



ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΑΤΡΩΝ

ΠΟΛΥΤΕΧΝΙΚΗ ΣΧΟΛΗ

**ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΗΛΕΚΤΡΟΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
& ΠΛΗΡΟΦΟΡΙΚΗΣ**

ΕΡΓΑΣΙΑ ΕΞΑΜΗΝΟΥ

ΓΙΑ ΤΟ ΜΑΘΗΜΑ

ΤΗΛΕΜΑΤΙΚΗ ΚΑΙ ΝΕΕΣ ΥΠΗΡΕΣΙΕΣ

**ΚΩΔΙΚΟΠΟΙΗΣΗ - ΣΥΜΠΙΕΣΗ
ΠΟΛΥΜΕΣΩΝ**

ΟΝΟΜΑ: ΤΑΛΙΕΡΗΣ ΚΑΝΣΤΑΝΤΙΝΟΣ

A.M. : 4373

ΔΙΔΑΣΚΩΝ: ΧΡΗΣΤΟΣ ΜΠΟΥΡΑΣ

ПАТРА 2015



ΠΕΡΙΕΧΟΜΕΝΑ

ΕΙΣΑΓΩΓΗ	3
ΚΕΦΑΛΑΙΟ 1 – ΑΝΑΛΟΓΙΚΑ ΚΑΙ ΨΗΦΙΑΚΑ ΔΕΔΟΜΕΝΑ	5
ΚΕΦΑΛΑΙΟ 2 - ΣΥΜΠΙΕΣΗ ΔΕΔΟΜΕΝΩΝ	8
ΚΕΦΑΛΑΙΟ 3: ΚΑΤΗΓΟΡΙΕΣ ΣΥΜΠΙΕΣΗΣ ΔΕΔΟΜΕΝΩΝ	10
ΚΕΦΑΛΑΙΟ 4: ΒΑΣΙΚΟΙ ΑΛΓΟΡΙΘΜΟΙ ΣΥΜΠΙΕΣΗΣ ΔΕΔΟΜΕΝΩΝ.....	11
ΚΕΦΑΛΑΙΟ 5: ΚΩΔΙΚΟΠΟΙΗΣΗ ΔΕΔΟΜΕΝΩΝ	16
ΚΕΦΑΛΑΙΟ 6: ΨΗΦΙΑΚΗ ΠΑΡΑΣΤΑΣΗ ΕΙΚΟΝΑΣ.....	23
ΣΥΜΠΕΡΑΣΜΑΤΑ	39
ΒΙΒΛΙΟΓΡΑΦΙΑ.....	41

ΕΙΣΑΓΩΓΗ

Τα δεδομένα, είναι γεγονότα, μηνύματα, κωδικοποιημένα ή όχι που αποτελούν ακατέργαστο υλικό. Όταν σχεδιάσθηκαν οι πρώτοι ηλεκτρονικοί υπολογιστές (ΗΥ), έπρεπε να βρεθεί ένα κοινό σημείο ανάμεσα στο τι μπορεί να κάνει και πώς στα αλήθεια δουλεύει ένας υπολογιστής από τη μια πλευρά, 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 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, δηλαδή το 1 bit. Όπως γνωρίζουμε, για την παράσταση ενός χαρακτήρα στον υπολογιστή απαιτούνται, συνήθως, 8 bits τα οποία αποτελούν και τη στοιχειώδη μονάδα αποθήκευσης, το χαρακτήρα (byte).

Κάθε λοιπόν δεδομένο αναπαριστάται με μια σειρά δυαδικών ψηφίων. Τα δεδομένα αυτά μπορούμε να τα μεταφέρουμε από ένα υπολογιστή σε ένα άλλο. Για να γίνει αυτό ορίζουμε διάφορες κωδικοποιήσεις κατάλληλες για την κάθε μορφή.

Γενικά η κωδικοποίηση δεδομένων χρησιμοποιείται για την μετάδοση των σημάτων (ψηφιακά ή αναλογικά) μια πηγής (ψηφιακής ή αναλογικής) μέσω ενός καναλιού μετάδοσης. Η κωδικοποίηση που επιλέγεται εξαρτάται από τις απαιτήσεις, το επικοινωνιακό σύστημα, το μέσο μετάδοσης και τον διαθέσιμο εξοπλισμό. Οι δυνατές κωδικοποιήσεις είναι

- Αναλογικά δεδομένα - Αναλογικά σήματα. Χρησιμοποιείται η τεχνική της διαμόρφωσης. Βασικές τεχνικές είναι η διαμόρφωση πλάτους (AM) , η διαμόρφωση συχνότητας (FM) και η διαμόρφωση φάσης (PM).

- Αναλογικά δεδομένα - Ψηφιακά σήματα. Χρησιμοποιείται συχνά η παλμοκωδική διαμόρφωση , τα δεδομένα μετά την δειγματοληψία τους κβαντίζονται και στην σειρά κωδικοποιούνται σε σειρά bits.

- Ψηφιακά δεδομένα - Αναλογικά σήματα. Χρησιμοποιούνται οι τεχνικές ψηφιακής διαμόρφωσης πλάτους (ASK), η ψηφιακή διαμόρφωση συχνότητας (FSK) και η διαμόρφωση φάσης (PSK).

