

# Σημειώσεις Πρότυπα Κωδικοποίησης δεδομένων I

## ΠΕΡΙΕΧΟΜΕΝΑ

<b>Σκοπός του Μαθήματος</b> .....	<b>3</b>
1. Η Αναπαράσταση των δεδομένων στον υπολογιστή.....	5
2. ISO.....	14
3. ΓΛΩΣΣΕΣ ΣΗΜΑΝΣΗΣ.....	16
3. ΓΛΩΣΣΕΣ ΣΗΜΑΝΣΗΣ.....	16
3.1 HTML.....	16
3.2 SGML.....	20
3.3 XML.....	21
3.4 TEI.....	26
4. Ήχος .....	28
4.1 Βασικές έννοιες του ήχου.....	28
4.2 Μορφότυπα ήχου.....	31
4.3 Μετάδοση φωνής .....	32
4.4 Κωδικοποίηση μορφής σήματος.....	32
4.5 Κωδικοποίηση πηγής.....	32
4.6 Μέθοδοι Αναγνώρισης / Σύνθεσης φωνής .....	33
4.7 Ποιότητα που επιτυγχάνεται.....	34
4.8 Μουσική και το πρότυπο MIDI.....	35
4.9 Εισαγωγή στο MIDI.....	36
4.10 Συσκευές MIDI.....	36
5. Ψηφιακή Εικόνα – Βασικές Έννοιες – Βασικοί Όροι .....	38
5.1 Αναλογικό σε Ψηφιακό. ....	38
5.2 Χωρική Διάταξη, Ανάλυση, Διακριτικότητα (Spatial Resolution).....	40
5.3 Άλλες σχετικές έννοιες - Pixels Per Inch (PPI), Dots Per Inch (DPI) & Lines Per Inch (LPI).....	41
5.4 Μεγέθη και Ανάλυση .....	42
5.5 Βάθος - Bit Depth.....	45
5.6 Η Έγχρωμη Εικόνα και το χρώμα στον Υπολογιστή, Βάθος Χρώματος.....	45
5.7 Μεγέθη Αρχείων.....	48
5.8 Ανύσματα (vectors) και bitmaps. ....	48
6. Τελικός Προορισμός – Output. ....	52
7. Σύλληψη της Ψηφιακής Εικόνας – Image Capture – Input. ....	53
8. Αποθήκευση, File types.....	55
9. ΣΥΜΠΙΕΣΗ ΔΕΔΟΜΕΝΩΝ .....	57
9.1 Αποθηκευτικός χώρος .....	57
9.2 Απαιτήσεις κωδικοποίησης .....	57
9.3 ΤΕΧΝΙΚΕΣ ΚΩΔΙΚΟΠΟΙΗΣΗΣ (ΠΗΓΗΣ, ΕΝΤΡΟΠΙΑΣ ΚΑΙ ΥΒΡΙΔΙΚΗ).....	60
9.4. Μερικές βασικές τεχνικές κωδικοποίησης .....	64
9.5 ΤΟ ΠΡΟΤΥΠΟ JPEG.....	70
9.6 Κβαντοποίηση (Quantization).....	75
9.7 Κωδικοποίηση βασισμένη στην εντροπία (Entropy Encoding).....	75
9.8 ΡΟΗ ΔΕΔΟΜΕΝΩΝ (ΡΟΗ ΗΧΟΥ).....	77
9.9 Mp3 .....	78
9.10 ΡΟΗ ΚΙΝΟΥΜΕΝΗΣ ΕΙΚΟΝΑΣ (VIDEO).....	79

9.11 MPEG-2.....	81
9.12 MPEG-4.....	87

## ***Σκοπός του Μαθήματος***

Τα συστήματα πολυμέσων γίνονται ένα αναπόσπαστο κομμάτι του ετερογενούς περιβάλλοντος υπολογιστών και επικοινωνιών στο οποίο ζούμε. Έχουμε δει μια εκρηκτική αύξηση των υπολογισμών, επικοινωνιών και εφαρμογών πολυμέσων την τελευταία δεκαετία με αποκορύφωμα τη δικτυακή επέκταση των υπερκείμενων και υπερμέσων στον παγκόσμιο ιστό. Οι εφαρμογές αυτές κάνουν την εμφάνιση τους με καταγιστικούς ρυθμούς και είναι στενά συνδεδεμένες με ένα αυξανόμενο αριθμό σχετικών προτύπων και τεχνολογιών που χρησιμοποιούνται για τη μορφοποίηση, την κωδικοποίηση και την απεικόνιση των εφαρμογών αυτών.

Οι ψηφιακές βιβλιοθήκες η ηλεκτρονική δημοσίευση είναι μία από τις εφαρμογές που χρησιμοποιούν πρότυπα κωδικοποίησης που βασίζονται σε κωδικοποιήσεις δεδομένων και μετα-δεδομένων. Η κωδικοποίηση των δεδομένων αυτών στηρίζονται σε πρότυπα κωδικοποίησης τα οποία μπορεί να είναι κωδικοποίηση χαρακτήρων κειμένου ή ήχου ή εικόνας ή βίντεο ή συνδυασμός αυτών. Τα μετα-δεδομένα στηρίζονται σε μια δομημένη περιγραφή των δεδομένων χρησιμοποιώντας διάφορες γλώσσες σήμανσης (HTML, SGML, XML) με σκοπό τα κωδικοποιημένα αρχεία (κειμένου, ήχου εικόνας βίντεο) να μπορούν να διαβαστούν σε δίκτυο ή τοπικά σε έναν υπολογιστή με ένα κατάλληλο λογισμικό εξέτασης ή αναζήτησης).

Σκοπός του μαθήματος αυτού είναι ο φοιτητής του τμήματος βιβλιοθηκονομίας να αποκομίσει μια σφαιρική γνώση του προτύπου κωδικοποίησης συγκεκριμένων δεδομένων (ψηφιακού υλικού ) που χρησιμοποιούνται ώστε να γίνουν απολύτως κατανοητές η σχετική δόμηση της τυπικής βιβλιογραφικής εγγραφής, η σχετική παραγωγή διατάξεων του πρότυπου MARC ή το διεθνές σύστημα ανταλλαγής (διαφορετικών διατάξεων MARC) UNIMARC καθώς και σχήματα κωδικοποίησης δεδομένων, που ενδιαφέρουν το έργο της πληροφόρησης, όπως είναι, για παράδειγμα οι σχετικές διατάξεις από το MARC21 για τα περιεχόμενα της συλλογής (holdings), του καθιερωμένου τύπους (authorities), την ταξινόμηση (classification), τις πληροφορίες για την κοινότητα (community information) κλπ.

*Η δομή του μαθήματος χωρίζεται στα ακόλουθα κεφάλαια:*

- 1. Η Αναπαράσταση των δεδομένων στον υπολογιστή**
- 2. ISO**
- 3. Εισαγωγή στις γλώσσες σήμανσης HTML, SGML, XML/TEI**
- 4. Κωδικοποίηση ήχου.**
- 5. Ψηφιακή Εικόνα**
- 6. Τελικός προορισμός της εικόνας**
- 7. Σύλληψη ψηφιακής εικόνας**
- 8. Αποθήκευση ψηφιακής εικόνας**
- 9. Συμπίεση Δεδομένων (ήχου, εικόνας, βίντεο)**

# 1. Η Αναπαράσταση των δεδομένων στον υπολογιστή

**Τα δεδομένα (data), είναι γεγονότα, μηνύματα, κωδικοποιημένα ή όχι που αποτελούν ακατέργαστο υλικό.** Όταν σχεδιάσθηκαν οι πρώτοι (ψηφιακοί) Η.Υ., έπρεπε να βρεθεί ένα κοινό σημείο ανάμεσα στο τι μπορεί να κάνει και πώς στ' αλήθεια δουλεύει ένας υπολογιστής από τη μια πλευρά (η εισαγωγή των δεδομένων), και στην ουσία της αριθμητικής από την άλλη, έτσι ώστε να μπορεί ο Η.Υ. να εκτελεί αριθμητικούς υπολογισμούς. Το κοινό αυτό σημείο βρέθηκε στη δυαδική αριθμητική. Πραγματικά τα ηλεκτρικά ή ηλεκτρονικά στοιχεία (λυχνίες / διακόπτες) από τα οποία αποτελείται ένας Η.Υ. είναι από τη φύση τους δυαδικά, καθώς μπορούν να βρίσκονται κάθε φορά σε μία από δύο μόνο καταστάσεις (περνάει ή δεν περνάει ρεύμα). Έτσι ένας διακόπτης προσομοιώνει ένα δυαδικό ψηφίο, που μπορεί να είναι μόνο 0 (ανοικτός διακόπτης) ή 1 (κλειστός διακόπτης). Με αυτό τον τρόπο όλες οι πληροφορίες που περνάνε από τον Η.Υ. καταχωρούνται σε δυαδική μορφή και υποβάλλονται σε επεξεργασία πάλι σε δυαδική μορφή. Η μικρότερη ποσότητα πληροφορίας που υπάρχει και που μπορεί να διαχειριστεί ο Η.Υ. είναι λοιπόν το δυαδικό ψηφίο, που στη γλώσσα των Η.Υ. λέγεται bit από το binary digit. Όπως γνωρίζουμε, για την παράσταση ενός χαρακτήρα στον υπολογιστή απαιτούνται, συνήθως, 8 bits τα οποία αποτελούν και τη στοιχειώδη μονάδα αποθήκευσης, το **χαρακτήρα (byte)**. Η χωρητικότητα των υπολογιστών προσδιορίζεται από τον αριθμό των bytes που μπορούν να αποθηκεύσουν. Ο αριθμός αυτός εκφράζεται σε πολλαπλάσια του byte τα **kilobytes, megabytes, gigabytes, ή terabytes**.

Η **λέξη (word)** είναι ένας από τους συντελεστές υπολογιστικής ισχύος που εξαρτάται από τον τύπο του υπολογιστή. Γενικά όσο μεγαλύτερη είναι η λέξη τόσο ισχυρότερος είναι ο υπολογιστής. Οι παλαιοί υπολογιστές είχαν λέξη των 8 bits οι προσωπικοί, των 16 bits οι μεσαίοι και των 32 bits οι μεγάλοι υπολογιστές. Ένας τυπικός pentium επεξεργαστής έχει λέξη των 64 bits και δυνατότητα να χειρίζεται 8 χαρακτήρες συγχρόνως.

1 kilobyte (KB)=2<sup>10</sup> bites= 1024 bytes

1 megabyte (MB) = 2<sup>10</sup> KB

$$1 \text{ gigabyte (GB)} = 2^{10} \text{ MB}$$

$$1 \text{ terabyte (TB)} = 2^{10} \text{ GB}$$

ένα byte ο υπολογιστής μπορεί να το χρησιμοποιήσει είτε σαν αριθμό είτε σα χαρακτήρα, ανάλογα με το τι οδηγίες έχει για το πως θα χρησιμοποιήσει αυτό το byte.

**Κώδικας χαρακτήρα είναι ένα σύνολο από δυαδικά ψηφία, συνήθως 6, 7, ή 8, τα οποία απαιτούνται για την κωδικοποίηση κάθε χαρακτήρα ξεχωριστά. Το σύνολο των χαρακτήρων αυτών (γράμματα, δεκαδικά ψηφία, ειδικά σύμβολα), αποτελεί το σύνολο χαρακτήρων του υπολογιστή.**

### ***Συστήματα Αναπαράστασης Χαρακτήρων (Κωδικοποίηση)***

Οι τεχνικές που χρησιμοποιούνται περισσότερο για την αναπαράσταση των δεδομένων στον υπολογιστή είναι:

- Κωδικοποίηση χαρακτήρων (ASCII, EBCDIC, UNICODE).
- Κωδικοποίηση αριθμητικών δεδομένων (BCD, συμπλήρωμα του 2, αριθμοί κινητής υποδιαστολής).

### ***Κώδικας ASCII (American Standard Code for Information Interchange)***

Σχεδιάστηκε αρχικά ως κώδικας των 8 bits όπου τα 7 χρησιμοποιήθηκαν για την κωδικοποίηση, ενώ το όγδοο χρησιμοποιήθηκε ως bit ελέγχου της ορθότητας των 7. Αυτό δίνει δυνατότητα 128 διαφορετικών συνδυασμών και επομένως 128 διαφορετικών χαρακτήρων. Με την εξέλιξη των μικροϋπολογιστών έγινε δυνατό να χρησιμοποιηθεί και το όγδοο bit για την κωδικοποίηση με αποτέλεσμα να έχουμε δυνατότητα για άλλους 128 χαρακτήρες. Οι επιπλέον 128 συνδυασμοί χρησιμοποιήθηκαν για άλλα σύμβολα. οποίοι χρησιμοποιήθηκαν για άλλα σύμβολα. Στη χώρα μας, οι συνδυασμοί αυτοί αξιοποιήθηκαν για την εισαγωγή των ελληνικών χαρακτήρων και λοιπών συμβόλων της γλώσσας.

Στο αρχικό σύνολο των 256 χαρακτήρων, οι χαρακτήρες, ανάλογα με τη σειρά τους, χρησιμοποιούνται για: 0-31 χαρακτήρες ελέγχου, 32-63 αριθμούς, κενά, σημεία στίξης, σύμβολα πράξεων, 64-95 κεφαλαία λατινικά γράμματα και ειδικά σύμβολα, 96-127 πεζά λατινικά γράμματα και ειδικά σύμβολα.

## **Κώδικας Χαρακτήρων UNICODE**

Το **Unicode Standard** είναι ένα σύγχρονο παγκόσμιο πρότυπο που χρησιμοποιείται για την κωδικοποίηση κειμένου στον υπολογιστή. Αποτελεί τμήμα του International Standard ISO (1993). Ο σχεδιασμός του βασίζεται στην απλότητα και τη συνοχή της κωδικοποίησης ASCII αλλά ξεπερνάει την περιορισμένη ικανότητά της να κωδικοποιεί μόνο το λατινικό αλφάβητο. Το Unicode Standard παρέχει τη δυνατότητα κωδικοποίησης όλων των χαρακτήρων που χρησιμοποιούνται από ένα μεγάλο αριθμό γλωσσών του κόσμου.

Για να κωδικοποιήσει τις χιλιάδες διαφορετικών χαρακτήρες που χρησιμοποιούνται στα αλφάβητα των διαφόρων γλωσσών, το Unicode Standard χρησιμοποιεί ένα κώδικα από 16 bits, ο οποίος παρέχει δυνατότητα διαφορετικών συνδυασμών για περισσότερους από 65.000 χαρακτήρες (για την ακρίβεια  $2^{16} = 65536$ ). Στο Unicode Standard, για να διατηρηθεί απλή και εφικτή η κωδικοποίηση των χαρακτήρων, κάθε χαρακτήρας κωδικοποιείται με 16 bits και δεν χρησιμοποιούνται άλλοι περίπλοκοι τρόποι ή συνδυασμοί κωδικοποίησης.

Το Unicode Standard παρέχει μια ενιαία μορφή κωδικοποίησης για το σύνολο των χαρακτήρων σε παγκόσμιο επίπεδο. Με τον τρόπο αυτό διευκολύνει τις συναλλαγές και την ανταλλαγή αρχείων κειμένου ανάμεσα στις χώρες με διαφορετικές γλώσσες. Οι χρήστες υπολογιστών που ασχολούνται με πολυγλωσσικό κείμενο – άνθρωποι των επιχειρήσεων σε όλο τον κόσμο, γλωσσολόγοι, ερευνητές, επιστήμονες και πολλοί άλλοι – θα ανακαλύψουν ότι το Unicode Standard απλοποιεί τη δουλειά τους σε μεγάλο βαθμό. Οι μαθηματικοί και οι τεχνικοί, που χρησιμοποιούν συχνά μαθηματικά σύμβολα και άλλους τεχνικούς χαρακτήρες, θα βρουν επίσης το Unicode Standard ιδιαίτερα πολύτιμο.

Το Unicode Standard ορίζει κωδικοποίηση όλων των χαρακτήρων που χρησιμοποιούνται στις πιο σημαντικές σημερινές γραπτές γλώσσες. Περιλαμβάνει τους Λατινικούς, τους Ελληνικούς, τους Κυριλλικούς, τους Αρμενικούς, τους Εβραϊκούς, τους Αραβικούς αλλά και πολλούς άλλους χαρακτήρες από λιγότερο διαδεδομένες αλλά "ζωντανές" γλώσσες. Επίσης καλύπτει και το ενοποιημένο σύνολο των Κινέζικων, Ιαπωνικών και Κορεατικών (CJK set) ιδεογραμμάτων. Συμπεριλαμβάνει επίσης τα σημεία στίξης, διακριτικά, μαθηματικά και τεχνικά σύμβολα, βέλη, τυπογραφικά σημεία κλπ.

Συνολικά, το Unicode Standard στη 2<sup>η</sup> έκδοσή του παρέχει σχεδόν **39.000**

κωδικοποιημένους χαρακτήρες από τα παγκόσμια αλφάβητα, τα σύνολα ιδεογραμμάτων και τις συλλογές ειδικών συμβόλων. Επίσης, περίπου 6.000 κωδικοί είναι δεσμευμένοι για ιδιωτική χρήση από τους δημιουργούς υλικού και λογισμικού. Ο τελευταίος τους χρησιμοποιούν για εσωτερικές διεργασίες των προϊόντων τους. Παραμένουν συνεπώς περίπου 20.000 αχρησιμοποίητοι κωδικοί για μελλοντική χρήση.

### ***Κώδικας Χαρακτήρων EBCDIC***

Το όνομα του κώδικα EBCDIC προέρχεται από τα αρχικά των αγγλικών λέξεων Extended Binary Coded Decimal Interchange Code. Είναι ένας από τους πιο διαδεδομένους κώδικες χαρακτήρων διεθνώς. Για την κωδικοποίηση κάθε χαρακτήρα χρησιμοποιούνται 8 bits και αυτό δίνει δυνατότητα 256 διαφορετικών συνδυασμών και επομένως 256 διαφορετικών χαρακτήρων.

Το σύστημα αυτής της κωδικοποίησης χρησιμοποιείται κυρίως σε μεγάλους υπολογιστές. Περιλαμβάνει αλφαβητικούς και αριθμητικούς χαρακτήρες, ειδικά σύμβολα καθώς και χαρακτήρες ελέγχου οι οποίοι δεν τυπώνονται. Κάθε χαρακτήρας έχει μια σειρά από 0 έως 255 μέσα στο σύνολο EBCDIC, όπου οι χαρακτήρες με σειρά 0-63 και 250-255 είναι μη εκτυπώσιμοι χαρακτήρες. Στο σχήμα 2-11 εμφανίζονται ορισμένοι από τους χαρακτήρες αυτούς, όπως αριθμητικοί, αλφαβητικοί, κεφαλαία-πεζά και μερικοί από τους ειδικούς χαρακτήρες με τους αντίστοιχους κώδικες σε δεκαδική και δεκαεξαδική μορφή καθώς και τη γραφική παράσταση.

### ***Αναπαράσταση Αριθμών***

#### ***Δυαδικός Κώδικας Δεκαδικών Ψηφίων ( Binary Coded Decimal)***

Ο Δυαδικός Κώδικας Δεκαδικών Ψηφίων (Binary Coded Decimal-BCD) χρησιμοποιείται για την αναπαράσταση αριθμών του δεκαδικού συστήματος αρίθμησης στον υπολογιστή. Κάθε ψηφίο του αριθμού μετατρέπεται σε αντίστοιχο αριθμό του δυαδικού συστήματος, π.χ. ο δεκαδικός αριθμός 7698 αντιστοιχεί στο δυαδικό 0111 0110 1001 1000. Είναι φανερό ότι για την κωδικοποίηση αυτή απαιτούνται 4 δυαδικά ψηφία για καθένα δεκαδικό ψηφίο αντίστοιχα.



## Παράσταση ακεραίων

Αντίθετα απ' ό τι γίνεται όταν γράφουμε στο χαρτί, όταν πρόκειται για αποθηκεύσουμε έναν αριθμό στον Η.Υ.(δηλ. στη μνήμη, στο δίσκο ή αλλού), είναι κάθε φορά προσδιορισμένο πόσος χώρος (σε byte) θα δεσμευτεί για τον αριθμό αυτό. Δηλαδή οι πληροφορίες καταχωρούνται κάθε φορά σε φόρμες συγκεκριμένου μεγέθους. Όταν το μέγεθος μιας φόρμας δεν επαρκεί για την καταχώρηση ενός, αριθμού, τότε έχουμε υπερχείλιση (overflow) που σημαίνει απώλεια ακρίβειας και πιθανόν διακοπή της εκτέλεσης του προγράμματος. Αν λοιπόν χρησιμοποιήσουμε για την αποθήκευση ενός ακεραίου 1 byte (=8 bit), μπορούμε να παραστήσουμε 2<sup>8</sup>=256 διαφορετικούς αριθμούς τους 0 + 255. Π.χ. ο αριθμός 01000111 αντιστοιχεί στο δεκαδικό 71. Αν χρησιμοποιήσουμε 2 byte (=16 bit) μπορούμε να παραστήσουμε 2<sup>16</sup>=65.536 διαφορετικούς αριθμούς τους 0 + 65.535. Στις συνήθεις όμως εφαρμογές συχνά χρειάζεται να δουλέψουμε με ακόμα μεγαλύτερους ακεραίους. Γι' αυτό χρησιμοποιείται επίσης και η φόρμα των 4 byte (=32 bit). Με αυτή μπορούμε να καλύψουμε το φάσμα από 0 μέχρι 4.294.967.296, το οποίο είναι αρκετά ικανοποιητικό, τουλάχιστον για ακεραίους.

## Αριθμοί με πρόσημο (Αρνητικοί)

Αν Θέλουμε, χρησιμοποιώντας 2 ή 4 byte, να μπορούμε να παραστήσουμε και αρνητικούς αριθμούς, θα πρέπει να μοιράσουμε το πλήθος των δυαδικών αριθμών που μας παρέχει κάθε φόρμα σε δύο ομάδες ίσου πλήθους, και να χρησιμοποιήσουμε τη μία από αυτές για να παραστήσουμε αρνητικούς αριθμούς. Τότε οι δυνατότητες κάθε φόρμας γίνονται:

1 byte	-128 + + 127
2 byte	-32.768 + +32.767 (INTEGER)
4 byte	-2.147.483.648 +2.147.483.647 (LONG)

Για παράδειγμα η γλώσσα προγραμματισμού Quick BASIC χρησιμοποιεί για την παράσταση των ακεραίων την φόρμα των 2 byte που την ονομάζει INTEGER και την φόρμα των 4 byte που την ονομάζει LONG. Η γλώσσα C τη φόρμα των 2 byte την ονομάζει short ή int και τη φόρμα των 4 byte την ονομάζει long. Είναι ενδιαφέρον, ότι η γλώσσα C διακρίνει κάθε μία από αυτές τις δύο φόρμες σε signed (με πρόσημο) και unsigned (χωρίς πρόσημο). Αυτό σημαίνει ότι π.χ. η φόρμα short unsigned της C

καταλαμβάνει 2 byte και θεωρεί τους αριθμούς που αποθηκεύονται με αυτή τη φόρμα σαν αριθμούς χωρίς πρόσημο (δηλ. με φάσμα τιμών 0 + 65.565). Αντίθετα η φόρμα short signed δέχεται αριθμούς με πρόσημο (δηλ. με φάσμα τιμών -32.768 + +32.767). Οι γλώσσες προγραμματισμού ονομάζουν τις διάφορες φόρμες με τις οποίες αποθηκεύονται οι αριθμοί και γενικότερα τα δεδομένα, τύπους δεδομένων (data types).

## **Αριθμοί Συμπλήρωμα του 2**

Ο κώδικας αυτός είναι ο πιο συνηθισμένος για την αναπαράσταση θετικών και αρνητικών αριθμών στον υπολογιστή. Το σημείο του αριθμού δηλώνει το πιο σημαντικό ψηφίο: εάν είναι 1, ο αριθμός είναι αρνητικός ενώ εάν είναι 0, ο αριθμός είναι θετικός.

Συγχρόνως όμως το ψηφίο αυτό διατηρεί και την αξία του ανάλογα με τη θέση του στον αριθμό. Θα εξετάσουμε το σύνολο των αριθμών που μπορούμε να παραστήσουμε με 8 δυαδικά ψηφία. Ανάλογα εργαζόμαστε και για 16 ή και περισσότερα. Αν είχαμε μόνο θετικούς αριθμούς, το εύρος θα ήταν από 0 έως  $2^8-1$  (0000 0000...1111 1111).

Εφόσον οι αριθμοί έχουν πρόσημο, το όγδοο bit παριστάνει το πρόσημο αυτό και συγχρόνως έχει τιμή:

$-1 \times 2^7 = -128$  αν το όγδοο bit είναι 1  $0 \times 2^7 = 0$  αν το όγδοο bit είναι 0. Παρατηρούμε ότι:  $1111\ 1111 = -128 = -2^7$  είναι ο μεγαλύτερος (κατ' απόλυτη τιμή)

αρνητικός,

$0111\ 1111 = +127 = 2^7-1$  είναι ο μεγαλύτερος θετικός αριθμός.

Έτσι, το εύρος δυαδικών με 8 bits είναι:

Για θετικούς (0 έως  $2^8-1$ ).

Για αρνητικούς και θετικούς αριθμούς ( $-128 = -2^7 \dots +127 = 2^7-1$ ).

Με παρόμοιους υπολογισμούς βρίσκουμε ότι το εύρος δυαδικών με 16 bits είναι:

Για θετικούς (0 έως  $2^{16}-1$ ).

Για αρνητικούς και θετικούς ( $-32768 = -2^{15} \dots +32767 = 2^{15}-1$ ).

Η κωδικοποίηση συμπλήρωμα του 2 χρησιμοποιείται για την εύκολη μετατροπή ενός θετικού

αριθμού στον αντίστοιχο αρνητικό και αντίστροφα. Έτσι, η αφαίρεση μετατρέπεται εύκολα σε

πρόσθεση, προσθέτοντας τον αντίθετο αριθμό, π.χ.  $10-5=10+(-5)$ .

Σύμφωνα με αυτή, για να βρούμε τον αντίθετο ενός αριθμού (Θετικού ή αρνητικού):

- α. Αντιστρέφουμε τα 0 του αριθμού σε 1 και τα 1 σε 0.
- β. Προσθέτουμε στον αριθμό που προέκυψε το 1.

Π.χ.

0110 (6)	10110000 (-80)	0001 (1)
1001 +	01001111 +	1110 +
1	1	1
1010 (-6)	01010000 (80)	1111 (-1)

Βλέπουμε ότι ο ίδιος δυαδικός αριθμός π.χ, ο 1010, αν τον δούμε σαν αριθμό χωρίς πρόσημο παριστάνει τον 10, ενώ αν τον δούμε στον αριθμό με πρόσημο παριστάνει τον -6. Η αντιστροφή των bit ενός δυαδικού αριθμού είναι μια από τις βασικές πράξεις που ο υπολογιστής μπορεί να εκτελέσει πολύ εύκολα και γρήγορα.

### **Η Αφαίρεση**

Η παράταση των ακεραίων αριθμών με πρόσημο με τη μέθοδο του συμπληρώματος ως προς δύο, έχει σαν άμεση συνέπεια ότι ο Η.Υ. αντί να κάνει αφαίρεση ενός αριθμού από έναν άλλο, κάνει πρόσθεση του ενός με τον αντίθετο (δηλ. το συμπλήρωμα) του άλλου. Π.χ.  $85 + (-32)$ . Δηλαδή ανάγει την αφαίρεση σε πρόσθεση. Αν μάλιστα σκεφθούμε ότι, έτσι κι' αλλιώς, ο πολλαπλασιασμός ανάγεται σε προσθέσεις και η διαίρεση σε αφαιρέσεις, προκύπτει ότι ο Η.Υ. ανάγει τις τέσσερις πράξεις της αριθμητικής στις βασικές γι' αυτόν πράξεις της πρόσθεσης και της αντιστροφής των ψηφίων ενός δυαδικού αριθμού. Η παράσταση των αρνητικών αριθμών με τη μέθοδο του συμπληρώματος ως προς δύο, έχει σαν έμμεση συνέπεια ότι αν μετά την πρόσθεση δύο δυαδικών αριθμών προκύψει αριθμός με πλήθος ψηφίων που υπερβαίνει την χρησιμοποιούμενη φόρμα (δηλ. με μήκος 9, 17 ή 33 ψηφία αντίστοιχα), το επιπλέον ψηφίο αγνοείται.

Π.χ.

01010101 (85) + 00100000 (32)	00100000 (32) + 11011111 1 11100000 (-32)	01010101 (85) + 11100000 (-32) 1)00110101 00110101 (53)
----------------------------------	--	--

### **Παράσταση αριθμών με κλασματικό μέρος (Μεθοδος κινητής υποδιαστολής)**

Για να παραστήσουμε στον Η.Υ. αριθμούς με κλασματικό μέρος, χρησιμοποιούμε τη μέθοδο που είναι γνωστή σαν μέθοδος κινητής υποδιαστολής (floating point), Σύμφωνα με αυτή μπορούμε να γράψουμε έναν αριθμό π.χ. τον 2562,867 σαν  $0,2562867 \times 10^4$ . Με τον ίδιο τρόπο ο αριθμός 0,0002562867 γράφεται σαν  $0,2562867 \times 10^{-3}$ . Ο συμβολισμός που χρησιμοποιείται συνήθως είναι  $0,2562867E+4$  για τον πρώτο αριθμό και  $0,2562867E-3$  για το δεύτερο αριθμό. Με αυτό τον τρόπο ο αριθμός χωρίζεται σε δύο μέρη: το πρώτο μέρος περιλαμβάνει τα σημαντικά του ψηφία, ενώ το δεύτερο (η δύναμη) δείχνει ποια είναι η θέση της υποδιαστολής σε σχέση με τα σημαντικά ψηφία του αριθμού. Στην πράξη, για την παράσταση των πραγματικών αριθμών στο εσωτερικό του υπολογιστή αποθηκεύονται ξεχωριστά τρία πράγματα: το πρόσημο του αριθμού, τα σημαντικά ψηφία του και από τη δύναμη μόνο ο εκθέτης, ο οποίος υπολογίζεται σαν εκθέτης του 2 και όχι του

10.

Οι αριθμοί με κινητή υποδιαστολή στηρίζονται στις ίδιες αρχές με βάση το 2. Στη μορφή αυτή ο αριθμός εκφράζεται ως γινόμενο ενός αριθμού μεταξύ του  $1/2$  και του 1 και

της κατάλληλης δύναμης του 2

Π.χ.

$$\begin{array}{ll} & 5 & & 6 \\ 11000 = 0.11 \times 10 & & 111000 = 0.111 \times 10 & \\ & -2 & & -5 \\ 0.0011 = 0.11 \times 10 & & 0.00000111 = 0.111 \times 10 & \end{array}$$

Η δύναμη του 2 προσδιορίζεται από τον αριθμό των θέσεων που πρέπει να μετακινηθεί η υποδιαστολή από την ορισμένη στην κινητή θέση. Το κλασματικό

μέρος του αριθμού λέγεται συντελεστής (mantissa) και η δύναμη του 2 εκθέτης (exponent). Η αναπαράσταση των αριθμών, με κινητή υποδιαστολή, εξαρτάται από τον τύπο του υπολογιστή καθώς και από την επιθυμητή ακρίβεια.

