όσον αφορά την παρουσία ή απουσία εικόνων B, την ανάλυση της χρωματικότητας καθώς και την επεκτασιμότητα ενός κωδικοποιημένου συρμού προς τα άλλα format εικόνων. Επιπροσθέτως, το πρότυπο MPEG-2 υποστηρίζει τέσσερα επίπεδα (levels) που ορίζουν την ανάλυση της εικόνας. Το χαμηλό επίπεδο (352 x 288) αντιστοιχεί στην προς τα πίσω συμβατότητα με το πρότυπο συμπίεσης MPEG-1 και στους video εγγραφείς (VCR), το κύριο επίπεδο (720 x 576) αντιστοιχεί στην τυποποιημένη ανάλυση για τη μετάδοση NTSC, το υψηλό-1440 επίπεδο (1440 x 1152) αποτελεί την HDTV και το υψηλό επίπεδο (1920 x 1152) βελτιστοποιεί την HDTV. (Εικόνα 3.10)

Levels	Profiles				
	Simple Main without B-frames 4:2:0	**Main** B-frames 4:2:0	**SNR Scalability** 4:2:0	**Spatial Scalability** 4:2:0	**Professional** 4:2:2
High 1920 x 1152	x	80 Mbit/s	x	x	100 Mbit/s
High -1440 1440 x 1152	x	60 Mbit/s	x	60 Mbit/s	80 Mbit/s
Main 720 x 576	15 Mbit/s	15 Mbit/s 90% of users	15 Mbit/s	x	20 Mbit/s
Low 352 x 288	x	4 Mbit/s	4 Mbit/s	x	x

Εικόνα 3.10 - *Επίπεδα που ορίζουν την ανάλυση μιας εικόνας στο πρότυπο MPEG-2*

Κλιμακωσιμότητα στο MPEG-2

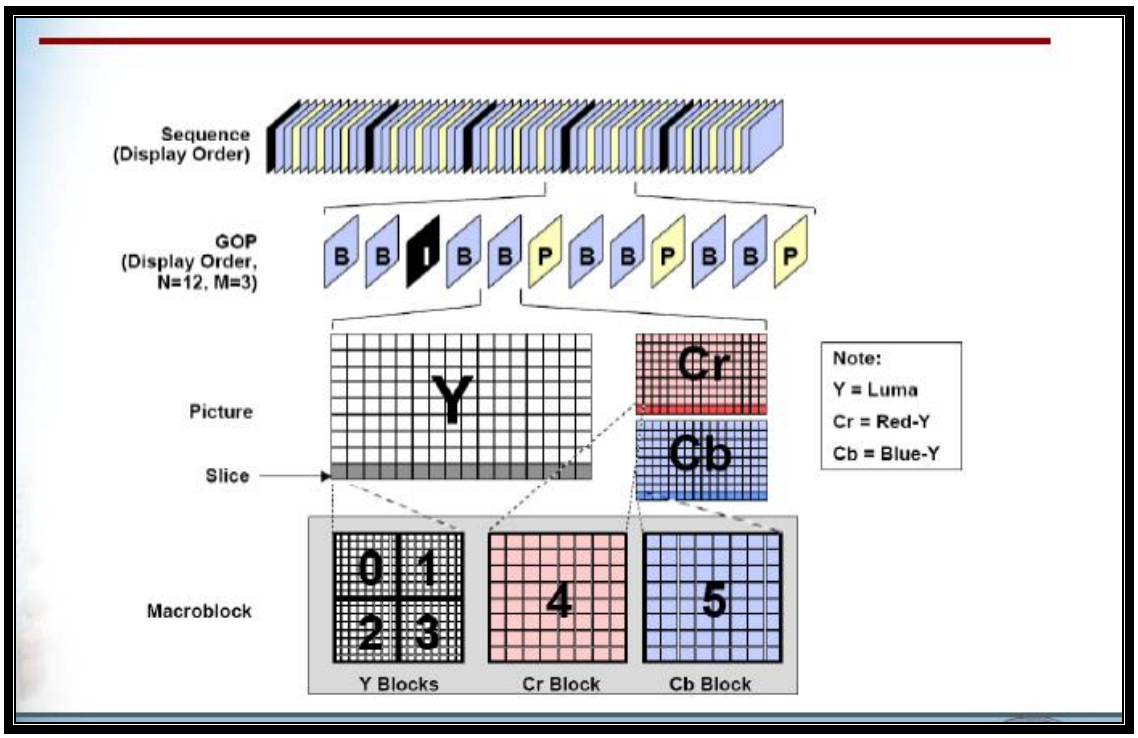
Χωρική κλιμακωσιμότητα (Spatial scalability): δύο επίπεδα video με διαφορετικές χωρικές αναλύσεις (κωδικοποίηση πυραμίδας).

Χρονική κλιμακωσιμότητα (Temporal scalability): δύο επίπεδα video με το ένα να κάνει χρονική πρόβλεψη ως προς το άλλο.

Τεμαχισμός δεδομένων (Data partitioning): χωρισμός σε κρίσιμα δεδομένα (π.χ. επικεφαλίδες, διανύσματα κίνησης κλπ.) και λιγότερο κρίσιμα (π.χ. ψηλοί DCT συντελεστές).

Κλιμακωσιμότητα εύρους (SNR scalability): δύο επίπεδα video με ίδια χωρική ανάλυση αλλά διαφορετική ποιότητα (προοδευτική ή κατά επίπεδα κωδικοποίηση).

Το πρότυπο συμπίεσης MPEG-2 αποτελεί τη βάση των προτύπων που έχουν αναπτυχθεί για την μετάδοση ψηφιακής τηλεόρασης, σε παγκόσμιο επίπεδο. (Εικόνα 3.11)



Εικόνα 3.11 *- Δομή του MPEG-2 bit stream*

3.5.3. Ατέλειες του MPEG-2

Παρόλο που κάθε τύπος μιας σειράς προγράμματος αποτελείται από ένα μοναδικό σύνολο από παραμέτρους εικόνας, συγκεκριμένες γενικεύσεις μπορούν να γίνουν, όσον αφορά τις ατέλειες, που ενδεχομένως να παρουσιαστούν ως αποτέλεσμα του προτύπου συμπίεσης MPEG-2.

Οι ατέλειες συμπεριλαμβάνουν τα ακόλουθα φαινόμενα:

«Μπλοκ εφέ»: Αυτά μπορούν να γίνουν ορατά, όταν το μάτι εντοπίσει ένα αντικείμενο που κινείται γρήγορα κατά μήκος της οθόνης. Τα μπλοκ κομμάτια παραμένουν σταθερά, ακόμα και όταν το αντικείμενο περνάει κινούμενο από μέσα τους. Αυτό το φαινόμενο προκαλείται από κακή εκτίμηση της κίνησης ή μικρή κατανομή bit στον κωδικοποιητή.

«Θόρυβος κουνουπιού»: Αυτό το πρόβλημα μπορεί να παρατηρηθεί στις άκρες κειμένων ή λογοτύπων που παρουσιάζονται στην οθόνη, αλλά και άλλων αντικειμένων που έχουν αιχμηρές άκρες. Οι αιχμηρές άκρες προκαλούν όρους υψηλής συχνότητας του μετασχηματισμού DCT, οι οποίοι κβαντίζονται λανθασμένα και απεικονίζονται στην οθόνη δημιουργώντας αυτό το πρόβλημα.

«Βρώμικο παράθυρο»: Το συγκεκριμένο πρόβλημα εμφανίζεται σα λωρίδα ή θόρυβος που παραμένει σταθερός καθώς αντικείμενα κινούνται από κάτω του. Σ' αυτήν την περίπτωση, ο κωδικοποιητής μπορεί να μη στέλνει αρκετά bit για να κωδικοποιηθεί το παραμένων λάθος στα P και B πλαίσια.

«Κυματιστός θόρυβος»: Αυτό το πρόβλημα γίνεται ορατό κατά τη διάρκεια αργών πλάνων σε σκηνές με πολλές λεπτομέρειες, όπως το πλήθος σε ένα στάδιο. Η άκομψη κβάντιση των όρων υψηλής συχνότητας προκαλείται από τέτοιες εικόνες.

Έτσι, συγκεκριμένοι τύποι κίνησης δεν ταιριάζουν στο MPEG-2 μοντέλο και παρουσιάζονται προβληματικοί. Αυτοί οι τύποι κίνησης περιλαμβάνουν:

- ✓ Ζουμ
- ✓ Περιστροφές
- ✓ Συμπαγή κινούμενα αντικείμενα

Επίσης, συγκεκριμένα στοιχεία εικόνας δεν μπορούν να προβλεφθούν σωστά. Αυτά τα στοιχεία είναι:

- ✓ Σκιές
- ✓ Αλλαγές στη φωτεινότητα
- ✓ Περιοχές με πολλές λεπτομέρειες
- ✓ Αθροιστικός θόρυβος

Προσπάθειες γίνονται για να ελαχιστοποιηθούν αυτές οι ατέλειες. Ένα από τα δυνατά σημεία του MPEG-2, είναι ότι επιτρέπει ποικιλία και καινοτομίες στη σχεδίαση του κωδικοποιητή.

3.6. Το πρότυπο κωδικοποίησης MPEG-4

Το MPEG-4 είναι ένα πρότυπο ISO/IEC που αναπτύχθηκε από την MPEG το 1998 και το οποίο έχει επιλεχθεί ως ένα από τα πρότυπα για μετάδοση πολυμέσων για τα ασύρματα δίκτυα 3ης και 4ης γενιάς. Το πρότυπο αυτό αναπτύχθηκε με σκοπό να γίνει το επόμενο πρότυπο των πολυμεσικών εφαρμογών, το οποίο θα υποστηρίζει πολύ χαμηλούς ρυθμούς μετάδοσης ψηφίων. Σε αντίθεση με το MPEG-1 και το MPEG-2 που προηγήθηκαν και που επικεντρώνονταν σε αποδοτικότερη συμπίεση (compression efficiency), το MPEG-4 δίνει έμφαση σε νέες λειτουργίες. Επιπρόσθετα, το πρότυπο αυτό δίνει στους χρήστες την ευκαιρία να επεξεργάζονται με διάφορους τρόπους τόσο τις ηχητικές πληροφορίες όσο και τις πληροφορίες μιας σκηνής video , αλλά και να συγχωνεύουν το συνθετικό και το φυσικό ήχο, καθώς και τις πληροφορίες video. Το MPEG-4 παρέχει υψηλά επίπεδα ευελιξίας και ελαστικότητας, έτσι ώστε να μπορεί να εκμεταλλευθεί τις αναπτυσσόμενες τεχνολογίες, ειδικά στα ασύρματα επικοινωνιακά συστήματα. Επιπλέον, το κωδικοποιημένο video MPEG-4, υποστηρίζει ένα ευρύ φάσμα ρυθμών μετάδοσης ψηφίων που εκτίνεται από 5 έως 64 kbits/s για κινητές εφαρμογές και μέχρι 4 Mbits/s για εφαρμογές τηλεόρασης. Το πρότυπο εισάγει νέες λειτουργίες video οι οποίες περιγράφονται στο MPEG-4 Proposal Package Description (MPEG AOE Group, 1995) και οι οποίες μπορούν να ομαδοποιηθούν σε τρεις κατηγορίες.

Λειτουργία αλληλεπίδρασης βασισμένης στο περιεχόμενο (content-based interactivity), αποδοτικότητας της κωδικοποίησης (coding efficiency) και καθολικής πρόσβασης (universal access). Αναλυτικότερα, η λειτουργία αλληλεπίδρασης βασισμένης στο περιεχόμενο περιλαμβάνει μια σειρά από διαδικασίες, όπως επεξεργασία του περιεχομένου, καθώς και διόρθωση της ακολουθίας των bit , υβριδική φυσική αλλά και τεχνητή κωδικοποίηση δεδομένων (Hybrid natural and synthetic data coding) και βελτιωμένη, χρονικά τυχαία πρόσβαση (Improved temporal random access). Στα δύο προηγούμενα πρότυπα MPEG (το MPEG-1 και το MPEG-2), υπήρχε η δυνατότητα για επεξεργασία της ακολουθίας δεδομένων video μόνο στο αρχικό επίπεδο. Επομένως, η συμπιεσμένη σειρά video έπρεπε να αποκωδικοποιηθεί στην αρχική της μορφή προτού κωδικοποιηθεί στην επιθυμητή συμπιεσμένη ακολουθία ψηφίων. Στο MPEG-4 υπάρχει η δυνατότητα για πρόσβαση και επεξεργασία των οπτικών αντικειμένων στο πεδίο που αυτά έχουν ήδη συμπιεστεί. Αυτό έχει ως αποτέλεσμα την ελάττωση της πολυπλοκότητας, των απαιτήσεων σε εύρος ζώνης και της καθυστέρησης, αλλά και καμία απώλεια σε ποιότητα. Επίσης παρέχει τα μέσα για αρμονική αλληλεπίδραση των αντικειμένων video διαφόρων προελεύσεων, τα οποία μπορεί να είναι είτε φυσικά, είτε τεχνητά. Έτσι, καθίσταται δυνατό να κωδικοποιηθεί

μια σειρά video με τέτοιο τρόπο ώστε να επιτρέπεται η ξεχωριστή αποκωδικοποίηση και ανακατασκευή των αντικειμένων σε μια σκηνή, αλλά και η επεξεργασία της αρχικής σκηνής μέσω απλών χειρισμών στην ακολουθία των ψηφίων.

Τελικά, σε ότι αφορά στη λειτουργία αλληλεπίδρασης βασισμένης στο περιεχόμενο, το MPEG-4 παρέχει επαρκείς μεθόδους για τυχαία πρόσβαση, εντός ενός περιορισμένου χρονικού διαστήματος αλλά και με καλή ανάλυση σε τμήματα (π.χ. πλαίσια ή αντικείμενα) μιας σειράς video. Το MPEG-4 διαθέτει εργαλεία και αλγορίθμους που προσφέρουν καλύτερη συμπίεση δίχως να επηρεάζουν την ποιότητα του video. Αυτό το καθιστά κατάλληλο για χρησιμοποίηση σε εφαρμογές όπως video-συνδιάσκεψη και ροή video πραγματικού χρόνου σε δίκτυα κινητής τηλεφωνίας. Επιπλέον, το MPEG-4 παρέχει μεθόδους για αποδοτική κωδικοποίηση των πολλαπλών όψεων μιας σκηνής.

Η λειτουργία καθολικής πρόσβασης που προτείνεται στο πρότυπο MPEG-4 καθορίζει τη σθεναρότητα σε επιρρεπή σε σφάλματα περιβάλλοντα (Robustness in error-prone environments) και την κλιμάκωση που βασίζεται στο περιεχόμενο (Content-based scalability). Το πρόβλημα με πολλούς αλγορίθμους συμπίεσης είναι ότι, όταν αυξάνεται ο λόγος συμπίεσης (compression ratio), αυξάνεται, επίσης, και ο κίνδυνος σφάλματος στα ψηφία της μεταδιδόμενης ακολουθίας. Τα επικοινωνιακά κανάλια των ασύρματων δικτύων υπόκεινται σε θόρυβο, διαλείψεις, ράδιο-σκίαση καθώς και σε παρεμβολές, παράγοντες, δηλαδή, που κάνουν δύσκολη την αξιόπιστη μετάδοση μιας κωδικοποιημένης ακολουθίας video. Το MPEG-4 παρέχει σθεναρότητα έναντι σφαλμάτων (error robustness) για εφαρμογές αποθήκευσης (storage applications) όπως video κατ' απαίτηση (video-on-demand), καθώς και για επικοινωνιακές εφαρμογές (communication-applications) όπως video συνδιασκέψεις (video-conferencing), σε επιρρεπή σε σφάλματα ετερογενή περιβάλλοντα (heterogeneous error-prone environments).

Εν ολίγοις το MPEG-4 είναι ένα πρότυπο παγκόσμιας εμβέλειας για τη δημιουργία πολυμέσων, τη διάθεση και την αναπαραγωγή τους στο internet . Το MPEG-1 έφερε επανάσταση στη βιομηχανία των CD, το MPEG-2 σχεδιάστηκε για την ανάπτυξη των DVD και το MPEG-4 έφερε επανάσταση στο διαδίκτυο. Χρησιμοποιείται στα κινητά μέχρι και τη δορυφορική τηλεόραση.

Το MPEG-4 σχεδιάστηκε μέσα στα πλαίσια του διεθνούς οργανισμού προτυποποίησης (**ISO),** που προδιέγραψε τα γνωστά πρότυπα MPEG-1,MPEG-2. Το MPEG-4 είναι περιλαμβάνει εκατοντάδες ερευνητές και τεχνικούς.

Τα πεδία στα οποία λειτουργεί το MPEG-4 όπως ήδη αναφέραμε είναι 2:

- Ψηφιακές τηλεοράσεις
- Αλληλεπιδραστικά πολυμέσα και γραφικά

Το MPEG-4 είναι ανοικτό πρότυπο είναι δύσκολο άλλη τεχνολογία να αναπτυχθεί σε τέτοιο βάθος και να υποστηριχτεί από τόσες εταιρείες καθώς είναι δουλειά πολλών χρόνων μοιρασμένη από εκατοντάδες εταιρείες. Είναι το μόνο πρότυπο, που μπορεί να εκμεταλλευτεί τις δυνατότητες, που προσφέρει η ψηφιακή επανάσταση.

Το MPEG-4 αναπτύσσει την συμπίεση του ήχου και του video επιτρέποντας τη μεταφορά των δεδομένων στο δίκτυο με υψηλή ποιότητα. Προσφέρει ένα πρότυπο για διάφορες μορφές πολυμέσων, όπως είναι το κείμενο, φωτογραφίες, κινούμενα σχέδια δυσδιάστατα και τρισδιάστατα αντικείμενα. Προσφέρει επίσης μια ποικιλία προφίλ για εξειδικευμένες εφαρμογές, έτσι ώστε οι χρήστες να διευκολύνονται, καθώς περιορίζονται στην απλή εφαρμογή των προφίλ αυτών και δεν είναι ανάγκη να προβούν σε εξειδικευμένες ενέργειες. Η τεχνολογία του MPEG-4 μεταφέρει εύκολα στο διαδίκτυο τις διάφορες εφαρμογές των πολυμέσων σε όχι απαραίτητα υψηλές ταχύτητες. Το λογισμικά, που εφαρμόζουν το MPEG-4 είναι ελεύθερα στο internet για χρήση με κάποιους κανόνες και άδειες χρήσης.