- Ψηφιακά δεδομένα - Ψηφιακά σήματα. Ο πιο απλός τρόπος είναι να αναπαρασταθούν τα δεδομένα με σήμα 2 επίπεδων τάσης ένα για το δυαδικό 0 και ένα άλλο για το δυαδικό 1.

Όμως τα δεδομένα ήχου, εικόνας, video είναι πολύ μεγάλα σε όγκο καθώς και άλλοι τύποι δεδομένων και για το λόγο αυτό είναι απαραίτητο να συμπιέζονται ώστε να μεταφέρονται πιο γρήγορα και με μεγαλύτερη ασφάλεια.

Η συμπίεση δεδομένων είναι η διαδικασία αυτή που στόχο έχει την ελάττωση του χώρου που χρειάζεται για την αποθήκευση ή τη μετάδοσή τους. Η διαδικασία της συμπίεσης εφαρμόζεται συστηματικά στα υπολογιστικά συστήματα που χρησιμοποιούν και επεξεργάζονται μεγάλο όγκο ψηφιακών δεδομένων.

ΚΕΦΑΛΑΙΟ 1 – ΑΝΑΛΟΓΙΚΑ ΚΑΙ ΨΗΦΙΑΚΑ ΔΕΔΟΜΕΝΑ

Τα μεγέθη και δεδομένα διακρίνονται σε δύο κατηγορίες. Τα αναλογικά και ψηφιακά δεδομένα.

Τα αναλογικά δεδομένα είναι τα δεδομένα που παίρνουν συνεχόμενες τιμές ενώ τα ψηφιακά αυτά που οι τιμές διακρίνονται σε δύο καταστάσεις 0 και 1.

Ο πραγματικός κόσμος είναι ένας αναλογικός κόσμος.

Όλα τα μεγέθη παίρνουν τιμές με άπειρη ακρίβεια. Π.χ. το ηλεκτρικό σήμα τάσης όπου κάθε στιγμή η τιμή του μπορεί να απαιτεί άπειρα δεκαδικά ψηφία για να περιγραφεί και μπορεί η τάση στα άκρα μιας πηγής να πάρει οποιαδήποτε τιμή από 0 έως X volt.

Ο ψηφιακός κόσμος των υπολογιστών, αντίθετα, μπορεί να χειριστεί μόνο πεπερασμένο αριθμό δεδομένων και με πεπερασμένη ακρίβεια.

Οι υπολογιστές και γενικά τα πληροφοριακά συστήματα χρησιμοποιούν ψηφιακά δεδομένα.

Για να μπορέσουμε να μεταφέρουμε τα δεδομένα μας λοιπόν επιβάλλεται η μετατροπή των αναλογικών σε ψηφιακά σήματα και το αντίθετο.

Ο όρος ψηφιακό σήμα αναφέρεται σε περισσότερες από μια έννοιες. Μπορεί να αναφέρεται σε ένα σήμα διακριτού χρόνου το οποίο μπορεί να πάρει συγκεκριμένες (διακριτές) τιμές στον άξονα του χρόνου. Σε αυτή τη περίπτωση μιλάμε για ένα σήμα το οποίο παράγεται μέσω μιας μεθόδου ψηφιακής διαμόρφωσης και θεωρείται περισσότερο ως αναλογικό σήμα (δηλ. ένα επεξεργασμένο αναλογικό σήμα για τη μετατροπή του σε ψηφιακό).

Ένα αναλογικό σήμα είναι μια χρονικά μεταβαλλόμενη τιμή δεδομένων ή όπως είναι ο ορισμός του, μια ομαλά μεταβαλλόμενη τιμή ηλεκτρικής τάσης ή έντασης

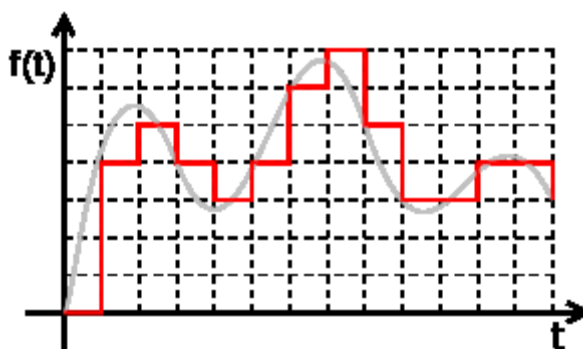
ρεύματος (δηλ. ένα σήμα με πλάτος χρονικά μεταβαλλόμενο) η οποία μπορεί να περιγραφεί από μια μαθηματική συνάρτηση, με το χρόνο να αποτελεί την ανεξάρτητη και την τιμή του σήματος, κάθε χρονική στιγμή, την εξαρτημένη μεταβλητή. Ένα διακριτό σήμα είναι το αποτέλεσμα που παίρνουμε μέσω της μεθόδου της δειγματοληπτικής μείωσης από το αρχικό αναλογικό σήμα : δηλαδή, η τιμή των δεδομένων σημειώνεται ανά τακτά χρονικά διαστήματα (π.χ. μικροδευτερόλεπτο) και όχι συνεχώς (όπως είναι εκ φύσεως τα μηχανικά κύματα).