Στην γλώσσα Quick BASIC χρησιμοποιούνται 2 φόρμες για αριθμούς με κλασματικό μέρος:

⑩ Η πρώτη ονομάζεται SINGLE και χρησιμοποιεί 4 byte. Από αυτά ένα bit χρησιμοποιείται για το πρόσημο του αριθμού, 23 για τα σημαντικά ψηφία του αριθμού και 8 για τον εκθέτη του 2. Καλύπτει το φάσμα από  $-3,402823 \times 10^{+38}$  μέχρι  $-1,40129 \times 10^{-45}$  για τους αρνητικούς και από  $1,40129 \times 10^{-45}$  μέχρι  $3,402823 \times 10^{+38}$  για τους θετικούς αριθμούς. Οι αριθμοί SINGLE είναι ακριβείς μέχρι και το έβδομο σημαντικό ψηφίο.

⑩ Η δεύτερη ονομάζεται DOUBLE και χρησιμοποιεί 8 byte (=64 bit). Από αυτά ένα bit χρησιμοποιείται για το πρόσημο του αριθμού, 52 για τα σημαντικά ψηφία του αριθμού και 11 για τον εκθέτη. Ο εκθέτης των αριθμών DOUBLE είναι από -324 μέχρι +308 και η ακρίβεια τους φθάνει μέχρι και το δέκατο πέμπτο σημαντικό ψηφίο.

Η γλώσσα C χρησιμοποιεί μια φόρμα των 4 byte που την ονομάζει float και μια φόρμα των 8 byte που την ονομάζει double. Το φάσμα που καλύπτουν αυτές οι φόρμες και η ακρίβεια τους είναι παρόμοια με αυτά της Quick BASIC.

#### Σημείωση:

*Οι υπολογιστές με επεξεργαστές τους Intel 8088, 8086 ή 80286 έχουν από την κατασκευή τους τη δυνατότητα να διαχειριστούν αριθμούς με μήκος μέχρι 16 bit (δηλ. 2 byte). Για την εκτέλεση πράξεων με αριθμούς σε μεγαλύτερες φόρμες οι υπολογιστές αυτοί χρησιμοποιούν είτε τεχνάσματα του λογισμικού, όπου η επεξεργασία γίνεται τμηματικά (δύο-δύο ή ένα-ένα byte), είτε ένα ειδικό τσιπ που εγκαθίσταται προαιρετικά και αναλαμβάνει αυτή τη δουλειά (μαθηματικός συνεπεξεργαστής). Στη δεύτερη περίπτωση η επεξεργασία των αριθμών 32 ή 64 bit είναι Θεαματικά ταχύτερη.*

### **Αλφαριθμητικά (Strings)**

Μια οποιαδήποτε ομάδα από χαρακτήρες στη σειρά (μία λέξη, μια πρόταση, μια γραμμή κειμένου ή μια διαδοχή συμβόλων ή γραμμάτων) την ονομάζουμε αλφαριθμητικό (string).

Όσον αφορά το ηώς αποθηκεύονται στον Η.Υ. τα αλφαριθμητικά, παρατηρούμε ότι δεν βολεύει η χρησιμοποίηση ενός αριθμού από φόρμες με καθορισμένο μέγεθος του bit.

Αυτό διότι κάθε χαρακτήρας που προστίεται ή διαγράφεται από το αλφαριθμητικό το αυξομειώνει κατά 1 byte. Αν λοιπόν κάναμε ότι και για τους αριθμούς, θα έπρεπε να είχαμε πάρα πολλές φόρμες, με αμφίβολα πάλι αποτελέσματα. Αν εξάλλου χρησιμοποιούμε μικρό αριθμό από φόρμες, τότε οδηγούμαστε σε σπατάλη χώρου στη μνήμη ή στο δίσκο. Η μέθοδος που χρησιμοποιείται συνήθως είναι, εκτός από τους αριθμούς ASCII των χαρακτήρων που αποτελούν το αλφαριθμητικό, να αποθηκεύεται πριν από αυτούς ένας αριθμό που δείχνει από πόσους χαρακτήρες (byte) αποτελείται το αλφαριθμητικό. Ο αριθμός αυτός λέγεται περιγραφέας και δεσμεύει ένα καθορισμένο αριθμό από byte.

Η γλώσσα προγραμματισμού Quick BASIC χρησιμοποιεί για την καταχώρηση του περιγραφέα 4 byte. Π.χ. για την αποθήκευση του αλφαριθμητικού \*ΑΛΦΑ\* χρειάζονται 16 byte. Δηλαδή ισχύει ο τύπος 4 byte + μήκος του αλφαριθμητικού (σε byte). Άμεση συνέπεια αυτού του τρόπου αποθήκευσης είναι ότι τα αλφαριθμητικά δεν μπορούν να έχουν μήκος μεγαλύτερο από τον αριθμό που μπορεί να χωρέσει ο περιγραφέας. Στην quick BASIC ένα αλφαριθμητικό μπορεί να έχει μήκος μέχρι 32.767 χαρακτήρες.

## 2. ISO

Ο Διεθνής Οργανισμός Τυποποίησης είναι μια μη κυβερνητική οργάνωση που ιδρύθηκε το 1947. Η αποστολή του ISO είναι να προωθηθεί η ανάπτυξη της τυποποίησης και των σχετικών δραστηριοτήτων στον κόσμο με σκοπό τη διευκόλυνση της διεθνούς ανταλλαγής των αγαθών και των υπηρεσιών, και την ανάπτυξη της συνεργασίας στις σφαίρες της πνευματικής, επιστημονικής, τεχνολογικής και οικονομικής δραστηριότητας. Στην κωδικοποίηση των διαφόρων γλωσσικών και γραφικών χαρακτήρων προτάθηκαν πολλοί 8-bit κώδικες. Η αντιμετώπιση διεθνώς της χαοτικής εισαγωγής αυτών των κωδικοποιημένων χαρακτηριστικών μορφών επιτεύχθηκε με την θέσπιση διαφόρων ISO οι σημαντικότερες είναι:

- ISO 646-1983, 7-bit character set

- ISO 6937-2, 8-bit κωδικοποιημένοι χαρακτήρες (λατινικοί, χαρακτήρες γραφικών)
- ISO 8859, Οικογένεια 8-bit γραφικών χαρακτήρων

Τέλος ένα σύστημα για να μπορέσει να αντιμετωπίσει τις απαιτήσεις των πολλών και ποικίλων χαρακτήρων και γραφικών αναγκών αντιμετωπίζεται μέσω προτύπου Unicode και ISO 10646

## 3. ΓΛΩΣΣΕΣ ΣΗΜΑΝΣΗΣ

### 3.1 HTML

Η γλώσσα HTML σχεδιάστηκε ώστε να είναι πολύ εύκολη στη χρήση. Δεν χρειάζεται να είσαι προγραμματιστής. Αν μπορείς απλά να γράψεις και να διορθώσεις ένα αρχείο κειμένου, τότε μπορείς να γράψεις σε HTML. Αν και τα αρχικά HTML σημαίνουν Hyper Text Markup Language, δεν πρόκειται στην πραγματικότητα για γλώσσα προγραμματισμού αλλά για κάτι πολύ πιο απλό. Είναι ένας τρόπος για να περιγράψουμε πως ένα σύνολο από κείμενο και εικόνες πρέπει να απεικονισθούν στο θεατή. Με έναν παρόμοιο τρόπο, ένας συντάκτης εφημερίδας σημειώνει οδηγίες για το πως θα σελιδοποιηθεί η εφημερίδα του (που θα μπουν τα κείμενα και που οι εικόνες).

#### Ανατομία μιας ιστοσελίδας

Μια συνηθισμένη ιστοσελίδα αποτελείται από ένα αρχείο HTML μαζί με κάποια αρχεία εικόνων που ίσως υπάρχουν σε αυτήν. Το αρχείο HTML περιέχει το κείμενο που πρέπει να εμφανιστεί και λειτουργεί επίσης ως "κόλλα" για να συγκρατεί το κείμενο και τις εικόνες στις σωστές τους θέσεις και για να τα εμφανίζει με το σωστό τρόπο. Η συγγραφή ενός HTML αρχείου περιλαμβάνει τη σύνθεση του κειμένου που θέλεις να εμφανιστεί καθώς και την εισαγωγή των απαιτούμενων ετικετών (tags) στις κατάλληλες θέσεις. Οι ετικέτες ξεκινούν με τον χαρακτήρα <, τελειώνουν με τον χαρακτήρα > και λένε στο πρόγραμμα πλοήγησης ιστοσελίδων να κάνει κάτι συγκεκριμένο, όπως να εμφανίσει ένα κείμενο σε πλάγια ή μαύρα, ή με μεγαλύτερη γραμματοσειρά, ή να εμφανίσει μια εικόνα, ή να δημιουργήσει έναν σύνδεσμο προς μια άλλη ιστοσελίδα. Αν και η HTML έχει πολλές ετικέτες, δεν χρειάζεται να τις γνωρίζεις όλες - μπορείς να φτιάχνεις σελίδες με λίγες μόνο από αυτές.

#### Ανατομία μιας ετικέτας HTML

Οι ετικέτες έχουν μια πολύ απλή δομή. Ανάμεσα στους χαρακτήρες < > περιλαμβάνεται το όνομα της ετικέτας (tag name) και όταν χρειάζεται κάποιες ιδιότητες (attributes). Οι περισσότερες ιδιότητες χρειάζονται και τις αντίστοιχες τιμές (values). Η γενική μορφή μιας ετικέτας, λοιπόν, είναι:



<tagname attribute1="value1" attribute2="value2"...>

Τα ονόματα των ετικετών και των ιδιοτήτων μπορούν να γράφονται είτε με κεφαλαία είτε με μικρά γράμματα, αλλά για κάποιες τιμές των ιδιοτήτων έχει σημασία το πως είναι γραμμένες.

Για παράδειγμα με την ετικέτα  ορίζεις την εισαγωγή του αρχείου εικόνας logo1.gif μέσω της ιδιότητας src.

### **Ετικέτες περιεχομένου (Container tags)**

Κάποιες ετικέτες όπως η <img> στέκονται μόνες τους. Στην πραγματικότητα δεν επηρεάζουν τίποτα τριγύρω τους. Άλλες όμως περιλαμβάνουν μια αρχική ετικέτα και μια τελική ετικέτα, επηρεάζοντας ο,τιδήποτε βρίσκεται ανάμεσά τους, ακόμη και άλλες ετικέτες. Αυτές οι ετικέτες λέγονται ετικέτες περιεχομένου (container tags). Για παράδειγμα, για να κάνεις κάποιο κείμενο μαύρο πρέπει να σημειώσεις που ξεκινάει το μαύρισμα και που τελειώνει. Αυτό γίνεται χρησιμοποιώντας τις <b> και </b> αντίστοιχα: Αυτό είναι απλό κείμενο. <b>Αυτό είναι μαύρο κείμενο.</b> Απλό ξανά.

Το παραπάνω HTML κείμενο θα εμφανιστεί από το πρόγραμμα πλοήγησης ως: Αυτό είναι απλό κείμενο. Αυτό είναι μαύρο κείμενο. Απλό ξανά.

Κάθε ετικέτα περιεχομένου τελειώνει με </όνομα ετικέτας> για οποιοδήποτε όνομα ετικέτας. Αντίθετα με τις ετικέτες έναρξης, οι ετικέτες τέλους δεν έχουν ιδιότητες.

### **Το "μυστικό" του Παγκόσμιου Ιστού!**

Η ετικέτα που δίνει στον παγκόσμιο ιστό τη μορφή που γνωρίζουμε, μια τεράστια συλλογή από διασυνδεδεμένες σελίδες, είναι η ετικέτα <a>. Η ετικέτα αυτή που είναι ετικέτα περιεχομένου ορίζει έναν σύνδεσμο προς μια άλλη σελίδα αρχείο HTML και είναι εύκολη στη χρήση. Ορίστε ένα παράδειγμα:

Επισκεφθείτε την <a href=http://www.manutd.com> επίσημη σελίδα</a> της Manchester United.

Το παραπάνω HTML κείμενο θα εμφανιστεί από το πρόγραμμα πλοήγησης ως: Επισκεφθείτε την επίσημη σελίδα της Manchester United.

Παρατήρησε ότι υπάρχει μια ετικέτα έναρξης (<a href=http://www.manutd.com>) και μια ετικέτα τέλους (</a>) ανάμεσα στις οποίες καθορίζεται ο δεσμός στον οποίο αν κάνεις κλικ θα μεταφερθείς στην ιστοσελίδα που περιγράφει η τιμή της ιδιότητας href.

Περισσότερα για τους συνδέσμους (links)

Μπορείς να βάλεις μια ετικέτα <img> μεταξύ των <a> και </a> ώστε ο χρήστης να μπορεί να κάνει κλικ πάνω στην εικόνα για να ακολουθήσει στο σύνδεσμο.

Για ένα σύνδεσμο σε μήνυμα e-mail η τιμή της href αρκεί να έχει τη μορφή "mailto:διεύθυνση e-mail". Για παράδειγμα:

Επικοινωνήστε με το <a href="mailto:peiramgym@her.forthnet.gr">σχολείο</a> μας.

Το παραπάνω HTML κείμενο θα εμφανιστεί από το πρόγραμμα πλοήγησης ως:

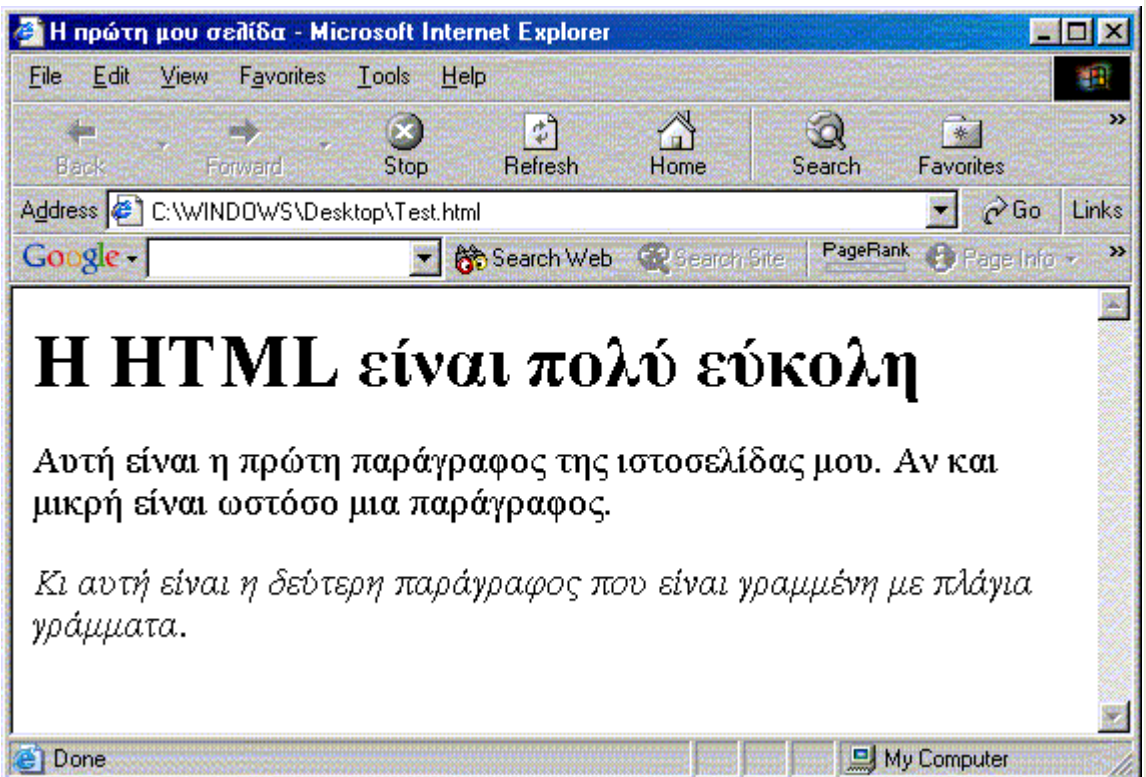
Επικοινωνήστε με το σχολείο μας.Ανατομία ενός αρχείου HTML

Ένα αρχείο HTML θα εμφανιστεί σωστά από ένα πρόγραμμα πλοήγησης εφ' όσον η επέκταση του ονόματός του είναι .htm ή .html και επιπλέον όλο το κείμενό του περιέχεται ανάμεσα στις ετικέτες περιεχομένου <html> και </html>. Ανάμεσα σε αυτές πρέπει να περιέχονται δύο τμήματα, η επικεφαλίδα και το σώμα, καθένα από τα οποία καθορίζεται από τις ετικέτες <head> - </head> και <body> - </body>. Το τμήμα επικεφαλίδας περιέχει πληροφορίες για τη σελίδα που αντλούν οι διάφορες μηχανές αναζήτησης (search engines) του διαδικτύου. Πιο συνηθισμένη είναι ο τίτλος της σελίδας που ορίζεται με τις ετικέτες <title> - </title>. Στο σώμα αναπτύσσεται όλο το περιεχόμενο της ιστοσελίδας.

Επισημαίνοντας ότι τα συνεχόμενα κενά και αλλαγές γραμμής σε ένα HTML αρχείο αγνοούνται (και μετατρέπονται σε ένα κενό) από το πρόγραμμα πλοήγησης, δες ένα απλό αλλά ολοκληρωμένο παράδειγμα που θα μπορούσε να είναι η πρώτη σου ιστοσελίδα:

<b><i>Αρχείο HTML</i></b>	<b><i>Εμφάνιση στο πρόγραμμα πλοήγησης</i></b>
---------------------------	--

```
<html>
<head>
<title>Η πρώτη
μου σελίδα</title>
</head>
<body>
<h1>Η HTML
είναι πολύ
εύκολη</h1>
<p>Αυτή είναι η
πρώτη παράγραφος
της ιστοσελίδας
μου. Αν και μικρή
είναι ωστόσο μια
παράγραφος.</p>
<p><i>Κι αυτή
είναι η δεύτερη
παράγραφος που
είναι γραμμένη με
πλάγια
γράμματα.</i></p>
</body>
</html>
```



## 3.2 SGML

Τα αρχικά SGML σημαίνουν Standard Generalised Mark-up Language. Η SGML είναι μια γλώσσα προσδιορισμού τύπων εγγράφων. Μπορεί να θεωρηθεί η SGML σαν ένα υπερσύνολο, και, για να δοθεί ένα παράδειγμα, η HTML είναι ένα υποσύνολο της SGML. Η SGML επινοήθηκε από την IBM προκειμένου να λυθεί το πρόβλημα της μη τυποποιημένης εμφάνισης κειμένων σε διάφορα υπολογιστικά συστήματα.

Ένα άλλο παράδειγμα: Κάποιος μπορεί να ορίσει τον τύπο ενός εγγράφου ως συνταγή. Το έγγραφο θα περιέχει ένα πρώτο μέρος, όπου θα παρουσιάζονται τα υλικά. Στο δεύτερο μέρος θα γίνεται μια εισαγωγή για τα σκεύη και ότι άλλο χρειαστεί στην όλη διαδικασία. Στο τρίτο μέρος, τελικά, θα περιγράφονται τα βήματα που χρειάζονται για να μαγειρέψουμε το αντικείμενό μας, και, τέλος μια ολοκληρωμένη εικόνα του τί κάναμε, για να δοθεί μια σφαιρική άποψη της πράξης μας.

Όλο αυτό ονομάζεται "Ορισμός του τύπου του εγγράφου" (Document Type Definition). Δεν περιγράφει το πως θα μοιάζει το τελικό έγγραφο, αλλά το τι μπορεί να περιέχει. Με αυτόν τον τρόπο, γράφει κανείς το έγγραφο του χωρίς να λογαριάζει την τελική του εμφάνιση, την εμφάνιση που θα δει ο αναγνώστης.

### Γενικά συμπεράσματα της γλώσσας SGML

- Ενοποίηση της Web τεχνολογίας με την αντικειμενοστραφή τεχνολογία
- Εξυπηρέτηση πολλαπλών ετερογενών βάσεων
- Μοντελοποίηση αντικειμένων
- Υποστήριξη μεγάλου αριθμού εξυπηρετούμενων

### 3.3 XML

#### Τι είναι η XML

Είναι μία μεταγλώσσα μέσα από την οποία μπορούμε να περιγράψουμε το περιεχόμενο και την δομή της πληροφορίας μας. Συσχετίζεται με τις προαναφερόμενες γλώσσες SGML/HTML σύμφωνα με τις παρακάτω προτάσεις:

**Η XML είναι μια μέθοδος για να μπουν δομημένα δεδομένα σε ένα αρχείο κειμένου.** Η XML προδιαγράφει ένα σύνολο κανόνων, οδηγιών και συμβάσεων για την κατασκευή μορφών κειμένου για δομημένα δεδομένα, τα οποία παράγονται και διαβάζονται εύκολα, είναι μη-διφορούμενα, επεκτάσιμα και υποστηρίζουν internationalization/localization.

**Η XML μοιάζει με την HTML αλλά δεν είναι HTML.** Το κοινό τους στοιχείο είναι ότι και οι δύο χρησιμοποιούν tags και attributes. Αλλά η XML είναι επεκτάσιμη, καθώς ορίζουμε τα δικά μας tags και attributes. Η HTML και η XML δεν αποτελούν «αδελφές» γλώσσες. Η XML είναι υποσύνολο της SGML ενώ η HTML εφαρμογή της.

**Τα έγγραφα XML πρέπει να είναι ευανάγνωστα και σαφή.** Εκτός από το software που επεξεργάζεται τα έγγραφα XML, πρέπει να μπορούν να διαβάζονται και από ανθρώπους.

**Η XML είναι μια οικογένεια από τεχνολογίες.** Το βασικό έγγραφο της γλώσσας είναι το specification XML 1.0, το οποίο βγήκε ως προϊόν του World Wide Web Consortium το 1998. Αυτό το specification πλαισιώνεται από έναν αριθμό περιφερειακών τεχνολογιών που ολοκληρώνουν την τεχνολογία της XML. Καταρχήν υπάρχει το **Xlink**, που είναι ένα πρότυπο για την εισαγωγή υπερ-συνδέσμων σε ένα έγγραφο XML. Έπειτα υπάρχει το **Xpointer** και το **Xfragments**, τα οποία περιγράφουν τη σύνταξη για τη δεικτοδότηση σε τμήματα των XML εγγράφων. Εξάλλου υπάρχει το παλιό καλό **CSS** για την δημιουργία style sheets σε έγγραφα XML, όπως ακριβώς δουλεύει και για την περίπτωση των HTML σελίδων. Εκτός από το CSS έχει αναπτυχθεί το **XSL**, το οποίο είναι επίσης μια γλώσσα για την περιγραφή style sheets ειδικά για XML έγγραφα. Ακόμα υπάρχει το **DOM**, ένα σύνολο από κλήσεις συστήματος, ουσιαστικά ένα API, για την επεξεργασία εγγράφων XML από γλώσσες προγραμματισμού (π.χ. Java). Μια αρκετά καινούρια

τεχνολογία είναι το **XML Namespaces** που προδιαγράφει τον τρόπο με τον οποίο μπορεί να συσχετιστεί ένα URL με κάθε tag και attribute ενός XML εγγράφου. Τέλος υπάρχει το **XML Schemas 1 και 2**, με το οποίο μπορεί να ορίσει κανείς επακριβώς το δικό του format βασισμένο στην XML. Υπάρχουν και αρκετά ακόμα άλλα σχετιζόμενα πρότυπα και εργαλεία υπό ανάπτυξη από το W3C.

**Η XML δημιουργεί μεγάλα αρχεία αλλά αυτό δεν αποτελεί πρόβλημα.** Ένα έγγραφο XML είναι στην ουσία κείμενο και φυσικά καταλαμβάνει μεγαλύτερο φυσικό χώρο από ότι ένα αντίστοιχο binary αλλά ο αποθηκευτικός χώρος και η μετάδοση μεγάλων αρχείων πλέον δεν αποτελεί μεγάλο πρόβλημα. Επιπλέον αντισταθμίζεται από τα πλεονεκτήματα των κειμένων έναντι των binaries.

**Η XML είναι καινούρια τεχνολογία, αλλά όχι και τόσο καινούρια.** Η ανάπτυξη της άρχισε το 1996 και το specification XML 1.0 βγήκε το 1998. Ωστόσο η SGML υπάρχει από τις αρχές της δεκαετίας του 80 και ως πρότυπο του ISO από το 1986. Και φυσικά υπάρχει και η HTML από το 1990. Οι κατασκευαστές της XML πήραν ουσιαστικά τα καλύτερα στοιχεία από την SGML και την πολύτιμη εμπειρία από την HTML και παρήγαγαν κάτι εξίσου δυνατό με την SGML, αλλά πολύ απλούστερο και ευκολότερο στη χρήση.

**Η XML είναι συμβατή με την SGML.** Κάθε valid XML document είναι και valid SGML. Φυσικά το αντίθετο δεν ισχύει καθώς η SGML είναι υπερ-σύνολο της XML.

**Η XML έχει δωρεάν άδεια χρήσης, είναι ανεξάρτητη πλατφόρμας και υποστηρίζεται ευρέως.** Το W3C φυσικά κατασκεύασε την XML για να τη διαθέσει στο ευρύ κοινό, δίνοντας λύσεις εκεί όπου η HTML και η SGML αποτυγχάνουν. Ως μορφή κειμένου δεν εξαρτάται από το περιβάλλον εκτέλεσης και μπορεί να μεταφερθεί μέσω δικτύου σε οποιοδήποτε μηχάνημα. Τέλος χαίρει την υποστήριξη μιας μεγάλης κοινότητας αποτελούμενης από ακαδημαϊκούς, βιομηχανικούς κύκλους και διεθνείς οργανισμούς.

## **Πού χρησιμοποιείται.**

Η χρησιμότητα της XML έγκειται στο γεγονός ότι μπορεί κανείς με ένα δομημένο τρόπο να περιγράψει ποία είναι τα δεδομένα του και να τα ιεραρχήσει χωρίς καμία απαίτηση για γνώση του τρόπου υλοποίησης και διαχείρισης των δεδομένων. Συγκεκριμένα, η χρήση της ενδείκνυται για περιπτώσεις όπου:

Υπάρχει ανάγκη διασύνδεσης διαφορετικών συστημάτων. Με την χρήση XML μπορούμε να διασυνδέσουμε συστήματα, δηλαδή να επιτρέψουμε την ανταλλαγή δεδομένων, την εκτέλεση πράξεων χωρίς να είναι απαραίτητο να επέμβουμε και να τροποποιήσουμε τον σχεδιασμό και τον τρόπο λειτουργίας αυτών των συστημάτων. Χαρακτηριστικά, μπορούμε να αντλήσουμε δεδομένα από απαξιωμένα συστήματα που κάνουν χρήση Cobol, να τα μορφοποιήσουμε σε XML και να επιτρέψουμε την επικοινωνία τους με άλλα συστήματα που χρησιμοποιούν σχεσιακές ή και ακόμα αντικειμενοστραφείς σχεδιάσεις (Oracle κ.α.).

Υπάρχει ανάγκη περιγραφής & προτυποποίησης δεδομένων. Με την χρήση XML μπορούμε να περιγράψουμε τα δεδομένα και την δομή τους με ένα απλό και ευέλικτο μηχανισμό. Η χρήση της ενδείκνυται για περιπτώσεις που επιθυμούμε την αποτύπωση της πληροφορίας με ένα μέσο που θα επιτρέπει την γρήγορη ανάλυση της χωρίς να επιβαρύνει το σύστημα. Για την συγκεκριμένη χρήση πρέπει να τονίσουμε ότι η XML δεν ενδείκνυται για χρήσεις που έχουν ιδιαίτερες απαιτήσεις ανάκτησης πληροφορίας<sup>1</sup>.

### **Από τι αποτελείτε ένα XML κείμενο.**

Από ένα Root Node.

Από ένα ή περισσότερα elements.

Ένα element μπορεί να περιέχει:

Sub Elements

Attributes

---

<sup>1</sup> Λέγοντας ανάκτηση, την αντιλαμβανόμαστε με βάση την σχεσιακή άλγεβρα όπου ως ανάκτηση ορίζεται η διαδικασία μορφοποίησης ενός συνόλου που πληροί ορισμένες συνθήκες.

## Παράδειγμα

<?xml version="1.0" ?>

<productDB> ←Root Node

<product prodid="223"> ←Element & Attribute

<prodname>RAM 128</prodname> ←Sub Element

<listprice>45000</listprice>

<dept deptname="RAM">174</dept>

</product>

<product prodid="224">

<prodname>PII333</prodname>

<listprice>20000</listprice>

<dept deptname="PII">179</dept>

</product>

<product prodid="225">

<prodname>PII4333</prodname>

<listprice>26000</listprice>

<dept deptname="PII">179</dept>

</product>

</productDB>



## Σχολιασμός

Το παράδειγμά μας καταγράφει το σύνολο της πληροφορίας που σχετίζεται με προϊόντα υπολογιστών. Συγκεκριμένα μας παρέχει πληροφορίες σχετικά με την κατηγορία που ανήκει το προϊόν, τον κωδικό της κατηγορίας, το όνομα του προϊόντος, τον κωδικό του καθώς και την τιμή του προϊόντος.

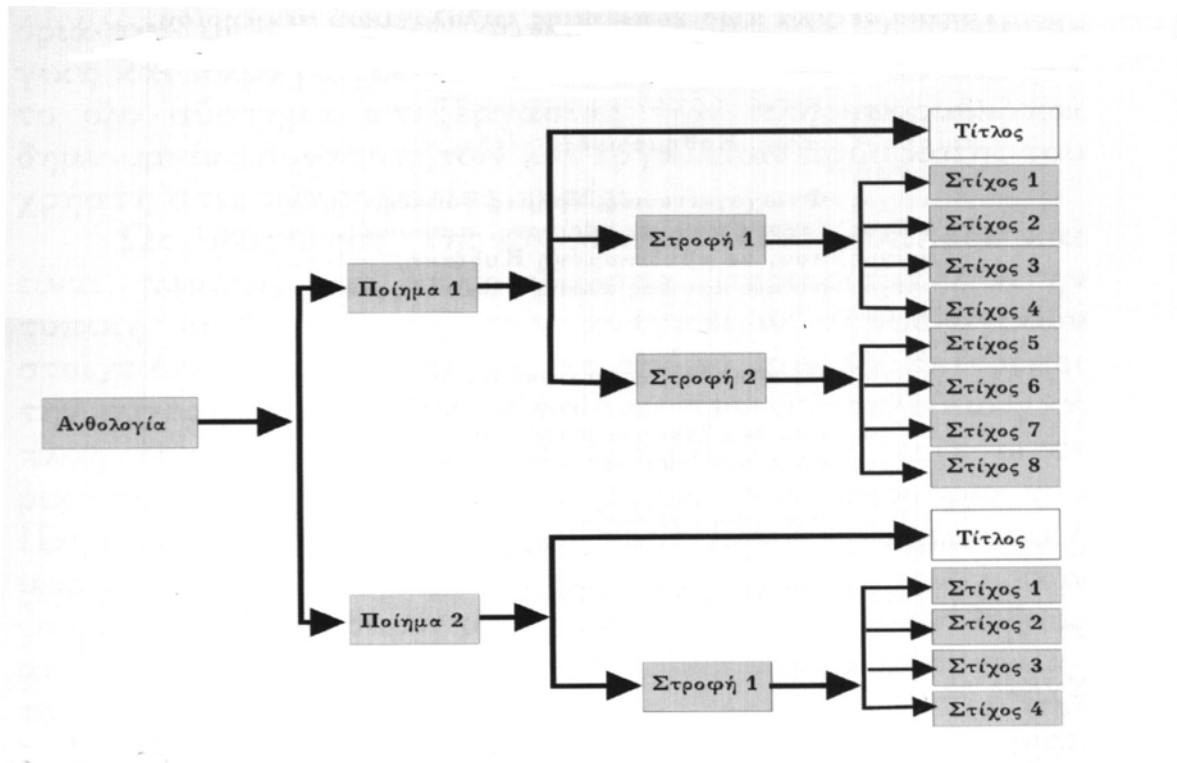
*Root Node:* Κάθε XML κείμενο μπορεί να έχει μόνο ένα root node. Στο παράδειγμα, root node είναι το element `<productDB>`.