Επανάσταση στην λειτουργικότητα

Το MPEG-4 αποτελεί επανάσταση στη λειτουργία των πολυμέσων. Εκτός των άλλων μπορεί να δουλεύει σε υπάρχουσες δομές, όπως σε περιβάλλοντα λειτουργίας του MPEG-2. Αυτή είναι μια πολύ σημαντική λειτουργία , γιατί με αυτό τον τρόπο δεν χρειάζεται να απορρίπτουμε παλιότερες τεχνολογίες, αλλά απλώς να τις βελτιώνουμε. Ακόμα με την κοινή χρήση του MPEG-2 και του MPEG-4 υπάρχει μια συμβατότητα του παλιού με το καινούριο και με τον τρόπο αυτό αποφεύγουμε την απώλεια δεδομένων. Η σύνδεση των προτύπων υποδηλώνει και τα προηγούμενα πρότυπα είναι διαχρονικά και μπορούν να ανταπεξέλθουν σε σύγχρονες ανάγκες. Επίσης οι τεχνικές συμπίεσης είναι περίτεχνες, δημιουργώντας μια παρουσίαση των πολυμέσων πιο αποτελεσματική και εύκολη στη χρήση. Το πρότυπο αυτό προσφέρει επαναστατικά εργαλεία για το περιεχόμενο, με το οποίο μεταφέρεται το ελάχιστο από τις γενικές και άχρηστες πληροφορίες. Αυτό επιτρέπει χωριστή αλληλεπίδραση με το κάθε αντικείμενο και σωστή μεταφορά σε όλα τα δίκτυα, όπως για παράδειγμα οι υπότιτλοι σε μια ταινία, οι οποίοι μπορεί να είναι αναγκαίοι για την κατανόηση μιας ταινίας, παραμένουν της ίδιας ποιότητας ακόμα και σε υψηλή συμπίεση. Επιλέγουμε λοιπόν πιο μέρος της πληροφορίας επιθυμούμε να διατηρηθεί σε καλύτερη ποιότητα. Έτσι τα δεδομένα, που δεν είναι αναγκαία, δέχονται μια μεγαλύτερη συμπίεση για την εξοικονόμηση του απαραίτητου χώρου. Αυτός είναι άλλωστε ήταν και ο σκοπός του προτύπου εξ αρχής.

Η *Internet Engineering Task Force(IETF)* είναι ανοικτή διεθνής κοινότητα από τις πιο μεγάλες όπου αποτελείται από σχεδιαστές δικτύου, διαχειριστές, πωλητές και ερευνητές που ενδιαφέρονται για την εξέλιξη της αρχιτεκτονικής του Internet. Η δουλειά τους είναι σχετική με το MPEG-4. Το χρησιμοποιούν σαν βασικό πρότυπο.

Ο βασικός στόχος της *Internet Streaming Media Alliance (ISMA)* είναι να προωθήσει τη χρήση ανοικτών προτύπων για τη ροή του περιεχομένου πολυμέσων όπως video , ήχο μέσω πρωτοκόλλων του διαδικτύου. Η ISMA έχει επιλέξει συγκεκριμένα ηχητικά και εικονικά προφίλ και επίπεδα του MPEG-4 και τα επαυξάνει με εξειδικευμένες τεχνικές μεταφοράς δεδομένων της IETF για να δημιουργήσει διαλειτουργικότητα για video στο Internet.

3.6.1. Αποδοχή του MPEG-4

Το MPEG-4 είναι ένα πρότυπο που γίνεται αποδεκτό από τα περισσότερα λογισμικά για την αναπαραγωγή πολυμέσων. Είναι πρότυπο που χρησιμοποιείται ευρέως στο Internet, κυρίως στις ψηφιακές βιβλιοθήκες πολυμέσων, ειδικά σε μορφή video. Έχει ήδη γίνει καθιερωμένο πρότυπο για όλες τις συσκευές, που χρησιμοποιούν πολυμεσικές εφαρμογές μικρής χωρητικότητας αλλά αξιόλογης ποιότητας, όπως είναι τα κινητά τρίτης γενιάς. Χρησιμοποιείται σε λογισμικά της Microsoft αλλά και πολλών άλλων εταιρειών. Εκτελείται σε εφαρμογές όπως Quicktime 6 και το Media Player της Microsoft που είναι από τα πιο διαδεδομένα και πιο εύχρηστα προγράμματα στο χώρο των υπολογιστών. Χρησιμοποιείται ακόμα και σε φορητές συσκευές των υπολογιστών, όπως είναι τα PDA, συσκευές ιδιαίτερα χρήσιμες για την εύκολη μεταφορά προσωπικών δεδομένων. Το MPEG-4 δεν περιορίζεται μόνο στο χώρο των υπολογιστών αλλά εμπλέκεται σε όλες τις συσκευές ψυχαγωγίας. Χρησιμοποιείται σε ψηφιακές τηλεοράσεις και τις ψηφιακές κάμερες σε συνδυασμό με τα προϋπάρχοντα πρότυπα MPEG-1 και MPEG-2.

3.6.2. Προφίλ και επίπεδα του MPEG-4

Το MPEG-4 αποτελείται από ένα μεγάλο αριθμό εργαλείων, τα οποία δεν είναι χρήσιμα σε όλες τις εφαρμογές. Υπάρχουν λοιπόν τα λεγόμενα προφίλ για την οργάνωση των εργαλείων αυτών , όπως για παράδειγμα το MPEG-4 Advanced Simple Visual Profile περιέχει ¼ αποκατάσταση της κίνησης, B-frames, αλλά δεν περιέχει σχηματοποιημένο κωδικοποιημένο video. Τα προφίλ αυτά επιτρέπουν στους χρήστες να διαλέξουν το κατάλληλο από μια ποικιλία εργαλείων για να εκτελέσουν τη

λειτουργία που αυτοί επιθυμούν. Τα προφίλ υπάρχουν σε ένα αριθμό επιπέδων, που παρέχουν έναν τρόπο να περιορίσουν την πολυπλοκότητα των υπολογιστών.

3.6.3. Τα βασικά μέρη του MPEG-4

Το MPEG-4 αποτελείται από στενά συνδεδεμένα μέρη που μπορούν να εφαρμοστούν χωριστά ή να συνδυαστούν με άλλα μέρη.
Η δομή του αποτελείται από τα:

- ✓ Συστήματα(Systems).
- ✓ Οπτική μορφή(Visual).
- ✓ Ήχο(Audio).
- ✓ Προσαρμογή(Conformance): διευκρινίζει πως να δοκιμάσουμε μια εφαρμογή του MPEG-4.
- ✓ Αναφορά Λογισμικού(Reference Software): μπορεί να χρησιμοποιηθεί για να ξεκινήσουμε την εφαρμογή του προτύπου.
- ✓ DMIF (Delivery Multimedia Integration Framework): προδιαγράφει μια διεπαφή μεταξύ εφαρμογής και δικτύου/ αποθετηρίου.
- ✓ Προδιαγραφή ενός βελτιωμένου κωδικοποιητή video .
- ✓ Η μεταφορά δεδομένων είναι η βασική αρχή του προτύπου αλλά δεν είναι προδιαγραμμένη. Το 8 μέρος προδιαγράφει πως να χαρτογραφήσουμε τη ροή δεδομένων του MPEG-4 σε μεταφορά στο internet.

3.6.4. Πλεονεκτήματα του MPEG-4

Ο καλύτερος τρόπος για να καταλάβουμε τις δυνατότητες του MPEG-4 είναι με τη σύγκριση με το MPEG-2. Στο MPEG-2 το περιεχόμενο δημιουργείται από πολλές πηγές όπως κινούμενο video , γραφικά, κείμενο. Μετά τη συγκρότηση τους σε ένα σχέδιο από εικονοστοιχεία, κωδικοποιούνται όλα, ανεξαρτήτως μορφής σε εικονοστοιχεία κινούμενου video. Από την πλευρά του καταναλωτή η αποκωδικοποίηση είναι ένα άμεσο εγχείρημα.

Το MPEG-2 είναι μία στατική μηχανή παρουσίασης έτσι για παράδειγμα σε ένα ρεπορτάζ των ειδήσεων το οποίο είναι ζωντανό, αν χρειαστεί να ξαναπροβληθεί, τα στοιχεία παραμένουν ίδια. Έχεις τη δυνατότητα να προσθέσεις γραφικά και άλλα στοιχεία αλλά δεν υπάρχει τρόπος να τα σβήσεις. Από την άλλη πλευρά το MPEG-4 είναι δυναμικό εκεί που είναι στατικό το MPEG-2. Διαφορετικά αντικείμενα μπορούν

να κωδικοποιηθούν και να μεταδοθούν χωριστά από τον κωδικοποιητή και σε δική τους ροή. Η σύνθεση τους πραγματοποιείται μετά την αποκωδικοποίηση. Για να επιτευχθεί αυτή η σύνθεση το MPEG-4 περιλαμβάνει μια ειδική γλώσσα περιγραφής σκηνών, αποκαλούμενη BIFS (Binary Format for Scenes). Η γλώσσα BIFS δεν περιγράφει μόνο το πότε και πού τα αντικείμενα θα εμφανιστούν στη σκηνή, αλλά περιγράφει και συμπεριφορά ακόμα και συμπεριφορά υπό συνθήκες, όπως για παράδειγμα αντικείμενα που αντιδρούν σε καταστάσεις. Αυτό παρουσιάζει την αλληλεπιδραστικότητα του MPEG-4. Όλα τα αντικείμενα μπορούν να κωδικοποιηθούν με δικό τους σχήμα κωδικών ενώ το video κωδικοποιείται σαν video , το κείμενο σαν κείμενο, τα γραφικά σαν γραφικά αντί να θεωρούνται όλα τα εικονοστοιχεία σαν κινούμενο video .

Όλοι οι κωδικοποιητές στο MPEG-4 βελτιώνονται για τους κατάλληλους τύπους δεδομένων και το MPEG-4 περιλαμβάνει αποτελεσματικούς κωδικοποιητές για ήχο, ομιλία, video και ακόμα και σύνθετο περιεχόμενο όπως κινούμενα πρόσωπα και σώματα.

Κεφάλαιο - 4 Αλλοίωση video με χρήση Matlab

Εισαγωγή

Για την υλοποίηση της πτυχιακής θα γράψουμε έναν κώδικα στο μαθηματικό εργαλείο MATLAB. Θα κωδικοποιήσουμε τα video που θα χρησιμοποιήσουμε σε MPEG στοιχεία για I και P πλαίσια και στη συνέχεια θα αποκωδικοποιήσουμε τα MPEG δεδομένα πάλι στην αρχική ταινία. Ο κώδικας είναι βασισμένος σε MPEG ύφος, δεν παράγει ακριβώς συμβατή ροή bits. Παρόλα αυτά οι σημαντικές απώλειες στα βήματα κβαντισμού αναπαράγονται πιστά γι 'αυτό και είναι ένα έγκυρο παράδειγμα από το οποίο μπορούμε να πειραματιστούμε και να έχουμε αποτελέσματα τις αντικειμενικής και υποκειμενικής ποιότητας ενός video.

Το MATLAB είναι ένα σύγχρονο ολοκληρωμένο μαθηματικό λογισμικό πακέτο που χρησιμοποιείται σε πανεπιστημιακά μαθήματα αλλά και ερευνητικές και άλλες εφαρμογές με επιστημονικούς υπολογισμούς . Το όνομά του προέρχεται από τα αρχικά γράμματα των λέξεων MATtrix LABoratory *(εργαστήριο πινάκων*). Το MATLAB είναι ένα διαδραστικό (interactive) πρόγραμμα για αριθμητικούς υπολογισμούς και οπτικοποίηση δεδομένων με δυνατότητες προγραμματισμού που το καθιστούν ένα ισχυρό και χρήσιμο εργαλείο στις μαθηματικές και φυσικές επιστήμες.

4.1. Δομή του κώδικα

Ο κώδικας για την πτυχιακή είναι οργανωμένος σε συναρτήσεις που βρίσκονται σε ένα ενιαίο m-αρχείο. Η συνάρτηση εκτελεί τέσσερα σημαντικά βήματα:

1. Φορτώνει το video για την κωδικοποίηση. (Τα video που χρησιμοποιήθηκαν είναι λιγότερα τον 30 sec.)
2. Κωδικοποιεί την ταινία και επιστρέφει τα δεδομένα σε MPEG μορφή.
3. Αποκωδικοποίηση των MPEG δεδομένων και ανακατασκευή του video από το οποίο παράχθηκε.
4. Αποθήκευση και των δύο εκδόσεων ,του video και των MPEG δεδομένων για προβολή και για ανάλυση.

Ο κώδικας χρησιμοποιεί μεταβλητές για να ελέγχουν μεγάλο μέρος της διαδικασίας κωδικοποίησης και έτσι μας επιτρέπει τον πειραματισμό. Μερικές από τις παραμέτρους που μπορούν εύκολα να ρυθμιστούν είναι η επιλογή του video για την

κωδικοποίηση, ο κβαντικός συντελεστή κλίμακας, το μοτίβο στα πλαίσια (π.χ.: IPPIPP), η διαδοχική ή λογαριθμική αναζήτηση στο διάνυσμα κίνησης, και η κίνηση αναζήτησης στο διάνυσμα μέγεθος του παραθύρου.

4.2. Παράδειγμα στο κωδικοποιημένο video

Στο παράδειγμα κωδικοποιήσαμε και αποκωδικοποιήσαμε 10 frame της ταινίας με την κλίμακα ποιότητας στο 31. Για όλα τα 10 frame , το αρχικό, το ανακατασκευασμένο και το λάθος εμφανίζονται στον πίνακα που ακολουθεί. Αυτά τα πλαίσιο δείχνουν μόνο το σήμα φωτεινότητας ακόμη και αν η ταινία ήταν έγχρωμη.



Εικόνα 4.1- *Frame 1 original – Compressed*



Εικόνα 4.2- *Frame 2 original – Compressed*



Εικόνα 4.3- *Frame 3 original – Compressed*



Εικόνα 4.4- *Frame 4 original – Compressed*



Εικόνα 4.5- *Frame 5 original – Compressed*



Εικόνα 4.6- *Frame 6 original – Compressed*



Εικόνα 4.7- *Frame 7 original – Compressed*



Εικόνα 4.8- *Frame 8 original – Compressed*



Εικόνα 4.9- *Frame 9 original – Compressed*



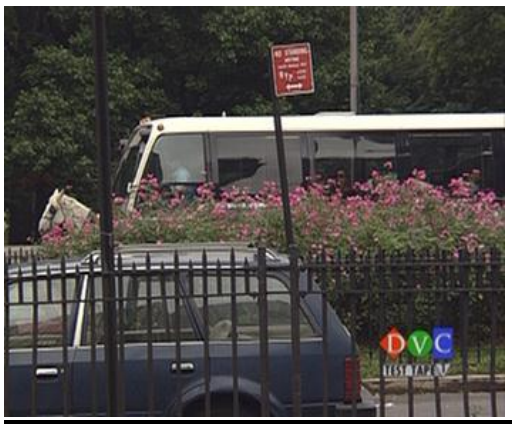
Εικόνα 4.10- *Frame 10 original – Compressed*

4.3. Το μέγεθος κλίμακας στην ποιότητα εικόνας

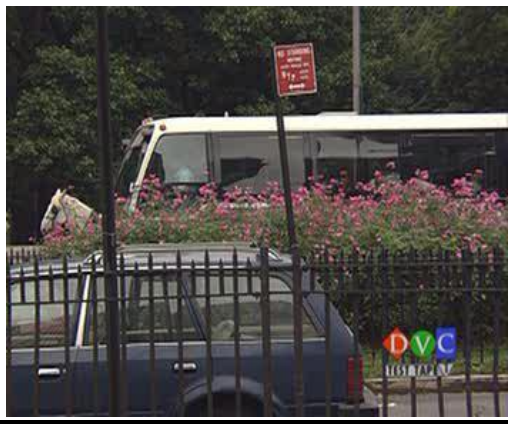
Στην συνέχεια κωδικοποιήσαμε το πλαίσιο I σε πολλά επίπεδα ποιότητας.



Εικόνα 4.11- *Αρχικό frame*



Scale 1



Scale 8



Scale 16



Scale 32



Εικόνα 4.12- κωδικοποίηση σε πολλά επίπεδα ποιότητας

Στην υψηλότερη ποιότητα, με κλίμακα 1, δεν μπορώ να δω καμία διαφορά μεταξύ του αρχικού καρέ και αυτού που ανακατασκευάστηκε . Με την κλίμακα 8, μπορώ να δω πολύ μικρές διαφορές σε σκιερά μέρη του πλαισίου, όπως τα παράθυρα του λεωφορείου. Καθώς αυξάνεται η κλίμακα περαιτέρω, τα αντικείμενα αρχίζουν να εμφανίζονται με προβλήματα και αρκετά θολά λόγω της υψηλής συχνότητας . Στην κατώτερη ποιότητα, με κλίμακα 112, υπάρχουν ενδιαφέροντα αποτελέσματα. Πέραν του ότι είναι πολύ θολή, το πλαίσιο φαίνεται να είναι ως επί το πλείστον αποχρώσεις του γκρι. Αυτό συμβαίνει επειδή δεν είναι τα χρώματα κορεσμένα αρκετά ώστε να είναι κβαντισμένες στο μηδέν κατά τη διάρκεια της κωδικοποίησης. Στη συνέχεια η ανοικοδόμηση αναπαράγει πιστά ένα μπλοκ με καμία πληροφορία χρωματισμού.

Κεφάλαιο 5 - Εκτίμηση της ποιότητας εικόνας video

5. Εισαγωγή

Τα ψηφιακά αρχεία με οπτικοακουστικό περιεχόμενο, δέχονται διάφορα είδη αλλοιώσεων τόσο κατά το στάδιο της συμπίεσης/κωδικοποίησης αλλά και κατά τη διάδοσή τους. Για παράδειγμα οι απωλέστηκες τεχνικές συμπίεσης και κωδικοποίησης που χρησιμοποιούνται σήμερα, με στόχο να μειώσουν το εύρος ζώνης (bandwidth) που απαιτείται για την αποθήκευση και μεταφορά του , έχουν ως συνέπεια την υποβάθμιση της αντιλαμβανόμενης ποιότητας κατά την διάρκεια της κωδικοποίησης. Επιπλέον, εξαιτίας της εμφάνισης διαφόρων παραμέτρων, τα δεδομένα που μεταφέρονται υφίστανται απώλειες και καθυστερήσεις, που συμβάλλουν στην υποβάθμιση της ποιότητας κατά την αναπαραγωγή του video, στον παραλήπτη της υπηρεσίας. Για αυτό καθίσταται πλέον αναγκαία η ανάπτυξη τεχνικών και μεθόδων για την εκτίμηση της ποιότητας κινούμενης εικόνας video. Οι τεχνικές που έχουν αναπτυχθεί ως σήμερα αναφέρονται κυρίως στην υποβάθμιση της ποιότητας που προέρχεται από την κωδικοποίηση του οπτικοακουστικού περιεχομένου.