Αν οι ατομικές τιμές του σήματος αντί να μετρηθούν επακριβώς, επάνω στον άξονα του χρόνου, είναι εναρμονισμένες με κάποια ορισμένη ακρίβεια, τότε η ροή δεδομένων που προκύπτει είναι το ψηφιακό σήμα. Η διαδικασία προσέγγισης αυτής της ακρίβειας (δηλ. μιας συγκεκριμένης τιμής), μέσα από ένα σταθερό αριθμό ψηφίων (δηλ. bit) ονομάζεται ψηφιοποίηση.

Σε γενικές γραμμές, ένα ψηφιακό σήμα είναι ένα ψηφιοποιημένο σήμα διακριτού χρόνου. Το διακριτό σήμα είναι το αποτέλεσμα της επεξεργασίας ενός αναλογικού σήματος με τη μέθοδο της δειγματοληπτικής μείωσης.

Η Ψηφιακή Επανάσταση έχει αυξήσει κατακόρυφα τη χρήση ψηφιακών σημάτων. Οι περισσότερες-αν όχι όλες, οι σύγχρονες συσκευές ειδικότερα αυτές που συνδέονται στους υπολογιστές, χρησιμοποιούν ψηφιακά σήματα για την αναπαράσταση σημάτων τα οποία παραδοσιακά αναπαριστώνταν ως σήματα συνεχούς χρόνου. Κινητά τηλέφωνα, συσκευές αναπαραγωγής βίντεο και ήχου, ψηφιακές φωτογραφικές μηχανές είναι μερικά παραδείγματα.

Σχεδόν σε όλες τις εφαρμογές, τα ψηφιακά σήματα αναπαριστώνται μέσω των δυαδικών αριθμών, έτσι ώστε να μπορούν να μετρηθούν σε bit. Επειδή 7 bits (δυαδικά ψηφία) μπορούν να καταγράψουν 128 διακριτές τιμές (0 έως 127), αυτό το σύστημα είναι το πιο ικανό από οποιοδήποτε άλλο για να εκφραστεί ένα τεράστιο πλήθος τιμών.



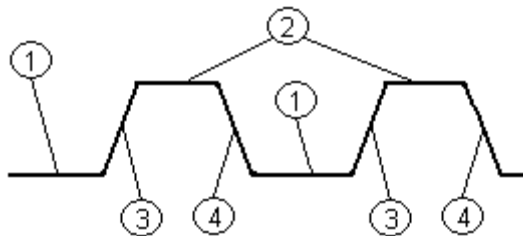
Μετατροπή αναλογικού σε ψηφιακό σήμα

Τόσο στους υπολογιστές όσο και σε οποιοδήποτε άλλο ψηφιακό σύστημα, η κυματομορφή του σήματος εναλλάσσεται μεταξύ δύο επιπέδων τάσης (0 και 4,8V) οι οποίες αναπαριστούν αντίστοιχα τις δύο τιμές του δυαδικού συστήματος (0 και 1). Έτσι, αναφερόμαστε σε αυτή τη κυματομορφή ως ψηφιακό σήμα. Παρ' όλο που πρόκειται για μια αναλογική κυματομορφή τάσεως, το ονομάζουμε ψηφιακό διότι εναλλάσσεται μεταξύ δύο σταθερών καταστάσεων.

Το σήμα του ρολογιού είναι ένα ειδικό ψηφιακό σήμα το οποίο χρησιμοποιείται για τον συγχρονισμό των ψηφιακών κυκλωμάτων. Στην διπλανή εικόνα φαίνεται η συγκεκριμένη κυματομορφή. Οι λογικές αλλαγές ενεργοποιούνται είτε από την αύξηση είτε από την μείωση του σήματος. Στο διπλανό διάγραμμα βλέπουμε ένα παράδειγμα του πρακτικού παλμού. Έτσι, όταν λέμε:

Αύξηση του σήματος : εννοούμε την διαδικασία της μετάβασης από χαμηλή σε υψηλή τάση.

Μείωση του σήματος : εννοούμε την διαδικασία της μετάβασης από υψηλή σε χαμηλή τάση.



Η κυματομορφή ενός ψηφιακού σήματος:

(1) χαμηλό επίπεδο τάσης, (2) υψηλό επίπεδο, (3) μετάβαση σε υψηλό επίπεδο τάσης, (4) μετάβαση σε χαμηλό επίπεδο

ΚΕΦΑΛΑΙΟ 2 - ΣΥΜΠΙΕΣΗ

ΔΕΔΟΜΕΝΩΝ

Με τον όρο συμπίεση δεδομένων (data compression) εννοούμε τη μετατροπή ενός ψηφιακού αρχείου σε μικρότερο αρχείο (που περιέχει μικρότερο αριθμό μπιτ) με τρόπο ώστε να είναι δυνατή η επαναμετατροπή του συμπιεσμένου αρχείου στο αρχικό. Για το σκοπό αυτό χρησιμοποιούνται πολλές μέθοδοι, οι οποίες χωρίζονται σε δύο μεγάλες κατηγορίες: τις μη απωλεστικές και τις απωλεστικές.