*Element:* Κάθε XML κείμενο περιέχει elements. Το `<product prodid="223">` είναι element που περιέχει ένα attribute το `prodid` και τρία subelement τα `<prodname>`, `<listprice>`, `<dept deptname="PII">`. Το `dept` περιέχει ένα attribute το `deptname`.

Συνεπώς για να μπορέσει κάποιος να διαβάσει την πληροφορία που σχετίζεται με το προϊόν με κωδικό 225 πρέπει να αναζητήσει το element 'product' που έχει τιμή στο attribute 'prodid' ίση με 225. Κατόπιν από τα sub elements 'prodname', 'listprice' θα πληροφορηθεί για το όνομα και την τιμή του προϊόντος. Τέλος από το element dept θα μάθει τον κωδικό της κατηγορίας του προϊόντος ενώ από το attribute 'deptname' του συγκεκριμένου element θα πληροφορηθεί για το όνομα της κατηγορίας που ανήκει το προϊόν.

### 3.4 TEI

Μία από τις σημαντικότερες, από την άποψη της επεξεργασίας και της ευρύτητας χρήσης, στην εκδοτική βιομηχανία για την υποστήριξη της ηλεκτρονικής παραγωγής του συμβατικού έντυπου τεκμηρίου και τις προβλέψεις του υλοποιήθηκε από τη γλώσσα κωδικοποίησης των φιλολογικών, γενικά, κειμένων , η γνωστή ως TEI (Text Encoding Initiative) η οποία είναι μία διάλεκτος της γλώσσας XML.



**Σχήμα 3.1.** Ενδεικτική ανάλυση δομής και τυπικής δυνατότητας κωδικοποίησης φιλολογικού κειμένου

```

<ανθολογία>
<ποίημα><τίτλος>Καθημερινές</τίτλος>
<στροφή>
<στίχος>Φτωχογειτονιές, έρημες γωνιές,</στίχος>
<στίχος>έρημες καρδιές, ψύχρες, παγωνιές,</στίχος>
<στίχος>που, σε μουνδιασμένη Κυριακή,</στίχος>
<στίχος>στέκει και σας κλαίει θλιμμένη μουσική!</στίχος>
</στροφή>
<στροφή>
<στίχος>Προσωπάκια που έφεξαν, δειλά</στίχος>
<στίχος>στόματα που η πίκρα τα σφαλά</στίχος>
<στίχος>που δεν ξέρουνε ποτές φιλί θερμό</στίχος>
<στίχος>άλλο από τον ύστατο ασπασμό,</στίχος>
</στροφή>
<!-- Περισσότερες στροφές ακολουθούν εδώ -->
</ποίημα>

<!-- Περισσότερα ποιήματα ακολουθούν εδώ -->
</Ανθολογία>

```

**Σχήμα 3.2.** Τυπικό δείγμα κωδικοποίησης φιλολογικού τεκμηρίου, σύμφωνα με τις προβλέψεις  
TEI

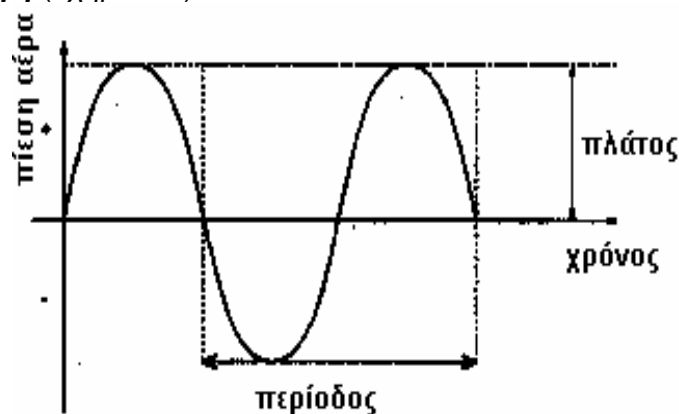
Μια ενδεικτική σχηματική απεικόνιση του τύπου της δομής και του τρόπου κωδικοποίησης ενός φιλολογικού τεκμηρίου, δίνεται στα σχήματα 3.1 και 3.2, όπου παρουσιάζεται τμήμα μιας τυπικής προς κωδικοποίηση δομής (στο πρώτο) και ένα δείγμα τυπικής κωδικοποίησης μέρους της δομής αυτής στο δεύτερο. Η λογική κωδικοποίησης και επισήμανσης των ποικίλων δομικών στοιχείων ενός τεκμηρίου, με ένα τυπικό και αναγνωρίσιμο, σε όλο το εύρος των πιθανών συστημάτων, τρόπο διαμορφώνει εξαιρετικά σημαντικές και, επί του παρόντος αδιερεύνητες προοπτικές. Οι προοπτικές αυτές αφορούν τόσο τη δυναμική διαμόρφωση τεκμηρίων, όσο και την πρόσβαση του ενδιαφερόμενου χρήστη σε οποιαδήποτε επίπεδο λεπτομέρειας και σε οποιαδήποτε πολυπλοκότητα συνδυασμών, σε ότι αφορά την αναζήτηση της πληροφορίας.

## 4. Ήχος

Οι μεθοδολογίες και τεχνικές του ήχου ασχολούνται με την επεξεργασία ηχητικών κυμάτων (ακουστικών σημάτων). Σημαντικά ζητήματα του χώρου αυτού είναι η **κωδικοποίηση**, η **αποθήκευση** και η **επεξεργασία** μουσικής και φωνής.

### 4.1 Βασικές έννοιες του ήχου

Ήχος είναι το φυσικό φαινόμενο και παράγεται από την ταλάντωση της ύλης. Καθώς η ύλη ταλαντώνεται, σημειώνονται διαφορές πίεσης στον περιβάλλοντα αέρα. Η εναλλαγή χαμηλής και υψηλής πίεσης μεταβιβάζεται διαμέσου του αέρα σε μορφή κύματος. Όταν το κύμα φτάσει στο ανθρώπινο αυτί, ακούμε τον ήχο. Το μοντέλο της ταλάντωσης καλείται **κυματομορφή** (σχήμα 4.1).



**Σχήμα 4.1 :** Ταλάντωση ενός κύματος πίεσης αέρα

Η κυματομορφή επαναλαμβάνεται σε τακτά χρονικά διαστήματα και αυτό το κομμάτι καλείται **περίοδος**. Γενικά, οι ήχοι που παρουσιάζουν αναγνωρίσιμη περιοδικότητα, τείνουν να είναι πιο μουσικοί από αυτούς που δεν είναι περιοδικοί. Παραδείγματα πηγών περιοδικού ήχου είναι: τα μουσικά όργανα, τα φωνήεντα, το σφύριγμα του ανέμου και τα τιπίσματα. Παραδείγματα μη-περιοδικού ήχου περιλαμβάνουν κρουστά όργανα χωρίς τονικότητα, το βήχα και το τρεχούμενο νερό.

## Συχνότητα

Η *συχνότητα* ενός ήχου είναι η αντίστροφη τιμή της περιόδου, αντιπροσωπεύει τον αριθμό των περιόδων ανά δευτερόλεπτο και μετράται σε **Χέρτζ (Hz)** ή **κύκλους ανά δευτερόλεπτο**. Πολλαπλάσιο του Hz είναι το 1 kHz = 1000Hz. Το εύρος συχνότητας διαιρείται ως εξής :

Υπόηχοι : 0 Hz – 20 Hz

Ακουστικοί ήχοι : 20 Hz – 20 kHz

Υπέρηχοι : 20 kHz – 10 THz

Τα συστήματα πολυμέσων, τυπικά, κάνουν χρήση του ήχου μόνο στα πλαίσια του εύρους της ανθρώπινης ακοής (*ακουστικό εύρος*): τα κύματα μέσα σε αυτό το εύρος συχνοτήτων λέγονται **ακουστικά σήματα**. Για παράδειγμα, η φωνή είναι ένα ακουστικό σήμα που παράγεται από τον άνθρωπο. Τα μουσικά σήματα έχουν εύρος συχνότητας μεταξύ 20 και 20000 Hz. Εκτός της φωνής και της μουσικής, εκλαμβάνουμε οποιοδήποτε άλλο ηχητικό σήμα ως θόρυβο.

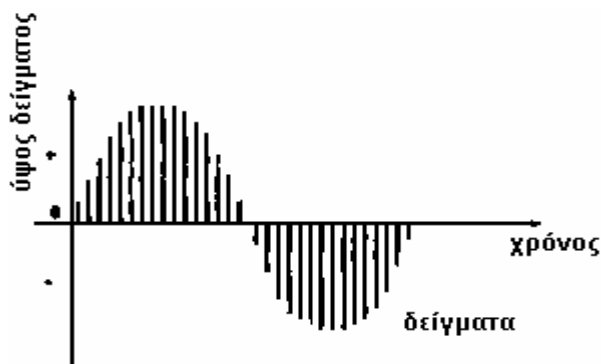
## Πλάτος ταλάντωσης

Ένας ήχος έχει και **πλάτος**, μια ιδιότητα που υποκειμενικά ακούγεται ως **ένταση**. Το πλάτος ενός ήχου είναι το μέτρο της μετατόπισης του κύματος της πίεσης του αέρα από το μέσο όρο ή την κατάσταση ηρεμίας.

## Αναπαράσταση του ήχου στον Υπολογιστή

Η ομαλή, συνεχής καμπύλη μιας ηχητικής κυματομορφής δεν παρίσταται κατευθείαν στον υπολογιστή. Ο υπολογιστής μετρά το πλάτος της κυματομορφής σε τακτά χρονικά διαστήματα και παράγει μια σειρά αριθμών. Κάθε τέτοια μέτρηση ονομάζεται **δείγμα**. Το σχήμα 4.2 απεικονίζει την περίοδο μιας κυματομορφής που το δείγμα της έχει ληφθεί ψηφιακά.

Ο μηχανισμός που μετατρέπει ένα ηχητικό σήμα σε ψηφιακά δείγματα είναι ο **αναλογικο-ψηφιακός μετατροπέας** ή **A.Ψ.Μ.(ADC)**. Η αντίστροφη διαδικασία τελείται από έναν **ψηφιακο-αναλογικό μετατροπέα** ή **Ψ.Α.Μ.(DAC)**.



*Σχήμα 4.2 : Κυματομορφή που έχει υποστεί δειγματοληψία*

### **Ρυθμός δειγματοληψίας**

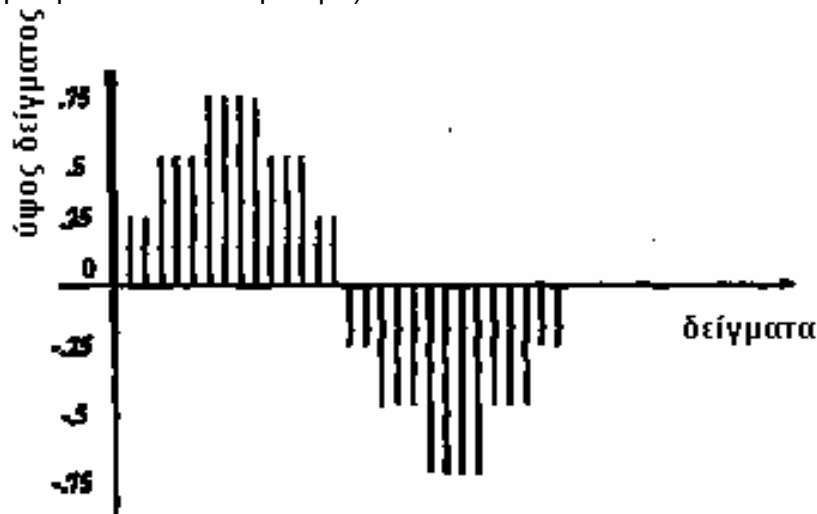
Ο ρυθμός με τον οποίον το δείγμα μια συνεχούς κυματομορφής (σχήμα. 4.2) λαμβάνεται, καλείται **ρυθμός δειγματοληψίας**. Όπως και οι συχνότητες, έτσι και οι ρυθμοί δειγματοληψίας μετρώνται σε Hz. Ο τυπικός ρυθμός δειγματοληψίας για τον οπτικό δίσκο (CD) είναι 44100 Hz, που σημαίνει ότι παίρνουμε δείγμα από την κυματομορφή 44100 φορές το δευτερόλεπτο. Φαινομενικά, αυτό βρίσκεται πάνω από το ακουστικό επίπεδο συχνότητας. Παρ' όλα αυτά, το εύρος ζώνης ( $20000 - 20 = 19980$  Hz) που μπορεί να αναπαραστήσει το δείγμα ενός ηχητικού σήματος που λαμβάνουμε ψηφιακά, είναι το πολύ ισοδύναμο με το μισό του οπτικού δίσκου (44100 Hz). Τα παραπάνω είναι εφαρμογή του **θεωρήματος δειγματοληψίας του Νίκουιστ (Nyquist)** - «Για ψηφιοποίηση χωρίς απώλειες, ο ρυθμός δειγματοληψίας πρέπει να είναι τουλάχιστον ο διπλός της μέγιστης συχνότητας του σήματος». Κατά συνέπεια, ένας ρυθμός δειγματοληψίας της τάξης των 44100 Hz μπορεί να αναπαραστήσει συχνότητα μέχρι 22050 Hz, τιμή κοντά στο άνω όριο του ακουστικού εύρους.

### **Κβαντοποίηση (Quantization)**

Όπως ακριβώς το δείγμα μιας κυματομορφής λαμβάνεται σε διακριτούς χρόνους, η τιμή του δείγματος είναι επίσης διακριτή. Η *ανάλυση* ή **κβαντοποίηση (quantization)** των τιμών ενός δείγματος εξαρτάται από τον αριθμό των δυαδικών ψηφίων που χρησιμοποιούνται στη μέτρηση του ύψους της κυματομορφής. Κβαντοποίηση 8

δυναμικών ψηφίων δίνει 256 πιθανές τιμές, ενώ κβαντοποίηση ποιότητας 16 δυναμικών ψηφίων σ' έναν οπτικό δίσκο, δίνει μέχρι 65536 τιμές.

Το σχήμα 3.3 παρουσιάζει μια κβαντοποίηση 3 δυναμικών ψηφίων. Το δείγμα της κυματομορφής με μέγεθος κβαντοποίησης 3 δυναμικών ψηφίων προσφέρει μόνο 8 πιθανές τιμές : 0,75 , 0,5 , 0,25 , 0 , -0,25 , -0,5 , -0,75 , -1. Το σχήμα της κυματομορφής γίνεται λιγότερο ευδιάκριτο με μικρότερο μέγεθος κβαντοποίησης, δηλαδή όσο μικρότερο είναι το μέγεθος της κβαντοποίησης, τόσο χειρότερη η ποιότητα του ήχου (το αποτέλεσμα μπορεί να είναι ένα βουητό).



Σχήμα 4.3 : Κβαντοποίηση τριών δυναμικών ψηφίων

## 4. 2 Μορφότυπα ήχου

Μια δειγματοληψία που χρησιμοποιεί 8 δυναμικά ψηφία για την αναπαράσταση του δείγματος και ρυθμό 8000 Hz θεωρείται αρκετά γρήγορη και ακριβής για είσοδο φωνής τηλεφωνικής ποιότητας. Ήχος ποιότητας οπτικού δίσκου παράγεται εάν ο στερεοφωνικός Ψ.Α.Μ. (μετατροπέας ψηφιακού σήματος σε αναλογικό) λειτουργεί στα 44100 δείγματα ανά δευτερόλεπτο με γραμμική, παλμοκωδική διαμόρφωση (**Pulse Code Modulation** ή **PCM**) 16 δυναμικών ψηφίων για κάθε τιμή δείγματος.

Τα παραπάνω παραδείγματα ποιότητας τηλεφωνικής γραμμής και οπτικού δίσκου δείχνουν ότι σημαντικές παράμετροι μορφοποίησης για τις προδιαγραφές του ήχου είναι : ο **ρυθμός δειγματοληψίας** και η **κβαντοποίηση του δείγματος**.

### 4.3 Μετάδοση φωνής

Η περιοχή της μετάδοσης φωνής έχει να κάνει με την αποτελεσματική κωδικοποίηση του σήματος φωνής / ήχου έτσι ώστε να επιτρέπεται η μετάδοση φωνής και ήχου σε δίκτυο. Ο στόχος είναι να δώσουμε στον παραλήπτη την ίδια ποιότητα ήχου με αυτή που αναπαρήχθη στην πλευρά του αποστολέα.

### 4.4 Κωδικοποίηση μορφής σήματος

Κωδικοποίηση αυτού του είδους δε λαμβάνει υπόψη συγκεκριμένες παραμέτρους και ιδιότητες της φωνής. Εδώ ο στόχος είναι να πετύχουμε την πιο αποδοτική κωδικοποίηση του ηχητικού σήματος. Ο ρυθμός μετάδοσης δεδομένων ενός κωδικοποιημένου-PCM στερεοφωνικού σήματος με απαιτήσεις ποιότητας οπτικού δίσκου(CD) είναι:

$$2 * \frac{44100}{s} * \frac{16bit}{8bits/sec} = 176.400 bytes/sec = 1.411.200 bits/sec$$

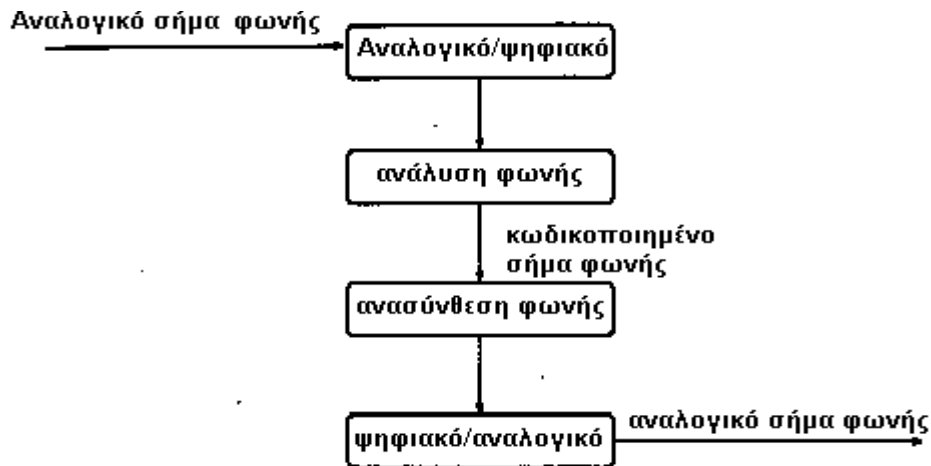
Για τηλεφωνική ποιότητα, σε αντίθεση με την ποιότητα οπτικού δίσκου, χρειάζεται ρυθμός μόνο 64 Kbit/sec. Χρησιμοποιώντας **Διαφορική Παλμοκωδική Διαμόρφωση (DPCM)** ο ρυθμός μπορεί να φτάσει στα 56 Kbits/sec χωρίς απώλεια ποιότητας. Η **Προσαρμόσιμη Διαφορική Παλμοκωδική Διαμόρφωση (ADPCM)** επιτρέπει την περαιτέρω μείωση του ρυθμού μετάδοσης στα 32 Kbits/sec.

### 4.5 Κωδικοποίηση πηγής

Ο κωδικοποιητής καναλιών φωνής (channel vo-coder) είναι ένα παράδειγμα συστήματος που λειτουργεί με αλγορίθμους κωδικοποίησης πηγής (*παραμετρικό σύστημα*) (σχήμα 4.4). Ρυθμός 3 Kbits/sec μπορεί να παραχθεί από έναν κωδικοποιητή φωνής όμως η ποιότητα δεν είναι πάντα ικανοποιητική.



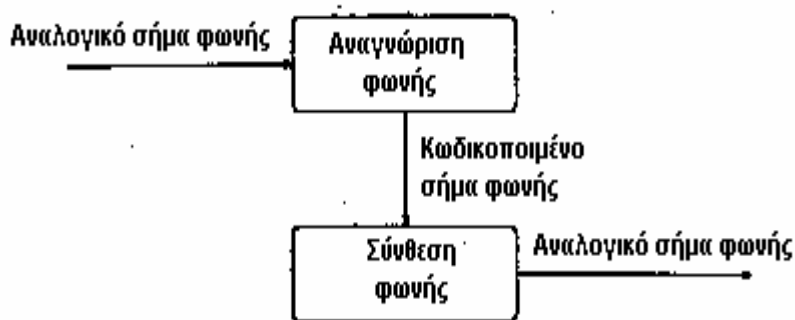
Μεγάλες προσπάθειες γίνονται για περαιτέρω μείωση του ρυθμού δεδομένων από τα 64 Kbits/sec στα 6 Kbits/sec, ώστε η ποιότητα του συμπιεσμένου σήματος, μετά την αποκωδικοποίηση, να ανταποκρίνεται σε αυτήν ενός μη συμπιεσμένου σήματος 64 Kbits/sec.



**Σχήμα 4.4 :** Κωδικοποίηση πηγής ενός παραμετρικού συστήματος: συστατικά ενός συστήματος μετάδοσης φωνής.

#### 4.6 Μέθοδοι Αναγνώρισης / Σύνθεσης φωνής

Έχουν γίνει προσπάθειες να μειωθεί ο ρυθμός μετάδοσης χρησιμοποιώντας μόνο μεθόδους αναγνώρισης / σύνθεσης. Η ανάλυση (αναγνώριση) φωνής γίνεται από τη μεριά του αποστολέα ενός συστήματος μετάδοσης φωνής και η σύνθεση (παραγωγή) φωνής ακολουθεί στη μεριά του παραλήπτη (σχήμα 4.5).

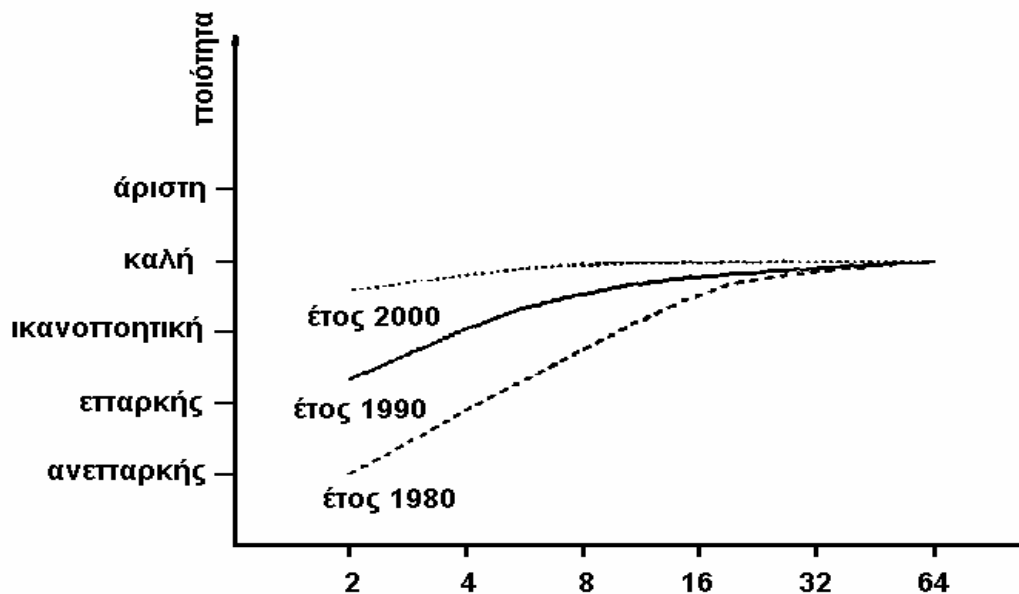


**Σχήμα 4.5 :** Συστήματα αναγνώρισης / σύνθεσης: συστατικά ενός συστήματος μετάδοσης φωνής.

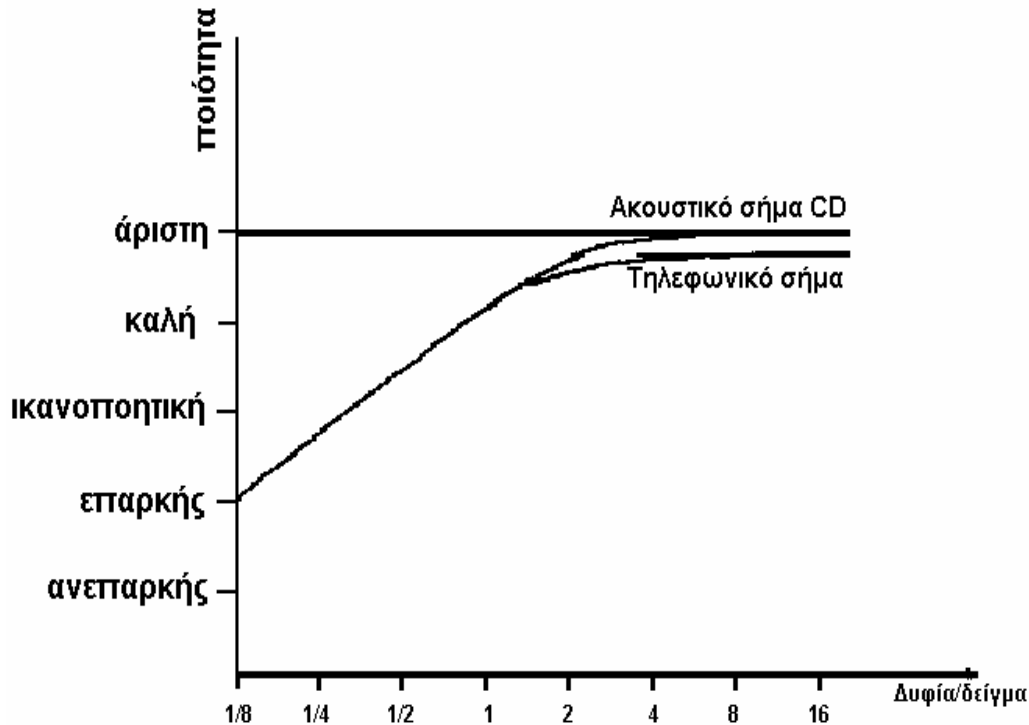
#### 4. 7 Ποιότητα που επιτυγχάνεται

Το θεμελιώδες ζήτημα που αφορά τη μετάδοση φωνής και ήχου, σχετικά με τα συστήματα πολυμέσων, είναι το πως θα πετύχουμε τον ελάχιστο ρυθμό δεδομένων για δεδομένη ποιότητα. Η δημοσιευμένη συνάρτηση του **Φλάναγκαν** (σχήμα 4.6) δείχνει τη σχέση μεταξύ της ποιότητας που επιτυγχάνεται στη συμπιεσμένη φωνή και του ρυθμού δεδομένων. Μπορούμε να υποθέσουμε πως για τηλεφωνική ποιότητα, ένας ρυθμός δεδομένων της τάξης των 8 Kbits/sec είναι αποδεκτός.

Το σχήμα 4.7 δείχνει τη σχέση της ποιότητας του ήχου με τον αριθμό των δυφίων ανά τιμή δείγματος. Για παράδειγμα, εξαιρετική ποιότητα οπτικού δίσκου μπορεί να επιτευχθεί με μια μείωση από τα 16 δυφία ανά τιμή δείγματος σε 2 δυφία ανά τιμή δείγματος. Αυτό σημαίνει ότι μόνο το 1/8 των αρχικών δεδομένων χρειάζεται να μεταδοθεί η ίδια ποιότητα σήματος.



Σχήμα 4.6: Σχέση της ποιότητας φωνής και του ρυθμού δεδομένων



**Σχήμα 4.7:** Σχέση της ποιότητας ήχου και του αριθμού των δυαδικών ψηφίων ανά δείγμα.

#### **4.8 Μουσική και το πρότυπο MIDI**

Γνωρίζουμε ότι κάθε ήχος μπορεί να αναπαρασταθεί σαν ένα ψηφιακά-κωδικοποιημένο ηχητικό σήμα που είναι μία αλληλουχία δειγμάτων, που το καθένα κωδικοποιείται με δυαδικά ψηφία. Αυτή η αλληλουχία μπορεί να είναι ο αποσυμπιεσμένη, όπως σε συμπυκνωμένους δίσκους ήχου ή συμπιεσμένη. Επίσης γνωρίζουμε ότι κάθε ήχος μπορεί να παρασταθεί με αυτόν τον τρόπο, περιλαμβανομένης και της μουσικής. Ένα χαρακτηριστικό αυτού του τρόπου αναπαράστασης είναι ότι δεν διατηρεί την σημασιολογική περιγραφή του ήχου. Εκτός και αν χρησιμοποιηθούν περίπλοκες τεχνικές αναγνώρισης, ο υπολογιστής δεν ξέρει αν μια αλληλουχία από βιτς παριστά, για παράδειγμα, ομιλία ή μουσική, ποιες νότες χρησιμοποιούνται και από ποιο όργανο.

Η μουσική μπορεί να περιγραφεί με ένα συμβολικό τρόπο. Στο χαρτί έχουμε τις παρτιτούρες. Οι υπολογιστές και ηλεκτρονικά μουσικά όργανα χρησιμοποιούν μια παρόμοια τεχνικά, και τα περισσότερα από αυτά χρησιμοποιούν την Ψηφική Διασύνδεση Μουσικού Οργάνου (Musical Instrument Digital Interface, MIDI), ένα πρότυπο που αναπτύχθηκε στις αρχές του 1980. Το πρότυπο MIDI ορίζει πως κωδικοποιούνται όλα τα

στοιχεία από τις παρτιτούρες μουσικής, όπως ακολουθίες από νότες, συνθήκες χρόνου και το όργανο που παίζει κάθε νότα.