Οι μέθοδοι εκτίμησης ποιότητας κινούμενης εικόνας χωρίζονται σε δύο κατηγορίες:

- ✓ Υποκειμενικές Μέθοδοι
- ✓ Αντικειμενικές Μέθοδοι

5.1. Υποκειμενικές Μέθοδοι

Οι υποκειμενικές μέθοδοι απαιτούν την ύπαρξη ακροατηρίου, εξειδικευμένου ή μη, το οποίο θα αξιολογεί την αντιλαμβανόμενη ποιότητα κινούμενης εικόνας, σε περιβάλλον με ελεγχόμενες συνθήκες θέασης. Υπάρχουν δύο τυποποιημένες μέθοδοι αντικειμενικής εκτίμησης, οι οποίες έχουν προταθεί από την ITU και είναι οι :

- ✓ Single Stimulus Continue Quality Evaluation (SSCQE)
- ✓ Double Stimulus Continue Quality Evaluation (DSCQE)

Στη SSCQE δεν χρησιμοποιείται σήμα αναφοράς, ουσιαστικά προσομοιάζει την αντίδραση που θα είχε ο θεατής σε τεχνουργήματα που οφείλονται λόγω μετάδοσης,

ενώ στην DSCQE χρησιμοποιείται σήμα αναφοράς, και το ακροατήριο καλείται να συγκρίνει μεταξύ του σήματος αναφοράς και του ποιοτικά υποβαθμισμένου σήματος.

- **SSCQE**

Η μέθοδος SSCQE προέκυψε από το ευρωπαϊκό ερευνητικό πρόγραμμα MOSAIC, με σκοπό κυρίως την εκτίμηση ψηφιακών οπτικοακουστικών σημάτων, χωρίς όμως να περιορίζεται η χρήση του μόνο εκεί. Αντιθέτως μπορεί να χρησιμοποιηθεί και σε αναλογικά σήματα. Το πρωτοποριακό της μεθόδου αυτής, σε σχέση με αυτές που προϋπήρχαν, είναι ότι αύξησε σημαντικά τη διάρκεια των οπτικοακουστικών πηγών που θέτονται σε δοκιμή. Δηλαδή χρησιμοποιούνται οπτικοακουστικά σήματα των οποίων η διάρκεια είναι μεγαλύτερη από 10 δευτερόλεπτα, που ήταν ουσιαστικά το όριο που χρησιμοποιούσαν οι προηγούμενοι μέθοδοι.

Η μέθοδος SSCQE χωρίζεται σε τρία στάδια:

Στάδιο 1

Το πρώτο στάδιο περιλαμβάνει την συνεχής εκτίμηση οπτικοακουστικής πηγής μονού ερεθίσματος και τη συγκέντρωση δεδομένων από τις στιγμιαίες εκτιμήσεις που πραγματοποιεί ο κάθε χρήστης, παρακολουθώντας τα οπτικοακουστικά σήματα δοκιμής.

Στάδιο 2

Το στάδιο αυτό χρησιμοποιείται για τη διόρθωση των διαβαθμίσεων των μετρήσεων που έχουν ληφθεί κατά το πρώτο στάδιο, σύμφωνα με την κλίμακα που προτείνεται και περιγράφεται u945 από την Internationa Telecommunication Union (ITU) στο έγγραφο ITU-R Recommendation BT(Broadcast Television) 500-7. Ουσιαστικά πρόκειται για μια απεικόνιση των μετρήσεων σε μια κλίμακα που αποτελείται από πέντε τιμές, όπου το 5 είναι το άριστα, και στον αντίποδα βρίσκεται το 0.

Στάδιο 3

Σε αυτό το τμήμα της υποκειμενικής εκτίμησης, προβλέπεται το φιλτράρισμα των αποτελεσμάτων με τη χρήση κατάλληλων φίλτρων, όπου θα θέτουν διαφορετικά βάρη στις ποικίλες εκτιμήσεις της ποιότητας, με σκοπό να προσομοιαστεί το ανθρώπινο οπτικό σύστημα (Human Visual System – HVS) και φαινόμενα όπως το forgiveness effects που επηρεάζουν τις αντίστοιχες μετρήσεις. Για παράδειγμα το forgiveness effects σχετίζεται με το γεγονός ότι ένα frame με κακή ποιότητα έχει ως αποτέλεσμα ο θεατής να θεωρεί ότι και τα επόμενα frame είναι κακής ποιότητας, χωρίς στην πραγματικότητα να ισχύει αυτό. Συνεπώς εισάγονται πολλοί ψυχολογικοί παράγοντες, οι οποίοι πρέπει να ληφθούν υπόψη ώστε να λαμβάνονται πραγματικά σωστά αποτελέσματα.

- **DSCQE**

Η DSCQE είναι μια παραλλαγή της SSCQE μεθόδου, όπου η διαφορά τους είναι η ύπαρξη του αρχικού οπτικοακουστικού περιεχομένου, πηγή αναφορά μέσω της οποίας ο τηλεθεατής καλείται να αξιολογήσει την αντιλαμβανόμενη ποιότητα που βλέπει. Η μέθοδος αυτή, εκτός από την εκτίμηση της ποιότητας, αποβλέπει και στην αξιολόγηση της αντιλαμβανόμενης πληροφορίας, αφού ο θεατής είναι σε θέση να συγκρίνει ανάμεσα στην αρχική μεταδιδόμενη πληροφορία και την τελικά λαμβανόμενη, την οποία μάλιστα καλείται να αξιολογήσει. Η DSCQE στην υλοποίηση της χρησιμοποιεί δύο οθόνες, όπου στη μία προβάλλεται το συμπιεσμένο/κωδικοποιημένο οπτικοακουστικό αρχείο, ενώ στην άλλη οθόνη προβάλει το αρχικό video αναφοράς με αυτό τον τρόπο, οι θεατές εκτιμούν την αντιλαμβανόμενη ποιότητα, συγκρίνοντας άμεσα την πιστότητα και την ποιότητα σε σύγκριση με το πρωτότυπο.

Με την μέθοδο DSCQE, βλέπουμε ότι η διαφορά στο ρυθμό κωδικοποίησης παίζει πολύ βασικό ρόλο στην τελική ποιότητα του οπτικοακουστικού αρχείου. Παρατηρούμε ότι όσο μικρότερος είναι ο ρυθμό κωδικοποίησης, τόσο πιο έντονη είναι η πτώση στην ποιότητα που αντιλαμβάνεται το κοινό, αλλά και τόσο μεγαλύτερο είναι το ποσοστό που συμφωνεί με αυτή την απώλεια στην ποιότητα. Η πιο αξιόπιστη μέθοδος που χρησιμοποιείται για την εξαγωγή αποτελεσμάτων από τις τεχνικές υποκειμενικής εκτίμησης είναι το Mean Opinion Score (MOS), το οποίο ουσιαστικά αποτελεί τον μέσο όρο των εκτιμήσεων που λαμβάνονται από όλους τους θεατές που μετέχουν στη διαδικασία.

Δυστυχώς όμως η οργάνωση, διεξαγωγή και πραγματοποίηση διαδικασιών τις υποκειμενικής εκτίμησης κινούμενης εικόνας είναι οικονομικά μη συμφέρουσα και χρονικά σπάταλη. Το γεγονός ότι απαιτούνται μια πληθώρα ειδικών συσκευών προβολής και καταγραφής περιορίζει την πραγματοποίηση τους σε ερευνητικούς σκοπούς και μόνο. Επιπλέον η συγκέντρωση κάποιου ικανοποιητικού αριθμού πλήθους που θα δεχτεί να συμμετέχει σε μια διαδικασία υποκειμενικής εκτίμησης, είναι ο βασικός ανασταλτικός παράγοντας των μεθόδων αυτών άλλα και κατά πόσο το πλήθος θα μπορεί να δώσει σωστές απαντήσεις .

Για αυτό το λόγο, πρόσφατα η επιστημονική έρευνα έχει επικεντρωθεί στην ανάπτυξη μεθόδων και τεχνικών οι οποίες θα μπορούν να προσομοιάζουν τα αποτελέσματα που λαμβάνονται από μια υποκειμενική μέθοδο, ώστε να απλοποιηθεί τόσο η διαδικασία, όσο και η διεξαγωγή αυτών. Με τις νέες αυτές μεθόδους, θα μπορούσαν πλέον ποικίλες εμπορικές εφαρμογές να συμπεριλάβουν την εκτίμηση της ποιότητας εικόνας video ως μέσο εξοικονόμησης πόρων.

Στα πλαίσια αυτών των αναγκών αναπτύχθηκαν οι αντικειμενικές μέθοδοι εκτίμησης ποιότητας εικόνας video που θα αναλύσουμε παρακάτω.

5.1.1. Πειραματικό μέρος υποκειμενικών μεθόδων

Για να υλοποιηθεί όσο το δυνατόν καλύτερα το πειραματικό μέρος των υποκειμενικών μεθόδων εκτίμησης video θα πρέπει να ορίσουμε κανόνες και βήματα ώστε να έχουμε όσο το δυνατών πιο έγκυρα αποτελέσματα. Γνωρίζοντας για την ανθρώπινη αντίληψη και τα ερεθίσματα που δέχεται επιλέγουμε κάποιες παραμέτρους που θα βοηθήσουν το κοινό ώστε να γίνει πιο κατανοητό το πείραμα.

Έτσι εισαγωγικά πριν δει ο κάθε θεατής το video ενημερώθηκε με ένα σύντομο power point για κάποια οπτικά χαρακτηριστικά των video ώστε να μπορούν να δώσουν όσο το δυνατών καλύτερες απαντήσεις.

- **Φωτεινότητα (Lightness)**

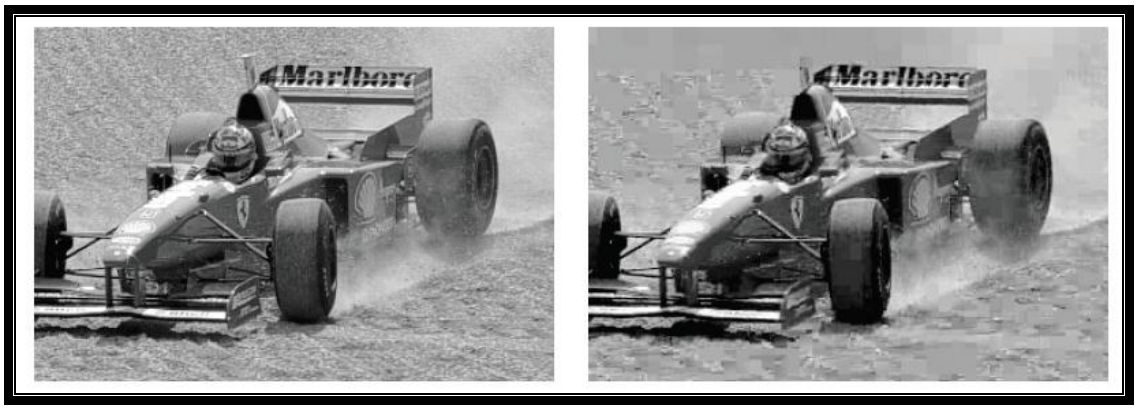
Η φωτεινότητα είναι η αντιστοίχιση σε μια διαβάθμιση από το πιο σκούρο στο πιο ανοιχτό. Με άλλα λόγια είναι η οπτική επίδραση του φωτισμού στα αντικείμενα ή σε σκηνές. (Εικόνα 5.1)



Εικόνα 5.1 - *Φωτεινότητα*

- **Προσδιορισμός των blocks (Blockiness)**

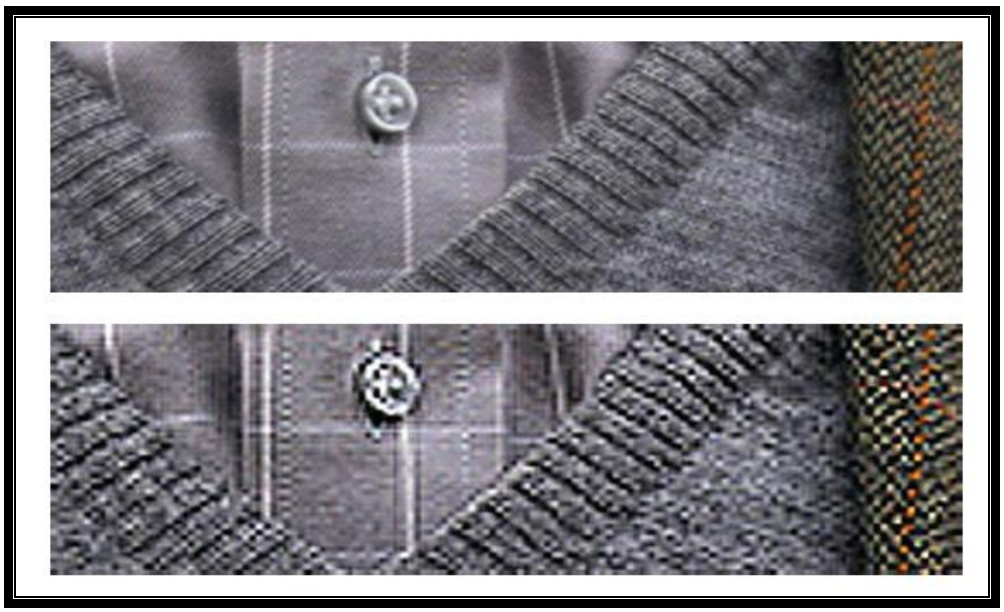
Με την παραπάνω έννοια αναφερόμαστε στο πόσο εύκολο είναι να δει κανείς τις άκρες των blocks. (Εικόνα 5.2)



Εικόνα 5.2 - *Blockiness*

- **Οξύτητα (Sharpness)**

Η παράμετρος αυτή χαρακτηρίζει πόσο διακριτά είναι τα σημεία όπου η φωτεινότητα αλλάζει. Πόσο οξύ είναι ένα σημείο ή πόσο «καθαρό». (Εικόνα 5.3)



Εικόνα 5.3 – *Οξύτητα*

- **Πιστότητα Χρώματος (Color Fidelity)**

Πιστότητα, ακρίβεια των χρωμάτων ενός video.

- **Ομαλότητα της κίνησης (Smoothness of Motion)**

Εάν υπάρχει το ενδεχόμενο, μετά από συμπίεση ενός Video να υπάρχει κάποια καθυστέρηση κατά την κίνηση για παράδειγμα ενός αντικειμένου.

5.1.2. Επιλογή video

Ακόμα ένα σημαντικό μέρος της πτυχιακής είναι η επιλογή video , έτσι επιλέχτηκαν πέντε video που παρουσιάζονται παρακάτω με τις σχετικές πληροφορίες τους.
Για να δούμε τις πληροφορίες ενός video χρησιμοποιήθηκε ο παρακάτω κώδικας στο Matlab.

```
Clip1 = VideoReader ('clip1.avi');
Info = get (Clip1);
```

Πρώτο video : Χρησιμοποιήθηκε ένα video σταθερό και σε εσωτερικό χώρο, ένα δελτίο ειδήσεων όπου έχουμε κίνηση μόνο του παρουσιαστή.



Info =

Duration	: 23.2000
Name	: 'clip1'
Type	: 'VideoReader'
User Data	: []
BitsPerPixel	: 24
FrameRate	: 30
Height	: 240
NumberOfFrames	: 580
VideoFormat	: 'RGB24'
Width	: 320
Format	: AVI
Display aspect ratio	: 4:3
Format	: RGB
Bit rate	: 46.1 Mbps
Bit depth	: 8 bits

Εικόνα 5.4 – *video 1 αρχική ποιότητα και πληροφορίες*



Info =

Duration	: 23.2000
Name	: 'clip1a.'
Type	: 'VideoReader'
User Data	: []
BitsPerPixel	: 24
FrameRate	: 25
Height	: 240
NumberOfFrames	: 580
VideoFormat	: 'RGB24'
Width	: 320
Format	: mpeg
Display aspect ratio	: 4:3
Format	: RGB
Bit rate	: 387 Kbps
Bit depth	: 8 bits
Compression mode	: Lossy
Bits/(Pixel*Frame)	: 0.202

Εικόνα 5.5 – *video 1 αλλοιωμένο και πληροφορίες*

Δεύτερο video : Χρησιμοποιήθηκε ένα video σε εξωτερικό χώρο με έντονο φυσικό φωτισμό.



Info =

Duration	: 27.4972
Name	: 'clip2.'
Type	: 'VideoReader'
User Data	: []
BitsPerPixel	: 24
FrameRate	: 29.9700
Height	: 270
NumberOfFrames	: 680
VideoFormat	: AVI
Width	: 480
Format	: AVI
Display aspect ratio	: 16:9
Format	: 'YUV'
Bit rate	: 489 Kbps
Bit depth	: 8 bits

Εικόνα 5.6 – *video 2 αρχική ποιότητα και πληροφορίες*



Info =

Duration	: 27.4972
Name	: 'clip2a.avi'
Type	: 'VideoReader'
User Data	: []
BitsPerPixel	: 24
FrameRate	:25
Height	: 270
NumberOfFrames	: 580
VideoFormat	: 'RGB24'
Format	: 16:9
Display aspect ratio	: 16:9
Format	: mpeg
Bit rate	: 510 Kbps
Bit depth	: 8 bits

Εικόνα 5.7 – *video 2 αλλοιωμένο και πληροφορίες*

Τρίτο video : Χρησιμοποιήθηκε ένα video σε εξωτερικό χώρο με κίνηση .