Αν θεωρήσουμε τα δεδομένα μας σαν μια ακολουθία τότε η συμπίεση ορίζεται όπως παρακάτω:

Συμπίεση (compression) μιας ακολουθίας δεδομένων, ονομάζουμε την ελάττωση του μεγέθους της ακολουθίας, ώστε να χρειάζεται λιγότερος χώρος για την αποθήκευση ή τη μετάδοσή της. Η διαδικασία της συμπίεσης εφαρμόζεται συστηματικά στα υπολογιστικά συστήματα που χρησιμοποιούν και επεξεργάζονται μεγάλο όγκο ψηφιακών δεδομένων.

Υπάρχουν πολλοί αλγόριθμοι για να καταφέρουμε να μειώσουμε το μέγεθος μίας ακολουθίας. Ένας αλγόριθμος συμπίεσης μπορεί να επιφέρει μεγάλη ελάττωση του μήκους σε ακολουθίες δεδομένων με κάποιο ιδιαίτερο χαρακτηριστικό, ενώ ο ίδιος να είναι αναποτελεσματικός για άλλες ακολουθίες, αφήνοντάς τις στο ίδιο μήκος ή ακόμα και μεγαλώνοντάς το σε μερικές περιπτώσεις.

Όταν έχουμε την ακολουθία των δεδομένων σε συμπιεσμένη μορφή, πρέπει να εφαρμοστεί η αντίστροφη διαδικασία της *αποσυμπίεσης* (decompression, extraction) προκειμένου τα δεδομένα να μπορούν να χρησιμοποιηθούν και πάλι. Η διαδικασία αυτή της αποσυμπίεσης των δεδομένων καθορίζει και τις κατηγορίες των μεθόδων συμπίεσης. Έτσι, υπάρχουν δύο κατηγορίες αλγορίθμων συμπίεσης, οι *απωλεστικοί* (lossy) και οι μη *απωλεστικοί* (lossless) αλγόριθμοι.

Στους απωλεστικούς αλγορίθμους, όταν γίνει η συμπίεση και μετά ακολουθήσει αποσυμπίεση των δεδομένων, η τελική ακολουθία των δεδομένων διαφέρει από την

αρχική. Αντίθετα στους μη απωλεστικούς αλγορίθμους, η διαδικασία συμπίεσης και αποσυμπίεσης επαναφέρει την αρχική ακολουθία. Αν πρέπει να μεταφερθούν δεδομένα με απόλυτη ακρίβεια, πιστότητα, χωρίς να αλλοιωθεί το περιεχόμενό τους, πρέπει να εφαρμοστεί μια μη απωλεστική μέθοδος συμπίεσης. Εφαρμογές ή συστήματα που μεταδίδουν αναλλοίωτες πληροφορίες από το ένα μέσο στο άλλο, για παράδειγμα κάρτες δικτύου ή modem, χρησιμοποιούν τεχνικές μη απωλεστικής συμπίεσης. Υπάρχουν όμως εφαρμογές όπου η μικρή διαφοροποίηση από την αρχική μορφή των δεδομένων δεν επιφέρει σημαντικές αλλαγές. Έτσι, οι περισσότερες εφαρμογές που έχουν να κάνουν με σύνθετες μορφές δεδομένων όπως είναι ο ήχος, η εικόνα, το video, όπου το τελικό αποτέλεσμα αξιολογείται από τον ανθρώπινο παράγοντα (αυτί, μάτι), μπορούν να κάνουν απωλεστική συμπίεση χωρίς πολλές φορές να υπάρχουν εμφανείς αλλοιώσεις στην ποιότητα των δεδομένων.

Πολλά αρχεία περιέχουν μεγάλα τμήματα τα οποία επαναλαμβάνονται, όπως ένα αρχείο κειμένου μπορεί να περιέχει πολλές φορές μια λέξη π.χ. Test οπότε αρκεί η λέξη να τοποθετηθεί σε έναν πίνακα με λέξεις και να αντικαθίσταται με έναν αριθμό που δείχνει τη θέση της στον πίνακα, ή περιέχουν δεδομένα που η απώλειά τους δεν προκαλεί σοβαρή μεταβολή του περιεχόμενου.

Η συμπίεση μπορεί να μειώσει τον όγκο των δεδομένων που στέλνονται ή αποθηκεύονται, με την ελαχιστοποίηση του ενυπάρχοντος πλεονασμού. Ο πλεονασμός παρουσιάζεται κατά τη δημιουργία των δεδομένων. Με τη διαδικασία της συμπίεσης η μεταφορά και η αποθήκευση γίνονται με πιο αποδοτικό τρόπο, ενώ παράλληλα διατηρείται η ακεραιότητα των δεδομένων.