#### **4.9 Εισαγωγή στο MIDI**

Το MIDI παριστά ένα σύνολο προδιαγραφών που χρησιμοποιείται για ανάπτυξη από όργανα, έτσι ώστε από διάφορους κατασκευαστές να μπορούν να ανταλλάξουν εύκολα μουσικές πληροφορίες. Το πρωτόκολλο MIDI είναι μια πλήρης γλώσσα μουσικής περιγραφής σε δυαδική μορφή. Κάθε λέξη που περιγράφει μία ενέργεια μιας μουσικής εκτέλεσης παίρνει ένα συγκεκριμένο δυαδικό κωδικό. Μία διασύνδεση MIDI αποτελείται από δύο διαφορετικά συστατικά.

Υλικό για σύνδεση του εξαρτήματος. Το υλικό MIDI καθορίζει την φυσική σύνδεση μουσικών οργάνων. Προσθέτει μία θύρα MIDI σε ένα όργανο, καθορίζει ένα καλώδιο MIDI (που συνδέει δύο όργανα), και επεξεργάζεται μουσικά σήματα που λαμβάνονται μέσω του καλωδίου.

Μία μορφή δεδομένων που κωδικοποιεί πληροφορίες που θα τις επεξεργαστεί το υλικό. Η μορφή δεδομένων δεν περιλαμβάνει την κωδικοποίηση των διακριτών τιμών δειγματοληψίας, όπως οι μορφές δεδομένων ήχου. Αντί αυτού, το MIDI χρησιμοποιεί μία συγκεκριμένη μορφή δεδομένων για κάθε όργανο, που περιγράφει πράγματα όπως η αρχή και το τέλος μιας παρτιτούρας, ή συχνότητα βάσης και η ένταση, εκτός των πληροφοριών που αφορούν το ίδιο το όργανο.

Η μορφή MIDI είναι ψηφιακή και τα δεδομένα ομαδοποιούνται σε μηνύματα MIDI. όταν ένας μουσικός παίζει ένα πλήκτρο, η διασύνδεση MIDI παράγει ένα μήνυμα MIDI που ορίζει την αρχή και το τέλος του κάθε κομματιού και την ένταση του. Αυτό το μήνυμα μεταδίδεται σε μηχανές που είναι συνδεδεμένες στο σύστημα. Μόλις ο μουσικός αφήσει το πλήκτρο, ένα άλλο σήμα (μήνυμα MIDI) δημιουργείται και μεταδίδεται.

#### **4.10 Συσκευές MIDI**

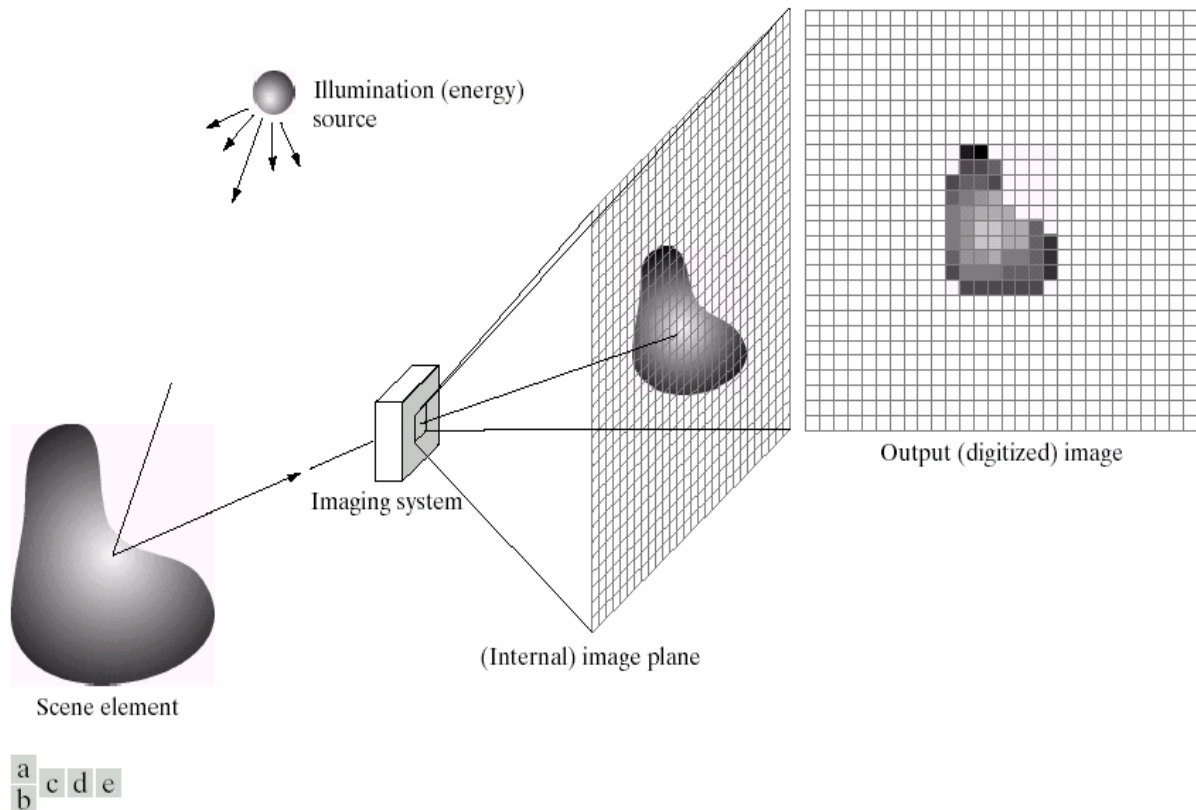
Ένα όργανο που συμφωνεί και με τα δύο συστατικά που ορίζονται από το πρότυπο MIDI (π.χ., συνθεσάιζερ) και είναι σε θέση να επικοινωνήσει με άλλες συσκευές MIDI μέσω καναλιών. Το πρότυπο MIDI καθορίζει 16 κανάλια. Μία συσκευή MIDI απεικονίζεται σε ένα κανάλι. Μουσικά δεδομένα που μεταδίδονται επάνω σε ένα κανάλι αναπαράγονται στο άκρο δέκτη του συνθεσάιζερ. Το πρότυπο MIDI αναγνωρίζει 128 όργανα με

ρυθμούς, περιλαμβανόμενων και εφφέ θορύβου (π.χ. η απογείωση ενός αεροπλάνου).  
Για παράδειγμα το 0 ) δηλώνει ένα πιάνο, 12 μια μαρίμπα, το 40 ένα βιολί και το 73 ένα φλάουτο.

## 5. Ψηφιακή Εικόνα – Βασικές Έννοιες – Βασικοί Όροι

### 5.1 Αναλογικό σε Ψηφιακό.

Η μετάβαση από τον αναλογικό κόσμο (ένα κόσμο με βασικό χαρακτηριστικό τη συνέχεια) στον ψηφιακό κόσμο (όπου τα μεγέθη απεικονίζονται μέσα από μια σειρά διακριτών δειγμάτων σε δεδομένες χρονικές στιγμές, συνήθως σε ίσες αποστάσεις μεταξύ τους), γίνεται με τη διαδικασία της ψηφιοποίησης (εικ. 5.1). Η διαδικασία της ψηφιοποίησης δεν είναι αντικείμενο του μαθήματος αν και καλύπτεται σε λεπτομέρειες στην προτεινόμενη βιβλιογραφία.

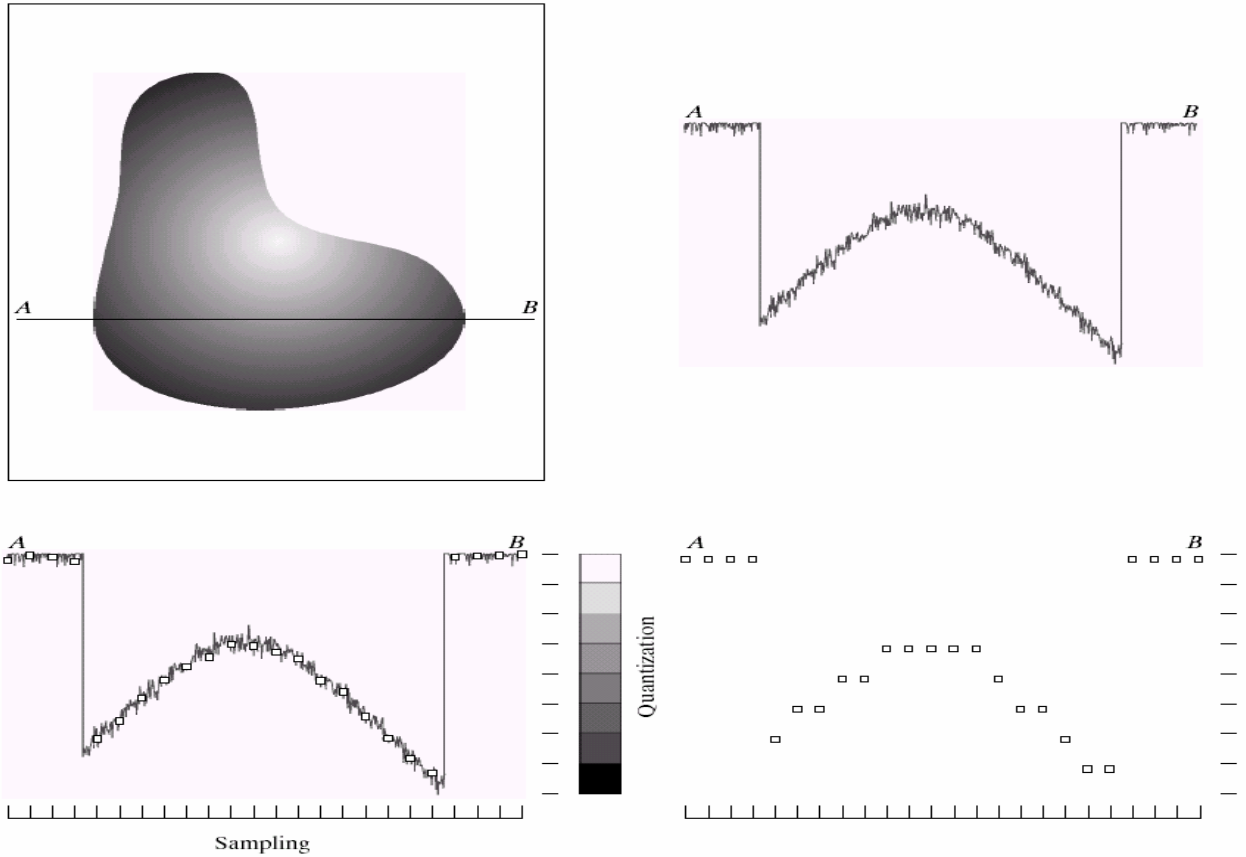


**FIGURE 2.15** An example of the digital image acquisition process. (a) Energy (“illumination”) source. (b) An element of a scene. (c) Imaging system. (d) Projection of the scene onto the image plane. (e) Digitized image.

#### Εικ. 5.1

Μια εικόνα ψηφιοποιείται μέσα από τις διαδικασίες της δειγματοληψίας και του κβαντισμού (εικ. 5.2). Με απλά λόγια, με δεδομένη συχνότητα λαμβάνονται και

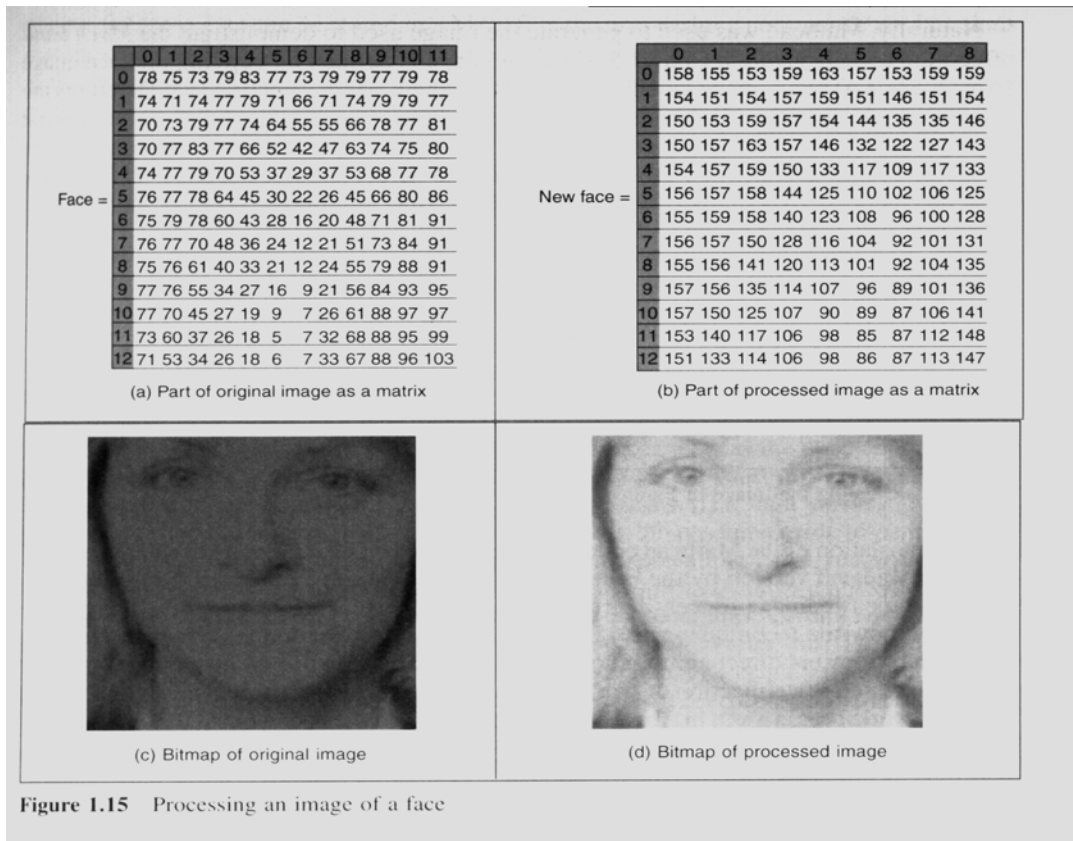
καταγράφονται δείγματα ενός φυσικού μεγέθους. Αν η συχνότητα δειγματοληψίας είναι μεγάλη σε σχέση με τον ρυθμό με τον οποίο το φυσικό μέγεθος μεταβάλλεται, μια αρκετά ακριβής απεικόνιση του μεγέθους αυτού προκύπτει μέσα από την μελέτη των δειγμάτων αυτών.



**Εικ. 5.2**

Στην περίπτωση της εικόνας, η διαδικασία της ψηφιοποίησης εκτελείται με φωτοευαίσθητες μονάδες καταγραφής όπως είναι τα ολοκληρωμένα CCD (Charged Coupled Diodes) ή τελευταία και CMOS. Και στις δύο περιπτώσεις, μια μικρή επιφάνεια αποτελείται από μεγάλο αριθμό βασικών στοιχείων σε διάταξη πλέγματος, που το κάθε ένα καταγράφει ένα στοιχείο ή σημείο της όλης εικόνας. Το κάθε βασικό στοιχείο που ονομάζεται pixel (picture element), μπορούμε να το δούμε σαν ένα μίνι-φωτόμετρο που μετράει την ένταση του φωτός που πέφτει πάνω στην μικρή επιφάνειά του. Μπορούμε να φανταστούμε τα pixels αυτά σαν κάτι αντίστοιχο του κόκκου στα συμβατικά φιλμ. Το κάθε pixel παράγει ένα επίπεδο τάσεως ανάλογο με την ένταση του προσπίπτοντος φωτός και σε μια διάρκεια χρόνου, μια ροή τάσης μεταβαλλόμενη ανάλογα με τις

μεταβολές του προσπίπτοντος φωτός ψηφιοποιείται από ένα μετατροπέα ADC (Analog to Digital Converter) και αποθηκεύεται σε μια κατάλληλη μονάδα μνήμης.



**Εικ. 5.3**

Στην περίπτωση που ένα σύστημα συλλαμβάνει μονοχρωματικές εικόνες με ακρίβεια 8-bit, κάθε ρixel καταγράφει σε οκτώ δυαδικά επίπεδα άρα μπορεί να διακρίνει  $2^8$  βαθμίδες άρα 256 διαφορετικές τιμές. Επομένως μπορεί να καταγράψει το μαύρο σαν 0, το άσπρο σαν 255 και ανάμεσά τους, 254 βαθμίδες (τόνους) του γκρι (εικ.5.3).

## **5.2 Χωρική Διάταξη, Ανάλυση, Διακριτικότητα (Spatial Resolution).**

Όλες οι ψηφιακές εικόνες αποτελούνται από εικονικά στοιχεία (picture elements, pixels). Η ποιότητα μιας εικόνας που μετράται σαν διακριτικότητα ή ανάλυση (resolution), σε μεγάλο βαθμό εξαρτάται από το συνολικό αριθμό από pixels που την αποτελούν. Είναι



προφανές πως όσο περισσότερα pixels διαθέτει μια εικόνα (με δεδομένο φυσικό μέγεθος), τόσο περισσότερη πληροφορία είναι διαθέσιμη για να την περιγράψει.

Ο συνολικός αριθμός από pixels βγαίνει πολλαπλασιάζοντας τα οριζόντια με τα κάθετα pixels μίας εικόνας.

Π.χ.  $640 \times 480 = 307,200$  pixels.

$1024 \times 680 = 696,320$  pixels δηλαδή περισσότερο από διπλάσια από την προηγούμενη εικόνα.

Σε μικρό μέγεθος, οι δύο παραπάνω εικόνες μπορεί να δείχνουν ίδιες αλλά αν μεγεθυνθούν αρκετά, οι διαφορές θα αρχίσουν να γίνονται ορατές.

Η δυνατότητα ανάλυσης κάθε συσκευής ψηφιοποίησης, όπως οι scanners και οι ψηφιακές μηχανές, μετράται σε Pixels ανά ίντσα ή PPI (Pixels Per Inch). Αν πάρουμε ένα επιτραπέζιο flat-bed ψηφιοποιητή (scanner), με αισθητήρα CCD μεγέθους 8,5 ιντσών το οποίο διαθέτει 2540 στοιχεία, τότε η μέγιστη διακριτικότητα (resolution) θα είναι 2540 διαιρεμένα δια 8,5 = 300 pixels ανά ίντσα (PPI). **ΚΑΚΩΣ** αυτό συχνά αναφέρεται σαν DPI (Dots Per Inch), μια μονάδα που όπως θα δούμε παρακάτω, αναφέρεται σε χαρακτηριστικά εκτύπωσης και δεν έχει άμεση μονοσήμαντη σχέση με το απόλυτο resolution.

### ***5.3 Άλλες σχετικές έννοιες - Pixels Per Inch (PPI), Dots Per Inch (DPI) & Lines Per Inch (LPI).***

Οι όροι που συναντάμε σχετικά με το resolution οδηγούν συχνά σε σύγχυση και είναι καλό να τους διαχωρίσουμε και επεξηγήσουμε.

#### **PPI – Pixels per inch:**

Αυτή είναι η μονάδα που δείχνει το resolution μίας ψηφιακής εικόνας. Είναι σωστό να συνδέεται με συσκευές ψηφιοποίησης καθώς έχει άμεση σχέση με τον διαθέσιμο αριθμό εικονοστοιχείων (pixels) που διαθέτει κάθε αισθητήρας οποιασδήποτε μορφής και τύπου. Οι συσκευές ψηφιοποίησης δημιουργούν pixels per inch.

### **DPI – Dots per inch:**

Αυτή είναι η μονάδα που δείχνει το resolution μίας μονάδας εξόδου (π.χ. ενός εκτυπωτή). Δείχνει πόσα σημεία μελάνης (χρώματος) έχει τη δυνατότητα μια τέτοια συσκευή να τοποθετεί πάνω σε μια δεδομένη διάσταση (πχ μία ίντσα) της επιφάνεια εκτύπωσης (π.χ. χαρτί). Δεν είναι άμεσα συνδεδεμένο με την τελική ποιότητα μίας εικόνας καθώς ένας εκτυπωτής των 300 DPI μπορεί να τυπώσει εικόνες τόσο των 300PPI όσο και των 900PPI αλλά και των 100PPI. Η διαφορά είναι το φυσικό μέγεθος στο οποίο θα τυπωθεί η κάθε μία καθώς και στις τρεις παραπάνω περιπτώσεις ο εκτυπωτής θα «τοποθετήσει» τα διαθέσιμα pixels σε πυκνότητα 300dots ανά ίντσα.

### **LPI – Lines per inch:**

Αυτή είναι η μονάδα που χρησιμοποιείται όταν μια ψηφιακή εικόνα πρόκειται να τυπωθεί με την μέθοδο του half-tone (θα το δούμε παρακάτω) που είναι συνηθισμένοι στον χώρο των εκδόσεων, εφημερίδων και εκτυπώσεων. Αναφέρεται στον αριθμό των γραμμών του πλέγματος που χρησιμοποιείται στην διαδικασία εκτύπωσης και τυπικές τιμές είναι τα 85LPI για εφημερίδες και 133LPI για βιβλία καλής ποιότητας.

## **5. 4 Μεγέθη και Ανάλυση**

Οι μονάδες PPI και DPI συσχετίζονται για να μας οδηγήσουν από το φυσικό μέγεθος πηγής (π.χ. το φυσικό μέγεθος μίας φωτογραφικής διαφάνειας ή αρνητικού) στο επιθυμητό φυσικό μέγεθος τελικού προϊόντος (π.χ. η εκτυπωμένη φωτογραφία, μια αφίσα κλπ).

Ένας scanner των 1000PPI δημιουργεί 1000pixels για κάθε ίντσα φυσικού μεγέθους πηγής που ψηφιοποιείται.

Ένας εκτυπωτής των 300 DPI τυπώνει 300 σημεία σε κάθε ίντσα κάθε διάστασης.

Μπορούμε να δούμε τα pixels σαν πλακάκια και το τελικό μέγεθος εκτύπωσης σαν ένα πάτωμα που θέλουμε να καλύψουμε με τα πλακάκια.

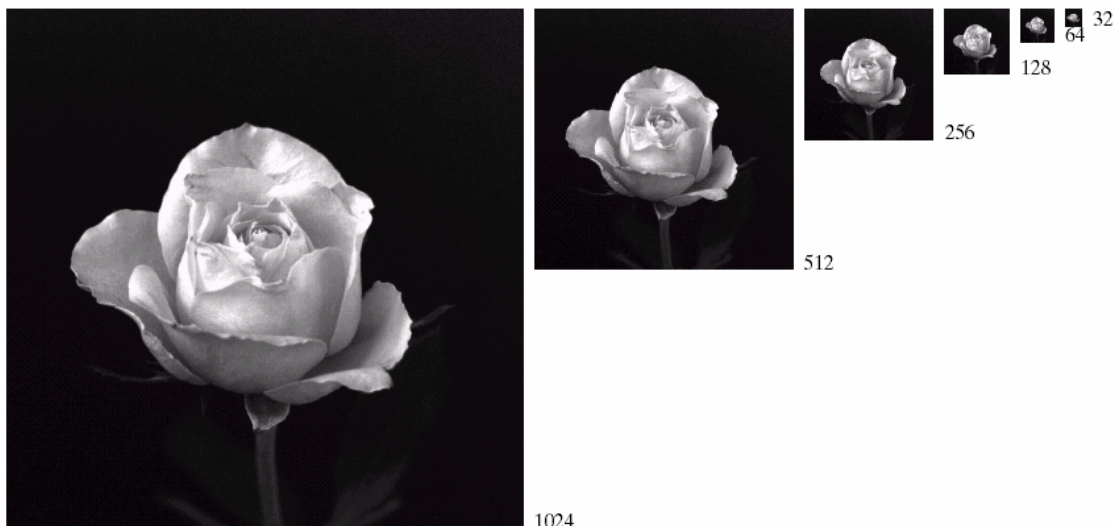
Αν έχουμε ένα πάτωμα 10x10 cm και πλακάκια 1x1cm, τότε, χρειαζόμαστε 100 πλακάκια για να «γεμίσουμε» το διαθέσιμο πάτωμα. Η πυκνότητα τοποθέτησης είναι 1 πλακάκι ανά cm. Αν όμως αποφασίσουμε να τοποθετήσουμε τα πλακάκια με συχνότητα 0,5 πλακάκια ανά cm, τότε με τα 100 πλακάκια που έχουμε θα καλύψουμε μια επιφάνεια 20x10cm άσχετα αν υπάρχει πολύς κενός χώρος ανάμεσα στο κάθε πλακάκι και τα διπλανά του. Με όμοιο τρόπο, αν έχουμε όχι 100 αλλά 50 πλακάκια, δεν επαρκούν για κάλυψη της επιφάνειας 10x10 με πυκνότητα 1πλακ/cm (καλύπτουν 10X5cm)

Συνεπώς:

Μια εικόνα των 300PPI μεγέθους 1x1 ίντσες, θα τυπωθεί σε ένα εκτυπωτή των 300DPI σε φυσικό μέγεθος 1x1 ίντσες.

Μια εικόνα των 600PPI μεγέθους 1x1 ίντσες, θα τυπωθεί σε ένα εκτυπωτή των 300DPI σε φυσικό μέγεθος 2x2 ίντσες.

Μια εικόνα των 150PPI μεγέθους 1x1 ίντσες, θα τυπωθεί σε ένα εκτυπωτή των 300DPI σε φυσικό μέγεθος 0,5x0,5 ίντσες.



**Εικ. 5.4**

Από το παραπάνω απλό παράδειγμα (εικ. 5.4) αρχίζει να γίνεται προφανές πως η τελική χρήση μίας εικόνας και η συσκευή εξόδου είναι το βασικό στοιχείο που πρέπει να λαμβάνεται υπόψη για να αποφασιστεί ποιο είναι το σωστό (ικανοποιητικό) μέγεθος στο οποίο πρέπει να ψηφιοποιηθεί μια εικόνα.

Αν για παράδειγμα θέλουμε μια εικόνα με τελικό φυσικό μέγεθος 20x25 ίντσες, η οποία θα τυπωθεί από ένα εκτυπωτή των 300 DPI και το αρχικό υλικό μας είναι μια διαφάνεια 1,5x1,5 ίντσες. Για να αποφασίσουμε ποια είναι η σωστή ποιότητα ψηφιοποίησης, διαλέγουμε τη μεγαλύτερη διάσταση του τελικού προϊόντος (25 ίντσες). Στις 25 ίντσες αυτές, με 300 σημεία σε κάθε ίντσα (DPI), χρειαζόμαστε  $25 \times 300 = 7,500 \text{ dots}$ . Για να πάρουμε τα σημεία αυτά από τη συσκευή ψηφιοποίησης, αυτή είναι απαραίτητο να διαθέτει τουλάχιστον μια ανάλυση του επιπέδου των:  $7,500 / 1,5 = 5000 \text{ PPI}$ .

Αντίστροφα τώρα για καλύτερη κατανόηση, ψηφιοποιώντας στα 5000PPI ένα αρνητικό των 1,5x1,5 ιντσών, παίρνουμε ένα αρχείο των  $5000 \times 1,5 = 7500 \text{ pixels}$ . Δηλαδή θα έχουμε ένα αρχείο των  $7,500 \times 7,500 \text{ pixels}$ .

Αν τυπώσουμε το αρχείο σε ένα εκτυπωτή των 300DPI, τότε το τελικό φυσικό μέγεθος που θα πάρουμε θα είναι  $7,500 / 300 = 25$  ίντσες σε κάθε διάσταση επομένως θα έχουμε μια εκτύπωση μέγιστου μεγέθους 25x25 ιντσών.

### **Συνοπτικά:**

Μέγεθος Αρχείου από scanner σε Pixels = (ανάλυση του scanner σε PPI) x (οριζόντιο φυσικό μέγεθος αρχείου,  $\chi$ ) x (ανάλυση του scanner σε PPI) x (κάθετο φυσικό μέγεθος αρχείου,  $\gamma$ ).

Τελικό φυσικό μέγεθος εκτύπωσης σε ίντσες:

Αν το αρχείο έχει συνολικά  $A$  pixels, τότε :  
στην οριζόντια διάσταση έχει  $A \cdot \chi / (\chi + \gamma)$  pixels,  
στην κάθετη διάσταση έχει  $A \cdot \gamma / (\chi + \gamma)$  pixels.

**Διάσταση x = (αριθμός pixels στην οριζόντια διάσταση) / (printer resolution σε Dpi)**

**Διάσταση y = (αριθμός pixels στην κάθετη διάσταση) / (printer resolution σε Dpi)**

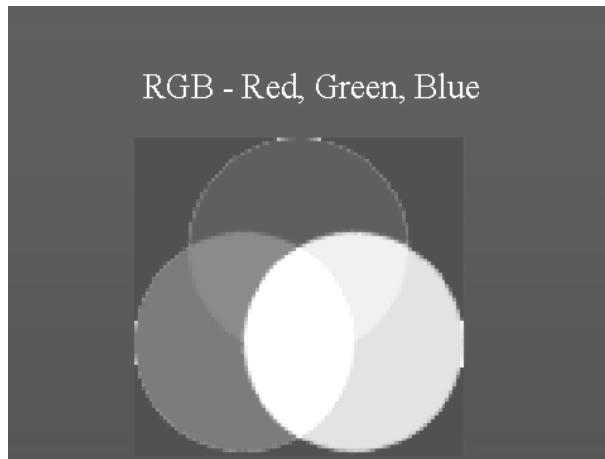
### **5. 5 Βάθος - Bit Depth.**

Το βάθος μας δείχνει πόσες πιθανές διαβαθμίσεις μπορεί να πάρει το κάθε pixel μίας εικόνας. Όπως είπαμε και παραπάνω, μια μονοχρωματική εικόνα με βάθος 2 bit έχει  $2^2 = 4$  διαβαθμίσεις, ενώ μια εικόνα των 3 bit θα έχει  $2^3 = 8$  διαβαθμίσεις. Μια εικόνα των 8 bit θα έχει  $2^8 = 256$  διαβαθμίσεις.

Σε γενικές γραμμές είναι αποδεκτό πως το ανθρώπινο μάτι αντιλαμβάνεται σαν απόλυτα συνεχή (continuous tone) τη διαβάθμιση από λευκό σε μαύρο για αναλύσεις από 8bit και πάνω.

### **5. 6 Η Έγχρωμη Εικόνα και το χρώμα στον Υπολογιστή, Βάθος Χρώματος.**

Η παραπάνω απεικόνιση εφαρμόζεται και στις έγχρωμες ψηφιακές εικόνες. Οι υπολογιστές δεν έχουν ουδεμία αίσθηση χρώματος. Για να απεικονίσουμε μια έγχρωμη εικόνα, είναι απαραίτητο να διαχωρίσουμε την εικόνα στα τρία βασικά συστατικά χρώματος (εικ.5.5), τα τρία πρωτεύοντα χρώματα, Κόκκινο, Πράσινο και Μπλε (RGB). Αν κάθε ένα από τα συστατικά αυτά απεικονίζεται με βάθος 8 bit, τότε το αρχείο που θα προκύψει θα έχει βάθος  $3 \times 8$  bit επομένως 24 bit. Οι δυνατότητες απεικόνισης διαφορετικών χρωματικών τόνων στα 24 Bit είναι  $256 \times 256 \times 256$  επομένως 16,777,216 διαφορετικοί τόνοι χρώματος.