Info =

Duration	: 24.5932
Name	: 'clip3'
Type	: 'VideoReader'
User Data	: []
BitsPerPixel	: 24
FrameRate	: 30
Height	: 270
NumberOfFrames	: 680
VideoFormat	: 'RGB24'
Width	: 480
Format	: AVI
Display aspect ratio	: 4:3
Format	: RGB
Bit rate	: 46.1 Mbps
Bit depth	: 8 bits

Εικόνα 5.8 – *video 3 αρχική ποιότητα και πληροφορίες*



Info =

Duration	: 24.5932
Name	: 'clip2a'
Type	: 'VideoReader'
User Data	: []
BitsPerPixel	: 24
FrameRate	: 25
Height	: 270
NumberOfFrames	: 680
VideoFormat	: 'RGB24'
Width	: 480
Format	: mpeg
Display aspect ratio	: 4:3
Format	: RGB
Bit rate	: 46.1 Mbps

Εικόνα 5.9 – *video 3 αλλοιωμένο και πληροφορίες*

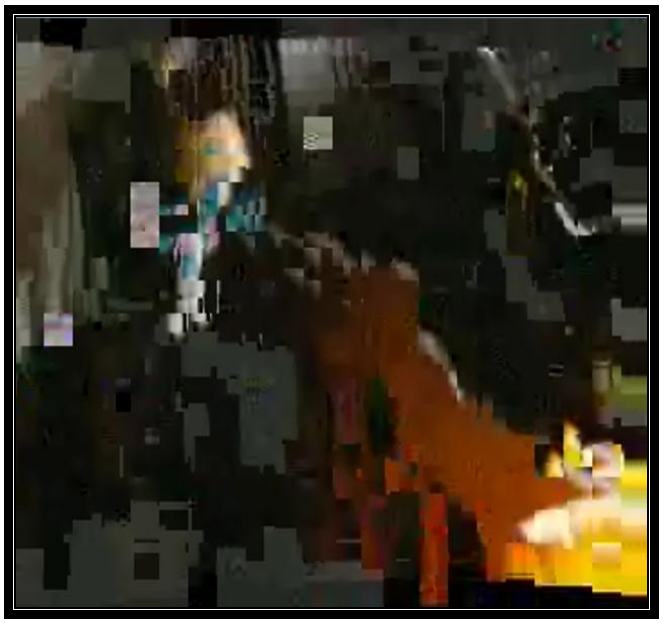
Τέταρτο video : Χρησιμοποιήθηκε ένα video με κινούμενα σχέδια .



Info =

Duration	: 16.9535
Name	: 'clip3'
Type	: 'VideoReader'
User Data	: []
BitsPerPixel	: 24
FrameRate	: 30
Height	: 288
NumberOfFrames	: 580
VideoFormat	: 'RGB24'
Width	: 424
Format	: mpg
Display aspect ratio	: 4:3
Format	: RGB
Bit rate	: 46.1 Mbps
Bit depth	: 8 bits

Εικόνα 5.10– *video 4 αρχική ποιότητα και πληροφορίες*

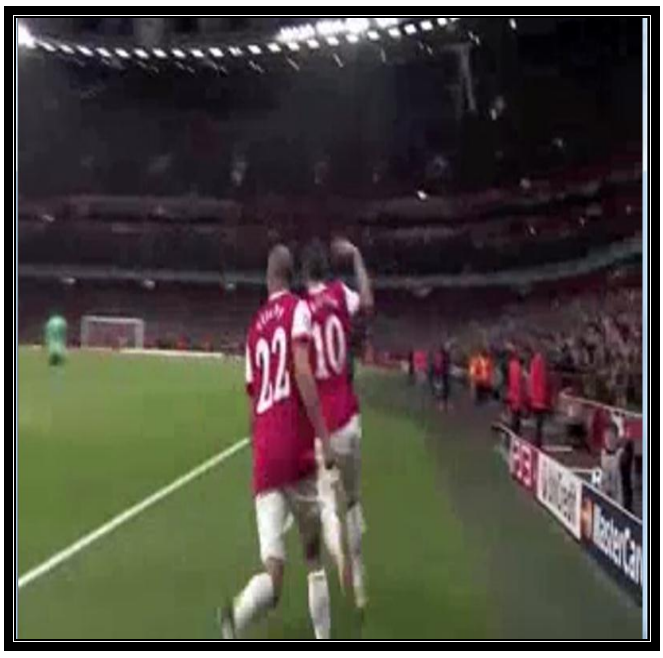


Info =

Duration	: 16.9535
Name	: 'clip3a'
Type	: 'VideoReader'
User Data	: []
BitsPerPixel	: 24
FrameRate	: 25
Height	: 288
NumberOfFrames	: 404
VideoFormat	: 'RGB24'
Width	: 424
Format	: mpeg
Display aspect ratio	: 4:3
Format	: RGB
Bit rate	: 46.1 Mbps
Bit depth	: 8 bits

Εικόνα 5.11 – *video 4 αλλοιωμένο και πληροφορίες*

Πέμπτο video : Χρησιμοποιήθηκε ένα video ποδοσφαιρικού αγώνα.



Info =

Duration	: 14.8667
Name	: 'clip5’
Type	: 'VideoReader'
User Data	: []
BitsPerPixel	: 24
FrameRate	: 30
Height	: 240
NumberOfFrames	: 223
VideoFormat	: 'RGB24'
Width	: 320
Format	: mp4
Display aspect ratio	: 16:9
Bit rate	: 46.1 Mbps
Bit depth	: 8 bits
Bits/(Pixel*Frame)	: 0.329

Εικόνα 5.12 – *video 5 αρχική ποιότητα και πληροφορίες*



Info =

Duration	: 14.8667
Name	: 'clip5a'
Type	: 'VideoReader'
User Data	: []
BitsPerPixel	: 24
FrameRate	: 25
Height	: 240
NumberOfFrames	: 214
VideoFormat	: 'RGB24'
Width	: 320
Format	: mpeg
Display aspect ratio	: 16:9
Format	: RGB
Bit rate	: 413 Kbps
Bit depth	: 8 bits
Bits/(Pixel*Frame)	: 0.359

Εικόνα 5.13 – *video 5 αλλοιωμένο και πληροφορίες*

Με την εντολή που δώσαμε παραπάνω στο Matlab για όλα τα video μας , πήραμε τα παραπάνω χαρακτηριστικά , παρακάτω ακολουθεί η σημασία τους

Duration - Συνολικό μήκος του αρχείου σε δευτερόλεπτα.
FrameRate - Ρυθμός frame του video σε δευτερόλεπτα.
NumberOfFrames - Συνολικός αριθμός των frame του video .
BitsPerPixel - Bits ανά pixel των δεδομένων video.
Height - Ύψος του frame του video σε pixels.
Width - Πλάτος του frametου video σε pixels.
VideoFormat - Μορφή video.
Name - Όνομα video.
Bit depth - Βάθος χρώματος.
Bit rate - ρυθμός μετάδοσης.

5.1.3 Διαμόρφωση ερωτηματολογίου

Ολοκληρώνοντας την επεξεργασία του video όπως είδαμε στο κεφάλαιο 4 , δημιουργώντας έτσι ένα δεύτερο video αλλοιωμένο για το κάθε ένα από τα παραπάνω. Το επόμενο μας βήμα είναι να δημιουργηθεί ένα τεστ όπου κάποιο δείγμα θεατών θα παρακολουθούσαν ξεχωριστά το αρχικό και το τελικό video και θα απαντούσαν στο ερωτηματολόγιο. Θα γίνει δηλαδή μια στατιστική μελέτη για την υποκειμενική ποιότητα των video που χρησιμοποιήθηκαν.

Αρχικά ο κάθε ένας που συμμετέχει στο τεστ καλείται να συμπληρώσει κάποια προσωπικά στοιχεία. Όπως φύλλο και ηλικία τα οποία θεωρήθηκαν απαραίτητα στοιχεία. Χρησιμοποιήθηκε η κλίμακα υποκειμενικού προσδιορισμού για να χαρακτηρισθεί η ποιότητα των video. Ξεκινώντας με ένα η καλύτερη ποιότητα και πέντε η χειρότερη .

Παρουσίαση του ερωτηματολογίου.

ΑΤΕΙ ΚΡΗΤΗΣ
ΤΜΗΜΑ ΕΦΑΡΜΟΣΜΕΝΗΣ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΠΟΛΥΜΕΣΩΝ

ΣΥΓΚΡΙΤΙΚΟ ΤΕΣΤ

Πειραματικό μέρος της πτυχιακής

ΜΕΛΕΤΗ ΚΑΙ ΣΥΓΚΡΙΣΗ ΜΕΘΟΔΩΝ ΕΚΤΙΜΗΣΗΣ ΥΠΟΚΕΙΜΕΝΙΚΗΣ ΠΟΙΟΤΗΤΑΣ ΤΗΛΕΟΠΤΙΚΗΣ ΕΙΚΟΝΑΣ

Επιμέλεια :Μαρία Σισμανίδου
Υπεύθυνος καθηγητής :Κος Γαρδίκης Γεώργιος

Στοιχεία συμμετέχοντα :

Ημερομηνία : _____ Ώρα : _____

- Φύλλο : Άντρας ____
  Γυναίκα ____
- Ηλικία : 18-20 ____
  20-25 ____
  25-30 ____
  30-40 ____
  40-άνω ____
- Θα παρακολουθήσω το πείραμα με οφθαλμολογικά γυαλιά/φακούς :
  Ναι ____ Όχι____
- Γνώση του αντικείμενου : Καθόλου ____
  Μέτρια ____
  Καλή ____

Εικόνα 5.14 – *Στοιχεία συμμετέχοντα*

Τα video που έχουν υποστεί αλλοιώσεις θα πρέπει να χαρακτηριστούν συμφώνα με την κλίμακα υποκειμενικού προσδιορισμοί της ποιότητας video στο παρακάτω ερωτηματολόγιο.

Συντελεστής	Ποιοτητα	Ενόχληση
Q=1	Εξαιρετική	Δεν αντιλαμβάνεται αλλοιώσεις
Q=2	Καλή	Αντιληπτές αλλά όχι ενοχλητικές
Q=3	Ικανοποιητική	Λίγο ενοχλητική
Q=4	Φτωχή	Ενοχλητική
Q=5	Κακή	Πολύ ενοχλητική

Κλίμακα υποκειμενικού προσδιορισμού

ΠΟΙΟΤΗΤΑ (Q)	1	2	3	4	5
VIDEO 1	___	___	___	___	___
VIDEO 2	___	___	___	___	___
VIDEO 3	___	___	___	___	___
VIDEO 4	___	___	___	___	___
VIDEO 5	___	___	___	___	___

ΣΑΣ ΕΥΧΑΡΙΣΤΩ ΓΙΑ ΤΗΝ ΣΥΜΜΕΤΟΧΗ

Εικόνα 5.15 – *Απαντήσεις συμμετέχοντα*

5.1.4 Διεξαγωγή/στατιστικά ερωτηματολογίου

Το πείραμα υλοποιήθηκε σε όσο το δυνατό κλειστό χώρο ώστε οι θεατές να έχουν συγκέντρωση και να υπάρχει ησυχία και με δύο οθόνες ώστε να μπορούν να βλέπουν ταυτόχρονα τα δύο video. Στα υποκειμενικά πειράματα που γίνονται από τις εταιρείες χρειάζονται χώρο με ιδανικές συνθήκες.

Σημαντικό ρολό έχουν και οι οθόνες όπως ήδη αναφέραμε που χρησιμοποιήθηκαν. Και αυτό γιατί για να υλοποιηθεί μια τέτοια σύγκριση και για να δοθούν έγκυρα αποτελέσματα, σίγουρα απαιτείται και ένα σύστημα με πολύ καλές προδιαγραφές. Οι οθόνες που χρησιμοποιήθηκαν ήταν 17¨ .

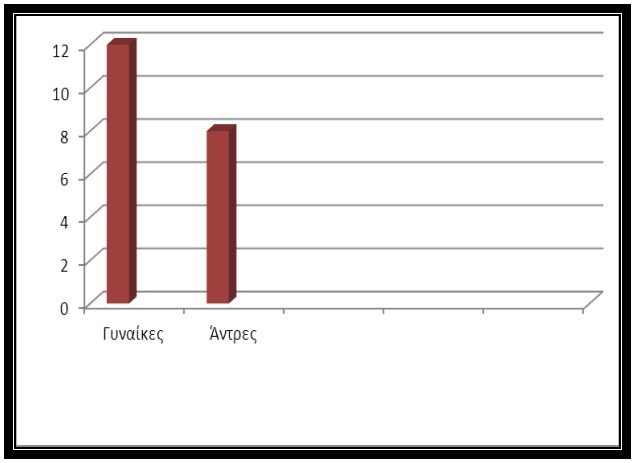
Η παρουσίαση των video έγινε σε συγγενείς φίλους και γνωστούς τους οποίους ευχαριστώ πολύ. Γινόταν ατομικά όπου ο θεατής έπρεπε να παρακολουθεί ένα ένα τα video. Πρώτα το κανονικό και στην συνέχεια αυτό που είχε υποστεί αλλοιώσεις. Στην συνέχεια ο θεατής έπρεπε να βαθμολογήσει την ποιότητα του video που είχε υποστεί αλλοιώσεις σε σχέση πάντα με το κανονικό, βάση της κλίμακας υποκειμενικού προσδιορισμού της ποιότητας του video.

Παρακάτω παραθέτονται τα αποτελέσματα σε πίνακες και διαγράμματα. Το προγράμματα που χρησιμοποιήθηκαν για την διεκπεραίωση της Στατιστικής Ανάλυσης είναι το Microsoft Office Excel 2010.

Στην συνέχεια υπάρχει μια γενική στατιστική ανάλυση, όπου φαίνεται πόσα άτομα πήραν μέρος στο συγκριτικό τεστ. Αναφέρουμε λοιπόν, ότι το πείραμα παρακολούθησαν 20 άτομα διαφορετικής ηλικίας και φύλλου. Αριθμός όχι και τόσο ικανοποιητικός για να μπορέσουμε να βγάλουμε αξιόλογα αποτελέσματα.

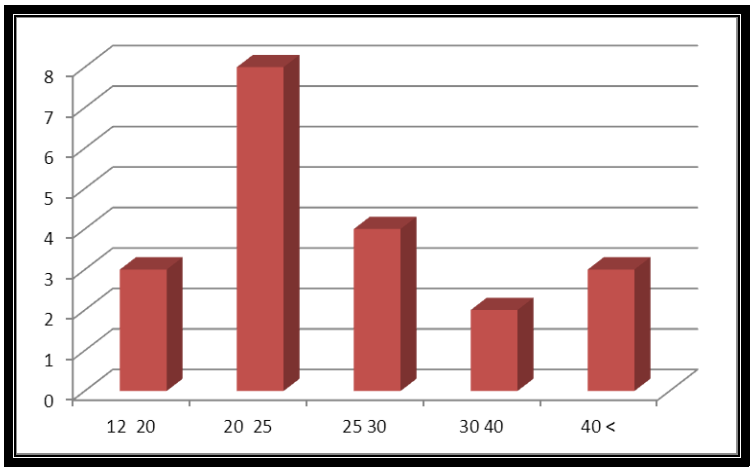
Στο πρώτο μέρος της Στατιστικής Ανάλυσης παρουσιάζονται τα αρχικά στοιχεία του ερωτηματολογίου με αντίστοιχα γραφήματα.

Αρχικά θα παρουσιαστούν τα αρχικά στοιχεία του ερωτηματολογίου, με πίνακες και με τα αντίστοιχα γραφήματα.



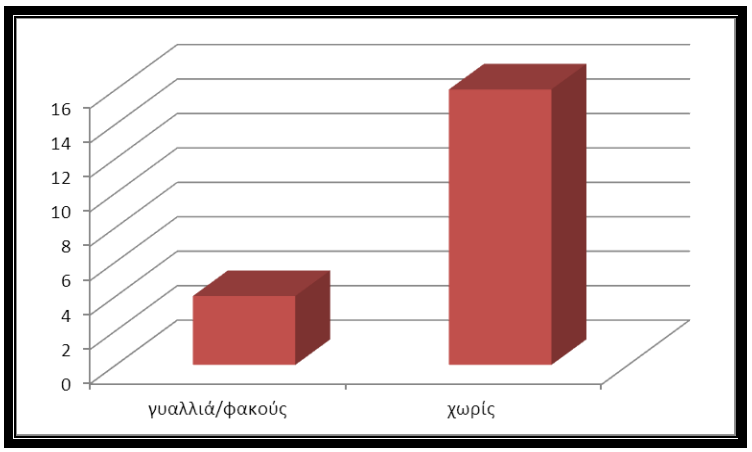
Στην εικόνα παρατηρούμε ότι το μεγαλύτερο ποσοστό συμμετοχής αντιστοιχεί σε κοπέλες και λιγότερο σε αγόρια.

Εικόνα 5.16 – *Φύλλο συμμετέχοντα*



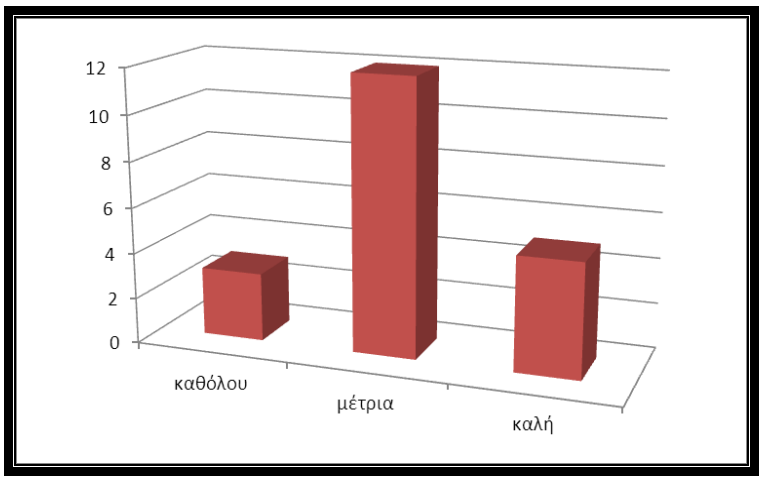
Σύμφωνα με το διάγραμμα το τεστ παρακολούθησαν άτομα ηλικίας 21-30 ετών. Ενώ λίγοι ήταν με τις υπόλοιπες ηλικίες .

Εικόνα 5.17– *Ηλικία συμμετεχόντων*



Οι πιο πολύ από τους συμμετέχοντες δεν φορούσαν γυαλιά ή φακούς .

Εικόνα 5.18– Συμμετοχή με γυαλιά ή μη



Από το τελευταίο γράφημα βλέπουμε ότι οι περισσότεροι που πήραν μέρος είχαν μέτρια γνώση Τα ίδια περίπου ποσοστά φαίνεται να έχουν εκείνοι που δεν είχαν καθόλου γνώση και εκείνοι που είχαν αρκετή.

Εικόνα 5.19 – *Γνώση συμμετεχόντων*

Στην συνέχεια γίνεται Στατιστική Ανάλυση όπου παρουσιάζονται τα στατιστικά που προέκυψαν από το τεστ για κάθε video. Κάτω από κάθε εικόνα σχολιάζονται τα αποτελέσματα.