ΚΕΦΑΛΑΙΟ 3: ΚΑΤΗΓΟΡΙΕΣ ΣΥΜΠΙΕΣΗΣ ΔΕΔΟΜΕΝΩΝ

Μη απωλεστική συμπίεση

Στην μη απωλεστική συμπίεση (lossless compression) διατηρείται η ακεραιότητα των δεδομένων. Τα αρχικά δεδομένα και τα δεδομένα μετά τη συμπίεση και την αποσυμπίεση είναι ακριβώς τα ίδια, επειδή σε αυτές τις μεθόδους ο αλγόριθμος συμπίεσης και ο αλγόριθμος αποσυμπίεσης είναι ακριβώς αντίστροφοι. Κατά τη διαδικασία δε χάνεται κανένα μέρος των δεδομένων. Τα πλεονάζοντα δεδομένα κωδικοποιούνται κατά τη συμπίεση και αποκωδικοποιούνται κατά την αποσυμπίεση. Αυτοί οι μέθοδοι χρησιμοποιούνται όταν δεν πρέπει να χαθεί ούτε ένα μπιτ δεδομένων όπως στην περίπτωση ενός αρχείου κειμένου ή ενός προγράμματος.

Τα γνωστά προγράμματα 7z, bz, zip και rar χρησιμοποιούν μη απωλεστική συμπίεση.

Απωλεστική συμπίεση

Η απώλεια δεδομένων μπορεί να μην είναι αποδεκτή σε αρχεία κειμένου ή ενός προγράμματος, είναι όμως αποδεκτή σε εικόνες και ταινίες. Ο λόγος είναι ότι τα μάτια μας και τα αφτιά μας δεν μπορούν να διακρίνουν πολύ μικρές αλλαγές. Για τέτοιες περιπτώσεις είναι κατάλληλες οι απωλεστικές μέθοδοι συμπίεσης (lossy data compression). Οι μέθοδοι αυτές είναι οικονομικότερες και απαιτούν λιγότερο χρόνο και χώρο όταν πρέπει να σταλούν εκατομμύρια μπιτ εικόνων και βίντεο το δευτερόλεπτο.

Χαρακτηριστικό παράδειγμα απωλεστικής συμπίεσης εικόνας είναι η μέθοδος JPEG (Joint Photographic Experts Group) για βίντεο η μέθοδος MPEG (Moving Pictures Experts Group) και για ήχο το πρότυπο mp3.

ΚΕΦΑΛΑΙΟ 4: ΒΑΣΙΚΟΙ ΑΛΓΟΡΙΘΜΟΙ ΣΥΜΠΙΕΣΗΣ ΔΕΔΟΜΕΝΩΝ

Αλγόριθμος RLE

Ο αλγόριθμος RLE (Run Length Encoding) αποτελεί μια από τις απλούστερες τεχνικές συμπίεσης. Σύμφωνα με τη μέθοδο αυτή, διατρέχεται η ακολουθία των bytes που αποτελούν τα δεδομένα προς συμπίεση και εντοπίζονται οι διαδοχικές επαναλήψεις του ίδιου χαρακτήρα. Στη συνέχεια αντικαθίστανται οι συνεχόμενες επαναλήψεις με το πλήθος τους, ακολουθούμενο από τον χαρακτήρα.



Αν συμπίεσουμε την ακολουθία AAABBEBEEEAADD με τη μέθοδο RLE, τότε η συμπίεσμένη μορφή της ακολουθίας θα είναι 3A3B3E2A3D, που έχει μήκος 10 χαρακτήρες και όχι 14 όπως η αρχική.

Αν προσέξουμε το σχήμα της παραπάνω συμπίεσης θα παρατηρήσουμε ότι αποδίδει ικανοποιητικά για ακολουθίες δεδομένων που έχουν συχνές επαναλήψεις χαρακτήρων. Αν για παράδειγμα συμπίεσουμε την ακολουθία «ABCDE» μήκους 5 bytes με τον αλγόριθμο RLE, η συμπίεσμένη μορφή της είναι «1A1B1C1D1E», με μήκος 10. Στην περίπτωση αυτή όχι μόνο δεν έγινε συμπίεση, αλλά διπλασιάστηκε το μήκος της ακολουθίας. Για να αποφευχθούν αυτές οι περιπτώσεις, ο RLE δε συμπιέζει τις ακολουθίες μη συνεχόμενων χαρακτήρων.

Στην πράξη ο αλγόριθμος διατρέχει την ακολουθία των δεδομένων και αναζητά διαδοχικές επαναλήψεις του ίδιου χαρακτήρα. Όταν βρεθεί μεμονωμένος χαρακτήρας

που δεν επαναλαμβάνεται, ή διπλή επανάληψη του, μένει ως έχει, καθώς δεν επωφελούμαστε από την κωδικοποίηση RLE που απαιτεί 2 χαρακτήρες. Όταν όμως βρεθούν διαδοχικές επαναλήψεις ενός χαρακτήρα, τις αντικαθιστά με ένα ζεύγος (αριθμός, χαρακτήρας), όπου «αριθμός» είναι το σύνολο των συνεχόμενων εμφανίσεων του χαρακτήρα.

Αλγόριθμος του Huffman

Αυτή η τεχνική συμπίεσης παρουσιάστηκε από τον D. Huffman το 1952 και επιτυγχάνει μεγάλα ποσοστά συμπίεσης εκμεταλλευόμενη τη στατιστική ανάλυση της ακολουθίας δεδομένων. Η ανάλυση αυτή είναι πιο σύνθετη από την απλή ανίχνευση των συνεχόμενων επαναλήψεων που κάνει ο RLE.