**Εικ. 5.5**

Περισσότερες πληροφορίες σχετικά με τον τρόπο που απεικονίζεται το χρώμα στον Υπολογιστή υπάρχουν στις Βοηθητικές Σημειώσεις μαθήματος με τίτλο: “Color Models”. Είναι απαραίτητο να γίνει κατανοητό το μέρος που αναφέρεται στα μοντέλα RGB και CMYK, οι έννοιες που αναφέρονται σε Πρωτεύοντα και συμπληρωματικά χρώματα, Προσθετικά και αφαιρετικά χρώματα, Εκπεμπόμενο και εξ αντανάκλασης χρώμα καθώς επίσης και για την έννοια των χρωματικών πεδίων Color Spaces.

Σήμερα ήδη είναι δυνατόν να ψηφιοποιούνται εικόνες σε περισσότερα bit ανά συστατικό χρώμα, όπως 10, 12 και 16 bit με αποτέλεσμα ακόμα μεγαλύτερης ποιότητας εικόνες.

Αν και το ανθρώπινο μάτι αντιλαμβάνεται σαν απόλυτα συνεχή (continuous tone) τη χρωματική διαβάθμιση για αναλύσεις από 24bit και πάνω, υψηλότερες αναλύσεις όπως οι παραπάνω έχουν σημαντική σημασία και αξία στην διαδικασία επεξεργασίας εικόνας κατά την οποία είναι δυνατόν να παραχθούν περισσότεροι ενδιάμεσοι τόνοι με μεγαλύτερη χρωματική ακρίβεια.



```

>> clear
>> [x,map]=imread('Fig6_361.jpg');
>> y=x(100:110,100:110);
>> imshow(y)
>> R=x(105:110,105:110,1);
>> G=x(105:110,105:110,2);
>> B=x(105:110,105:110,3);
>> R

R =

    32    31    30    30    30    31
    32    32    31    31    31    31
    32    32    32    31    31    30
    32    32    32    32    31    30
    31    30    32    32    31    31
    31    30    32    31    30    30

>> G

G =

    33    32    31    31    31    32
    33    33    32    32    32    32
    33    33    33    32    32    31
    35    34    33    33    32    32
    35    34    33    32    31    31

>> B

B =

    35    34    33    33    33    34
    35    35    34    34    34    34
    35    35    35    34    34    33
    35    35    35    35    34    33
    36    35    35    35    34    34
    36    35    35    34    33    33

```

**Εικ. 5.6** Ανάλυση εικόνας στα τρία χρωματικά πρότυπα RGB

## **5. 7 Μεγέθη Αρχείων.**

Η ανάλυση μίας ψηφιακής εικόνας οδηγεί και σε δεδομένα μεγέθη ψηφιακών αρχείων. Ο τρόπος υπολογισμού, για ασυμπίεστα αρχεία είναι πολύ απλός.

**Συνολικός αριθμός pixel X (βάθος χρώματος διαιρούμενο δια 8) = συνολικός αριθμός bytes**

Παράδειγμα, μια εικόνα 1200x1200 pixels των 24bit:

Μέγεθος ψηφιακού αρχείου=1200x1200x(24/8)=2,880,000bytes, δηλαδή περίπου 2,8MB

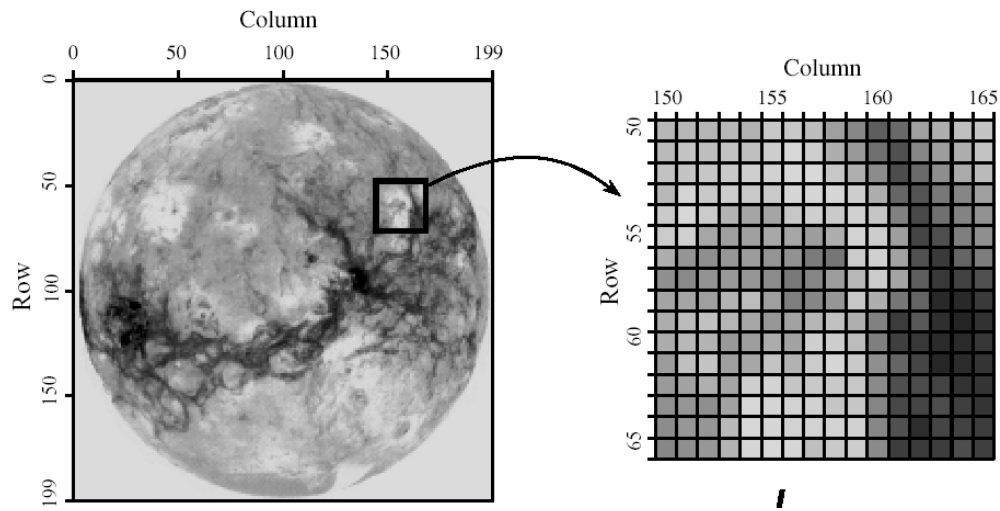
Στην καθημερινή πρακτική, καθώς είναι συνηθισμένο να συναντάμε τιμές σε Mpixels (MegaPixels), ένας εύκολος πρακτικός τρόπος είναι να πολλαπλασιάσουμε τον συνολικό αριθμό από Mpixels με το 3 οπότε και παίρνουμε περίπου το μέγεθος σε MB. Αν έχουμε για παράδειγμα μια εικόνα των 1,5 εκατ pixels, τότε αυτή θα μας δώσει ένα αρχείο των 4,5MB.

Ας μην ξεχνάμε πάντως πως οι διάφορες τεχνικές συμπίεσης έχουν σαν αποτέλεσμα το τελικό μέγεθος ενός ψηφιακού αρχείου να μην είναι άμεσα συσχετιζόμενο με την ανάλυσή του σε pixels.

## **5. 8 Ανύσματα (vectors) και bitmaps.**

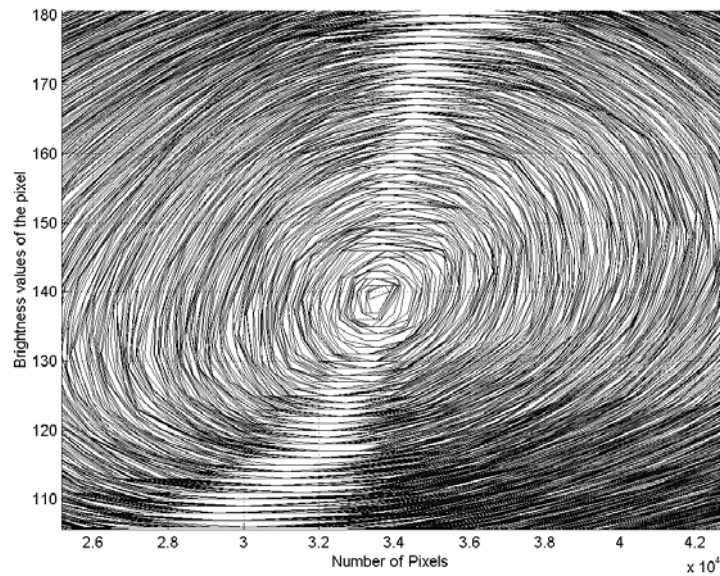
Το παρόν μάθημα θα ασχοληθεί με αρχεία βασισμένα σε bits, τα οποία ονομάζονται και bitmapped. Σε ένα αρχείο της μορφής αυτής, μια εικόνα, όπως είδαμε, χωρίζεται σε στοιχεία, pixels. Το κάθε Pixel περιγράφεται με συντεταγμένες χ και ψ που μονοσήμαντα ορίζουν τη θέση του και είτε μία τιμή (αν πρόκειται για μονοχρωματική εικόνα) είτε τρεις τιμές ( μία για κάθε πρωτεύον χρώμα) που καθορίζουν το χρώμα του. Κάθε Pixel έχει μόνο ένα χρωματικό τόνο και η εικόνα ανασυντίθεται από τον συσχετισμό όλων των pixel της. Ένα pixel με τιμή 125 δίνει ένα μέσο γκρι τόνο (εικ. 5.7) ενώ ένα pixel με R255, G255, B0 δίνει κίτρινο χρώμα (κόκκινο + πράσινο = κίτρινο). Για μια Bitmapped εικόνα είναι απαραίτητο να αποθηκεύονται όλες οι παραπάνω πληροφορίες για κάθε διακριτό pixel. Μια και η ελάχιστη καταγραφόμενη πληροφορία έχει το μέγεθος του ενός pixel, πλάγιες γραμμές παρουσιάζουν το φαινόμενο του aliasing (μοιάζουν με σκάλα).





**Εικ. 5.7**

Αντίθετα, μία ανυσματική εικόνα (εικ.5.8) χρησιμοποιεί διανύσματα για να περιγράψει σχήματα χωρίς να αποθηκεύει επιμέρους σημεία της. Ένα τετράγωνο επομένως θα απεικονίζεται με τον μαθηματικό τύπο του τετραγώνου και μια περιγραφή του χρώματος που αυτό περιέχει. Οι ανυσματικές εικόνες απεικονίζουν γραμμικά σχέδια με μεγάλη ακρίβεια αλλά δεν έχουν δυνατότητες απεικόνισης διαβαθμίσεων χρώματος και είναι φτωχά στην απόδοση συνεχών τόνων. Για το λόγο αυτό χρησιμοποιούνται κατά κόρον σε σχεδιαστικές εφαρμογές ενώ τα bitmapped αρχεία είναι τα προτιμώμενα στην περίπτωση της ψηφιακής εικόνας.



**Εικ. 5.8** Παράδειγμα Διανυσματικής Εικόνας

**Ανακεφαλαιωτικά** στα computer graphics υπάρχουν δύο κύριες κατηγορίες εικόνων, οι vector και οι raster.

► **Vector** (διανυσματικές)

Είναι οι εικόνες που δημιουργούνται από τον μαθηματικό σχεδιασμό γραμμών και καμπύλων γραμμών, που ονομάζονται vectors. Τα αντίστοιχα προγράμματα για τη δημιουργία των vector εικόνων ονομάζονται vector - programs (π.χ. CorelDRAW!, Illustrator κ.ά.) και χρησιμοποιούνται κυρίως για κείμενο, λογότυπα ή σχήματα, όταν θέλουμε να έχουμε πολύ καθαρές γραμμές.

► **Raster** εικόνες

Raster εικόνες είναι εκείνες που δημιουργούνται από κάποιο πλέγμα (grid ή raster), από πολύ μικρά τετράγωνα που ονομάζονται pixels. Προγράμματα όπως το Photoshop και το Photostyler ονομάζονται raster - based programs και τα χρησιμοποιούμε κυρίως όταν θέλουμε να δουλέψουμε με φωτογραφίες,

που έχουν εισαχθεί στον υπολογιστή με scanners ή από Photo-CD, καθώς και με εικόνες που έχουν δημιουργηθεί από προγράμματα ζωγραφικής.

Η κυριότερη διαφορά μεταξύ των εικόνων vector και raster είναι η σχέση που έχουν με την ανάλυση (resolution). Οι μεν vector εικόνες δεν επηρεάζονται από την ανάλυση, γι' αυτό ονομάζονται και resolution independent, ενώ οι raster εικόνες έχουν άμεση σχέση με την ανάλυση και γι' αυτό ονομάζονται και resolution dependent.

## 6. Τελικός Προορισμός – Output.

Η ψηφιακή επεξεργασία σε μια εικόνα εκτελείται πάντα με δεδομένο τελικό σκοπό χρησιμοποίησης της. Ο τελικός προορισμός μίας εικόνας αποτελεί σημαντικό παράγοντα που επηρεάζει αποφάσεις και επιλογές που έχουν να κάνουν με όλη την αλυσίδα επεξεργασίας. Σαν προϊόν εξόδου πρέπει να διαχωρίσουμε τα:

Τελικό προϊόν, το προϊόν στο οποίο η διαδικασία όχι μόνο τελειώνει αλλά και η ίδια η αλυσίδα επεξεργασίας και χειρισμού του προϊόντος τελειώνει. Πχ, μια αφίσα, μια τυπωμένη φωτογραφία σε σελίδα περιοδικού, βιβλίου κλπ.

Δοκίμιο (proof), ένα ενδιάμεσο προϊόν εξόδου, που έχει σα σκοπό τον ποιοτικό έλεγχο του αποτελέσματος της επεξεργασίας ή πιθανόν αποτελεί το τελικό στάδιο παράδοσης του προϊόντος από ένα συνεργάτη στον επόμενο που θα συνεχίσει τη διαδικασία. Παράδειγμα, ένα proof, μια εκτύπωση όσο πιο κοντά στο τελικό προϊόν είναι εφικτό που συνοδεύει ένα CD με την τελική εικόνα προς το ειδικό τυπογραφείο στο οποίο θα γίνει η ένθεση μίας εικόνας σε μια σελίδα περιοδικού.

Όταν αναφερόμαστε στην αξιολόγηση προϊόντων εξόδου είναι απαραίτητο να έχουμε υπόψη μας πως η “τέλεια απόδοση” πρέπει να συνδυάζεται με την “βέλτιστη απόδοση” όπως αυτή εφαρμόζεται στην ορισμένη συσκευή εξόδου και για την δεδομένη εφαρμογή. Για κάθε διαφορετικό είδος τελικού προϊόντος, διαφορετικοί παράγοντες οδηγούν στο “βέλτιστο προϊόν”. Δεν είναι οι απαιτήσεις μίας υπερπολυτελούς έκδοσης οι ίδιες με εκείνες της εκτύπωσης μίας εικόνας σε οικιακό desktop printer.

**ΚΑΝΟΝΑΣ:** Να διαχειριζόμαστε ΑΚΡΙΒΩΣ τόση πληροφορία όση είναι απαραίτητη για κάθε εφαρμογή. Σαν αποτέλεσμα, αναλύοντας τις ειδικές απαιτήσεις και χαρακτηριστικά μίας χρήσης ης εικόνας μας πρέπει εξ αρχής να επιλέγουμε κατάλληλες τιμές για τα παρακάτω:

Output resolution,

Μέγεθος αρχείου, τόσο σε φυσικές διαστάσεις παρουσίασης/εκτύπωσης (Xcm x Ycm) αλλά και σε μέγεθος αποθήκευσης (MB),

Βάθος χρώματος (8, 12, 16, 24, 36, 48bit)

Όλα τα παραπάνω καθορίζουν την ποιότητα του τελικού προϊόντος αλλά πάντα πρέπει να ευθυγραμμίζονται με την τελική χρήση.

Εξετάζοντας προσεκτικά τις απαιτήσεις της εξόδου μπορούμε να επιλέξουμε τις τιμές εκείνες που θα μας επιτρέψουν την σωστή επεξεργασία ενός αρχείου από την ίδια την είσοδο του στην αλυσίδα επεξεργασίας.

Επίσης είναι απαραίτητο πάντα να έχουμε υπόψη μας πως υπάρχουν δεδομένες Χρωματικές Διαφορές ανάμεσα στην εικόνα που βλέπουμε στο monitor του Η/Υ που είναι RGB και την τελική εκτυπωμένη εικόνα που είναι προϊόν εκτύπωσης σε τετραχρωμία CMYK. Η σχέση των χρωμάτων διατηρείται μέσα από τη χρήση των λεγομένων χρωματικών προφίλ (περισσότερα παρακάτω). Αποφάσεις που λαμβάνονται μόνο με βάση την εικόνα στο Monitor μπορούν να οδηγήσουν σε σοβαρά λάθη. Ενώ είναι εύκολο να απορρίψουμε μία κακή εκτύπωση σε ένα inkjet και να ξανατυπώσουμε την εικόνα, δεν είναι το ίδιο απλό αν το λάθος έγινε στην εκτύπωση ενός βιβλίου ή περιοδικού, ιδιαίτερα όταν έχουν εκτυπωθεί ήδη μερικές εκατοντάδες ή χιλιάδες αντίτυπα.

## **7. Σύλληψη της Ψηφιακής Εικόνας – Image Capture – Input.**

Μια αναλογική εικόνα μετατρέπεται σε ένα ΠΛΕΓΜΑ ΒΑΣΙΚΩΝ ΣΤΟΙΧΕΙΩΝ (pixels) που κάθε ένα χαρακτηρίζεται από έναν αριθμό. Περιγράφουμε όλα τα στοιχεία της εικόνας καθώς και πληροφορίες υποστήριξης. Το pixel είναι σαν ένα συρτάρι που μπορεί να χωρέσει πληροφορία ανάλογη του μεγέθους του. Αυτό δεν σημαίνει πως η πληροφορία θα είναι και σωστή, ακριβής ή με κάποιο νόημα. Επομένως το πόσα pixels έχει τη δυνατότητα να καταγράψει μια συσκευή δεν δείχνει αυτόματα και την ποιότητα του τι καταγράφεται.

Η ψηφιοποίηση γίνεται με την διαδικασία της δειγματοληψίας και του κβαντισμού. Η δειγματοληψία εδώ είναι ένα φαινόμενο στον χώρο (pix/inch, PPI) και όχι στον χρόνο. Ο κβαντισμός καθορίζει τα διακριτά επίπεδα του χαρακτηρισμού.

Μια εικόνα επομένως είναι:

2000 x 1500 pix στα 24bit (sampling, δειγματοληψία, κβαντισμός). Δηλαδή έχει 2000 x 1500 στοιχεία που περιγράφουν το περιεχόμενο της εικόνας και κάθε στοιχείο περιέχει ένα αριθμό των 24bit, δηλαδή μία από περίπου 16 εκατομμύρια διαφορετικές τιμές. Αν η εικόνα είναι έγχρωμη, αυτές οι περίπου 16εκ τιμές  $(2^8)^3 = 16.777.216$  περιγράφουν 256 διαφορετικές αποχρώσεις κάθε ενός από τα τρία βασικά χρώματα RGB.

Αισθητήρας φωτός είναι ένας Μετατροπέας φωτός σε ηλεκτρικό φορτίο. Διαφορετικά είδη αισθητήρων που συναντάμε σε ψηφιακές μηχανές είναι:

CCD ή CMOS

Full Frame CCD: Δίνουν Fill Factor 70%

Interline CCD 30% Fill factor

(No shutter) fast video

CMOS (Complementary Metal Oxon Semiconductor)

All in one (έχει ηλεκτρονικά πάνω στο chip)

Δύο διαδικασίες απαραίτητες (εναλλακτικές)

WHITE BALANCE: Εξίσωση του ΛΕΥΚΟΥ και αποφυγή casts (αποχρώσεων)

SCENE BALANCE: Ολοκλήρωση στο ΓΚΡΙ

ΠΩΣ ΔΟΥΛΕΥΕΙ ΤΟ ΦΙΛΜ:

Υπάρχουν τρία φωτοευαίσθητα LAYERS: RGB: color

Negative:	Προσπίπτον φως	Χρώμα πάνω στο αρνητικό
	G	Clear in M
	R	Clear in C
	B	Clear in Y

Print σε χαρτί με τα ίδια επίπεδα (layers):

G: τρύπα σε M                      C+Y=G

R: τρύπα σε C                        M+Y=R

B: τρύπα σε Y                        C+M=B

Διαφορετικές λύσεις για CCD:

Single CCD. Για να απεικονιστεί μια έγχρωμη εικόνα είναι απαραίτητο να χρησιμοποιηθούν ομάδες από τέσσερα pixels με ειδικό φιλτράρισμα για κάθε ένα ώστε να αποτυπωθούν τα τρία χρώματα (R,G,B). Επομένως η πραγματική ανάλυση είναι το ένα τέταρτο των συνολικών pixels καθώς ανά τέσσερα μας δίνουν ένα έγχρωμο πραγματικό pixel.

G	B
R	G

Το ανθρώπινο μάτι είναι πιο ευαίσθητο στο πράσινο και γι'αυτό επιλέχθηκε να χρησιμοποιούνται δύο Pixels για το πράσινο για κάθε ένα B & R. (Bayer Pattern).

## 8. Αποθήκευση, File types

Εικόνες που δημιουργούνται από προγράμματα επεξεργασίας φωτογραφίας όπως το Photoshop, το Photopaint κ.α., ή από σαρωτές (scanners) είναι χαρτογραφικές και αποτελούνται από "τετραγωνάκια" pixels παραταγμένα το ένα δίπλα στο άλλο, καθέτως και οριζοντίως. Κάθε pixel φέρει τις απαραίτητες πληροφορίες, ανάλογα με το είδος του bitmap, ώστε όλα μαζί να συνθέτουν μια ορθογώνια εικόνα. Τα bitmaps ανάλογα με το βάθος χρώματος (color depth) που υποστηρίζουν μπορεί να είναι 1bit "γραμμικά" (line art ασπρόμαυρα), οπότε κάθε pixel φέρει την πληροφορία άσπρο ή μαύρο. Μπορεί να είναι 8bit grayscale οπότε κάθε pixel φέρει την πληροφορία μιας από τις 256 βαθμίδες της κλίμακας του γκρι από το μαύρο (0) έως το λευκό (255). Μπορεί να είναι 24bit RGB οπότε κάθε pixel φέρει πληροφορίες βαθμίδας (0-255) φωτεινότητας και για τα τρία χρώματα Red, Green, Blue που χρησιμοποιούν η οθόνη του υπολογιστή, οι σαρωτές και το ανθρώπινο μάτι για να αντιλαμβάνονται τα χρώματα. Μπορεί να είναι CMYK, ένα

εξειδικευμένο στην εκτύπωση με μελάνια χρωματικό μοντέλο, οπότε κάθε pixel φέρει πληροφορίες για το ποσοστό που περιέχει από καθένα από τα 4 χρώματα: (C) Cyan, (M) Magenta, (Y) Yellow και (K) black. Μπορεί ακόμα να είναι Lab (CIE L-a-b) όπου κάθε pixel φέρει πληροφορίες για τρεις παραμέτρους: (L) luminance-φωτεινότητα, (a) απόχρωση από green σε red και (b) απόχρωση από blue σε yellow. Όσο μεγαλύτερο βάθος χρώματος υποστηρίζει ένα bitmap τόσο μεγαλύτερο αρχείο δημιουργεί κατά την αποθήκευσή του. Υπάρχουν πολλές μορφές αρχείων για bitmap που εξυπηρετούν διαφορετικούς σκοπούς και βάθη χρώματος και είναι ανεξάρτητες (universal) από προγράμματα και λειτουργικά συστήματα όπως .tif, .jpg, .gif, .bmp, pict, psd, pcd, κα.

### **GIF**

Graphics Interchange Format. Μορφή bitmap που αναπτύχθηκαν από την CompuServe για να καταλαμβάνει μικρό αποθηκευτικό χώρο και να μεταφέρεται γρήγορα μέσω δικτύων. Πολύ χρήσιμο στο internet, μπορεί να συνθέτει εικόνες με 256 και λιγότερα χρώματα. Υποστηρίζει μόνο 256 χρώματα.

### **TIFF**

Μορφή αρχείων Bitmap (Tagged Image File Format) για σχεδιαστικές εφαρμογές, υποστηρίζεται από όλα τα προγράμματα επεξεργασίας εικόνας. Μεταφέρει πληροφορίες RGB, CMYK, και Lab αλλά όχι duotones.

### **BMP**

Τα φυσικά αρχεία εικόνων bitmap για τα Windows. Υποστηρίζονται και από άλλες εφαρμογές. Υποστηρίζουν μέχρι 24 bit.

### **Postscript**

Είναι η γλώσσα "περιγραφής σελίδας" που ανέπτυξε η Adobe και που χρησιμοποιείται για να πραγματοποιηθεί η εκτύπωση σε εκτυπωτές postscript όπως πολλά μοντέλα laser printers ή οι εικονοθέτες λιθογραφικού φιλμ.



## 9 ΣΥΜΠΙΕΣΗ ΔΕΔΟΜΕΝΩΝ

### 9.1 Αποθηκευτικός χώρος

Τα ασυμπίεστα γραφικά χρειάζονται πολύ χώρο για ν' αποθηκευτούν. Υπάρχουν περιπτώσεις όπου ασυμπίεστο βίντεο δεν μπορεί ν' αποθηκευτεί λόγω έλλειψης χώρου σε CD σημερινής τεχνολογίας. Η ύπαρξη συμπίεσης είναι απαραίτητη, γιατί για την μεταφορά δεδομένων ασυμπίεστου βίντεο από άκρη σε άκρη σε κάποιο ψηφιακό δίκτυο χρειάζεται να παρέχεται πολύ υψηλό εύρος. Τα περισσότερα συστήματα πολυμέσων χειρίζονται συμπιεσμένα γραφικά και οι όροι κωδικοποίηση (coding) και συμπίεση (compression) μεταχειρίζονται ως ταυτόσημοι.

Υπάρχουν πολλές τεχνικές συμπίεσης που είναι εν μέρει ανταγωνιστικές και εν μέρει συμπληρωματικές. Οι σημαντικότερες τεχνικές συμπίεσης είναι η JPEG (για απλές φυσικές εικόνες), η MPEG (για βίντεο και ήχο), η DVI (για βίντεο, ήχο και εικόνα) και η H.261 (για βίντεο). Παρακάτω παρουσιάζονται τα πλεονεκτήματα και τα μειονεκτήματα των τεχνικών συμπίεσης. Μέσα από τις κοινές ικανότητες και τις διαφορές τους θα διαπιστωθεί η καταλληλότητα των τεχνικών αυτών στα διάφορα συστήματα πολυμέσων.

### 9.2 Απαιτήσεις κωδικοποίησης

Το βίντεο, ο ήχος (audio) και οι εικόνες (images) έχουν σημαντικά μεγαλύτερες απαιτήσεις αποθήκευσης σε σχέση μ' ένα κείμενο. Δεν είναι μόνο ο τεράστιος αποθηκευτικός χώρος που απαιτείται, αλλά και ο ρυθμός δεδομένων για επικοινωνία συνεχών μέσων είναι επίσης σημαντικός. Από τα παρακάτω παραδείγματα φαίνεται η ποιοτική διαφορά ανάμεσα στη μεταφορά απλού κειμένου και στα γεμάτα κίνηση δεδομένα του βίντεου και γίνεται φανερή η ανάγκη για συμπίεση. Για να συγκρίνουμε τον αποθηκευτικό χώρο και τις απαιτήσεις εύρους ζώνης διάφορων μέσων (κείμενο, γραφικά, εικόνα), τα ακόλουθα στοιχεία βασίζονται σε μια τυπική ανάλυση των 640X480 pixels σε μια οθόνη:

Για την αναπαράσταση κειμένου, δύο bytes χρησιμοποιούνται για κάθε χαρακτήρα. Κάθε χαρακτήρας παρουσιάζεται χρησιμοποιώντας 8X8 pixels, το οποίο είναι ικανοποιητικό για την αναπαράσταση ASCII χαρακτήρων.

Στην αναπαράσταση διανυσματικών γραφικών, μια τυπική ακίνητη εικόνα αποτελείται από 500 γραμμές. Κάθε γραμμή ορίζεται από την οριζόντια θέση, την κάθετη θέση και πεδίο ιδιοτήτων μεγέθους 8 bits. Ο οριζόντιος άξονας αναπαριστάνεται χρησιμοποιώντας 10 bits ( $\log_2(640)$ ) και ο κάθετος άξονας κωδικοποιείται χρησιμοποιώντας 9 bits ( $\log_2(480)$ ).

Σ' ένα απλό τρόπο παρουσίασης ένα pixel ενός χάρτη δυαδικών ψηφίων μπορεί να αναπαρασταθεί από 256 διαφορετικά χρώματα, άρα ένα byte/pixel είναι απαραίτητο.

Τα επόμενα παραδείγματα αναφέρουν λεπτομερώς συνεχή μέσα (continuous media) και ορίζουν την ποσότητα αποθήκευσης που απαιτείται για ένα δευτερόλεπτο ενός playback:

Η ποιότητα ενός ασυμπίεστου σήματος ήχου τηλεφώνου (audio signal of telephony quality) δειγματολαμβάνεται (sampled) στα 8 kHz και κβαντοποιείται (quantized) με 8 bits ανά δείγμα. Αυτό οδηγεί σ' ένα εύρος ζώνης των 64 kbits / sec και μια απαίτηση αποθήκευσης των 64 kbits για αποθήκευση ενός δευτερολέπτου από playback.

Η ποιότητα ενός ασυμπίεστου στερεοφωνικού σήματος ήχου (uncompressed stereo audio signal) ενός CD, δειγματολαμβάνεται με ρυθμό 44.1 kHz και κβαντοποιείται με 16 bits/δείγμα. Όμως οι απαιτήσεις αποθήκευσης είναι  $(44.1 \text{ kHz} \times 16 \text{ bits}) = 705.6 \times 1000 \text{ bits}$  για να αποθηκεύσουν ένα sec of playback και οι απαιτήσεις ρυθμοαπόδοσης (throughput) είναι  $705.6 \times 1000 \text{ bits/second}$ .

Σύμφωνα με το Ευρωπαϊκό PAL, το βίντεο ορίζεται από 625 γραμμές και 25 πλαίσια ανά δευτερόλεπτο. Η φωτεινότητα και τα σήματα των διαφόρων χρωμάτων κωδικοποιούνται ξεχωριστά. Το αποτέλεσμα της ροής ψηφιακών δεδομένων μεταφέρεται χρησιμοποιώντας μια πολύπλοκη τεχνική.

Στην περίπτωση του βίντεο η επεξεργασία ασυμπίεστων ροών δεδομένων σ' ένα ολοκληρωμένο σύστημα πολυμέσων οδηγεί σε δευτερεύουσες απαιτήσεις αποθήκευσης. Οι υπάρχουσες τεχνικές συμπίεσης είναι η JPEG για εικόνα, η MPEG για βίντεο και ήχο και η H.261 για βίντεο conference. Επίσης υπάρχουν τεχνικές για τη συμπίεση της ομιλίας και της μουσικής. Τέλος η τεχνική DVI χρησιμοποιείται για συμπίεση εικόνων και για συνεχή μέσα. Το DVI μπορεί να χρησιμοποιήσει δύο

διαφορετικούς τρόπους για την κωδικοποίηση βίντεο: Βίντεο σε επίπεδο παρουσίασης (presentation-level video-PLV) και βίντεο πραγματικού χρόνου (Real-Time Video-RTV). Η συμπίεση στα συστήματα πολυμέσων είναι ένα θέμα που επιβάλλει περιορισμούς. Η ποιότητα των κωδικοποιημένων και αργότερα των αποκωδικοποιημένων δεδομένων πρέπει να είναι όσο το δυνατόν καλύτερη. Για να γίνει αποτελεσματική οικονομικά η εφαρμογή θα πρέπει η πολυπλοκότητα της τεχνικής συμπίεσης που χρησιμοποιείται να είναι κατά το δυνατόν πολύ μικρή. Η διαδικασία ενός αλγορίθμου δεν θα πρέπει να υπερβαίνει ορισμένα χρονικά όρια.

Για κάθε τεχνική συμπίεσης, υπάρχουν απαιτήσεις που διαφέρουν από τις υπόλοιπες τεχνικές. Μπορούμε να κάνουμε μια διάκριση ανάμεσα στις απαιτήσεις μιας εφαρμογής που τρέχει σε μορφή διαλογική («dialogue mode») και σε μορφή ανάκτησης («retrieval mode»). Η διαλογική μορφή συνίσταται στην αλληλεπίδραση που υπάρχει ανάμεσα σε ανθρώπινους χρήστες και εφαρμογές πολυμέσων, ενώ η μορφή ανάκτησης συνίσταται στην ανάκτηση πολυμεσικών πληροφοριών από μια βάση δεδομένων πολυμέσων. Κάποιες τεχνικές συμπίεσης, όπως η rxc64 είναι περισσότερο κατάλληλη για εφαρμογές μεθόδου διαλόγου (dialogue mode). Άλλες τεχνικές όπως DVI, PLV βελτιστοποιούνται όταν εφαρμόζονται σε εφαρμογές μεθόδου ανάκτησης (retrieval mode).