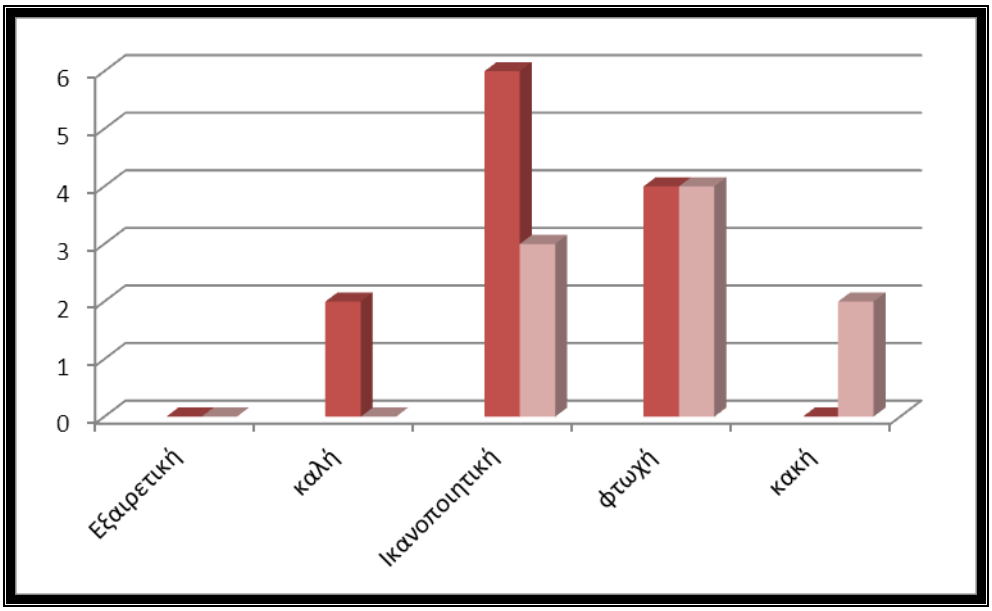
Πρώτο video : Σταθερό σε εσωτερικό χώρο.



Εικόνα 5.20 - Frame 102 αλλοιωμένο video 1



Εικόνα 5.21 - Frame 105 αρχικό video 1



Εικόνα 5.22 – *Διάγραμμα 1^{ο} video 1^{η} στήλη γυναίκες 2^{η} άντρες*

Είναι προφανές ότι , τα αποτελέσματα που αφορά το 1^{ο} video είναι σχετικά καλά έως και φτωχά. Κανένα άτομο δεν τσέκαρε εξαιρετικά. Η αλλοίωση που δέχτηκε το video μπορούμε να πούμε ότι δεν είναι τόσο ενοχλητική στο σύνολο αλλά σε σημεία και ο λόγος είναι επειδή το video ήταν σε σταθερό χώρο. Δέχτηκε αλλοιώσεις σε συγκεκριμένα σημεία και για μικρό χρονικό διάστημα που όπως φαίνεται δεν ενοχλήθηκαν και τόσο οι θεατές, γιατί πέρα από τον παρουσιαστή το υπόλοιπο πλάνο σχεδόν παρέμενε σταθερό. Μεγάλη διάφορα απόψεων βλέπουμε και στα δυο φύλλα αλλά και στην ηλικία.

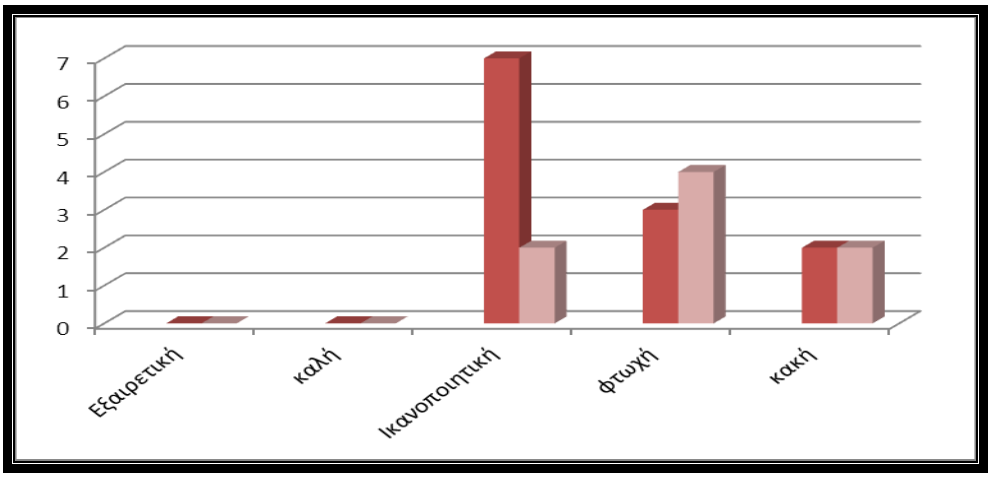
<u>Δεύτερο video</u> : video σε εξωτερικό χώρο με έντονο φυσικό φωτισμό.



Εικόνα 5.23 - *Frame 108 αρχικό video 2*



Εικόνα 5.24 - *Frame 108 αλλοιωμένο video 2*



Εικόνα 5.25 – *Διάγραμμα 2^{ο} video 1^{η} στήλλη γυναίκες 2^{η} άντρες*

Στο δεύτερο video βλέπουμε ότι οι απόψεις μοιράζονται σε ικανοποιητική , φτωχή και κακή. Λόγο του φυσικού φωτισμού σε εξωτερικό χώρο οι αλλοιώσεις πολλές φόρες δεν έγιναν τόσο εμφανείς ή αντιληπτές λόγο φυσικού φωτισμού, παρά μόνο σε συγκεκριμένα κομμάτια. Και σε αυτό το video οι γυναίκες και τα άτομα μεγαλύτερης ηλικίας είναι πιο ελαστικά και ο λόγος ίσως είναι ότι δεν είναι τόσο εξοικειωμένη με την τεχνολογία άρα δεν έχουν και υψηλές απαιτήσεις.

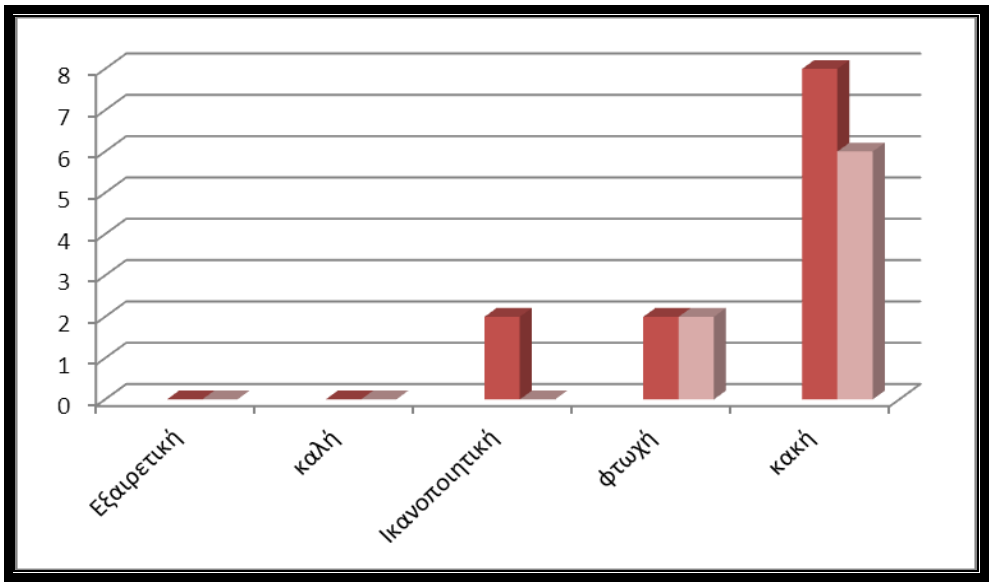
Τρίτο video : Χρησιμοποιήθηκε ένα video σε εξωτερικό χώρο με κίνηση.



Εικόνα 5.26 - *Frame 50 αρχικό video 3*



Εικόνα 5.27 - *Frame 50 αλλοιωμένο video 3*



Εικόνα 5.28 – *Διάγραμμα 3^{ου} video 1^{η} στήλλη γυναίκες 2^{η} άντρες*

Εδώ τα ποσοστά δείχνουν ότι ενοχληθήκαν οι θεατές αφού μοιράζονται σε φτωχή ποιότητα και κακή. Αυτό συμβαίνει γιατί τα χρώματα στο video ήταν θολά και με την κίνηση που είχαν τα αντικείμενα έκαναν τα pixel αρκετά αλλοιωμένα κατά την διάρκεια του video. Λίγα είναι τα ποσοστά που χαρακτηρίζουν την ποιότητα του video έστω ικανοποιητική και τα ποσοστά αυτά προήλθαν από μεγαλύτερη ηλικία.

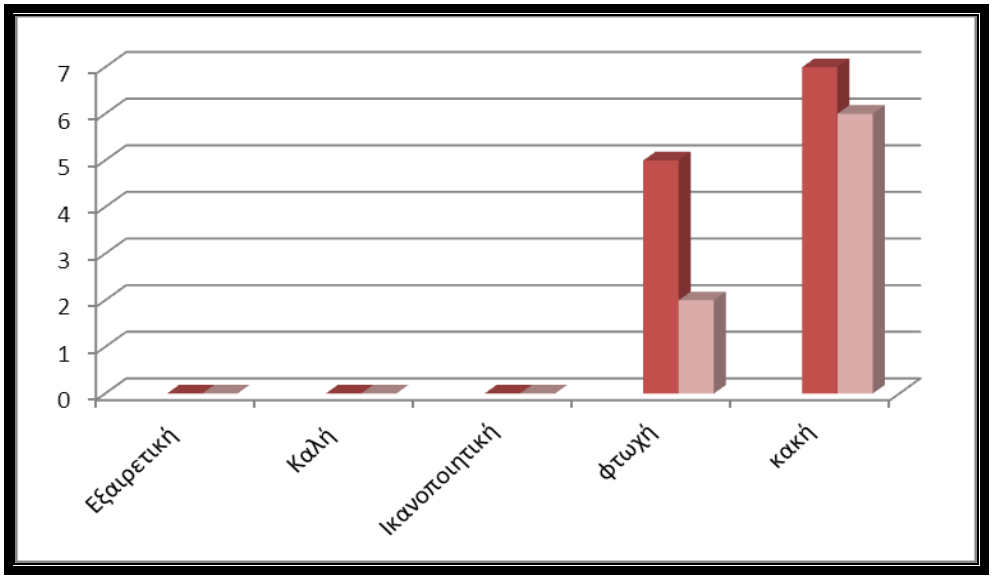
Τέταρτο video : Χρησιμοποιήθηκε ένα video με κινούμενα σχέδια .



Εικόνα 5.29 - *Frame 250 αρχικό video 4*



Εικόνα 5.30 - *Frame 250 αλλοιωμένο video 4*



Εικόνα 5.31 – Διάγραμμα 4^{ου} video 1^{η} στήλλη γυναίκες 2^{η} άντρες

Για το τέταρτο video, τα ποσοστά έχουν "μοιραστεί "σε ποιότητα από κακή έως και φτωχή. Αυτό έχει συμβεί γιατί στο συγκεκριμένο video τα χρώματα είναι έντονα και η διαφορά τους είναι πολύ εμφανής. Δημιουργήθηκαν αλλοιώσεις όπως είναι φυσικό όχι μονό στα χρώματα του video άλλα και στην κίνηση αντικειμένων του.

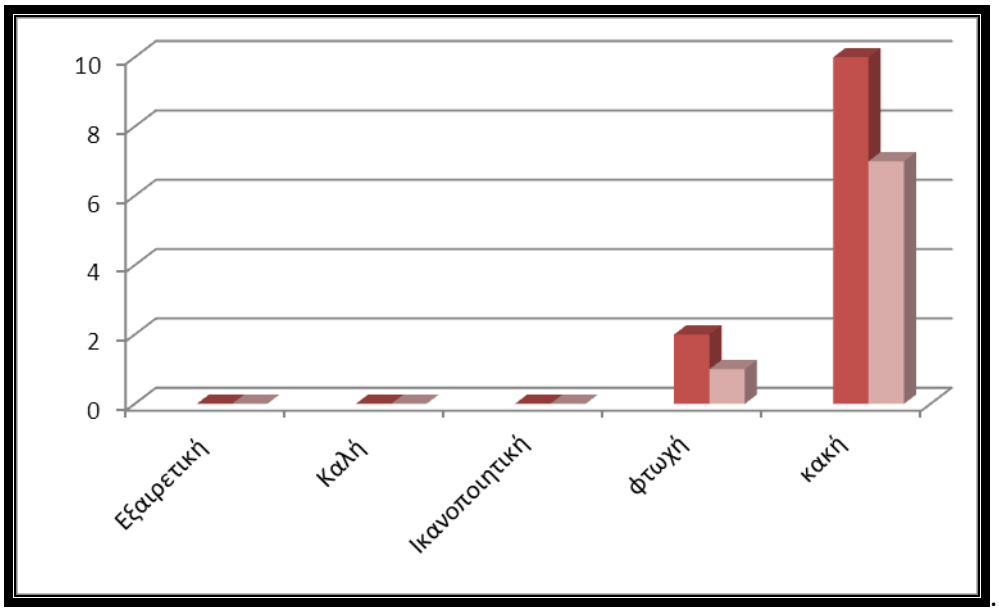
<u>Πέμπτο video</u> : Χρησιμοποιήθηκε ένα video ποδοσφαιρικού αγώνα



Εικόνα 5.32 - *Frame 120 αρχικό video 5*



Εικόνα 5.33 - *Frame 122 αλλοιωμένο video 5*



Εικόνα 5.34 – *Διάγραμμα 5^{ου} video 1^{η} στήλλη γυναίκες 2^{η} άντρες*

Στο συγκεκριμένο video , η κίνηση των αντικειμένων είναι αρκετά γρήγορη και το συγκεκριμένο video επιλέχτηκε για να δούμε κατά πόσο μπορεί το ανθρώπινο μάτι να αντιληφθεί τυχόν καθυστέρηση και αλλοιώσει Pixel στο αλλοιωμένο video. Οι περισσότεροι το χαρακτήρισαν γενικά κακό. Ελάχιστα ήταν τα ποσοστά των απαντήσεων που το θεωρούν έστω φτωχό. Καθώς επίσης όλος ο αριθμός των συμμετεχόντων κυμάνθηκαν στις ίδιες απαντήσεις.

Μελετώντας τα παραπάνω παρατηρήθηκε ότι ήταν ελάχιστες οι φορές όπου κάποιο από απ' αυτά τα αρχικά στοιχεία φαίνεται να επηρεάζουν τα αποτελέσματα. Σίγουρα υπήρχε διαφορά βαθμολόγησης ανάλογα με τον κάθε άνθρωπο, και σίγουρα σημαντικό ρόλο παίζουν σε αυτή την περίπτωση το πόσο γνώστης είναι κάποιος αλλά και η ηλικία γιατί το ανθρώπινο μάτι δεν έχει την ίδια ευαισθησία σε όλες τις ηλικίες αλλά και με την τεχνολογική εξέλιξη οι μικρότερες ηλικίες έχουν πιο πολλές απαιτήσεις.

Και με το κλείσιμο αυτής της υποενότητας να ευχαριστήσω για ακόμα μια φόρα όσους με βοήθησαν να βγάλω τα συγκεκριμένα στατιστικά αποτελέσματα.

5.2. Αντικειμενικές Μέθοδοι

Για τον λόγο ότι η οργάνωση, διεξαγωγή και πραγματοποίηση υποκειμενικών μεθόδων για την εκτίμηση ποιότητας video είναι χρονοβόρα και δαπανηρή, αφού απαιτείται ειδικός εξοπλισμός αλλά και ένα πλήθος θεατών για να συμμετέχουν στη

διαδικασία, για το λόγο αυτό η χρήση υποκειμενικών μεθόδων περιορίζεται μόνο για ερευνητικούς σκοπούς.

Η γρήγορη εξέλιξη της τεχνολογίας καθίστα πλέον αναγκαία την υιοθέτηση τεχνικών απλών, οι οποίες θα μπορούν εύκολα και γρήγορα να παρέχουν αξιόπιστες μετρήσεις της αντιλαμβανόμενης ποιότητας κινούμενης εικόνας. Για αυτόν τον σκοπό αναπτύχθηκαν οι αντικειμενικές μέθοδοι εκτίμησης ποιότητας εικόνας Video , οι οποίες αξιοποιώντας μεγέθη τα οποία μπορούν να μετρηθούν / εκτιμηθούν αντικειμενικά, εξομοιώνουν τις μετρήσεις των αντίστοιχων υποκειμενικών μεθόδων και μας παρέχουν εύκολα και σύντομα αξιόπιστες μετρήσεις.

Για τη λύση προβλημάτων που σχετίζονται σχετικά με θέματα εκτίμησης της ποιότητας κινούμενης εικόνας, έχει συσταθεί η ομάδα VQEG (Video Quality Experts Group), η οποία απαρτίζεται από επιστήμονες ποικίλων ειδικοτήτων, με σκοπό την διεξαγωγή έρευνας στο τομέα της αντιλαμβανόμενης ποιότητας. Αρχικός στόχος της ομάδας ήταν η δημιουργία μιας κοινής μεθόδου μετρήσεως της αξιοπιστίας των διαφόρων μεθόδων εκτίμησης κινούμενης εικόνας, ώστε να είναι δυνατή η αξιολόγησή τους, αλλά και η μεταξύ τους σύγκριση. Οι αντικειμενικές μέθοδοι εκτίμησης εικόνας video ταξινομούνται με βασικό κριτήριο την διαθεσιμότητα / απαίτηση του αρχικού / αψεγάδιαστου video ως οντότητα αναφοράς, με βάση το οποίο θα γίνει και εκτίμηση της ποιότητας εικόνας. Επιπλέον το video αναφοράς θεωρείται ότι είναι άριστης ποιότητας.