Συμφώνα με την κωδικοποίηση του προτύπου ASCII, σε κάθε χαρακτήρα αντιστοιχεί ένα byte, που είναι ο κωδικός ASCII του χαρακτήρα αυτού. Αν η κωδικοποίηση γίνεται με τον νέο κώδικα Unicode, χρειάζονται 2 bytes για κάθε χαρακτήρα. Το συνολικό μέγεθος της ακολουθίας σε bytes είναι ίσο με το πλήθος των χαρακτήρων επί 1 ή 2 bytes ανά χαρακτήρα (ανάλογα με την κωδικοποίηση).

Στην κωδικοποίηση Huffman, επιτυγχάνεται συμπίεση καθώς ο κώδικας που αντιστοιχεί σε κάθε χαρακτήρα δεν έχει σταθερό μήκος αλλά μεταβλητό. Η κωδικοποίηση για τους πιο συχνά εμφανιζόμενους χαρακτήρες είναι μικρότερη από ό,τι για τους λιγότερο συχνά εμφανιζόμενους, με αποτέλεσμα το συνολικό μέγεθος να είναι μικρότερο, καθώς για τους συχνότερους χρησιμοποιούνται μόνο 2-3 bits και όχι 8 ή 16.

Ο αλγόριθμος συμπίεσης Huffman αποτελείται από τα εξής στάδια:

Βήμα 1: Μετράμε τη συχνότητα του κάθε χαρακτήρα στην ακολουθία

Βήμα 2: Ταξινομούμε τις συχνότητες εμφάνισης σε μια λίστα κατά φθίνουσα τάξη

Βήμα 3: Κατασκευάζουμε ένα «δένδρο» για την κωδικοποίηση ξεκινώντας με τους συχνότερα εμφανιζόμενους χαρακτήρες

Βήμα 4: Αντιστοιχίζουμε τα δυαδικά '0' και '1' σε κάθε κόμβο του δέντρου: Αρχίζοντας από την ρίζα του δέντρου, προσθέτουμε '0' για κάθε αριστερό παιδί και '1' για κάθε δεξί. Οι χαρακτήρες που κωδικοποιούνται είναι τα φύλλα στην βάση του δέντρου. Αρχίζοντας από την κορυφή (ρίζα) του δέντρου, διατρέχοντας το μοναδικό μονοπάτι προς κάθε φύλλο, συλλέγουμε 0 ή 1 και ορίζουμε τον κώδικα για το χαρακτήρα που αντιστοιχεί στο φύλλο αυτό.

Έστω ότι θέλουμε να κωδικοποιήσουμε τη λέξη «ΑΛΛΟΣ».

Βήμα 1,2: Δημιουργούμε μια λίστα με τους χαρακτήρες ταξινομώντας τους με σειρά εμφάνισης. Έτσι έχουμε τη λίστα {«Λ»,«Α»,«Ο»,«Σ»} αφού το «Λ» εμφανίζεται δύο φορές και οι υπόλοιποι χαρακτήρες από μία (τους οποίους και τοποθετούμε στη λίστα με τη σειρά εμφάνισης τους).

Βήμα 3: Ξεκινώντας από τη ρίζα, δημιουργούμε δύο παιδιά: το «Λ» (το γράμμα που εμφανίζεται πρώτο στη λίστα) και έναν εσωτερικό κόμβο. Στη συνέχεια από τον κόμβο αυτό δημιουργούμε πάλι δύο παιδιά, ένα με τον επόμενο χαρακτήρα στη λίστα, δηλαδή το «Α» και έναν εσωτερικό κόμβο. Διαγράφουμε το «Λ» από τη λίστα. Επαναλαμβάνουμε την διαδικασία μέχρι να τελειώσουν όλοι οι χαρακτήρες της λίστας.

Τελικά όλοι οι χαρακτήρες «Λ», «Α», «Ο», και «Σ» αποθηκεύονται σε φύλλα του δέντρου.

Βήμα 4: Αντιστοιχίζουμε τα δυαδικά 0 και 1, διατρέχοντας το μοναδικό μονοπάτι προς κάθε φύλλο-χαρακτήρα.

Τελικά, η ακολουθία «ΑΛΛΟΣ» που είχε αρχικό μήκος $5 \times 8 = 40$ bits, συμπιεσμένη γίνεται: **1000110111**, συνολικού μήκους 11 bits, άρα περίπου 4 φορές

μικρότερη. Βέβαια, στην πράξη, στο τέλος κάθε ακολουθίας προστίθεται και το σχήμα της κωδικοποίησης, ώστε να μπορεί να γίνει η αποσυμπίεση.

Η ακολουθία κωδικοποίησης που παράγεται με τον προηγούμενο αλγόριθμο είναι μοναδική, ώστε να μπορεί να γίνεται η αποσυμπίεση. Έτσι, αν διαθέτουμε το σχήμα της κωδικοποίησης (πίνακας στο πλάι) και την κωδικοποιημένη ακολουθία, τότε μπορούμε να κάνουμε εύκολα την αποσυμπίεση κωδικοποιημένης ακολουθίας **1000110111**. Διαβάζεται από αριστερά προς τα δεξιά το πρωτο 1. Δεν υπάρχει χαρακτήρας που να αντιστοιχεί σε αυτό. Διαβάζεται και το 0. Ο χαρακτήρας A αντιστοιχεί στο 10, άρα αντικαθίσταται το 10 με αυτόν. Στη συνέχεια διαβάζουμε το 0, και ο μοναδικός χαρακτήρας που αρχίζει με 0 είναι ο Λ. Όμοια και για τα υπόλοιπα γράμματα.