Σε μια εφαρμογή μεθόδου διαλόγου (dialogue mode application), οι ακόλουθες απαιτήσεις οι οποίες βασίζονται σε ανθρώπινης αντίληψης χαρακτηριστικά πρέπει να ισχύουν:

Η καθυστέρηση απ' άκρη σ' άκρη δεν πρέπει να υπερβαίνει τα 150 milliseconds (για συμπίεση και αποσυμπίεση). Μια καθυστέρηση των 50 milliseconds θα πρέπει να υπολογίζεται στην υποστήριξη εφαρμογών μεθόδου διαλόγου. Ο αριθμός των 50 milliseconds αναφέρεται σε μια καθυστέρηση η οποία εισήχθη μόνο από συμπίεση και αποσυμπίεση. Η συνολική απ' άκρη σ' άκρη καθυστέρηση περιέχει κάθε επιπλέον καθυστέρηση που συμβαίνει στη διάδοση μέσω του δικτύου.

Σε μια εφαρμογή μεθόδου ανάκτησης (retrieval mode application) προκύπτουν οι παρακάτω απαιτήσεις:

Γρήγορη προς τα εμπρός και προς τα πίσω ανάκτηση δεδομένων με ταυτόχρονη εμφάνιση στην οθόνη θα πρέπει να είναι δυνατή. Αυτό το γεγονός συνεπάγεται γρήγορη αναζήτηση πληροφοριών σε βάσεις δεδομένων πολυμέσων.

Τυχαία προσπέλαση σε απλές εικόνες και πλαίσια ήχου μιας ροής δεδομένων θα πρέπει να είναι δυνατή κάνοντας έτσι το χρόνο προσπέλασης μικρότερο από 0,5 δευτερόλεπτα.

Η προσπέλαση αυτή θα πρέπει να είναι γρηγορότερη από ένα συμβατικό ακουστικό σύστημα CD για να υποστηρίζει το χαρακτήρα της εφαρμογής.

Αποσυμπίεση εικόνων, βίντεο ή ήχου θα πρέπει να είναι δυνατή χωρίς δεσμούς σε άλλες ομάδες δεδομένων. Αυτό επιτρέπει την τυχαία προσπέλαση και διόρθωση κειμένου.

Για τις δύο μεθόδους **dialogue** και **retrieval** εφαρμόζονται οι ακόλουθες απαιτήσεις:

Για να υποστηρίξεις κλιμακωτό βίντεο σε διαφορετικά συστήματα, είναι απαραίτητο να ορίσεις ένα μορφότυπο ανεξάρτητο από το μέγεθος του πλαισίου και το ρυθμό του πλαισίου βίντεο.

Πολλές διαβαθμίσεις δεδομένων ήχου και βίντεο θα πρέπει να υποστηρίζονται. Συνήθως αυτό οδηγεί σε διαφορετικές ποιότητες. Γι' αυτό το λόγο οι ρυθμοί δεδομένων πρέπει να προσαρμόζονται στις απαιτήσεις των διαφόρων ειδικών συστημάτων.

Θα πρέπει να είναι δυνατό να συγχρονιστούν δεδομένα ήχου και βίντεο τόσο μεταξύ τους όσο και με άλλα μέσα.

Για να γίνει δυνατή μια οικονομική λύση, η κωδικοποίηση θα πρέπει να πραγματοποιείται χρησιμοποιώντας λογισμικό (για μια φτηνή και χαμηλής ποιότητας λύση) ή VLSI chips (για μια υψηλής ποιότητας λύση).

Θα πρέπει να είναι δυνατή η παραγωγή δεδομένων σε ένα σύστημα πολυμέσων και η αναπαραγωγή αυτών σε κάποιο άλλο σύστημα. Η τεχνική συμπίεσης θα πρέπει να είναι συμβατή.

### **9.3 ΤΕΧΝΙΚΕΣ ΚΩΔΙΚΟΠΟΙΗΣΗΣ (ΠΗΓΗΣ, ΕΝΤΡΟΠΙΑΣ ΚΑΙ ΥΒΡΙΔΙΚΗ)**

Οι τεχνικές συμπίεσης ανήκουν σε διαφορετικές κατηγορίες όπως φαίνεται στον πίνακα 9.1. Στα συστήματα πολυμέσων μπορούμε να διακρίνουμε τρεις τεχνικές: τις κωδικοποιήσεις πηγής και εντροπίας και την υβριδική κωδικοποίηση. Η κωδικοποίηση εντροπίας είναι μια διαδικασία χωρίς απώλειες (lossless process), ενώ η κωδικοποίηση πηγής είναι μια διαδικασία με απώλειες (lossy process). Τα περισσότερα συστήματα

πολυμέσων χρησιμοποιούν την υβριδική κωδικοποίηση η οποία είναι ένας συνδυασμός και των δύο.

Η κωδικοποίηση εντροπίας χρησιμοποιείται ανεξάρτητα από τα ειδικά χαρακτηριστικά των μέσων. Η ροή των δεδομένων που πρόκειται να συμπιεστεί θεωρείται σαν μια απλή ψηφιακή ακολουθία και η σημασιολογία των δεδομένων αγνοείται. Η κωδικοποίηση εντροπίας είναι ένα παράδειγμα κωδικοποίησης χωρίς απώλειες καθώς η διαδικασία αποσυμπίεσης αναδημιουργεί τα δεδομένα πλήρως. Η κωδικοποίηση σειρών μεγάλου μήκους (run-length) είναι ένα παράδειγμα κωδικοποίησης εντροπίας η οποία χρησιμοποιείται για συμπίεση δεδομένων στα συστήματα αρχείων. Η κωδικοποίηση πηγής λαμβάνει υπόψη την σημασιολογία των δεδομένων. Ο βαθμός της συμπίεσης που μπορεί να επιτευχθεί εξαρτάται από το περιεχόμενο των δεδομένων. Στην περίπτωση των τεχνικών συμπίεσης με απώλειες, υπάρχει μία μονόδρομη σχέση ανάμεσα στις αρχικές ροές δεδομένων και στις κωδικοποιημένες ροές δεδομένων. Οι ροές δεδομένων είναι παρόμοιες αλλά όχι ίδιες. Διαφορετικές τεχνικές κωδικοποίησης πηγής κάνουν εκτεταμένη χρήση των χαρακτηριστικών του ειδικού μέσου. Ένα παράδειγμα είναι η κωδικοποίηση πηγής ήχου όπου ο ήχος μετατρέπεται από time-dependent σε frequency-dependent αλληλουχία ήχου, ακολουθώντας την κωδικοποίηση του μορφότυπου. Αυτός ο μετασχηματισμός ουσιαστικά μειώνει το πλήθος των δεδομένων. Τα formants ορίζονται σαν τα μέγιστα στο φάσμα φωνής. Στις περισσότερες περιπτώσεις 3 με 5 formants είναι επαρκή ν' ανακατασκευάσουν το αρχικό σήμα στο πεδίο του χρόνου. Το κυριότερο πρόβλημα είναι η σωστή αναπαραγωγή των μεταβάσεων μεταξύ των επιμέρους μονάδων φωνής στο πεδίο του χρόνου.

Υπάρχουν τεχνικές που διεξάγουν ένα μετασχηματισμό από το πεδίο του χώρου (spacial domain) σε πεδία δύο συχνοτήτων χρησιμοποιώντας το μετασχηματισμό του διακριτού συνημιτόνου (discrete cosine transformation (DCT)). Χαμηλές συχνότητες ορίζουν το μέσο χρώμα και η πληροφορία των υψηλών συχνοτήτων περιέχει the sharp edges. Παρ' όλα αυτά οι χαμηλές συχνότητες είναι πολύ πιο σημαντικές από τις υψηλότερες οι οποίες είναι ένα σημαντικό χαρακτηριστικό που χρησιμοποιείται στην συμπίεση που είναι βασισμένη στο DCT.

Κωδικοποίηση εντροπίας	Κωδικοποίηση Run - Length	
	Κωδικοποίηση Huffman	
	Αριθμητική Κωδικοποίηση	
Κωδικοποίηση Πηγής	Πρόβλεψη	DPCM
		DM
	Μετασχηματισμός	FFT
		DCT
	Στρωματοποιημένη Κωδικοποίηση	Bit Position
		Subsampling
		Sub-Band _Coding
Κβαντοποίηση διανύσματος (Vector Quantization)		
Υβριδική Κωδικοποίηση	JPEG, MPEG, H.261	
DVI,RTV,DVI,PLV		

**Πίνακας 9.1:** Ταξινομήσεις τεχνικών κωδικοποίησης/ συμπίεσης στα συστήματα πολυμέσων

Ο πίνακας 9.1 δείχνει παραδείγματα κωδικοποίησης και τεχνικές συμπίεσης που είναι εφαρμόσιμες σε εφαρμογές πολυμέσων σε σχέση με τις κωδικοποιήσεις εντροπίας,

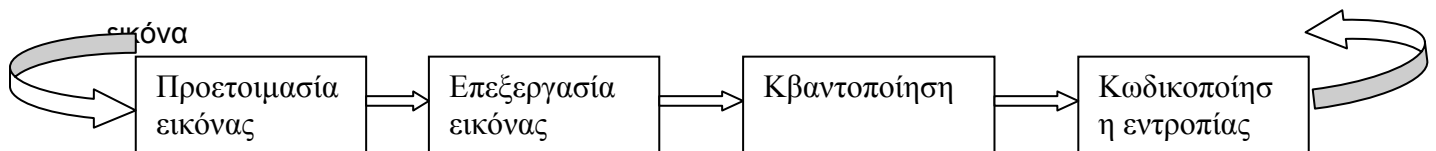
πηγής και υβριδικής. Για την κατανόηση των τριών τεχνικών κωδικοποίησης ορίζονται τα παρακάτω βήματα.

Το σχήμα 9.1 δείχνει μια τυπική ακολουθία από λειτουργίες που διεξάγεται κατά την συμπίεση εικόνων, βίντεο και ήχου.

Ασυμπίεστη  
Συμπιεσμένη  
εικόνα

### Σχήμα 9.1

(Κυριότερα βήματα συμπίεσης δεδομένων)



Τα ακόλουθα τέσσερα βήματα περιγράφουν την συμπίεση μιας εικόνας.

Το στάδιο της προετοιμασίας περιλαμβάνει μετατροπή από αναλογική σε ψηφιακή και παράγει μια κατάλληλη ψηφιακή αναπαράσταση της πληροφορίας. Μια εικόνα διαιρείται σε blocks των 8x8 pixels και αναπαριστάται από ένα σταθερό αριθμό bits/pixel.

Το στάδιο της επεξεργασίας είναι πραγματικά το πρώτο βήμα της διαδικασίας συμπίεσης το οποίο χρησιμοποιεί έμπειρους αλγορίθμους. Ένας μετασχηματισμός από το πεδίο χρόνου στο πεδίο συχνότητας μπορεί να διεξαχθεί χρησιμοποιώντας DCT. Στην περίπτωση συμπίεσης κινούμενου βίντεο η κωδικοποίηση χρησιμοποιεί ένα διάλυμα κίνησης για κάθε block 8x8 pixel.

Η διαδικασία της κβαντοποίησης είναι το αποτέλεσμα του προηγούμενου βήματος. Αυτό ορίζει την granularity της απεικόνισης από πραγματικούς αριθμούς σε ακεραίους. Αυτή η διαδικασία έχει σαν αποτέλεσμα την μείωση της ακρίβειας. Μπορεί επίσης να θεωρηθεί ισοδύναμη του m-low και του a-low το οποίο εφαρμόζεται στα δεδομένα ήχου. Στα μετασχηματισμένα πεδία, οι συντελεστές διακρίνονται σύμφωνα με το νόημα τους.

Για παράδειγμα, μπορεί να γίνει κβαντοποίηση χρησιμοποιώντας διαφορετικό αριθμό bits ανά συντελεστή.

Η κωδικοποίηση εντροπίας είναι συνήθως το τελευταίο βήμα. Αυτό συμπιέζει μια ακολουθιακή ψηφιακή ροή δεδομένων χωρίς απώλειες

Η κωδικοποίηση και η κβαντοποίηση μπορούν να επαναληφθούν πολλές φορές σε αναδρομικές ανακυκλώσεις όπως στην περίπτωση της προσαρμοστικής, διαφορικής, παλμοκωδικής διαμόρφωσης (Adaptive Differential Pulse Code Modulation) ή ADPCM). Το σχήμα 6.1 δείχνει τη διαδικασία συμπίεσης που ακολουθείται για την συμπίεση εικόνας. Η ίδια διαδικασία εφαρμόζεται για την συμπίεση βίντεο και ήχου. Η αποσυμπίεση είναι η αντίστροφη διαδικασία της συμπίεσης.

Οι κωδικοποιητές και οι αποκωδικοποιητές μπορούν να εφαρμοστούν με ποικίλους τρόπους. Στις συμμετρικές εφαρμογές όπως είναι για παράδειγμα οι εφαρμογές διαλόγου έχουμε περίπου το ίδιο κόστος στην κωδικοποίηση και την αποκωδικοποίηση. Στην περίπτωση ασύμμετρων τεχνικών η διαδικασία αποκωδικοποίησης είναι λιγότερο δαπανηρή από την διαδικασία κωδικοποίησης. Αυτό συμβαίνει σε περιπτώσεις που η διαδικασία συμπίεσης συμβαίνει μια φορά και η αποσυμπίεση συμβαίνει συχνά και πρέπει να γίνει γρήγορα. Για παράδειγμα ένα εκπαιδευτικό πρόγραμμα δημιουργείται μια φορά και χρησιμοποιείται από πολλούς μαθητές, άρα αποκωδικοποιείται πολλές φορές. Το συμπέρασμα είναι ότι η κωδικοποίηση δεν χρειάζεται να είναι σε πραγματικό χρόνο, ενώ η αποκωδικοποίηση είναι σε πραγματικό χρόνο. Εκμεταλλευόμαστε αυτή την ασύμμετρη διαδικασία για ν' αυξήσουμε την ποιότητα των εικόνων.

#### **9.4. Μερικές βασικές τεχνικές κωδικοποίησης**

Οι υβριδικές τεχνικές συμπίεσης είναι ένας συνδυασμός των γνωστών αλγορίθμων και τεχνικών συμπίεσης και μπορούν να εφαρμοστούν στα πολυμέσα. Όλες οι υβριδικές τεχνικές που εμφανίζονται στον πίνακα 6.1 χρησιμοποιούν κωδικοποίηση εντροπίας (entropy encoding).

Τα δεδομένα του δείγματος μπορεί να περιέχουν ακολουθίες των ίδιων bytes. Η **run-length** κωδικοποίηση (κωδικοποίηση σειρών μεγάλου μήκους) αντικαθιστά τα επαναλαμβανόμενα bytes με έναν αριθμό που ισούται με το πλήθος εμφάνισης των bytes. Ένα ειδικό flag (σημαία), έστω "!", σηματοδοτεί την διαδικασία κωδικοποίησης.



(Δύο συνεχόμενα !! ερμηνεύονται ως δεδομένα και όχι ως flag). Έτσι αν ένα byte εμφανίζεται τουλάχιστον τέσσερις φορές, σε μια ροή δεδομένων, η ακολουθία αυτή θα αντικατασταθεί στα συμπιεσμένα δεδομένα με το ειδικό flag, το πλήθος των εμφανίσεων και το ίδιο το byte ως εξής :

**ασυμπίεστα δεδομένα** : ABCCCCCCCCCDEFGGG

**run-length κωδικοποίηση** : ABC!8DEFGGG

Με τον τρόπο αυτό επιτυγχάνεται σημαντική μείωση των δεδομένων αφού 4 ως 259 ίδια bytes μπορούν να συμπιεστούν σε 3 bytes μόνο.

Στην περίπτωση της **διανυσματικής κβαντοποίησης** (vector quantization), η ροή των δεδομένων διαιρείται σε ομάδες (blocks) των n bytes ( $n > 0$ ). Ένας προκατασκευασμένος πίνακας περιέχει κάποιο σύνολο προτύπων (patterns). Σε κάθε block αντιστοιχίζεται μία εγγραφή του πίνακα με το πιο ταιριαστό πρότυπο (pattern). Ένας αποκωδικοποιητής χρησιμοποιεί τον ίδιο πίνακα για ν' αναπαράγει κατά προσέγγιση την αρχική ροή δεδομένων. Για παράδειγμα για την συμπίεση κειμένου (text), πρότυπα με μεγάλη συχνότητα εμφάνισης (όπως στις γλώσσες υψηλού επιπέδου τα τερματικά σύμβολα "Begin", "end", "if") μπορούν να αντικαθίστανται από 1 byte.

Μετά από έναν χαρακτήρα διαφυγής , ο οποίος υποδηλώνει ότι έπεται ένα κωδικοποιημένο πρότυπο, ακολουθεί ένας δείκτης προς την αντίστοιχη λέξη.

Μια παρόμοια τεχνική μπορεί να εφαρμοστεί σε εικόνα, video και ήχο. Στην περίπτωση αυτή είναι αποδοτικότερο να αναζητείται για κάθε block ένα προσεγγιστικό και όχι απαραίτητα ακριβές πρότυπο.

Στη διατομική κωδικοποίηση (diatomic encoding) που είναι μια παραλλαγή της κωδικοποίησης σειρών μεγάλου μήκους (run-length) τα συχνότερα εμφανιζόμενα ζεύγη χαρακτήρων (π.χ. στην αγγλική γλώσσα τα 'E ', 'T ', 'TH', ' A', 'S ', 'RE', 'IN' και 'HE') μπορούν να αντικατασταθούν με 1 byte.

Ωστόσο, διαφορετικοί χαρακτήρες θα μπορούσαν να κωδικοποιηθούν με διαφορετικό πλήθος bits, ανάλογα με την συχνότητα εμφάνισής τους. Το αλφάβητο του **Morse** βασίζεται σ' αυτήν την ιδέα. Συχνοεμφανιζόμενοι χαρακτήρες κωδικοποιούνται με μικρά strings, ενώ οι σπανίως εμφανιζόμενοι με μακρύτερα. Οι πιο γνωστές τεχνικές που υιοθετούν παρόμοια στατιστικά μοντέλα είναι η **Huffman** και η **Αριθμητική κωδικοποίηση**.

Δοθέντων των χαρακτήρων που πρόκειται να κωδικοποιηθούν και της αντίστοιχης πιθανότητας εμφάνισής τους, ο αλγόριθμος Huffman προσδιορίζει τον βέλτιστο κώδικα, αυτόν δηλ. που χρησιμοποιεί τον μικρότερο αριθμό από δυαδικά ψηφία. Για την

καλύτερη κατανόηση της κωδικοποίησης Huffman είναι χρήσιμη η κατασκευή ενός δυαδικού δέντρου (βλπ. Σχήμα 5.2).

Τα φύλλα του δέντρου παριστάνουν τους προς κωδικοποίηση χαρακτήρες (A, B, C, D, E). Κάθε κόμβος περιέχει την πιθανότητα εμφάνισης καθενός απ' τους χαρακτήρες που ανήκουν στο αντίστοιχο υπο-δέντρο. Για τους αρχικούς χαρακτήρες οι πιθανότητες είναι :

$$p(A)=0.16, p(B)=0.51, p(C)=0.09, p(D)=0.13, p(E)=0.11$$

Πρώτα συνδυάζονται οι χαρακτήρες με τις μικρότερες πιθανότητες, δηλ. οι C και E. Η πιθανότητα εμφάνισης του ενός ή του άλλου είναι  $p(C) + p(E)=0.20=p(CE)$ . Απομένουν οι κόμβοι :

$$p(A)=0.16, p(B)=0.51, p(CE)=0.20, p(D)=0.13$$

Ομοίως, συνδυάζουμε τους χαρακτήρες A και D με  $p(AD)=0.29$ . Απομένουν οι κόμβοι :

$$p(AD)=0.29, p(B)=0.51, p(CE)=0.20$$

Τις μικρότερες πιθανότητες έχουν οι κόμβοι AD και CE, οι οποίοι συνδυαζόμενοι δίνουν τον κόμβο ADCE με πιθανότητα 0.49. Απομένουν οι κόμβοι :

$$p(ADCE)=0.49, p(B)=0.51$$

οι οποίοι συνδυάζονται στην ρίζα του δέντρου .

Σημειώτεον πως η καταχώρηση του 0 ή 1 σε κάθε ακμή είναι αυθαίρετη. Έτσι τα δεδομένα μπορούν να δώσουν διαφορετικούς Huffman κώδικες. Το αποτέλεσμα είναι ο παρακάτω κώδικας, ο οποίος αποθηκεύεται σε κάποιον πίνακα :

$$w(A)=001, w(B)=1, w(C)=011, w(D)=000, w(E)=010$$

Σε αντίθεση με τον Αλγόριθμο Huffman, η Αριθμητική κωδικοποίηση δεν κωδικοποιεί κάθε χαρακτήρα ανεξάρτητα. Ένα σύμβολο κωδικοποιείται με βάση τα προγενέστερα δεδομένα. Πρακτικά ο μέσος ρυθμός συμπίεσης είναι ο ίδιος και στις δύο μεθόδους.

Επιπλέον υπάρχουν τεχνικές κωδικοποίησης, που μετασχηματίζουν τα δεδομένα (Transformation encoding) σε διαφορετικό πεδίο (domain) απ' αυτό που βρίσκονται. Για παράδειγμα στην **ανάλυση Fourier** τα δεδομένα μετασχηματίζονται από το πεδίο του χρόνου στο πεδίο των συχνοτήτων. Ο πιο αποτελεσματικός μετασχηματισμός για την συμπίεση εικόνας είναι ο **Μετασχηματισμός Διακριτού Συνημιτόνου - Discrete Cosine Transformation (DCT)** αλλά και ο **Fast-Fourier Μετασχηματισμός (FFT)**.

Στην **διαφορική κωδικοποίηση (differential encoding)**, η συμπίεση βασίζεται στον υπολογισμό των διαφορών της τρέχουσας ροής δεδομένων απ' την προηγούμενη ροή. Για τα διάφορα μέσα (media), η τεχνική αυτή εφαρμόζεται ως εξής :

Για τις *εικόνες* υπολογίζονται οι διαφορές μεταξύ γειτονικών pixels ή ομάδων pixels. Μια ομογενής περιοχή, όσον αφορά στην φωτεινότητα και τη χρωμικότητα, χαρακτηρίζεται από έναν μεγάλο αριθμό από 0, τα οποία μπορούν να συμπιεστούν περαιτέρω με την run-length κωδικοποίηση.

Για το *video* η χρήση της διαφορικής κωδικοποίησης στο πεδίο του χρόνου, συμπιέζει τις περιοχές της τρέχουσας εικόνας οι οποίες έχουν μεταβληθεί σε σχέση με την προηγούμενη. Σε εφαρμογές όπως video telephone ή κατά την μετάδοση τηλεοπτικών ειδήσεων, το background παραμένει απaráλλαχτο για μεγάλο χρονικό διάστημα και συνεπώς οι διαφορές μεταξύ διαδοχικών εικόνων είναι μικρές.

Οι τεχνικές για τον *ήχο* συχνά εφαρμόζουν **Διαφορική Παλμοκωδική Διαμόρφωση (DPCM)** σε διαδοχικά δείγματα τα οποία έχουν υποστεί Παλμοκωδική Διαμόρφωση. Δεν είναι απαραίτητο να αποθηκεύονται όλα τα δυαδικά ψηφία για κάθε δείγμα. Αρκεί να παρασταθεί το πρώτο δείγμα πλήρως, και τα επόμενα ως διαφορές από τα προηγούμενα.

Η διαμόρφωση **Δέλτα** είναι μια τροποποίηση της DPCM. Για την κωδικοποίηση των διαφορών χρησιμοποιείται μόνο ένα bit, το οποίο υποδηλώνει τότε το σήμα αυξάνεται ή μειώνεται και συνεπώς το αποτέλεσμα είναι λίγο ασαφές.

Κάποιες τεχνικές συμπίεσης προσαρμόζουν (adapt) μια ειδική μέθοδο κωδικοποίησης σε συγκεκριμένα δεδομένα κατά την διάρκεια της συμπίεσής τους (on the fly). Η πιο γνωστή απ' τις τεχνικές αυτές είναι η **Adaptive DPCM (ADPCM, Προσαρμοσμένη Διαφορική Παλμοκωδική Διαμόρφωση)**, η οποία αποτελεί επέκταση της DPCM.

Στην DPCM οι διαφορές κωδικοποιούνται με λίγα μόνο δυαδικά ψηφία (π.χ. 4). Έτσι, κωδικοποιούνται επακριβώς είτε οι απότομες μεταβολές (δηλ. τα 4 δυαδικά ψηφία αναπαριστούν δυαδικά ψηφία με την μεγαλύτερη σημαντικότητα ) ή οι μικρές αλλαγές (δηλ. τα 4 δυαδικά ψηφία αναπαριστούν δυαδικά ψηφία με την μικρότερη σημαντικότητα). Στην πρώτη περίπτωση ο ήχος με χαμηλή συχνότητα δεν θα είναι ικανοποιητικός ενώ στην δεύτερη θα υπάρχουν απώλειες στις υψηλές συχνότητες

Η ADPCM προσαρμόζει έναν ειδικό συντελεστή αναλογα με την σημαντικότητα μιας συγκεκριμένης ροής δεδομένων. Υψηλή συχνότητα σήματος συνεπάγεται υψηλή τιμή και για τον συντελεστή, και αντιστρόφως. Ο κωδικοποιητής προσδιορίζει μια συγκεκριμένη

τιμή για τον συντελεστή και διαιρεί με αυτόν το DPCM δείγμα και ο αποκωδικοποιητής πολλαπλασιάζει τα συμπιεσμένα δεδομένα με τον ίδιο συντελεστή. Ο συντελεστής μπορεί να είναι το μέγεθος του βήματος των αλλαγών του σήματος .

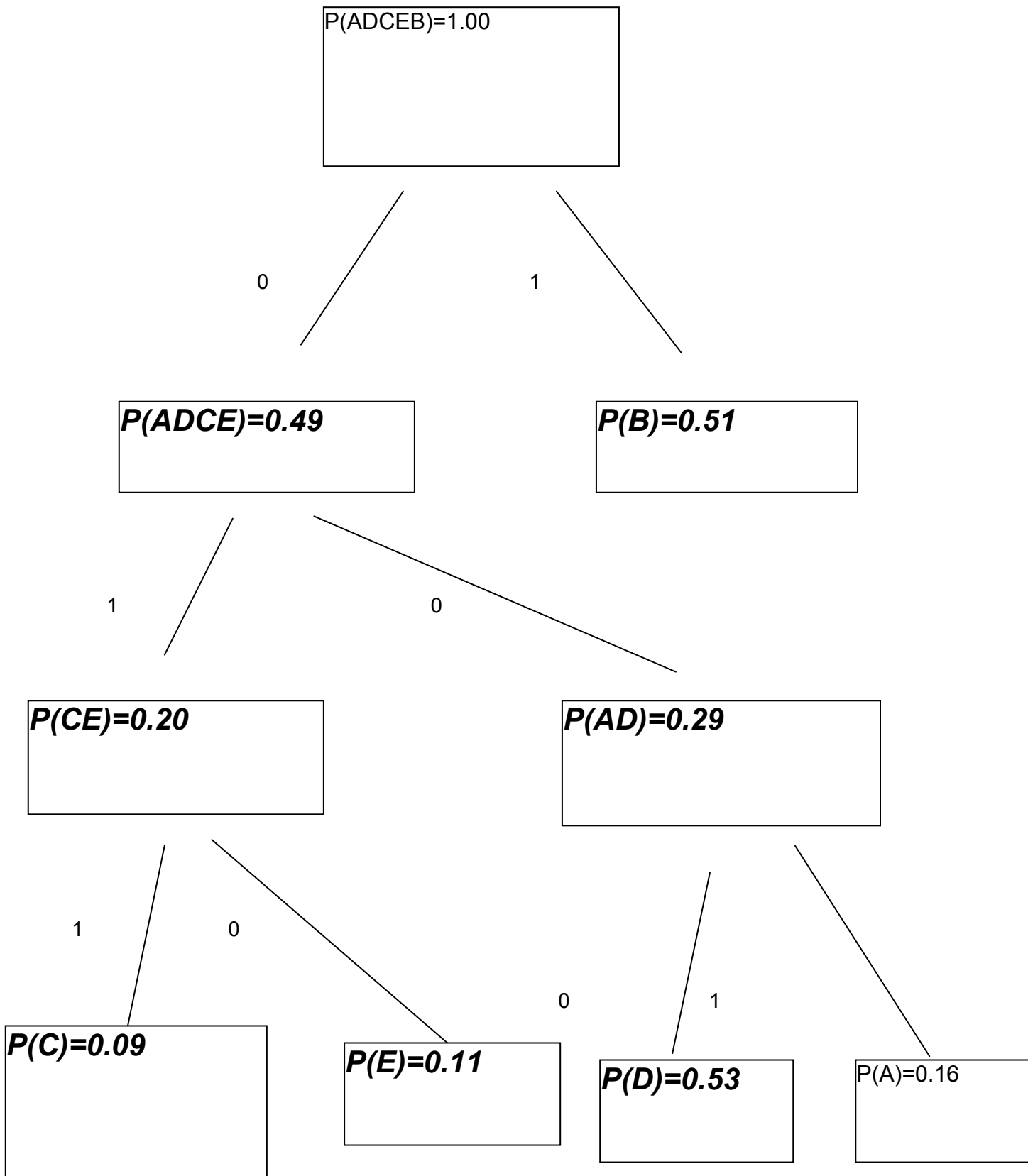
Ένα ηχητικό σήμα με απότομες και συχνές αυξομειώσεις της συχνότητας, δεν είναι κατάλληλο για ADPCM κωδικοποίηση. Η CCITT έχει θεσπίσει για την ADPCM στην σειρά G.700 την χρησιμοποίηση 32Kbps για τηλεφωνικές εφαρμογές δηλ. αναπαράσταση του κάθε δείγματος με 4 δυαδικά ψηφία και ρυθμό δειγματοληψίας 8KHz.

Τέλος εκτός από τις προαναφερθείσες τεχνικές συμπίεσης, υπάρχουν και άλλες ευρέως χρησιμοποιούμενες στις μέρες μας. Για παράδειγμα η **CLUT (Color Look-Up Tables)** πετυχαίνει μεγάλη μείωση δεδομένων στην συμπίεση video. Η τεχνική αυτή εφαρμόζεται κυρίως σε κατανεμημένα πολυμεσικά συστήματα.