Οι πιο γνωστές μέθοδοι που προτάθηκαν αρχικά αλλά και που χρησιμοποιήθηκαν ευρέως είναι Peak Signal to Noise Ratio (PSNR)

5.2.1. Peak Signal to Noise Ratio (PSNR)

Ο όρος PSNR (Peak Signal to Noise Ratio) είναι ένα μέτρο εκτίμησης της επίδρασης των λαθών που συμβαίνουν σε ένα σύστημα μετάδοσης σήματος video, αλλά και της συνολικής απόδοσης ενός τέτοιου συστήματος. Μπορεί να χρησιμοποιηθεί για τη μέτρηση της παραμόρφωσης μιας εικόνας ή πλαισίου λόγω των λαθών στην μετάδοση, σε σύγκριση με την αρχική εικόνα η πλαίσιο. Αναφέρονται διάφοροι τρόποι ορισμού του , ένας από τους οποίους δίνεται από τη σχέση:

$$PSNR = 10 \cdot \log_{10} \left(\frac{MAX_I^2}{MSE} \right)$$
$$= 20 \cdot \log_{10} \left(\frac{MAX_I}{\sqrt{MSE}} \right)$$

όπου *I MAX* είναι η μέγιστη τιμή του pixel μιας εικόνας. Όταν το κάθε pixel αναπαρίσταται με 8-bits τότε η τιμή αυτή είναι 256. Ο όρος PSNR εκφράζεται σε dB και οι συνηθισμένες τιμές για μια συμπιεσμένη εικόνα είναι μεταξύ 30 και 40 dB. Για να υπολογίσουμε τον PSNR σε μια μονοχρωματική εικόνα, όπως στις grayscale εικόνες που μελετάμε στην εφαρμογή μας, αρχικά υπολογίζουμε το Μέσο Τετραγωνικό Σφάλμα (MSE – Mean Squarred Error) ανάμεσα στην αρχική και την τελική μας εικόνα με τον παρακάτω τύπο :

$$MSE = \frac{1}{m\,n} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i,j) - K(i,j)]^2$$

όπου για το κάθε frame έχουμε :

- ✓ m είναι το πλάτος της εικόνας
- ✓ n είναι το ύψος της εικόνας
- ✓ i είναι η στήλη που βρίσκεται το pixel
- ✓ j είναι η γραμμή που βρίσκεται το pixel
- ✓ και με I(i,j) και K(i,j) το αρχικό video και το παραμορφωμένο video αντίστοιχα

Όταν συγκρίνουμε ένα αρχικό video και ένα video που έχει υποστεί συμπίεση το PSNR συνήθως κυμαίνεται από 30 έως 50 db. Όσο πιο μεγάλο το PSNR τόσο λιγότερη η παραμόρφωση .

Αν και το PSNR είναι ο πιο γνωστός τρόπος για την μέτρηση εκτίμησης της ποιότητας εικόνας έχουν αναπτυχτεί και πολύ άλλοι αξιόπιστοι αλγόριθμοι όπως οι

- ✓ CZD πανόμοια λειτουργιά με αυτήν του PSNR αφού μετράει διαφορές μεταξύ pixel.
- ✓ VQM αξιόπιστος αλγόριθμος αφού προσεγγίζει κατά πολύ την ανθρωπινή όραση.

Το PSNR είναι αλγόριθμος πλήρους αναφοράς δηλαδή απαιτείται και το αρχικό και το τελικό video και η εφαρμογή προϋποθέτει ότι οι δόκιμες αναφέρονται στο ίδιο περιεχόμενο και οι κωδικοποιήσεις έγιναν με τον ίδιο αλγόριθμο συμπίεσης. Επίσης υπολογίζεται μόνο για την φωτεινότητα επειδή η ανθρωπινή όραση είναι πιο ευαίσθητη σε διαφορές ποσότητες φωτός έτσι μετρώντας μόνο την φωτεινότητα δεν λαμβάνει υπόψη τις απώλειες από την μετατροπή RGB σε YUV.

Παρακάτω βλέπουμε τέσσερις εικόνες, που εκ πρώτης όψεως φαίνονται μερικές σχεδόν ίδιες διαφέρει όμως η ποιότητα τους και το PSNR .



Q=90, PSNR 45.53dB

Εικόνα 5.35 – *psnr 45.53*



Q=30, PSNR 36.81dB

Εικόνα 5.36 – *psnr 36.81*



Q=10, PSNR 31.45dB

Εικόνα 5.37 – *psnr 31.45*

Οι εικόνες παραπάνω δεν προέρχονται από τα video που χρησιμοποιήθηκαν για το πειραματικό μέρος της πτυχιακής. Μας βοηθούν όμως να πάρουμε μια αρχική εικόνα για το τι είναι το PSNR αλλά και να δούμε ότι καθώς η τιμή του PSNR πέφτει τόσο η ποιότητα μιας εικόνας γίνεται κακή.

5.2.2. Πειραματικό μέρος

Για να γίνει ακόμα πιο κατανοητό το PSNR παρακάτω θα δούμε τις τιμές του σε διάφορες κλίμακες για το video 3 που χρησιμοποιήθηκε για το πειραματικό μέρος της πτυχιακής.

Στην υψηλότερη ποιότητα, με κλίμακα 1, το PSNR είναι 40.7325 δεν μπορούμε να δούμε καμία διαφορά μεταξύ του αρχικού frame και αυτού που ανακατασκευάστηκε. Με την κλίμακα 8, μπορώ να δω πολύ μικρές διαφορές σε ελάχιστα μέρη του frame το PSNR του είναι 33.1528. Καθώς αυξάνεται η κλίμακα περαιτέρω, τα αντικείμενα αρχίζουν να εμφανίζονται με προβλήματα και αρκετά θολά, αυτό πέρα από την υποκειμενική αντίληψη του καθενός για την ποιότητα γίνεται εμφανής και από τις τιμές του PSNR. Έτσι για την κλίμακα 16, όπου πλέον η κατώτερη ποιότητα είναι εμφανής επιβεβαιώνεται και από την τιμή του PSNR που είναι 29.5041. Στην ποιότητα, με κλίμακα 32 , 64 και 112, υπάρχουν ενδιαφέροντα αποτελέσματα. Πέραν του ότι είναι πολύ θολή, το frame φαίνεται να είναι ως επί το πλείστον αποχρώσεις του γκρι και δεν έχει τα χρώματα που είχε το αρχικό frame. Για αυτό άλλωστε το PSNR είναι 26.4183 , 23.8254 και 22.1889 και όπως αναφέραμε στην προηγούμενη ενότητα οι συνηθισμένες τιμές είναι μεταξύ 30 και 40 dB .



Scale 1

- ✓ Mean Square Error = 5.4933
- ✓ Peak Signal to Noise Ratio = 40.7325

Scale 8

- ✓ Mean Square Error = 31.4627
- ✓ Peak Signal to Noise Ratio = 33.1528

Scale 16

- ✓ Mean Square Error = 72.8910
- ✓ Peak Signal to Noise Ratio = 29.5041



Scale 32

- ✓ Mean Square Error = 148.3383
- ✓ Peak Signal to Noise Ratio = 26.4183

Scale 64

- ✓ Mean Square Error = 269.4905
- ✓ Peak Signal to Noise Ratio = 23.8254

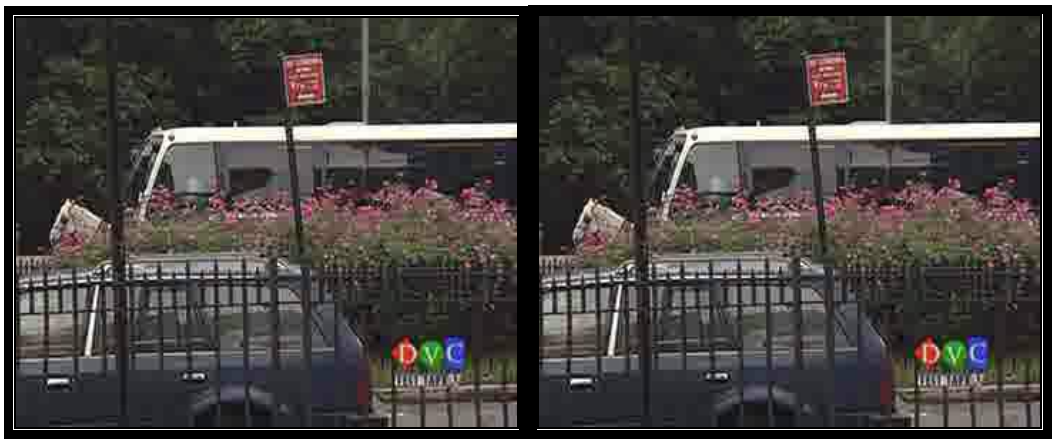
Scale 112

- ✓ Mean Square Error = 392.8149
- ✓ Peak Signal to Noise Ratio = 22.1889

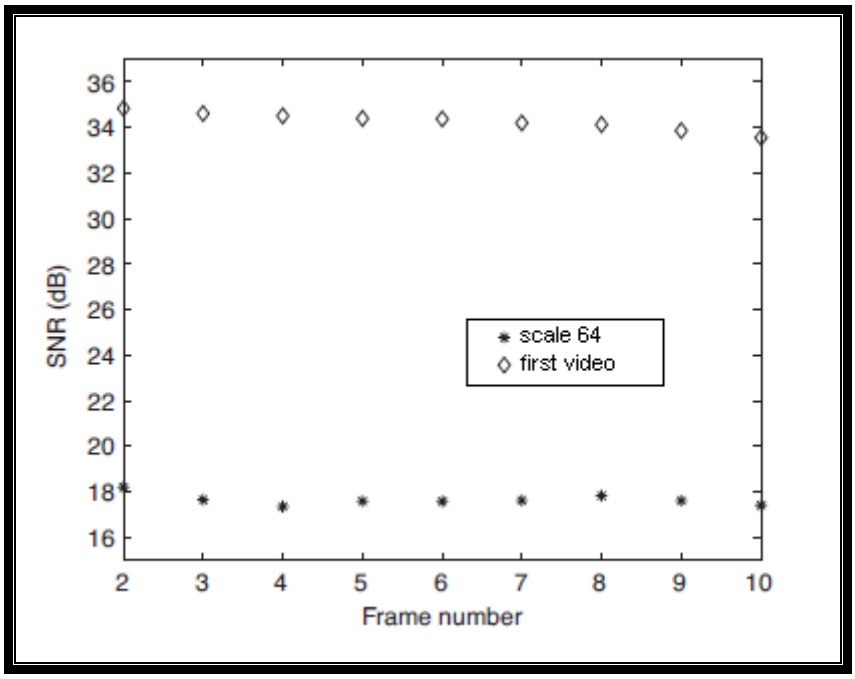
Εικόνα 5.38 – *PSNR και MSE σε διάφορες κλίμακες για το Fame 50 του video 3*

Εδώ πρέπει να προσέξουμε ότι η σχέση του με το MSE είναι αντίστροφη και ότι το PSNR μετρείται σε λογαριθμική κλίμακα. Το PSNR είναι ένα μέτρο ομοιότητας μίας εικόνας ή μιας σειράς εικόνων και παράγεται μετρώντας τη διαφορά των pixel μεταξύ της αρχικής και της δεύτερης εικόνας. Το PSNR εκφράζεται σε decibels. Επίσης έχει σχέση με τη μαθηματική ομοιότητα δύο εικόνων. Οι τιμές του βελησνεκούς του PSNR είναι μεταξύ απείρου (για πανομοιότυπες εικόνες) και 0 (για εικόνες που δεν έχουν καμία ομοιότητα). Όπως θα περιμέναμε, το PSNR μειώνεται όσο η ποιότητα της εικόνας μειώνεται.

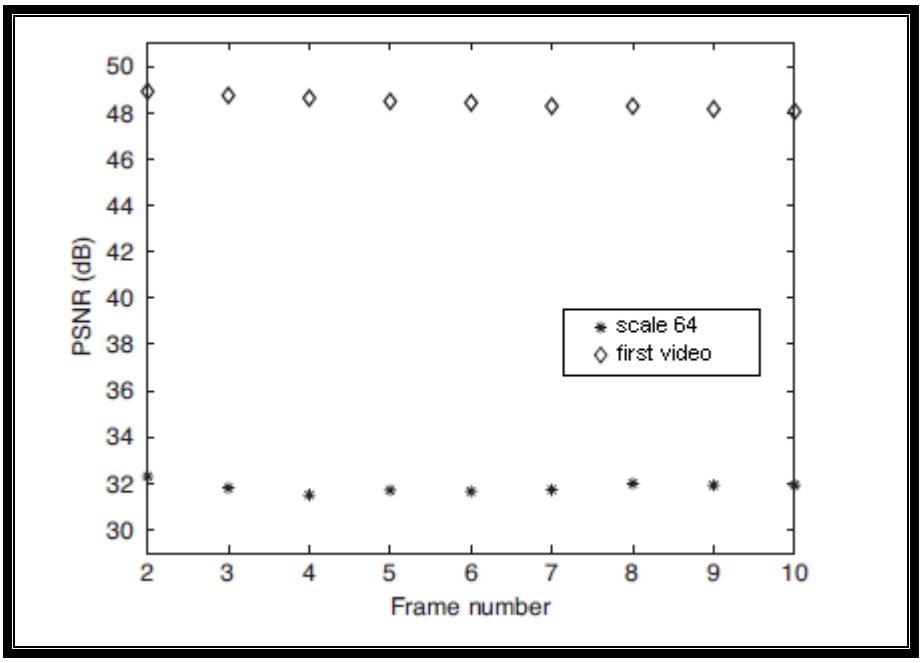
Παρακάτω θα βρούμε το PSNR με τη χρησιμοποίηση δεδομένων video τα οποία είναι .avi , .mpeg4 , .mpg . Οι εικόνες που εξάγει είναι RGB και το μέγεθος τους 352 × 240, 320 x 240, 352 x 288 ,480 x 270 pixel. Η πρώτη εικόνα είναι intra και αποθηκεύεται έτσι όπως είναι. Οι επόμενες εικόνες είναι predictive. Τα frame χαμηλής ποιότητας και υψηλότερης ποιότητας από το video παρουσιάζονται στην εικόνα 5.44 (a) (b), αντίστοιχα. Όπως μπορείτε να δείτε η ποιότητα τον frame έχει αλλοιώσεις σε συγκεκριμένα σημεία όχι όμως τόσο εμφανείς. Αυτό είναι εμφανές από τις τιμές του SNR και τη μέγιστη σήματος προς θόρυβο αναλογία (PSNR) τα οποία παρουσιάζονται στην εικόνα 5.45 και 5.46, αντίστοιχα. Αυτές οι τιμές υπολογίζονται μόνο για εννέα frame.



Εικόνα 5.39 – *(α)scale 1 και (β)scale 16*



Εικόνα 5.40 - *Οι τιμές είναι για 2 έως 10 frame του video 3 οι τιμές SNR είναι σε db.*



Εικόνα 5.41 - Οι τιμές είναι για 2 έως 10 frame του video 3 οι τιμές PSNR είναι σε db.

Ο κώδικας για να βρούμε το PSNR βρίσκεται στο κεφάλαιο 6 (υποενότητα 6.2). Για να εξάγουμε όμως από τα video τα frame χρησιμοποιήθηκε ο παρακάτω κώδικας σε Matlab.

```
Clear all;

Clip = mmreader ('clip4.avi');    %Video που θέλουμε να τρέξουμε
VidFrames = read (Clip);          %Πληροφορίες του video
NumFrames = get (Clip, 'numberOfFrames');   %Αριθμός Frame συνολικά στο video
For k =250: 300   %Δίνουμε frame που θέλουμε να εξάγουμε, χρησιμοποιήθηκε
διαφορετικές τιμές για το κάθε video
Mov (k).cdata = vidFrames (:,:,:, k); % Αποθήκευση πίνακα στον  Mov πίνακα
Mov (k).colormap = []   % Χωρίς colormap ο νέος πίνακας

if (rem(k,10)==0)
A=mov(k).cdata;
[r c h]=size(A);
for i=50:100;
    for j=25:50;
        for l=1:2;
            B(i,j,l)=A(i,j,l)*1.5;
           mov(k).cdata= B(i,j,l);
         end
     end
end
end

imagename=strcat(int2str(k), '.jpeg');  %Αντιγράφει την εικόνα και την κάνει jpeg
imwrite(mov(k).cdata, strcat(' framek,imagename));  %Την  αποθηκεύει με το όνομα frame κ
οπού το κ μας δίνει τον  αριθμό frame
%extractComponents(mov(k).cdata);

end
```

Στην συνέχεια ακολουθεί η εύρεση του psnr για διάφορα frame των video.

Video 1



Frame 273
MSE= 25.5130
PSNR= 34.0632

Frame 274
MSE= 23.1718
PSNR= 34.4812

Frame 275
MSE= 22.8936
PSNR= 34.5337

Frame 276
MSE=18.3534
PSNR= 35.4936



Frame 277
MSE=589.4480
PSNR= 20.4263

Frame 278
MSE=586.0681
PSNR= 20.4513

Frame 279
MSE=586.0681
PSNR= 20.4801

Εικόνα 5.42 – *Ακολουθία frame από 273-279 για το video 1*



Frame 282
MSE=591.2036
PSNR= 20.4134



Frame 283
MSE=589.8377
PSNR= 20.4235



Frame 284
MSE=596.7240
PSNR= 20.3731

Frame 285
MSE=596.7240
PSNR= 21.3932

Frame 286
MSE=590.4564
PSNR= 20.4189

Εικόνα 5.43 – *Ακολουθία frame από 282-286 για το video 1*

Για την παραπάνω ακολουθία του video 1 αρχίζει να "χάνει" την ποιότητα του από τα frames 277 και να συνεχίζει. Εκεί όπου πειράχθηκαν οι τιμές των pixel ώστε να έχουμε ένα παραμορφωμένο video. Οι αλλαγές πλάνου παρουσιάζονται στην εικόνα 5.42 και 5.43 μαζί με τα γειτονικά τους καρέ, ώστε να κατανοηθεί η αιτία του σφάλματος.

Για το παραπάνω video ο μέσος όρος την υποκειμενικής ποιότητας του πειραματικού μέρους ήταν ικανοποιητική προς φτωχή στην αντικειμενική ποιότητα το psnr του video είναι 30.4936 και MSE 32.5678.

Δείκτης Ποιότητας (n)	**Οπτική Ενόχληση**
5	Δεν γίνεται αντιληπτή (εξαιρετική εικόνα)
4	Αντιληπτή αλλά όχι Ενοχλητική
3	Λίγο Ενοχλητική
2	Ενοχλητική
1	Πολύ Ενοχλητική

Τιμές Δείκτης Ποιότητας

Από τον μαθηματικό τύπο που δίδεται, έχουμε:

$PSNR = 10X \log_{10}(2^n-1)^2/MSE$

$30(db)= 10 \times \log_{10} (2^n-1)^2/ 0.32 =>$ <u>n=3 Λίγο Ενοχλητική</u>

Επομένως, για n=3 από τον πίνακα βλέπουμε ότι η οπτική ενόχληση στον χρήστη γίνεται αντιληπτή και επομένως στην συγκεκριμένη περίπτωση η ποιότητα της εικόνας είναι ενοχλητική αφού και στην υποκειμενική εκτίμηση ο μέσος όρος ήταν λίγο ενοχλητική προς ενοχλητική. Και στις δυο περιπτώσεις το αποτέλεσμα είναι σχεδόν ίδιο.