Συμπίεση LZW

Πολλά αρχεία, ιδιαίτερα αρχεία χαρακτήρων, έχουν συμβολοσειρές που επαναλαμβάνονται συχνά, για παράδειγμα το άρθρο «τον». Για την αναπαράσταση της λέξης αυτής με χρήση του κώδικα ASCII χρειάζονται 5 bytes, αν συμπεριλάβουμε τα κενά πριν και μετά τη λέξη. Αν αντιστοιχίσουμε έναν αριθμό σε ολόκληρη τη λέξη (π.χ. το 256 που έχει 2 bytes μήκος), τότε με 2 bytes αντικαθιστούμε παντού τα 5 bytes του άρθρου «τον».

Αυτή είναι η προσέγγιση που ακολουθείται στον αλγόριθμο LZW, ο οποίος παρουσιάστηκε από τους J. Ziv και A. Lempel το 1977, και τροποποιήθηκε από τον T. Welch το 1984.

Ο αλγόριθμος ξεκινά φτιάχνοντας ένα λεξικό όλων των ακολουθιών χαρακτήρων (συμβολοσειρών) που εμφανίζονται στο υπό συμπίεση αρχείο. Στο λεξικό αυτό, που ονομάζεται και πίνακας αναφορών, αποθηκεύεται για κάθε συμβολοσειρά ο αριθμός των εμφανίσεων της και δίπλα ένας μοναδικός κωδικός που αντιστοιχεί σε αυτή. Στη φάση της συμπίεσης αντικαθιστούμε κάθε συμβολοσειρά με τον κωδικό της, ο οποίος έχει πολύ μικρότερο μέγεθος από αυτή. Επιπλέον, στο τέλος προσθέτουμε και το λεξικό, ώστε να μπορεί να γίνει η αποσυμπίεση. Για την αποσυμπίεση γίνεται κατ' αρχήν ανάγνωση του λεξικού και αντικατάσταση των κωδικών με τις αντίστοιχες αρχικές συμβολοσειρές. Ο αλγόριθμος LZW χρησιμοποιείται με επιτυχία για αρχεία με μεγάλη κανονικότητα και συχνές επαναλήψεις μεγάλων συμβολοσειρών.

Τα αρχεία τύπου «zip» δημιουργούνται με συμπίεση βασισμένη στον αλγόριθμο LZW.