Στα πλαίσια του ISO υπάρχουν τα εξής group που ασχολούνται με την κωδικοποίηση Video και ήχου :

**JPEG (Joint Photographic Experts Group)**, σχετικό με αλγορίθμους συμπίεσης σταθερών (still) εικόνων.

**MPEG (Moving Photographic Experts Group)**, για την συμπίεση Video.



**Σχήμα 9.2 :** Παράδειγμα κωδικοποίησης Huffman

## **9.5 ΤΟ ΠΡΟΤΥΠΟ JPEG**

Η μέθοδος συμπίεσης JPEG (Joint Photographic Experts Group) είναι ιδανική για φωτογραφίες και για άλλες εικόνες «φυσικού» χρώματος. Οι εικόνες που έχουν συμπειστεί με την μέθοδο JPEG μπορούν να έχουν εκατομμύρια χρώματα, και το μέγεθος αρχείου εξαρτάται κυρίως από το μέγεθος της εικόνας και όχι από τον αριθμό των χρωμάτων. Το JPEG έγινε πρότυπο του ISO το 1992. Το πρότυπο αυτό εφαρμόζεται σε έγχρωμες και ασπρόμαυρες εικόνες.

Το πρότυπο εκπληρώνει τις παρακάτω απαιτήσεις για να εγγυάται απώτερη διανομή και εφαρμογή:

Η εκτέλεση JPEG πρέπει να είναι ανεξάρτητη από το μέγεθος της εικόνας.

Η εκτέλεση JPEG πρέπει να εφαρμόζεται σ' όλες τις εικόνες και λόγο διαστάσεων (aspect ratio) pixels.

Η αναπαράσταση των χρωμάτων πρέπει να είναι ανεξάρτητη από την ειδική εκτέλεση.

Το περιεχόμενο της εικόνας μπορεί να είναι σύνθετο, με οποιαδήποτε στατιστικά χαρακτηριστικά.

Ο ορισμός του προτύπου JPEG πρέπει να είναι βέλτιστος σχετικά με τον παράγοντα συμπίεσης και την ποιότητα της εικόνας που επιτυγχάνεται.

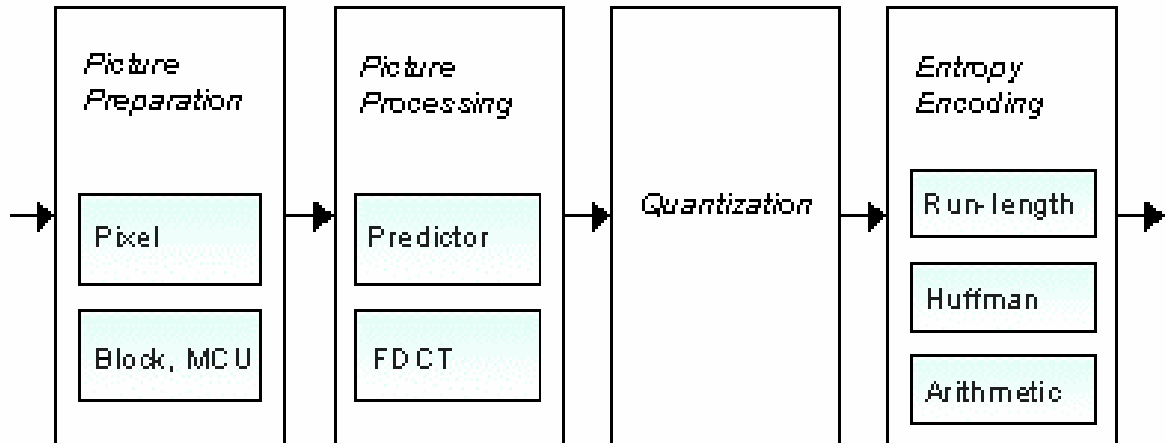
Η πολυπλοκότητα επεξεργασίας πρέπει να επιτρέπει μια λύση λογισμικού (software solution) να εκτελεστεί σε όσο το δυνατό περισσότερους διαθέσιμους επεξεργαστές. Επιπρόσθετα, η χρήση ειδικού hardware θα έπρεπε να αυξάνει την ποιότητα εικόνας. Η σειριακή αποκωδικοποίηση (γραμμή προς γραμμή) και προοδευτική (progressive) αποκωδικοποίηση θα έπρεπε να είναι δυνατές.

Ο χρήστης μπορεί να επιλέξει την ποιότητα εικόνας που αναπαράγεται, τον χρόνο εκτέλεσης συμπίεσης και το μέγεθος της συμπεισμένης εικόνας επιλέγοντας κατάλληλες ατομικές παραμέτρους.

Οι εφαρμογές δεν χρειάζεται να συμπεριλαμβάνουν και τον κωδικοποιητή και τον αποκωδικοποιητή, στις περισσότερες εφαρμογές μόνο το ένα χρειάζεται. Η κωδικοποιημένη ροή πληροφορίας έχει μια σταθερή μορφοποίηση ανταλλαγής δεδομένων (interchange format) που περιλαμβάνει κωδικοποιημένα δεδομένα εικόνας καθώς επίσης και επιλεγμένες παραμέτρους και πίνακες της διαδικασίας

κωδικοποίησης. Η μορφοποίηση αυτή σε κανονική μορφή περιλαμβάνει όλη την απαραίτητη πληροφορία που χρειάζεται για αποκωδικοποιήσει χωρίς καμία προηγούμενη γνώση της διαδικασίας κωδικοποίησης.

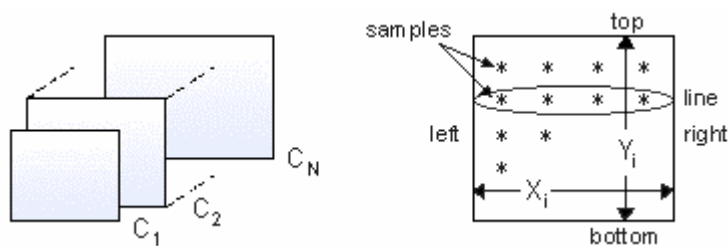
Το σχήμα 93. δείχνει τα βήματα της κωδικοποίησης JPEG.



**Σχήμα 93:** Τα βήματα της κωδικοποίησης JPEG

### 9.5.1 ΠΡΟΕΤΟΙΜΑΣΙΑ ΕΙΚΟΝΑΣ (IMAGE PREPARATION)

Για το πρώτο βήμα της προετοιμασίας της εικόνας, το JPEG ορίζει ένα πολύ γενικό μοντέλο εικόνας. Με αυτό το μοντέλο είναι δυνατό να περιγραφούν οι περισσότερες γνωστές αναπαραστάσεις δισδιάστατων εικόνων.

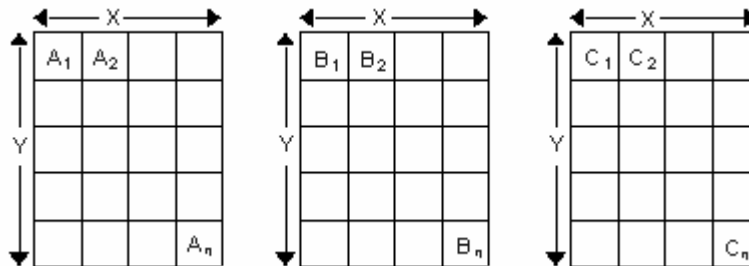


**Σχήμα 9.4**

Η εικόνα πηγής αποτελείται τουλάχιστον από 1 και το πολύ από 255 επίπεδα ή συνιστώσες (σχήμα 9.4). Κάθε συνιστώσα  $C_i$  μπορεί να έχει διαφορετικό αριθμό pixels στον οριζόντιο ( $X_i$ ) και στον κάθετο ( $Y_i$ ) άξονα. (ο δείκτης υποδηλώνει τον

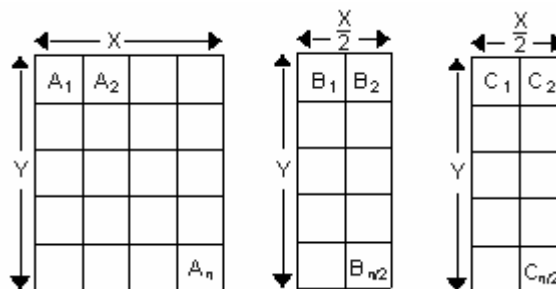
αριθμό επιπέδου). Σ'αυτές τις συνιστώσες μπορεί να εκχωρούνται τα τρία χρώματα RGB, YIQ ή σήματα YUV.

Η ανάλυση των ατομικών συνιστωσών μπορεί να είναι ίδια ή διαφορετική. Το σχήμα 9.5 παρουσιάζει την πρώτη περίπτωση και το σχήμα 9.6 την δεύτερη.



**Σχήμα 9.5:**

Σχήμα 9.5: Οι τρεις συνιστώσες μιας εικόνας, κάθε μια έχει την ίδια ανάλυση και έναν ορθογώνιο πίνακα  $C_i$  μεγέθους  $X_i \times Y_i$  pixels. Οι τρεις τιμές  $X_i$  και  $Y_i$  είναι ίδιες.



**Σχήμα 9.6:**

μια εικόνα με μισό αριθμό στηλών στο 2<sup>ο</sup> και 3<sup>ο</sup> επίπεδο σε σχέση με το 1<sup>ο</sup> επίπεδο:

$$Y_1 = Y_2 = Y_3 \text{ και } X_1 = 2X_2 = 2X_3$$

Μια ασπρόμαυρη (grayscale) εικόνα αποτελείται συνήθως από μια μόνο συνιστώσα ενώ μια έγχρωμη παράσταση RGB αποτελείται από τρεις συνιστώσες με την ίδια ανάλυση (ίδιος αριθμός γραμμών  $Y_1 = Y_2 = Y_3$  και ίδιος αριθμός στηλών  $X_1 = X_2 = X_3$ ).

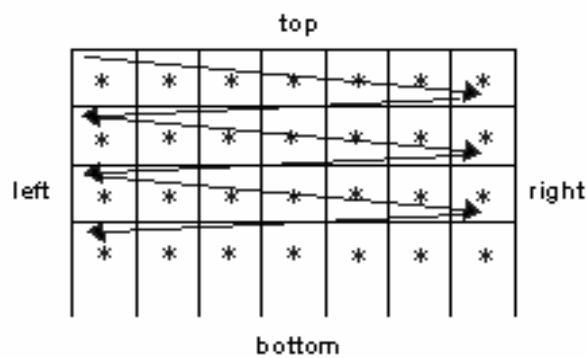
Κάθε pixel αναπαριστάνεται με  $p$  δυαδικά ψηφία που παίρνουν τιμές από 0 έως  $2^p - 1$ . Όλα τα εικονοστοιχεία των συνιστωσών μιας εικόνας κωδικοποιούνται με τον ίδιο αριθμό δυαδικών ψηφίων.

Οι διαστάσεις της συμπιεσμένης εικόνας ορίζονται με νέες τιμές  $X$  (μέγιστη  $X_i$ ),  $Y$  (μέγιστη  $Y_i$ ),  $H_i$  και  $V_i$ . Τα  $H_i$  και  $V_i$  είναι οι οριζόντιοι και κάθετοι λόγοι



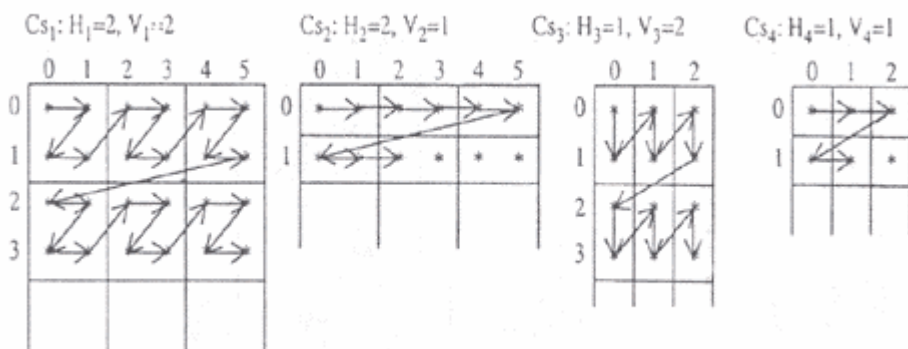
δειγματοληψίας (sampling ratios) που προσδιορίζονται για κάθε συνιστώσα  $i$ . Τα  $H_i$  και  $V_i$  παίρνουν τιμές από 1 έως 4.

Για τη συμπίεση, η εικόνα διαρείται σε μονάδες δεδομένων. Ο τρόπος συμπίεσης χωρίς απώλειες χρησιμοποιεί ένα δυαδικό ψηφίο σαν μονάδα δεδομένων, ενώ ο τρόπος συμπίεσης με απώλειες χρησιμοποιεί blocks των 8 δυαδικών ψηφίων. Οι μονάδες δεδομένων επεξεργάζονται συνιστώσα ανα συνιστώσα, όπως φαίνεται στο σχήμα 5.7. Η σειρά επεξεργασίας των μονάδων δεδομένων μιας συνιστώσας είναι από αριστερά προς τα δεξιά και από πάνω προς τα κάτω. Αυτός ο τρόπος ονομάζεται σειρά μη παρεμβαλλόμενων δεδομένων (**non-interleaved data ordering**).



σχήμα 5.7

Αν χρησιμοποιείται ο **non-interleaved data ordering** για μια εικόνα RGB υψηλής ανάλυσης, η οθόνη θα δείξει αρχικά μόνο την κόκκινη συνιστώσα, μετά την μπλε και τελικά την πράσινη με αποτέλεσμα να ανασχηματίζονται τα αρχικά χρώματα της εικόνας. Λόγω της περιορισμένης ταχύτητας επεξεργασίας του κωδικοποιητή JPEG, είναι πολλές φορές προτιμότερο να παρεμβάλλουμε (**interleave**) τις μονάδες δεδομένων (data units) σχήμα 9.8).



**Σχήμα 9.8:** Μονάδες παρεμβαλλόμενων δεδομένων : παράδειγμα με 4 συστατικά όπως ορίζονται από το πρότυπο JPEG.

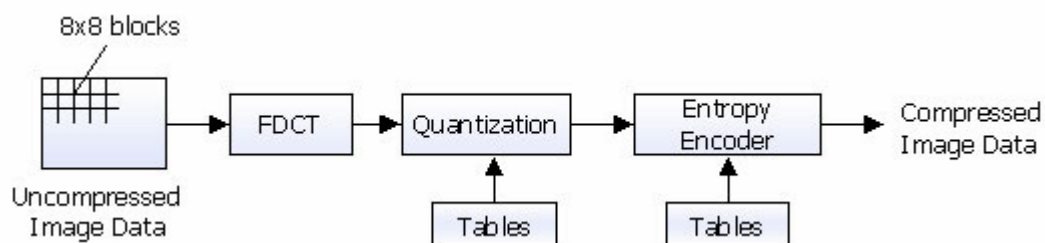
Οι **interleaved** μονάδες δεδομένων των διαφόρων συνιστωσών συνδυάζονται σε MCUs (Minimum Coded Units). Αν όλες οι συνιστώσες έχουν την ίδια ανάλυση ( $X_i \times Y_i$ ), το MCU αποτελείται ακριβώς από μια μονάδα δεδομένων για κάθε συνιστώσα. Σε περίπτωση που οι συνιστώσες έχουν διαφορετικές αναλύσεις, η κατασκευή των MCU γίνεται πιο πολύπλοκη (σχήμα 9.8). Για κάθε συνιστώσα προσδιορίζονται περιοχές μονάδων δεδομένων. Κάθε συνιστώσα αποτελείται από τον ίδιο αριθμό περιοχών (πχ στο σχήμα 9.8 κάθε συνιστώσα αποτελείται από 6 περιοχές).

Το σχήμα 9.8 δείχνει ένα παράδειγμα με τέσσερις συνιστώσες. Η πρώτη συνιστώσα έχει την μεγαλύτερη ανάλυση και στις δυο διαστάσεις και η τέταρτη συνιστώσα την μικρότερη. Τα βέλη δείχνουν την κατεύθυνση της δειγματοληψίας των μονάδων δεδομένων για κάθε συνιστώσα. Τα MCU κατασκευάζονται με την παρακάτω σειρά:

### 9.5.2 LOSSY SEQUENTIAL DCT-BASED MODE

Μετά την προετοιμασία της εικόνας, τα μη συμπιεσμένα δείγματα εικόνας ομαδοποιούνται σε μονάδες των  $8 \times 8$  εικονοστοιχείων και περνιούνται στον κωδικοποιητή. Η σειρά των μονάδων δεδομένων ορίζεται από τα MCU.

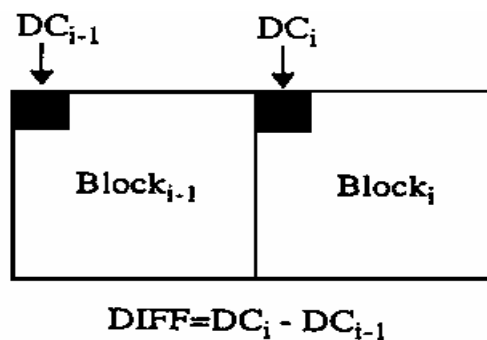
Στο baseline mode (baseline process), το πρώτο βήμα στην επεξεργασία εικόνας είναι ο μετασχηματισμός διακριτού συνημιτόνου (Discrete Cosine Transformation), όπως δείχνει το σχήμα 9.9. Οι τιμές των εικονοστοιχείων μετατοπίζονται (shifted) στη περιοχή  $[-128..127]$  με κέντρο το μηδέν. Αυτές οι νέες τιμές μονάδων δεδομένων ορίζονται με  $S_{yx}$  όπου x και y παίρνουν τιμές από 1 έως 7.



**Σχήμα 9.9**

## 9. 6 Κβαντοποίηση (Quantization)

Η εφαρμογή JPEG παρέχει πίνακες (tables) με 64 εισόδους. Κάθε είσοδος χρησιμοποιείται για την κβαντοποίηση ενός από τους 64 DCT– συντελεστές. Υπάρχει η πιθανότητα να επηρεαστεί η σχετική σπουδαιότητα (relative significance) διαφορετικών συντελεστών και μάλιστα σε συγκεκριμένες συχνότητες μπορεί να δοθεί μεγαλύτερη σημασία από άλλες. Αυτοί οι συντελεστές θα πρέπει να προσδιοριστούν με βάση τα χαρακτηριστικά της αρχικής (source) εικόνας. Η κατάλληλη συμπίεση επιδρά στην έξοδο της επιτυγχανόμενης ποιότητας της εικόνας.



**Σχήμα 9.10:** Προπαρασκευή των DCT DC-συντελεστών για την κωδικοποίηση εντροπίας. Περιλαμβάνει τον υπολογισμό της διαφοράς μεταξύ γειτονικών τιμών.

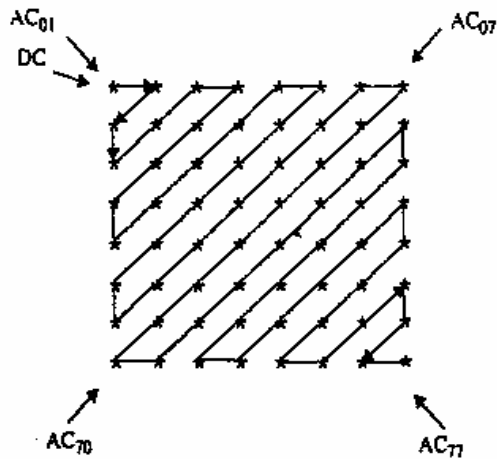
Κάθε είσοδος του πίνακα είναι μια ακέραιη τιμή 8 bit που ονομάζεται  $Q_{vu}$ . Η διαδικασία της κβαντοποίησης γίνεται λιγότερο ακριβής καθώς το μέγεθος των εισόδων του πίνακα αυξάνεται. Η κβαντοποίηση και η απο-κβαντοποίηση (de-Quantization) πρέπει να χρησιμοποιούν τους ίδιους πίνακες.

## 9. 7 Κωδικοποίηση βασισμένη στην εντροπία (Entropy Encoding)

Κατά τη διάρκεια του αρχικού βήματος της κωδικοποίησης της εντροπίας, οι κβαντοποιημένοι DC–συντελεστές αναπτύσσονται χωριστά από τους AC – συντελεστές .

- Οι DC–συντελεστές προσδιορίζουν το βασικό χρώμα των μονάδων δεδομένων (data units). Επειδή μεταξύ γειτονικών μονάδων δεδομένων η απόκλιση από το χρώμα είναι αρκετά μικρή ένας DC– συντελεστής είναι κωδικοποιημένος ως η διαφορά μεταξύ του τρέχοντος DC–συντελεστή και του προηγούμενου.

- Η DCT εξελικτική διαδικασία (processing order) διευκρινίζει πως οι συντελεστές με χαμηλότερες συχνότητες (τυπικά με υψηλότερες τιμές )



κωδικοποιούνται πρώτοι, ακολουθούμενοι από αυτούς με υψηλές συχνότητες (με τυπικά μικρή, σχεδόν μηδενική τιμή). Το αποτέλεσμα είναι μια εκτεταμένη σειρά από πανομοιότυπα bytes δεδομένων , επιτρέποντας πολύ αποδοτική κωδικοποίηση εντροπίας.

**Σχήμα 9.11:** Προπαρασκευή των DCT AC-συντελεστών για την κωδικοποίηση εντροπίας με ζικ-ζακ αυξανόμενη συχνότητα

Το JPEG ορίζει την Huffman και την Αριθμητική κωδικοποίηση ως μεθόδους Κωδικοποίησης βασισμένες στην Εντροπία (Entropy Encoding). Για την DCT μέθοδο με απώλειες (Lossy Sequential DCT-based mode) μόνο η Huffman κωδικοποίηση επιτρέπεται. Πρώτα εφαρμόζεται κωδικοποίηση μεγάλων σειρών (run length encoding) των μηδενικών τιμών. Έπειτα, μη μηδενικοί AC-συντελεστές (σχήμα 9.11), καθώς επίσης και DC-συντελεστές, μετασχηματίζονται σε μια φασματική αναπαράσταση (spectral representation) για να συμπίεσουν τα δεδομένα περισσότερο. Ο αριθμός των απαιτούμενων δυαδικών ψηφίων εξαρτάται από την τιμή του συντελεστή. 1 μέχρι 10 δυαδικά ψηφία για την αναπαράσταση ενός μη μηδενικού AC-συντελεστή και 1 ως 11 για έναν DC. Το αποτέλεσμα είναι μια αναπαράσταση σύμφωνα με το μορφότυπο *ISS* (Intermediate Symbol Sequence).

Το σημαντικότερο πλεονέκτημα της Huffman κωδικοποίησης έναντι της Αριθμητικής είναι η ελεύθερη υλοποίηση (free implementation) (σχήμα 9.12), γιατί δεν προστατεύεται με ευρεσιτεχνία.

Μειονέκτημα είναι το γεγονός ότι η εφαρμογή πρέπει να εξασφαλίζει κωδικοποιημένους πίνακες αφού το JPEG δεν ορίζει κανένα από αυτά.



**Σχήμα 9.12** Παρουσίαση ακολουθιακής εικόνας που χρησιμοποιεί την DCT μέθοδο με απώλειες

### **9.8 ΡΟΗ ΔΕΔΟΜΕΝΩΝ (ΡΟΗ ΗΧΟΥ)**

Το πρότυπο MPEG καθορίζει μια σύνταξη για ροή δεδομένων ήχου και εικόνας που παρεμβάλλονται. Μια ροή δεδομένων ήχου αποτελείται από πλαίσια τα οποία διαιρούνται σε ηχητικές μονάδες πρόσβασης. Κάθε ηχητική μονάδα πρόσβασης συντίθεται από σχισμές (slots) τεσσάρων ή ενός byte κάθε μια. Το πλαίσιο αποτελείται από ένα προκαθορισμένο αριθμό ηχητικών δειγμάτων. Το σημαντικότερο είναι η *ηχητική μονάδα πρόσβασης*, που είναι η μικρότερη δυνατή ακολουθία συμπιεσμένων δεδομένων που μπορεί να αποκωδικοποιηθεί ανεξάρτητα από άλλα δεδομένα. Οι ηχητικές μονάδες πρόσβασης οδηγούν σε ένα χρόνο παιχνιδιού ήχου 8 milliseconds σε 48kHz ή 8.7 milliseconds σε 44.1kHz και 12 milliseconds σε 32kHz ποιότητα. Στην περίπτωση στερεοφωνικού σήματος, δεδομένα και από τα δυο κανάλια συγχωνεύονται σε ένα πλαίσιο.

### **9.9 Mp3**

το MP3 : είναι ένας τρόπος κωδικοποίησης (ακριβέστερα συμπίεσης) της μουσικής έτσι ώστε να καταλαμβάνει περίπου το 1/10 το χώρου που καταλαμβάνει σήμερα ένα μουσικό κομμάτι χωρίς αντιληπτή διαφορά στην ποιότητά του. Αυτό έχει σαν αποτέλεσμα να 10-πλασιάζεται η χωρητικότητα ενός μουσικού CD και από 74-80 λεπτά που είναι σήμερα να γίνεται 740-800 λεπτά (δηλ. περίπου 200 κομμάτια σε ένα και μόνο CD !). Η κωδικοποίηση (συμπίεση) κατά MP3. Αναλύει την μουσική κάθε φορά, βρίσκει ποιοι είναι οι ισχυροί και ποιοι η ασθενείς ήχοι και παράγει ένα αρχείο το οποίο αποτελείται μόνον από τους ισχυρούς, δηλ. μόνον από τους ήχους εκείνους που ακούγονται. Είναι φανερό ότι ο τρόπος αυτός κωδικοποίησης είναι απωλεστικός" δηλαδή το παραγόμενο (συμπιεσμένο) μουσικό κομμάτι έχει λιγότερη (δηλ. χειρότερη) πληροφορία από το αρχικό (ασυμπιεστο). Το πόση απώλεια πληροφορίας θα έχουμε μεταξύ αρχικού (ασυμπιεστού) και τελικού (συμπιεσμένου) μουσικού κομματιού εξαρτάται από τις παραμέτρους κωδικοποίησης : Όσο περισσότερο συμπιέσουμε την μουσική μας, τόσο μικρότερο αρχείο θα παράγουμε αλλά και τόσο χειρότερη θα είναι η ποιότητα του συμπιεσμένου κομματιού. Συγκεκριμένα ακολουθείται το πρότυπο κωδικοποίησης εντροπίας Huffman.

## 9.10 ΡΟΗ ΚΙΝΟΥΜΕΝΗΣ ΕΙΚΟΝΑΣ (VIDEO)

Μια ροή δεδομένων κινούμενης εικόνας περιλαμβάνει 6 επίπεδα:

Στο υψηλότερο επίπεδο, το *επίπεδο συχνότητας*, γίνεται διαχείριση ενταμίευσης δεδομένων. Μια ροή δεδομένων θα πρέπει να έχει χαμηλές απαιτήσεις σε αποθηκευτικό χώρο. Για τον λόγο αυτό, στην έναρξη του *επιπέδου ακολουθίας*, υπάρχουν 2 εγγραφές: η συχνότητα με την οποία έρχονται τα δεδομένα και η ανάγκη που υπάρχει σε αποθηκευτικό χώρο για την αποκωδικοποίηση τους. Στην σχήμα επεξεργασίας, ένας *επαληθευτής ενταμίευσης video* χρησιμοποιείται μετά την κβαντοποίηση. Ο ρυθμός ροής δεδομένων που προκύπτει χρησιμοποιείται για να επαληθεύσει την καθυστέρηση που προκαλείται από την αποκωδικοποίηση των δεδομένων. Ο επαληθευτής ενταμίευσης video επιδρά στην διαδικασία κβαντοποίησης αποτελεί ένα είδος βρόγχου ελέγχου. Διαδοχικές ακολουθίες μπορούν να έχουν μεταβλητό ρυθμό δεδομένων. Κατά την αποκωδικοποίηση διαδοχικών ακολουθιών δεν υπάρχει άμεση συσχέτιση μεταξύ του τέλους μιας ακολουθίας και της αρχής της επόμενης. Η βασικές παράμετροι του αποκωδικοποιητή επανεξετάζονται και εκτελείται διαδικασία επανεκίνησης του αποκωδικοποιητή με κάποιες αρχικές τιμές.

Το επόμενο επίπεδο είναι αυτό των *ομάδων εικόνων*. Το επίπεδο αυτό αποτελείται, κατ' ελάχιστο, από ένα I-πλαίσιο (frame), το οποίο αποτελεί και το πρώτο πλαίσιο του video. Τυχαία προσπέλαση σε αυτή την εικόνα είναι πάντοτε επιτρεπτή. Σ' αυτό το επίπεδο είναι δυνατό να ξεχωρίσει κανείς την σειρά των εικόνων σε μια ροή δεδομένων καθώς και την σειρά τους κατά την διαδικασία επεικόνισης. Η πρώτη εικόνα μίας ροής δεδομένων θα πρέπει να είναι ένα I-πλαίσιο. Για τον λόγο αυτό, ο αποκωδικοποιητής αποκωδικοποιεί και αποθηκεύει πρώτα το πλαίσιο αναφοράς. Κατά την σειρά εμφάνισεως μπορεί να προκύψει ένα πλαίσιο τύπου B πριν από ένα πλαίσιο τύπου I.

Σειρά εμφάνισης πλαισίων:

Τύπος πλαισίου		B	B	I	B	B	P	B	B	P	B	B	P
Αριθμός πλαισίου	0	1	2	3	4	5	6	7	8	9	10	11	

Σειρά αποκωδικοποίησης πλαισίων:

Τύπος πλαισίου	I	B	B	P	B	B	P	B	B	P	B	B
Αριθμός πλαισίου	2	0	1	5	3	4	8	6	7	11	9	10

Το *επίπεδο εικόνας* περιέχει μια ολόκληρη εικόνα. Η χρονική αναφορά καθορίζεται από ένα αριθμό εικόνων. Αξιοσημείωτο είναι ότι υπάρχουν πεδία που καθορίζονται σε αυτό το επίπεδο τα οποία δεν χρησιμοποιούνται άμεσα στο πρότυπο συμπίεσης MPEG καθώς είναι σχεδιασμένα για μελλοντικές επεκτάσεις.

Το επόμενο επίπεδο είναι το *επίπεδο φέτας* (slice). Κάθε τεμάχιο αποτελείται από ένα αριθμό μακροτμημάτων που μπορεί να ποικίλλουν από την μια εικόνα στην άλλη. Επιπροσθέτως, καθορίζεται η DCT κβαντοποίηση κάθε μακροτμήματος.

Το πέμπτο επίπεδο είναι το *επίπεδο μακροτεμαχίου* (macro block). Περιλαμβάνει το άθροισμα των χαρακτηριστικών κάθε μακροτεμαχίου, όπως περιγράφηκε ανωτέρω.

Το κατώτερο επίπεδο είναι το *επίπεδο τεμαχίου* που περιγράφηκε πιο πάνω.

Μια ακόμα λειτουργία του προτύπου MPEG είναι ότι καθορίζει τον συνδυασμό πολλών ροών δεδομένων σε μία. Η ίδια ιδέα επιδιώχτηκε να χρησιμοποιηθεί και στο πρότυπο DVI για να καθορισθεί το μορφότυπο ήχου/εικόνας AVSS (Audio/Video Support System). Το πιο σημαντικό στοιχείο αυτής της διεργασίας είναι η πολύπλεξη, η οποία περιλαμβάνει συντονισμό ροών δεδομένων εισόδου και εξόδου, συντονισμό ρολογιών συγχρονισμού και διαχείριση ενταμιευτών.