Στην συνέχει θα δούμε τα αποτελέσματα και των δύο πειραμάτων για όλα τα video με τον ίδιο τρόπο.

<u>Video 4</u>



Frame 121
MSE=240.7146
PSNR= 24.3158



Frame 122
MSE=220.4265
PSNR= 24.6982

Frame 123
MSE=1.3667e+003
PSNR= 16.7741

Frame 124
MSE=1.7904e+003
PSNR= 15.6012

Frame 125
MSE=427.7063
PSNR= 21.8193



Frame 125
MSE=427.7063
PSNR= 21.8193

Εικόνα 5.44 – *Ακολουθία frame από 121-125 για το video 4*



Frame 271
MSE=493.4395
PSNR= 21.1985

Frame 272
MSE=484.9617
PSNR= 21.2737

Frame 273
MSE=759.7078
PSNR= 19.3243



Frame 274
MSE=658.4089
PSNR= 19.9458

Frame 275
MSE=600.4242
PSNR= 20.3462

Frame 276
MSE=562.2111
PSNR= 20.6318

Εικόνα 5.45 – *Ακολουθία frame από 271- 276 για το video 4*

Video 5



Frame 101
MSE=1.7194e+003
PSNR= 15.7770



Frame 102
MSE=1.3499e+003
PSNR= 16.8276

Frame 103
MSE=1.1979e+003
PSNR= 17.3467

Frame 104
MSE=2.1767e+003
PSNR= 14.7528

Frame 105
MSE=1.4967e+003
PSNR= 16.3795



Frame 105
MSE=2.4864e+003
PSNR= 14.1751

Frame 107
MSE = 2.2029e+003
PSNR= 14.7009

Frame 108
MSE=1.7188e+003
PSNR= 15.7787

Εικόνα 5.46 – *Ακολουθία frame από 101 – 108 για το video 5*

Video 2



Εικόνα 5.47 – *Ακολουθία frame από 351 – 359 για το video 2*



Εικόνα 5.48 – *Ακολουθία frame από 592– 599 για το video 2*



Εικόνα 5.49 *–psnr 40.2561 frame 632*



Εικόνα 5.50 *–psnr 18.2671 frame 632*

Με τον τρόπο που είδαμε στην σελίδα 105 για το video 2 κάνουμε και στα υπόλοιπα και έτσι έχουμε τον παρακάτω πίνακα :

Video	***Υποκειμενική***	***Αντικειμενική***
Video 1	*Λίγο ενοχλητική - ενοχλητική*	Ενοχλητική
Video 2	*Ενοχλητική*	*Ενοχλητική*
Video 3	*πολύ ενοχλητική*	*πολύ ενοχλητική*
Video 4	*Ενοχλητική- πολύ ενοχλητική*	*πολύ ενοχλητική*
Video 5	*πολύ ενοχλητική*	*πολύ ενοχλητική*

Σύμφωνα με τον παραπάνω πίνακα βλέπουμε ότι τα αποτελέσματα συμπίπτουν. Άλλωστε αυτός είναι ο σκοπός των αντικειμενικών μεθόδων , αξιοποιώντας μεγέθη τα οποία μπορούν να μετρηθούν / εκτιμηθούν αντικειμενικά, εξομοιώνουν τις μετρήσεις των αντίστοιχων υποκειμενικών μεθόδων και παρέχουν εύκολα και σύντομα αξιόπιστες μετρήσεις.

Οι αντικειμενικές μέθοδοι εκτίμησης εικόνας video ταξινομούνται με βασικό κριτήριο την διαθεσιμότητα / απαίτηση του αρχικού / αψεγάδιαστου video ως οντότητα αναφοράς, με βάση το οποίο θα γίνει και εκτίμηση της ποιότητας εικόνας. Επιπλέον το video αναφοράς θεωρείται ότι είναι άριστης ποιότητας.

Η πλειοψηφία των προτεινόμενων αντικειμενικών μεθόδων εκτίμησης ποιότητας εικόνας video βασίζονται σε ένα πλαίσιο εκτίμησης σφάλματος ανάμεσα στην ποιοτικά υποβαθμισμένη κινούμενη εικόνα και στην αντίστοιχη αρχική αψεγάδιαστη υψηλής ποιότητας. Ουσιαστικά οι μέθοδοι αυτοί αποτέλεσαν τα πρώτα βήματα στην ανάπτυξη μεθόδων εκτίμησης ποιότητας κινούμενης εικόνας, αφού ουσιαστικά εκμεταλλευόμενοι το γεγονός ότι το video αποτελείται από μια αλληλουχία εικόνων, δανείζονται τις μεθόδους που χρησιμοποιήθηκαν για την εκτίμηση ποιότητας ψηφιακής εικόνας. Συνεπώς οι μέθοδοι αυτοί συγκρίνουν καρέ-καρέ το περιεχόμενο του αρχικού και του ποιοτικά υποβαθμισμένου video , υιοθετώντας μεθόδους που είχαν ήδη προταθεί στα πλαίσια της ποιότητας εικόνας.

Παρόλα αυτά τέτοιοι αλγόριθμοι έχουν εντόνως κατηγορηθεί ότι δεν μετρούν ικανοποιητικά / επιτυχώς την αντιλαμβανόμενη ποιότητα. Πιο συγκεκριμένα υπάρχει μια πληθώρα λόγων που μπορούν αν αιτιολογήσουν γιατί οι προαναφερθέντες μέθοδοι δεν παρέχουν σωστές και αξιόπιστες μετρήσεις της αντιλαμβανόμενης ποιότητας:

Η τιμή του εκάστοτε Pixel που χρησιμοποιούν οι παραπάνω μέθοδοι για την εξαγωγή εκτιμήσεων της ποιότητας κινούμενης εικόνας, μπορεί να μην αντιπροσωπεύουν στην πραγματικότητα την φωτεινή/χρωματική ένταση που το ανθρώπινο μάτι λαμβάνει.

Η ευαισθησία του ανθρωπίνου οπτικού συστήματος ενδεχομένως να είναι διαφορετική για τα ποικίλα τεχνουργήματα που μπορεί να παρουσιαστούν στην εικόνα, εξαιτίας της ποιοτικής υποβάθμισης. Επιπλέον η ευαισθησία αυτή πολύ πιθανό να εξαρτάται από το περιεχόμενο του καρέ αλλά και από τη χωρική και χρονική δραστηριότητα του περιεχομένου της κινούμενης εικόνας. Για παράδειγμα, για σκηνές με έντονη δραστηριότητα και κίνηση, η ευαισθησία του ματιού στο να εντοπίσει τεχνουργήματα και αλλοιώσεις μειώνεται σημαντικά σε σύγκριση με κινούμενες εικόνες στατικού περιεχομένου. Συνεπώς οι παραπάνω μέθοδοι αγνοούν εντελώς αυτά τα κριτήρια.

Τέτοιο μέθοδοι ουσιαστικά μετρούν το βαθμό του σφάλματος που παρατηρείτε ανάμεσα στην αρχικά και την ποιοτικά υποβαθμισμένη κινούμενη εικόνα. Όμως δύο σήματα που παρουσιάζουν τον ίδιο βαθμό σφάλματος, μπορεί να περιέχουν εντελώς διαφορετικά είδη λαθών. Επιπλέον η απλή άθροιση των σφαλμάτων (όπως ακριβώς γίνεται στο MSE) μπορεί να αποτελεί σημαντική αυθαιρεσία, αφού δεν γνωρίζουμε εάν το οπτικό σύστημα του ανθρώπου αντιλαμβάνεται τα σφάλματα με τον ίδιο τρόπο.

Για αυτούς τους λόγους, τα τελευταία χρόνια παρατηρήθηκε μια έντονη προσπάθεια για την ανάπτυξη μεθόδων, οι οποίες θα προσομοιάζουν τον τρόπο με τον οποίο λειτουργεί το οπτικό σύστημα του ανθρώπου και θα παρέχουν σωστά και αξιόπιστα αποτελέσματα.

Κεφάλαιο 6 – Κώδικας

6.1. Κώδικας κεφαλαίο 4

1. Συνάρτηση mpegproj()

```
Function mpegproj

Fprintf ('\Mpeg Project\n')   %Εμφανίζει στην οθόνη ότι του έχουμε γράψει στα εισαγωγικά
NF = 10;    %Αριθμός των frame που θα επεξεργαστούμε, για 0 θα χρησιμοποιούσαμε όλοι την ταινία
Fprintf ('%i frames\n', NF)   %Εμφανίζει στην οθόνη ότι του έχουμε γράψει στα εισαγωγικά
Mov = getmov (10);    %Κάνει το mov ίσο με την συνάρτηση 2 getmov και nf = 10.
Tic   % Μέτρηση της απόδοσης χρησιμοποιώντας χρονόμετρο.
Mpeg = encmov (mov); %Κάνει το mpeg ίσο με την συνάρτηση 3 encmov και mov.
Fprintf ('Encode time: %s\n', sec2timestr (toc)) %Εμφανίζει στην οθόνη.
Tic % Μέτρηση της απόδοσης χρησιμοποιώντας χρονόμετρο.
mov2 = decmpeg (mpeg);    %Κάνει το mov2 ίσο με την συνάρτηση 4 decmpeg και mpeg.
Fprintf ('Decode time: %s\n', sec2timestr (toc))   %Εμφανίζει στην οθόνη.
Save lastmov mov mpeg mov2  %Αποθήκευση στον δίσκο.
```

2. Συνάρτηση getmov()

```
Function movdata = getmov (NF)   % Επιλέγουμε την ταινία για την επεξεργασία. Κάθε ταινία πρέπει να αποθηκευτεί  στο δίσκο σε ένα αρχείο .mat  .Το αρχείο  .mat περιέχει μια μεταβλητή που ονομάζεται "mov" . Στις συναρτήσεις  loadFileYuv και loadFileY4m γίνεται μετατροπή από YUV σε  Μορφή Y4M στις ταινίες.

 Load clip1
% load clip2
% load clip3
% load clip4
% load clip5

% Οι ταινίες είναι RGB άρα 4-D πίνακας ( ύψος , πλάτος , color component, frame ).
If NF == 0    %Αν το NF είναι ίσο με 0.
   NF = length (mov);   %Τότε το NF είναι ίσο με το μήκος της ταινίας που του δώσαμε.
End
Movdata = repmat (uint8 (0), [size (mov (1).cdata), NF]);   %Δημιουργούν έναν πίνακα.Μετατροπή σε unsigned ακέραιος 0 έως 255.
For i = 1: NF % Για ι από 1 έως NF με βήμα 1.
   Movdata (: , :,  i) = mov(i).cdata;   %Χρήση του πρώτου frame για να πάρει διαστάσεις από τον πίνακα .
End
```

3. Συνάρτηση encmov ()

```
Function mpeg = encmov (mov)

 %Μοτίβο  Πλαίσιο (επαναλαμβάνεται για ολόκληρη την ταινία).
Fpat = 'IPPPP'; % Τυπικά IPPPP
% fpat = 'I';  I για όλα τα  frames
% fpat = ['I', repmat ('P', 1, size (mov, 4))]; % P για όλα τα frames
%βρόχος για όλα τα frames.
k = 0;
Pf = [];
Progress bar
for i = 1:size(mov,4)  % I  από 1 έως mov  όπου επιστρέφει  το μήκος της διάστασης που
έχουμε καθορίσει δηλαδή  το 4.
      f = double (mov (:,:,:,i));   %Το f γίνεται ίσο με τον πίνακα mov αφού έχει μετατροπή σε
double.
       f = rgb2ycc (f);    %Μετατροπή των frame σε YCbCr.
       k = k + 1;
            If k > length (fpat)
            k = 1;
            End
        Ftype = fpat (k);
    % encode frame
        [Mpeg {i}, pf] = encframe (f, ftype, pf);   %Γίνεται ίσο με την συνάρτηση 4 encframe.
        Progress bar (i/ (2*size (mov, 4)))
   End
```

4. Συνάρτηση encframe()

```
Function [mpeg, df] = encframe (f, ftype, pf)   %Συνάρτηση encframe με παραμέτρους f
ftype και pf.

[M, N, i] = size (f);
Mbsize = [M, N] / 16;
Mpeg = struct ('type', [],'mvx', [],'mvy', [],'scale', [],'coef', []);
Mpeg (mbsize (1), mbsize (2)).type = [];
% Loop macroblocks.
Pfy = pf (: 1);
DF = zeros (size (f));
For m = 1: mbsize (1)
   For n = 1: mbsize (2)

      % encode macroblock.
      x = 16*(m-1)+1 : 16*(m-1)+16;
      y = 16*(n-1)+1 : 16*(n-1)+16;
      [mpeg(m,n),df(x,y,:)] = encmacroblock(f(x,y,:),ftype, pf,pfy,x,y);
   End % macroblock loop.
End % macroblock loop.
```

5. Συνάρτηση getblocks()

```
Function b = getblocks (mb)

b = zeros ([8, 8, 6]);
% τέσσερα luma blocks.
b(:,:,1) = mb( 1:8,  1:8,  1);
b(:,:,2) = mb( 1:8,  9:16, 1);
b(:,:,3) = mb( 9:16, 1:8,  1);
b(:,:,4) = mb( 9:16, 9:16, 1);
%Μέσος  όρος γειτονικών.
b(:,:,5) = 0.25 * ( mb(1:2:15,1:2:15, 2) + mb(1:2:15,2:2:16, 2) ...
            + mb(2:2:16,1:2:15, 2) + mb(2:2:16,2:2:16, 2) );
b(:,:,6) = 0.25 * ( mb(1:2:15,1:2:15, 3) + mb(1:2:15,2:2:16, 3) ...
            + mb(2:2:16,1:2:15, 3) + mb(2:2:16,2:2:16, 3) );
```

6. Συνάρτηση rgb2ycc()

```
Function ycc = rgb2ycc (rgb)

%Μετατροπή πίνακα .
m = [ 0.299     0.587     0.144;
    -0.168736 -0.331264  0.5;
     0.5      -0.418688 -0.081312];
%Δεδομένα πίνακα.
[nr,nc,c] = size (rgb);
%Ανασχηματίζει και πολλαπλασιάζει τον πίνακα .
Rgb = reshape (rgb, nr*NC, 3);
% Μετατροπή color coding.
Ycc = m * rgb';
Ycc = ycc + repmat ([0; 0.5; 0.5], 1, nr*NC);
Ycc = reshape (ycc', nr, nc,3);
```

7. Συνάρτηση encmacroblock()

```
Function [mpeg, emb] = encmacroblock (mb, ftype, pf, pf, x, y)
%Συντεταγμένες κβαντοποίησης .
Persistent q1 q2
if is empty(q1)
   q1 = qintra;
   q2 = qinter;
End
%Κλιμάκωση ποιότητας.
Scale = 31;
% Init mpeg struct.
mpeg.type = 'I';
mpeg.mvx = 0;
mpeg.mvy = 0;
% Βρίσκει motion vectors.
If ftype == 'P'
```

```
   mpeg.type = 'P';
   [Mpeg, emb] = getmotionvec (mpeg, mb, pf, pfy, x, y);
   Mb = emb; % Set macroblock to error for encoding.
   q = q2;
Else
   q = q1;
End
% Παίρνει luma και chrom blocks
b = getblocks (mb);
% encode blocks
For i = 6:-1:1
   Mpeg. Scale (i) = scale;
   Coef = dct2 (b(:,:,i));
   Mpeg. Coef (:,:,i) = round( 8 * coef ./ (scale * q) );
End
 %Αποκωδικοποίηση  macroblock αναφορά σε μελλοντικό  P frame
dmb = decmacroblock(mpeg,pf,x,y);
```

8. Συνάρτηση getmotionvec()

```
Function [mpeg, emb] = getmotionvec (mpeg, mb, pf, pf, x, y)

%Γίνεται αναζήτηση μόνο στο Y
Mby = mb (:,:,1);
[M,N] = size (pfy);
 % Λογαριθμική αναζήτηση
Step = 8;   %Αρχικό μέγεθος βήμα για λογαριθμική αναζήτηση
Dx = [0 1 1 0 -1 -1 -1  0  1];
Dy = [0 0 1 1  1  0 -1 -1 -1]; % [origin, right, right-up, up, left-up,
                    %       left, left-down, down, right-down]
Mvx = 0;
Mvy = 0;
While step >= 1
   Minsad = INF;
   For i = 1: length (dx)
      TX = x + mvx + dx (i)*step;
      If (TX (1) < 1) || (M < TX (end))
         Continue
      End
      Ty = y + mvy + Dy (i)*step;
      If (ty (1) < 1) || (N < ty (end))
         Continue
      End
      Sad = sum (sum (abs (mby-pfy (TX, ty))));
      If sad < minsad
         ii = i;
         Minsad = sad;
      End
   End
   Mvx = mvx + dx (ii)*step;
```

```
   Mvy = mvy + Dy (ii)*step;
   Step = step / 2;
  End
Mpeg. Mvx = mvx; Mpeg. Mvy = mvy;
Emb = mb - pf(mvx, y+mvy,:); % Error macroblock
```

9. Συνάρτηση decmpeg()

```
Function mov = decmpeg (mpeg)

Movsize = size (mpeg {1});
Mov = repmat (uint8 (0), [16*movsize (1:2), 3, length (mpeg)]);
% Loop frames
Pf = [];
For i = 1: length (mpeg)
   % Αποκωδικοποίηση     frame
   f = decframe (mpeg {i}, pf);
   % Κρυφό previous frame
   Pf = f;
   % Μετατροπή frame σε  RGB
   f = ycc2rgb (f);
   % σιγουρεύουμε το ότι η ταινία  8 bit
   f = min (max (f,0), 255);
   Mov (:,:,:,i) = uint8(f);
   Progress bar ((length (mpeg))/(2*length(mpeg)))
End
```

10. Συνάρτηση decframe()

```
Function Fr = decframe (mpeg, pf)

Mbsize = size (mpeg);
M = 16 * mbsize (1);
N = 16 * mbsize (2);
 % Loop macroblocks
Fr = zeros (M, N,3);
For m = 1: mbsize (1)
   For n = 1: mbsize (2)
      % Κατασκευή frame
      x = 16*(m-1)+1 : 16*(m-1)+16;
      y = 16*(n-1)+1 : 16*(n-1)+16;
      fr(x,y,:) = decmacroblock(mpeg(m,n),pf,x,y);
   End % macroblock loop
End % macroblock loop
```

11. Συνάρτηση putblocks()

```
Function mb = putblocks (b)
```

```
Mb = zeros ([16, 16, 3]);
% Τέσσερα luma blocks
mb( 1:8,  1:8,  1) = b(:,:,1);
Mb ( 1:8,  9:16, 1) = b(:,:,2);
Mb ( 9:16, 1:8,  1) = b(:,:,3);
Mb ( 9:16, 9:16, 1) = b(:,:,4);
% Δύο chrom blocks
z = [1 1; 1 1];
Mb (:,:, 2) = kroon(b(:,:,5),z);
Mb (:,:,3) = kroon(b(:,:,6),z);
```

12. Συνάρτηση ycc2rgb()

```
Function rgb = ycc2rgb (ycc)

%Μετατροπή πίνακα
m = [ 0.299     0.587     0.144;
    -0.168736 -0.331264  0.5;
     0.5      -0.418688 -0.081312];
m = inv (m);
% Παίρνει τα δεδομένα του video
[Nr, NC, c] = size (ycc);
%πολλαπλασιάζει τον πίνακα
Ycc = reshape (ycc, nr*NC, 3);
%Μεταμορφώνει την  χρωματική κωδικοποίηση
Rgb = ycc - repmat ([0, 0.5, 0.5], nr*NC, 1);
Rgb = m * rgb';
%Ανασχηματίζει  στο αρχικό μέγεθος
Rgb = reshape (rgb', nr, NC, 3);
```

13. Συνάρτηση decmacroblock()

```
Function mb = decmacroblock (mpeg, pf, x, y)

%κβαντοποίηση συντεταγμένων
Persistent q1 q2
If is empty (q1)
   q1 = qintra;
   q2 = qinter;
End
Mb = zeros (16,16,3);
% Πρόβλεψη με motion vectors
If mpeg. Type == 'P'
   Mb = pf(x+mpeg.mvx, y +mpeg. Mvx,:);
   q = q2;
Else
   q = q1;
End
% αποκωδικοποίηση blocks
```

```
For i = 6:-1:1
   Coef = mpeg. Coef (:,:, i) .* (mpeg. Scale (i) * q) / 8;
   B (:,:,i) = idct2(coef);
End
 % κατασκευή macroblock
Mb = mb + putblocks (b);
```

14. Συνάρτηση qinter()

```
Function q = qinter
%Πίνακας κβαντοποίησης για    P η B frames
% q = repmat (16, 8, 8);
q = 16;
```

15. Συνάρτηση qintra()

```
Function q = qintra
% Πίνακας κβαντοποίηση για I frames
q = [ 8 16 19 22 26 27 29 34;
    16 16 22 24 27 29 34 37;
    19 22 26 27 29 34 34 38;
    22 22 26 27 29 34 37 40;
    22 26 27 29 32 35 40 48;
    26 27 29 32 35 40 48 58;
    26 27 29 34 38 46 56 69;
    27   9 35 38 46 56 69 83 ];
```

16. Συνάρτηση idct2()

```
Function y = idct2(x)

% Perform 2-D IDCT
%Χρησιμοποιείται dctmtx
Persistent d
If is empty (d)
   d = dctmtx (8);
End
y = d' * x * d;
%Υπολογισμός DCT   y = idct(x); στήλες   y = idct(y')'; γραμμές
```

6.2. Κώδικας κεφάλαιο 5

```
Clear;
N = 8;     % Μέγεθος block είναι N x N pixels
N2 = 2*N;
W = 16;    %Το μέγεθος αναζήτησης W x W pixels
quantize scale = 4;    % Χρησιμοποιείται μόνο για την πρώτη ροή
InFile = 'Clip1.avi';     % AVI αρχείο από το MATLAB
StrIndx = regexp (inFile,'\.');
If strcmpi ('avi', inFile (strIndx+1: strIndx+3))
M = aviread (inFile);   % Υπάρχουν 114 frames 320x240x3
F = int16 (1:10);        % frames   encode
Else
F = int16 (40:49);      % frames encode
M.cdata = cell (1, length (F));
For k = 1: length (F)
strIndx1 = regexp (inFile,'0');
inFile1 = strcat (inFile (1:strIndx1 (1)), num2str (F (k)), inFile (strIndx: end));
M (k).cdata = imread (inFile1);
end
end

% διαιρείται το  μέγεθος της εικόνας κατά 8
[X, Y, Z] = size (M (1).cdata);
If mod(X, 8)~=0
Height = floor(X/8)*8;
Else
Height = X;
End
If mod(Y, 8)~=0
Width = floor(Y/8)*8;
Else
Width = Y;
End
Depth = Z;
Clear X Y
If Depth == 3
A = rgb2ycbcr (M (1).cdata);    % Μετατροπή RGB σε YCbCr και διατηρεί μόνο Y
Y_ref = A(:,:,1);
else
A = M (1).cdata;
Y_ref = A;
End
% το πλαίσιο αναφοράς αριστερά & δεξιά και από πάνω & κάτω
Y_ref = double (padarray (Y_ref,[W/2 W/2],'replicate'));
% αποθήκευση SNR και PSNR
Base_snr = zeros (1, length (F)-1);
Enhanced _snr = zeros (1, length (F)-1);
Base_snr = zeros (1, length (F)-1);
```

```
Enhanced_psnr = zeros (1, length (F)-1);
%κωδικοποίηση μονόχρωμου video
 For f = 2: length (F)
If Depth == 3
B = rgb2ycbcr (M (f).cdata);
Y_current = B (:,:, 1);
Else
Y_current = M (f).cdata;
End
Y_current = double (padarray (y current,[W/2 W/2],'replicate'));
For r = N: N: Height
Rblk = floor(r/N);
For c = N: N: Width
Cblk = floor (C/N);
D = 1.0e+10;
For u = -N: N
For v = -N: N
d = y current(r+1: r+N, c+1: c+N)-y ref(r+u+1: r+u+N,c+v+1:c+v+N);
d = sum (abs (d (:)));    % block distance
between pixels
if d < D
D = d;
End
End
End

% compensated difference coding
Temp = Y_current(r+1:r+N,c+1:c+N)-Y_ref(r+1+y1:r+y1+N,c+1+x1:c+x1+N);
TemP = dct2 (temp);     % DCT of difference
s = sign (TemP);
TemP1 = s .* round(abs(TemP)/(16*quantizer scale) *(16*quantizer scale); %
quantize/dequantize DCT
Temp = idct2 (TemP1); % IDCT
Base(r-N+1:r,c-N+1:c) = y ref(r+1+y1:r+y1+N,c+1+x1:c+x1+N)+temp;  % reconstructed block -
base quality
Delta_DCT = TemP - TemP1;   % incremental DCT
s1 = sign (delta_ DCT);    % extract the sign of incremental DCT
Delta_DCT = s1.* round (abs (Delta_DCT)/4)*4;
temp1 = idct2 (TemP1 + delta DCT);
Enhanced(r-N+1: r, c-N+1: c) = y ref(r+1+y1: r+y1+N, c+1+x1: c+x1+N) +temp1;
End
End
% υπολογισμός SNRs PSNRs
Base_snr (f-1) = 20*log10 (std2 (Y_current (N+1: Height+N, N+1: Width+N))
/std2 (Y_current (N+1: Height+N, N+1: Width+N)-Base));
Enhanced_psnr (f-1) = 20*log10 (std2 (Y_current (N+1: Height+N, N+1:
Width+N))/std2 (Y_current (N+1: Height+N, N+1: Width+N)-Enhanced));
Base_snr (f-1) = 20*log10 (255/std2 (Y_current (N+1: Height+N, N+1: Width+N)-Base));
Enhanced_psnr (f-1) = 20*log10 (255/std2(Y_current(N+1:Height+N,N+1:Width+N)-
Enhanced));
```

```
% Αντικαταστήσει τα προηγούμενα πλαίσια από τα ανακατασκευασμένα
Y_ref = Base;
Y_ref = double (padarray (Y_ref,[W/2 W/2],’replicate’));
End
Figure, plot (F (2: end), Base snr,’k*’,’LineWidth’,1 ), hold on
Plot (F (2: end), Enhanced snr,’kd’,’LineWidth’,2), title(’SNR (dB)’)
Axis ([F (2) F (end) min (Base snr)-2 max (Enhanced snr) +2])   %Σκηνή clip1.avi
Legend (’Base Quality’, ‘Enhanced Quality’,0)
X label(’Frame #’), y label(’SNR (dB)’), hold off
Figure, plot (F (2: end), Base_snr,’k*’,’LineWidth’,1), hold on
Plot (F (2: end), Enhanced_psnr,’kd’,’LineWidth’,2), title(’PSNR (dB)’)
Axis ([F(2) F(end) min(Base psnr)-2 max(Enhanced psnr)+2]) %Σκηνή clip1.avi
Legend (’Base Quality’, ‘Enhanced Quality’,0)
Xlabel (’Frame #’), ylabel (’PSNR (dB)’), hold off
```

Λεξιλόγιο

Aliasing: Τα «σκαλάκια» που παρατηρούμε μέσα από μόνιτορ σε διαγώνιες ή καμπυλωτές γραμμές.
Artifact: Θόρυβος που δημιουργείται σε μια εικόνα ή στον ήχο. Στην ψηφιακή εικόνα, τα artifacts μπορούν να προέλθουν από την υπερφόρτωση στην συσκευή ή εφαρμογή εισόδου με υψηλό σήμα ή από υπερβολική ή λάθος χρήση συμπιεστή.
Aspect ratio: Ο λόγος του μήκους της εικόνας προς ο ύψος της. Αναφέρεται και ως ratio και το συναντούμε είτε ως λόγο είτε ως αποτέλεσμα της διαίρεσης, π.χ. 4:3 ή 1.33
AVI: Σύντμηση των λέξεων Audio-Video Interleaved. Ένας αλγόριθμος που έχει κατασκευαστεί από την Microsoft και χρησιμοποιείται για τον συγχρονισμό και την συμπίεση αναλογικού ήχου και εικόνας.
Binary: Το δυαδικό σύστημα που χρησιμοποιείται για την κωδικοποίηση ψηφιακών πληροφοριών. Αποτελείται από δύο στοιχεία: το 1 και το 0.
Bit rate: Η μέτρηση του μεγέθους ενός αρχείου σε σχέση με μια σταθερή χρόνου.
Chroma: Το μέρος του σήματος βίντεο που περιέχει την πληροφορία χρώματος.
Chromatic Aberration: Μετακινήσεις χρώματος και χρωματικά λάθη στην εικόνα που μπορούν να προκύψουν είτε από προβληματικό φακό είτε από πρόβλημα της κάμερας να καταγράψουν και τα τρία χρώματα. Οι κάμερες που διαθέτουν ένα CCD είναι πιο πιθανό να παρουσιάσουν αυτού του είδους το πρόβλημα.
Chrominance: Οι πληροφορίες απόχρωσης (hue) και κορεσμού χρωμάτων (saturation) που βρίσκονται σε ένα σήμα βίντεο.
Cinepak : Ένας τρόπος συμπίεσης αρκετά κακής ποιότητας που μπορεί να χρησιμοποιηθεί μέσα από το QuickTime. Χρησιμοποιείται ως επί των πλείστον σε CD-ROM και το internet.
Codec: Ένωση των λέξεων Compressor / Decompression. Είναι ο επιλεγόμενος αλγόριθμος που χρησιμοποιείται για την συμπίεση ή την αποσυμπίεση εικόνας και ήχου.
Color Depth: Το βάθος χρώματος, ο αριθμός των χρωμάτων που χρησιμοποιούνται για την προβολή μιας εικόνας ή βίντεο. Όσο μεγαλύτερος ο αριθμός bits ανά pixel που χρησιμοποιούνται για την αποθήκευση ενός χρώματος, τόσο μεγαλύτερο το βάθος χρώματος.
Component video: Το σήμα βίντεο που αποτελείται από τρία ξεχωριστά σήματα χρωμάτων. Συνήθως αναφερόμαστε σε αυτό όταν μιλάμε για το σήμα Y, R-Y, B-Y (Luminance/φωτεινότητα, κόκκινο μείον φωτεινότητας, μπλε μείον φωτεινότητας). Έχουμε ακόμα τα RGB (Κόκκινο, Πράσινο, Μπλε) και YCbCr (Luminance, Chroma μείον Μπλε, Chroma μείον Κόκκινο).
Composite Video: Είναι το σύνθετο σήμα εικόνας που περιέχει όλες τις πληροφορίες μαζί (Luminance, Chroma και παλμό συγχρονισμού).
Depth of Field: Το βάθος πεδίου είναι η περιοχή, η ζώνη κατά την οποία η εικόνα δεν είναι θολή.
DV: Ένα format ψηφιακής εικόνας κατασκευασμένο από μια ομάδα κατασκευαστών, η οποία χρησιμοποιεί έναν MJPEG codec για να ψηφιοποιήσει 25 καρέ το δευτερόλεπτο εικόνα διαστάσεων 720x576.
DVI: Digital Video Input. Αναφέρεται σε μεταφορά ψηφιακού σήματος εικόνας με μέγιστη οριζόντια ανάλυση τα 1600 pixel.
FPS: Frames per Second, είναι ο ρυθμός εναλλαγής των καρέ ανά δευτερόλεπτο. Στο φιλμ έχουμε 24, στο PAL / SECAM 25 και στο NTSC 30 που ουσιαστικά είναι 29.97.
Frame: Είναι το καρέ στα γαλλικά ή το κάδρο. Μια πλήρης εικόνα φιλμ ή video. Περιέχει δυο πεδία (fields). Χρειάζονται τουλάχιστο 18 καρέ ανά δευτερόλεπτο για τα φανεί η κίνηση στο ανθρώπινο μάτι και 24 καρέ αν θέλουμε και ήχο να την συνοδεύει.

Frame rate: Ο ρυθμός εναλλαγής των καρέ ανά δευτερόλεπτο κατά τη διάρκεια της εγγραφής και αναπαραγωγής φιλμ ή βίντεο.
GOP: Ακρωνύμιο των Group of Pictures. Πρόκειται για ορολογία που συναντούμε στο MPEG. Ένα GOP εμπεριέχει διαφόρων ειδών συμπιεσμένα καρέ: τύπου I και P&B.
Lossless: Υποδηλώνει πως η μέθοδος συμπίεσης που επιλέχθηκε δεν μειώνει την ποιότητα της εικόνα που συμπιέζεται.
Lossy: Υποδηλώνει πως η μέθοδος συμπίεσης που επιλέχθηκε μειώνει την ποιότητα της εικόνα που συμπιέζεται.
Luminance: Η ένταση της φωτεινότητας των γκρίζων σημείων μιας εικόνας σε ένα σήμα βίντεο.
MJPEG: Ένας codec πολύ καλής ποιότητας με μικρές απώλειες στην εικόνα που όμως ζητά μεγάλο αποθηκευτικό χώρο και συνήθως ειδικό hardware για την χρήση του.
Pixel: Συντομογραφία των λέξεων PIXture Element που σημαίνει ένα στοιχείο της εικόνας.
Polarizer: Φίλτρο που τοποθετείται στον φακό και πολώνει το φως που το διαπερνά. Χρησιμοποιείται για την μείωση των ανακλάσεων και έτσι για τον κορεσμό των χρωμάτων.
Resolution: Πρόκειται για την ανάλυση και αφορά τις διαστάσεις του τηλεοπτικού κάδρου. Οι διαστάσεις του κάδρου μιας DV εικόνας είναι 720x576 που σημαίνει μήκος 720 pixel με ύψος 576 pixel.
RGB color space: Είναι ο τρόπος περιγραφής του χρώματος ανάλογα με τα ποσοστά κόκκινου, πράσινου και μπλε που υπάρχουν σε αυτό. Είναι ο πιο συνηθισμένος τρόπος περιγραφής χρωμάτων στους υπολογιστές σε ότι αφορά τα μόνιτορ.

Βιβλιογραφία

Κεφάλαια βιβλίων

[1] Σ. Ν. Δημητριάδης Α.Σ. Πομπότσης Ε.Γ Τριανταφύλλου Τεχνολογία πολυμέσων πράξη και θεωρία ‘ 2004.

[2] Behrouz A. Forouzan. «Εισαγωγή στην Επιστήμη των Υπολογιστών», Επιμέλεια Γιώργος Στεφανίδης, Αλέξανδρος Χατζηγεωργίου, Εκδόσεις Κλειδάριθμος.

[3]Σήματα, Συστήματα και Κυκλώματα Από τον Ηρακλής Γ. Δημόπουλος [σελ. 25].

[4]Transporting compressed digital video Από τον Xuemin Chen [σελ.38 - 39].

[5]Image and video compression standards: algorithms and architectures Από τον Vasudev Bhaskaran [Κεφ. 2].

[6]MPEG-4 facial animation: the standard, implementation and applications από τον Igor S. Pandzic , Ryoichi Komiya, Robert Forchheimer [σελ.32]

[7] The MPEG handbook: MPEG-1, MPEG-2, MPEG-4 από τον/την John Watkinson

[8] Video compression techniques από τον/την Wolfgang Effelsberg , Ralf Steinmetz

Ιστοσελίδες

[1]http://en.wikipedia.org/wiki/Sampling_signal_processing/

[2]http://www.jhu.edu/signals/sampling/index.html/

[3]http://en.wikipedia.org/wiki/Nyquist%E2%80%93Shannon_sampling_theorem/

[4]http://www.mpeg.org/

[5]http://www.it.uom.gr/project/MultimediaTechnologyNotes/chap2b_5.htm

[6]http://www.cs.cf.ac.uk/Dave/Multimedia/

[7]http://tech.ebu.ch/docs/techreview/trev_271-evain.pdf/

[8]http://en.wikipedia.org/wiki/Peak_signal-to-noise_ratio

[9]http://www.image.ntua.gr/projects/meleti172KTP/node/11

[10]http://www.it.uom.gr/project/MultimediaTechnologyNotes/chap2b_5.htm

[11]http://www.wcl.ece.upatras.gr/dalas/lib/exe/fetch.php?media=diplomat_thesis.pdf

[12]http://www.multilingualarchive.com/ma/enwiki/el/Analog_television

[13]http://el.wikipedia.org/wiki/ΨηφιακήΤηλεόραση/

[14] http://www2.ucy.ac.cy/~georgios/bookfiles/MATLABbook.pdf