Γι αυτό τον λόγο η ροή δεδομένων που καθορίζεται από το ISO 11172 διαιρείται σε *πακέτα* (racks). Ο αποκωδικοποιητής παίρνει την πληροφορία που χρειάζεται για δέσμευση πόρων από αυτήν την πολυπλεγμένη ροή δεδομένων. Ο μέγιστος ρυθμός ροής δεδομένων περιγράφεται στην αρχή του πρώτου πακέτου κάθε ISO 11172 ροής δεδομένων.

Ο ορισμός αυτός της ροής δεδομένων υπονοεί ότι για τα δεδομένα που αποθηκεύονται σε δευτερεύοντα/βοηθητικά μέσα είναι πιθανό να διαβαστεί πρώτα μια τέτοια επικεφαλίδα (εάν κριθεί απαραίτητο, με τυχαία προσπέλαση). Σε



περιπτώσεις αμφίδρομης επικοινωνίας που χρησιμοποιούνται δίκτυα επικοινωνιών, όπως το τηλέφωνο ή οι τηλεδιασκέψεις, ο χρήστης λαμβάνει πάντα πρώτα την πληροφορία που του παρέχει η επικεφαλίδα. Σε μια τηλεδιάσκεψη, η χρησιμοποίηση ροής δεδομένων τύπου MPEG μπορεί να μην είναι η κατάλληλη γιατί ένας νέος χρήστης μπορεί να θέλει να συμμετάσχει σε μια υπάρχουσα τηλεδιάσκεψη μετά το πέρας της αρχικοποίησης των ροών δεδομένων. Έτσι, δεν θα μπορέσει να πάρει τις πληροφορίες της επικεφαλίδας γιατί δεν θα του είναι διαθέσιμες.

Για μια ροή δεδομένων δημιουργημένη σύμφωνα με το ISO 11172, το πρότυπο MPEG, παρέχει χρονοσφραγίδες που είναι απαραίτητες για συγχρονισμό. Οι χρονοσφραγίδες αυτές αναφέρονται στη συσχέτιση μεταξύ των πολυπλεγμένων ροών δεδομένων μόνο και όχι σε άλλες υπάρχουσες ροές δεδομένων του ISO 11172.

Τέλος, θα πρέπει να αναφερθεί ότι το πρότυπο MPEG δεν ορίζει κωδικοποίηση σε πραγματικό χρόνο. Το MPEG καθορίζει την διαδικασία αποκωδικοποίησης αλλά όχι τον ίδιο τον αποκωδικοποιητή.

## **9.11 MPEG-2**

Η ποιότητα μιας κινούμενης εικόνας (video) σύμφωνα με το πρότυπο MPEG είναι κοντά στο μέγιστο ρυθμό ροής δεδομένων του 1.5 Mbits/sec. Αυτό το βέλτιστο μέγεθος έχει να κάνει με την ποιότητα και όχι με την απόδοση. Περαιτέρω εξελίξεις στον τομέα της κωδικοποίησης video, βασίζονται σ' ένα ρυθμό μετάδοσης δεδομένων της τάξης των 40 Mbits/sec. (το γνωστό MPEG-2).

Το MPEG –2 έχει σαν στόχο μια υψηλότερη ανάλυση εικόνας ανάλογη του προτύπου CCIR 601 και οδεύει προς μια ακόμα καλύτερη ποιότητα απαραίτητη για τη HDTV. Αξίζει να επισημανθεί ότι οι πληροφορίες που αναφέρονται στα πρότυπα MPEG-2 και MPEG-4 περισυλλέχθηκαν από συγγραφείς δημοσιεύσεων τύπου και πολλές προσωπικές συζητήσεις με τα μέλη της ομάδας που έφτιαξε το MPEG πρότυπο.

Η ομάδα προτυποποίησης του MPEG ανέπτυξε ένα νέο πρότυπο, το πρότυπο MPEG-2, το οποίο καθορίζει την κωδικοποιημένη ροή δεδομένων με τέτοιο τρόπο ώστε να παρέχεται υψηλής ποιότητας ψηφιακό βίντεο. Σαν πρότυπο το οποίο

παρέχει συμβατότητα στο παλαιότερο MPEG-1, με το να υποστηρίζει πλεκτά (interlaced) μορφώττυπα βίντεο, καθώς και πολλά άλλα προηγμένα χαρακτηριστικά συμπεριλαμβανομένων και αυτό της υποστήριξης HDTV.

Σαν ένα γενικό και διεθνές πρότυπο, το MPEG-2 καθορίστηκε έτσι ώστε να παρέχει επεκτάσιμα προφίλ κάθε ένα από τα οποία θα υποστηρίζει χαρακτηριστικά που είναι απαραίτητα για ομάδες εφαρμογών. Το πρότυπο MPEG-2 κατασκευάστηκε με τέτοιο τρόπο έτσι ώστε να υποστηρίζει ψηφιακή μετάδοση βίντεο με εύρος από 2 έως 80 Mbits/sec. μέσω καλωδίου, δορυφόρου και άλλων καναλιών μετάδοσης, καθώς και να υποστηρίζει ψηφιακή αποθήκευση και άλλες εφαρμογές επικοινωνίας.

Παράμετροι όπως το *κυρίως προφίλ* και το *υψηλό προφίλ* είναι κατάλληλες για υποστήριξη HDTV μορφωτύπων .

Ανώτερο επίπεδο 1920 εικονοστοιχεία /γραμμή.		< 80 Mbit/s			< 100 Mbit/s
Ανώτερο επίπεδο 1440 1440 εικονοστοιχεία (pixel)/γραμμή 1152 γραμμές.		< 60 Mbit/s		< 60 Mbit/s	< 80 Mbit/s
Κύριο επίπεδο 720 εικονοστοιχεία/γραμμή 576 γραμμές.	< 15 Mbit/s	< 15 Mbit/s	<15Mbit/s		< 20 Mbit/s
Κατώτερο επίπεδο 352 εικονοστοιχεία (pixels)/γραμμή 288 γραμμές.		< 4 Mbit/s	< Mbit/s		
ΣΤΡΩΜΑΤΑ Και ΠΡΟΤΥΠΑ	Απλό προφίλ Χωρίς πλαίσια τύπου B	Κύριο Προφίλ Πλαίσια τύπου B 4:2:0	SNR Κλιμακωτό Προφίλ Πλαίσια τύπου	ΡΙΠΟΙΔΕΣ Κλιμακωτό Προφίλ Πλαίσια τύπου	Ανώτερο Προφίλ Πλαίσια τύπου B 4:2:0 ,ή

	4:2:0 Όχι κλιμάκωση	Όχι κλιμάκωση	B 4:2:0 SNR κλιμάκωση	B 4:2:0 SNR κλιμακωτό ,ή ριποειδές κλιμακωτό	4:2:2 SNR Κλιμακωτό , ή ριποιδές κλιμακωτό.
--	---------------------------	------------------	--------------------------------	--	--

**Πίνακας 9.4:** Τα προφίλ του MPEG-2 και τα επίπεδα με τα πιο σημαντικά χαρακτηριστικά τους. Οι περιοχές του πίνακα χωρίς στοιχεία δεν καθορίζονται ως σημεία συμβατότητας.

Οι ειδικοί στο πρότυπο MPEG φρόντισαν στο να επεκταθούν τα χαρακτηριστικά του κυρίως προφίλ με το να καθορίσουν ένα ιεραρχικό/κλιμακωτό προφίλ. Στόχος αυτού του προφίλ είναι να υποστηρίξει εφαρμογές όπως επίγεια τηλεόραση (τόσο κοινή όσο και HDTV), συστήματα video για δίκτυα μεταγωγής πακέτου, συμβατότητα με παλαιότερες εφαρμογές και πρότυπα όπως το MPEG-1 και H.261 και άλλες εφαρμογές που απαιτούν κωδικοποίηση πολλών επιπέδων .

Για παράδειγμα, ένα τέτοιο σύστημα μπορεί να δώσει στον καταναλωτή την επιλογή να χρησιμοποιήσει ένα φορητό δέκτη για να αποκωδικοποιήσει το τηλεοπτικό σήμα συνηθισμένης ευκρίνειας ή έναν μεγαλύτερο, σταθερό δέκτη για να αποκωδικοποιήσει το σήμα της τηλεόρασης υψηλής ευκρίνειας (HDTV) από το ίδιο κανάλι μετάδοσης.

Όλα τα προφίλ παρουσιάζονται στον πίνακα 6.4. Ο οριζόντιος άξονας παρουσιάζει τα πρότυπα με αύξουσα σειρά υποστηριζόμενης λειτουργικότητας. Ο κάθετος άξονας αναφέρει τα προφίλ με αυξημένες παραμέτρους, όπως μικρότερα και μεγαλύτερα μεγέθη πλαισίων.

Για παράδειγμα, το Κύριο Προφίλο στο Χαμηλό Επίπεδο καθορίζει 352 εικονοστοιχεία/ γραμμή, 288 γραμμές/πλαίσιο και 30 εναλλαγές εικόνας το δευτερόλεπτο, κατά τις οποίες τα πλαίσια τύπου B επιτρέπεται να προκύψουν και ταυτόχρονα απαγορεύεται ο ρυθμός μετάδοσης δεδομένων να ξεπεράσει τα 4 Mbit/sec. Το Κύριο Προφίλ στο ανώτατο επίπεδο καθορίζει 1920 εικονοστοιχεία/γραμμή, 1152 γραμμές/πλαίσιο και 60 εναλλαγές εικόνας το δευτερόλεπτο με ρυθμό μετάδοσης δεδομένων που δεν πρέπει να ξεπερνάει τα 80 Mbit/sec.

Το πρότυπο MPEG-2 χρησιμοποιεί μια δομή παρεμφερή μ' αυτή του ιεραρχικού μοντέλου JPEG. Η ιεραρχία αποτελείται από την κλιμάκωση (scaling) των συμπιεσμένων, κινούμενων εικόνων. Δηλαδή, το video κωδικοποιείται με διαβαθμίσεις ποιότητας. Η αλλαγή των διαστάσεων των εικόνων μπορεί να επιδρά σε διαφορετικές παραμέτρους:

Η χωρική κλιμάκωση επιτρέπει την αποσυμπίεση ακολουθιών εικόνων με διαφορετικές αναλύσεις στον κάθετο και οριζόντιο άξονα. Μία μόνο ροή δεδομένων μπορεί, για παράδειγμα, να περιλαμβάνει εικόνες με 353 X 288 εικονοστοιχεία (H.261 CIF μορφότυπο), 360 X 240 εικονοστοιχεία, 704 X 575 εικονοστοιχεία (ένα μορφότυπο σύμφωνο με το CCIR 601) ή 1250 γραμμές με μια απεικόνιση σε αναλογία 16:9 (Ευρωπαϊκή τηλεόραση υψηλής ευκρίνειας). Αυτές οι αναλύσεις αναφέρονται στο συστατικό φωτεινότητας. Τα συστατικά χρωμικότητας υποδειγματοληπτούνται (subsampling) με αναλογία 1:2. Αυτό μπορεί να εφαρμοσθεί χρησιμοποιώντας μια διάταξη πυραμίδας για τα επίπεδα του συντελεστή DCT. Ούτως, ένας DCT 8 X 8, ένας DCT 7 X 7, ένας 6 X 6 DCT και άλλοι μετασχηματισμοί μπορεί να πραγματοποιηθούν. Από τεχνικής απόψεως, μόνο βήματα πολλαπλασία του παράγοντα 2 είναι χρήσιμα.

Η κλιμάκωση του ρυθμού μετάδοσης δεδομένων επιτρέπει την αναπαραγωγή σε χαμηλό ρυθμό εναλλαγής πλαισίων (εικόνων) ή γρήγορη κίνηση προς τα εμπρός σε συνεχή ρυθμό εναλλαγής εικόνας. Στο πρότυπο MPEG-1 αυτό καθορίζεται με την χρήση πλαισίων τύπου D. D πλαίσια δεν επιτρέπονται στο πρότυπο MPEG-2. Γι αυτό, η γρήγορη κίνηση προς τα εμπρός μπορεί να υλοποιηθεί με πλαίσια τύπου I εάν υπάρχει ομοιόμορφη κατανομή μεταξύ τους μέσα στην ροή δεδομένων ολόκληρου του αποσπάσματος video και όχι μόνο σε μία ομάδα εικόνων.

Η κλιμάκωσης του εύρους μπορεί να ερμηνευτεί ως διαφορετική ανάλυση διαφορετικών εικονοστοιχείων ή ως διαφορετική κβαντοποίηση των DCT συντελεστών. Αυτό οδηγεί σε μια κωδικοποίηση διαφόρων επιπέδων και στην δυνατότητα προοδευτικής αναπαραστάσης εικόνων. Η προοδευτική κωδικοποίηση δεν είναι καθόλου σημαντική για την παρουσίαση δεδομένων video. Παρ' όλα αυτά, θα πρέπει να παρέχεται η δυνατότητα να πάρει κανείς μεμονωμένες εικόνες από μια ακολουθία τους, σαν σταθερή εικόνα, από ένα ρεύμα ροής δεδομένων και αυτό είναι το σημείο στο οποίο η προοδευτική κωδικοποίηση μπορεί να έχει ενδιαφέρον. Η

προοδευτική κωδικοποίηση μπορεί επίσης να χρησιμοποιηθεί για καταμερισμό δεδομένων έτσι ώστε τα σημαντικά δεδομένα να μπορούν να μεταδοθούν με καλύτερη διόρθωση λάθους απ' ό,τι τα λιγότερο σημαντικά.

Η κλιμάκωση είναι μια βασική προέκταση από το πρότυπο MPEG-1 στο πρότυπο MPEG-2. Επιπρόσθετα, το πρότυπο MPEG-2 αγκαλιάζει τις τελευταίες εξελίξεις στον κόσμο μετάδοσης ISDN μεγάλου εύρους ζώνης. Ο Ασύγχρονος Τρόπος Μετάδοσης (ATM – Asynchronous Transfer Mode) αποτελεί την υλοποίηση του B-ISDN, βασιζόμενος στην μεταφορά μικρών πακέτων, γνωστών ως *κελιών (cells)*. Η πιθανή απώλεια ενός κελιού ATM, το οποίο περιέχει κωδικοποιημένα δεδομένα τύπου MPEG-2, λήφθηκε υπόψη κατά την ανάπτυξη του προτύπου MPEG-2. Σ' αυτή την περίπτωση οι παρενέργειες σ' άλλες εικόνες και μέρη του βίντεο πρέπει να ελαχιστοποιηθούν. Πρέπει, ακόμα, να είναι δυνατόν να καθοριστεί η αλληλουχία διαφορετικών τύπων εικόνων (I, P, και B πλαισίων) η οποία ελαχιστοποιεί την καθυστέρηση απ' άκρη σ' άκρη για ένα ρυθμό δεδομένων-στόχο.

Η ομάδα ανάπτυξης του MPEG ανέπτυξε το MPEG-2 πρότυπο ήχου για χαμηλού ρυθμού κωδικοποίησης πολυκάναλο ήχο. Το πρότυπο MPEG-2V, για κωδικοποίηση ήχου, προμηθεύει έως πέντε κανάλια πλήρους εύρους (αριστερό, δεξιό, κεντρικό και δύο κανάλια για περιφερειακό (surround) ήχο) και, επιπρόσθετα, ένα ακόμα χαμηλής συχνότητας κανάλι βελτιστοποίησης και/ή έως επτά πολυγλωσσικά κανάλια ή κανάλια σχολιασμού. Το πρότυπο MPEG-2, για ήχο, επεκτείνει την κωδικοποίηση μονοφωνικού και στερεοφωνικού ήχου του MPEG-1. Τα πρότυπα ήχου παρέχουν δειγματοληψία ελαττωμένη κατά το ήμισυ (16 KHz, 22.05 KHz και 24 KHz) ώστε να παρέχουν βελτιωμένη ποιότητα για ρυθμούς μετάδοσης σε 64 kbits/sec. ή πιο κάτω για κάθε κανάλι.

Το MPEG-2 πρότυπο πολυκάναλης ηχητικής κωδικοποίησης παρέχει προς τα πίσω συμβατότητα με το υπάρχον πρότυπο ήχου MPEG-1. Το MPEG οργανώνει επίσημες υποκειμενικές δοκιμές των προτεινόμενων κωδικοποιητών/αποκωδικοποιητών πολυκάναλου ήχου κωδικοποίησης MPEG-2 και μέχρι τρεις μη συμβατούς προς τα πίσω κωδικοποιητές/αποκωδικοποιητές. Αυτοί οι μηχανισμοί κωδικοποίησης λειτουργούν με ρυθμούς μετάδοσης δεδομένων, από 256 έως 448 Kbits/sec.

Αξιοσημείωτο είναι ότι για να παρέχει το πρότυπο ακριβή περιγραφή, στο κάτωθι κείμενο, ακολουθείται η αναπαράσταση και η ορολογία που καθορίζει το

πρωτότυπο πρότυπο MPEG-2. Το πρότυπο MPEG-2 απευθύνεται σε βίντεο αλλά και στον συσχετιζόμενο ήχο. Παρέχει, στο MPEG-2 σύστημα, ένα τρόπο ώστε να ορίζει το πως ο ήχος, το βίντεο και άλλα δεδομένα συνδυάζονται σ' ένα ή σε πολλαπλά ρεύματα ροής τα οποία είναι κατάλληλα για αποθήκευση και μετάδοση. Γι' αυτό, επιβάλλει συντακτικούς και σημασιολογικούς κανόνες οι οποίοι είναι απαραίτητοι και επαρκείς για να συγχρονίσουν την αποκωδικοποίηση και την παρουσίαση βίντεο και ηχητικών πληροφοριών, εξασφαλίζοντας παράλληλα ότι τα κωδικοποιημένα δεδομένα στους ενταμιευτές του αποκωδικοποιητή δεν θα υπερχειλίσουν ή οι ενταμιευτές θ' αδειάσουν. Τα ρεύματα ροής περιλαμβάνουν χρονοσφραγίδες που αφορούν την αποκωδικοποίηση, παρουσίαση και παράδοση αυτών των δεδομένων.

Σ' ένα πρώτο στάδιο, η βασική προσέγγιση της πολύπλεξης προσθέτει σε κάθε ξεχωριστό ρεύμα πληροφορίες επιπέδου συστήματος και κάθε ανεξάρτητο ρεύμα πακετάρεται για να παραχθεί το *στοιχειώδες πακέτο ροής (PES – Packetized Elementary Stream)*. Σε ένα μεταγενέστερο βήμα, τα PES συνδυάζονται ώστε να σχηματίσουν μία *Προγραμματική ή Ροή Μεταφοράς*. Και τα δύο ρεύματα είναι σχεδιασμένα να υποστηρίζουν ένα μεγάλο αριθμό γνωστών και αναμενόμενων εφαρμογών και διατηρούν μια σημαντική ικανότητα προσαρμογής που μπορεί να είναι απαραίτητη, ενώ παράλληλα παρέχουν δυνατότητα υποστήριξης επικοινωνίας και διαμεσότητα μεταξύ διαφορετικών συσκευών.

Η Προγραμματική ροή είναι παρόμοια με το ρεύμα του προτύπου MPEG. Έχει σαν στόχο ένα σχετικά ελεύθερο λαθών περιβάλλον. Τα πακέτα Προγραμματικής ροής μπορεί να ποικίλλουν σε μήκος. Η πληροφορία συγχρονισμού σε μία τέτοια ροή μπορεί να χρησιμοποιηθεί σαν εργαλείο για την επίτευξη μίας σταθερής καθυστέρησης από άκρο σε άκρο (end-to-end delay), καλύπτοντας το μονοπάτι από την είσοδο του κωδικοποιητή στην έξοδο του αποκωδικοποιητή.

Η ροή μεταφοράς συνδυάζει τα στοιχειώδη πακέτα ροής (PESs ) και μία ή περισσότερες χρονικές βάσεις σε ένα μόνο ρεύμα. Η ροή μεταφοράς είναι σχεδιασμένη για χρήση σε απωλεστικά ή θορυβώδη μέσα. Τα αντίστοιχα πακέτα έχουν μήκος 188 bytes, στο οποίο συμπεριλαμβάνεται και μια επικεφαλίδα μήκους 4 bytes. Η ροή μεταφοράς είναι κατάλληλη για μετάδοση ψηφιακής τηλεόρασης και βιντεοτηλεφώνου με χρήση οπτικής ίνας, δορυφόρο, καλώδο, ISDN, ATM και άλλα δίκτυα και ακόμα για αποθήκευση ψηφιακού βίντεο σε κασέτα και άλλες συσκευές.

Μια μετατροπή ανάμεσα σε μία προγραμματική ροή και μία ροή μεταφοράς είναι δυνατή και λογικό να μπορεί να γίνει. Αξιοσημείωτο είναι το ότι το MPEG-2 ορίζει για την διαχείριση των ενταμιευτών του σταθερές συνολικής καθυστέρησης λιγότερο από ένα δευτερόλεπτο για δεδομένα ήχου και βίντεο, μια τιμή που είναι πολύ υψηλή (όχι αποδεκτή από τον άνθρωπο) για αλληλεπιδραστικές εφαρμογές.

Μια τυπική ροή βίντεο τύπου MPEG-2 έχει μη σταθερό ρυθμό μετάδοσης. Με την χρήση ενταμιευτή βίντεο που καθορίζεται σε αυτό το πρότυπο, είναι δυνατό να επιβληθεί επίσης ένας σταθερός ρυθμός μετάδοσης που να οδηγεί σε μεταβαλλόμενη ποιότητα.

Το πρότυπο MPEG-2 στο καθεστώς CD (αρχικής επιτροπής) στο τέλος του 1993, απαιτούσε 3 μήνες για να γίνει ένα DIS (διεθνές αποδεκτό αρχικό πρότυπο) και μετά απαιτήθηκαν άλλοι 6 μήνες ψηφοφορίας πριν γίνει ένα IS (διεθνές πρότυπο). Αρχικά, υπήρχαν σχέδια για να καθοριστεί ένα πρότυπο MPEG-3 που να πλησιάζει την ψηφιακή τηλεόραση υψηλής ευκρίνειας HDTV. Παρ' όλα αυτά, κατά την διάρκεια ανάπτυξης του προτύπου MPEG-2, ανακαλύφθηκε ότι η μεγέθυνση κλίμακας μπορούσε να καλύψει σε επαρκή βαθμό τις απαιτήσεις της HDTV. Κατά συνέπεια, μεταγενέστερες προσπάθειες για την καθιέρωση του προτύπου MPEG-3 απορρίφθηκαν.

## **9.12 MPEG-4**

Η εργασία σε ένα καινούριο πρότυπο MPEG προσανατολισμένο σε ένα χαμηλό ρυθμό δεδομένων για κωδικοποίηση ηχητικών/οπτικών προγραμμάτων ξεκίνησε τον Σεπτέμβριο του 1993 στο ISO/IEC JTC1. Ήταν σχεδιασμένο να πάρει την έγκριση του καθεστώτος της αρχικής επιτροπής (CD) μέσα στο 1995 ή 1996.

Αυτή η εργασία θα απαιτούσε την εξέλιξη ριζικά καινούριων αλγοριθμικών τεχνικών, συμπεριλαμβανομένης κωδικοποίησης εικόνων βασισμένης σε μοντέλα (model-based image coding) για την αλληλεπίδραση ανθρώπων με πολυμεσικά περιβάλλοντα καθώς και τη κωδικοποίηση ομιλίας χαμηλού ρυθμού μετάδοσης για χρήση σε περιβάλλοντα όπως το Ευρωπαϊκό Σύστημα Κινητής Τηλεφωνίας (GSM).

## Σχόλια

Όλες οι σημαντικές τεχνικές κωδικοποίησης που χρησιμοποιούνται στα πολυμεσικά συστήματα καταλήγουν να αποτελούνται από συνδυασμούς διαφορετικών γνωστών αλγορίθμων.

Το πρότυπο JPEG πρέπει να θεωρείται το μέλλον σε μελλοντικά πρότυπα κωδικοποίησης ακίνητων εικόνων. Συνδυάζει μια αξιοθαύμαστη ποικιλία από διαφορετικούς τρόπους κωδικοποίησης με ένα υψηλό βαθμό ελευθερίας. Για παράδειγμα, μπορεί να υπάρχουν μέχρι 255 συστατικά ή επίπεδα, μια εικόνα μπορεί να αποτελείται από μέχρι 65535 γραμμές και κάθε γραμμή να έχει μέχρι και 65525 εικονοστοιχεία. Ως ένα μέτρο επάρκειας, τα απαιτούμενα δυαδικά ψηφία ανά εικονοστοιχείο μπορεί να συντίθενται. Δηλαδή, η μέση τιμή να καθορίζεται από τον λόγο του αριθμού των κωδικοποιημένων bits και τον αριθμό των εικονοστοιχείων της εικόνας. Οι παρακάτω δηλώσεις εφαρμόζονται στην κωδικοποίηση ακίνητων εικόνων κατα DTC:

0.25 έως 0.50 δυαδικά ψηφία ανά εικονοστοιχείο – μέτρια έως καλή ποιότητα, επαρκής για ορισμένες εφαρμογές.

0.50 έως 0.75 δυαδικά ψηφία ανά εικονοστοιχείο – καλή έως πολύ καλή ποιότητα, επαρκής για πολλές εφαρμογές.

έως 1.50 δυαδικά ψηφία ανά εικονοστοιχείο – εξαιρετική ποιότητα, επαρκής για τις περισσότερες εφαρμογές.

1.50 έως 2.00 δυαδικά ψηφία ανά εικονοστοιχείο. Στις πιο πολλές περιπτώσεις, δεν μπορεί κανείς να ξεχωρίσει την τελική εικόνα από την πρωτότυπη. Είναι όμως επαρκής για όλες σχεδόν τις εφαρμογές, ακόμα και για τις πιο υψηλές απαιτήσεις.

Κατά τον απωλεστικό τρόπο συμπίεσης, ένας λόγος συμπίεσης 2:1 επιτυγχάνεται παρόλη την αξιοθαύμαστη απλότητα της τεχνικής. Σήμερα, η υλοποίηση του JPEG προτύπου είναι διαθέσιμη τόσο ως λογισμικό, όσο και ως υλικό. Παρ' όλα αυτά, στις πιο πολλές περιπτώσεις, μόνο ο βασικός τρόπος (baseline mode) και προκαθορισμένες μορφές της εικόνας υποστηρίζονται. Το JPEG χρησιμοποιείται συχνά σε εφαρμογές πολυμέσων για τις οποίες απαιτείται υψηλή ποιότητα. Η πρωταρχική λειτουργία του είναι η συμπίεση μιας εικόνας. Όμως, χρησιμοποιείται



επίσης και σαν JPEG με κίνηση (MotionJPEG) για συμπίεση βίντεο σε εφαρμογές όπως ιατρική απεικόνιση.

Το H.261 είναι ένα καθιερωμένο πρότυπο το οποίο υποστηρίζεται ευρέως από παροχείς τηλεπικοινωνιακών υπηρεσιών. Εξ' αιτίας της πολύ περιορισμένης ανάλυσης που παρέχει, στην μορφή QCIF, και του περιορισμένου ρυθμού κίνησης πλαισίων, τα εργαλεία των κωδικοποιητών και αποκωδικοποιητών H.261 έχουν πιθανόν λίγα, εάν όχι καθόλου, τεχνικά προβλήματα. Αυτό ισχύει απόλυτα εάν ο συμβιβασμός κίνησης (motion compensation) και το οπτικό φίλτρο χαμηλής σάρωσης (low-pass filter) δεν είναι στοιχεία της υλοποίησης, παρόλο που η ποιότητα δεν είναι πάντα ικανοποιητική σε αυτή την περίπτωση. Εάν η εικόνα είναι κωδικοποιημένη σε μορφή τύπου CIF σε 25 ή 30 εικόνες το δευτερόλεπτο χρησιμοποιώντας συμπίεση κίνησης η ποιότητα είναι αποδεκτή. Το H.261 χρησιμοποιείται κυρίως σε εφαρμογές που έχουν διαδικασίες διαλόγου σε ένα περιβάλλον δικτύου: βίντεοτηλέφωνο και τηλεδιάσκεψη. Το αποτέλεσμα συνεχούς ρυθμού μετάδοσης είναι αξιοσημείωτα κατάλληλο για τα σημερινά ευρείας περιοχής δίκτυα (WANs) που λειτουργούν με ISDN και μισθωμένες γραμμές.

Το πρότυπο MPEG είναι ένα πολλά υποσχόμενο πρότυπο για μελλοντική κωδικοποίηση ψηφιακού βίντεο και ήχου. Ενώ η ομάδα ανάπτυξης του JPEG έχει ένα σύστημα το οποίο μπορεί επίσης να χρησιμοποιηθεί για βίντεο, είναι εντούτοις πολύ προσανατολισμένο προς την εναλλαγή (animation) σταθερών εικόνων παρά προς τη χρησιμοποίηση ιδιοτήτων των κινούμενων εικόνων. Προς το παρόν, η ποιότητα του MPEG βίντεο μπορεί να συγκριθεί με αυτήν του VHS βίντεο, μετρώντας ένα ρυθμό μετάδοσης δεδομένων 1.2 Mbits/sec., ρυθμός κατάλληλος για τους οδηγούς CD-ROM. Ο αλγόριθμος συμπίεσης είναι πολύ καλός για ανάλυση 360 X 240 εικονοστοιχείων. Είναι φανερό ότι υψηλότερες αναλύσεις μπορεί επίσης να αποκωδικοποιηθούν (για παράδειγμα μια ανάλυση με 625 γραμμές), αλλά η ποιότητα επηρεάζεται. Το MPEG-2, καθορίζει μια ροή δεδομένων συμβατή με το MPEG-1, αλλά παρέχει ρυθμούς μετάδοσης δεδομένων μέχρι 40 Mbits/sec. Αυτό αποδεικνύει μια βελτίωση της τρέχουσας διαθέσιμης ποιότητας των κωδικοποιημένων δεδομένων με MPEG.

Το πρότυπο MPEG καθορίζει επίσης μία ροή ήχου, με δειγματοληψία διαφόρων ρυθμών, έως την ποιότητα του DAT των 16 bits ανά δείγμα. Επιπλέον, το πρότυπο MPEG βελτιώθηκε με το να κάνει χρήση του μοντέλου ανάκτησης σε περιοχές

εφαρμογών όπως τα συστήματα διδασκαλίας βασισμένα σε CD-ROMs και στην αλληλεπιδραστική τηλεόραση. Αυτή η, ενσωματωμένη στο MPEG-2, βελτιστοποίηση θα επιτρέψει ποιότητα τηλεόρασης και υψηλής ποιότητας ψηφιακή τηλεόραση με κόστος ένα μεγαλύτερο ρυθμό δεδομένων. Το πρότυπο MPEG-4 θα παρέχει έναν πολύ μεγάλο λόγο συμπίεσης βίντεο και συνοδευτικού ήχου.