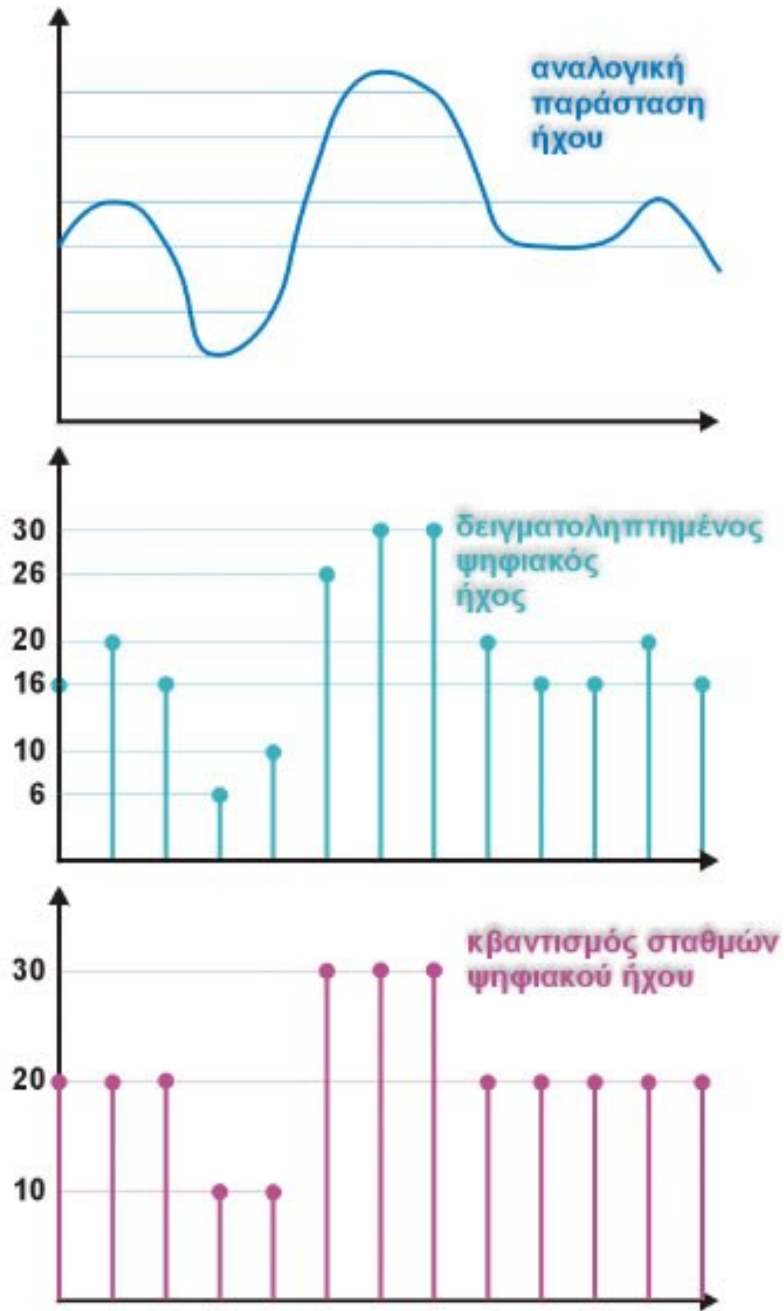
ΚΕΦΑΛΑΙΟ 5: ΚΩΔΙΚΟΠΟΙΗΣΗ ΔΕΔΟΜΕΝΩΝ

Ψηφιακή Παράσταση Ήχου

Η ψηφιοποίηση του αναλογικού ήχου γίνεται με την περιοδική λήψη δειγμάτων από το αναλογικό σήμα πολλές φορές το δευτερόλεπτο, η οποία λέγεται *δειγματοληψία* (sampling). Ο αριθμός των δειγμάτων που παίρνουμε ανά δευτερόλεπτο, ώστε ο ψηφιακός ήχος να έχει την ίδια ποιότητα με τον αναλογικό, καθορίζεται από τη μέγιστη συχνότητα που εμφανίζει ο αναλογικός ήχος μας. Ο αριθμός αυτός πρέπει να είναι τουλάχιστο ίσος με το διπλάσιο της μέγιστης συχνότητας του ήχου, σύμφωνα με το θεώρημα του Shannon. Το ανθρώπινο αυτί αντιλαμβάνεται ήχους συχνοτήτων από 20Hz έως 20KHz. Έτσι, για να έχουμε πιστή αναπαραγωγή του αναλογικού ήχου χρειάζονται πάνω από 40.000 δείγματα ανά δευτερόλεπτο.

Το πλήθος των δειγμάτων ή αλλιώς ο *ρυθμός* (rate) ή *συχνότητα δειγματοληψίας* (sampling frequency), δεν είναι το μόνο στοιχείο που καθορίζει την κωδικοποίηση του ψηφιακού ήχου. Το κάθε δείγμα αντικατοπτρίζει την ένταση του ήχου για τη στιγμή της δειγματοληψίας στην οποία αντιστοιχεί. Στον αναλογικό ήχο, οποιαδήποτε τιμή έντασης είναι επιτρεπτή. Στον ψηφιακό ήχο όμως, το *πλήθος των bits* (number of bits), που χρησιμοποιούμε για την αποθήκευση του κάθε δείγματος, καθορίζει και τον αριθμό των διαφορετικών τιμών εντάσεως που μπορεί να εμφανιστεί. Συνεπώς, κοντινές, αλλά διαφορετικές τιμές αναλογικής έντασης αντιστοιχούν στην ίδια ψηφιακή τιμή. Το φαινόμενο αυτό λέγεται *κβαντισμός* (quantization) των σταθμών έντασης του ήχου. Είναι φανερό ότι όσο πιο πολλά bits χρησιμοποιούμε για την αποθήκευση του κάθε δείγματος, τόσο πιο πιστή αναπαράσταση του αναλογικού ήχου πετυχαίνουμε. Αυτός ο τρόπος κωδικοποίησης λέγεται *παλμοκωδική κωδικοποίηση* (Pulse Code Modulation).

Η παρακάτω εικόνα δείχνει πως ακριβώς γίνεται η κωδικοποίηση του ήχου σε ψηφιακά δεδομένα:



Ήχος σε ποιότητα CD χρειάζεται δείγματα των 16 bits (2 bytes) και ρυθμό δειγματοληψίας 44,1 KHz (χιλιάδες δείγματα ανά δευτερόλεπτο) για κάθε κανάλι. Έτσι για ψηφιακό στερεοφωνικό ήχο (2 κανάλια ήχου) ποιότητας CD έχουμε 176.400 bytes/δευτερόλεπτο· ένα τραγούδι διάρκειας πέντε λεπτών καταλαμβάνει περίπου 50MB.

Αν οι ανάγκες μας σε ποιότητα ήχου είναι μεγάλες, τότε η παλμοκωδική κωδικοποίηση έχει πολύ μεγάλες αποθηκευτικές απαιτήσεις. Σε συνήθεις εφαρμογές πολυμέσων δεν έχουμε απαιτήσεις στερεοφωνικού ήχου ή ακουστικών συχνοτήτων άνω

των 4-5KHz (συχνότητα δειγματοληψίας 11KHz) και αφιερώνουμε 1 byte για κάθε δείγμα. Έτσι για το παραπάνω τραγούδι των 5 λεπτών χρειαζόμαστε το 1/16 των 50MB.

Το επόμενο βήμα για να μειώσουμε περισσότερο το μέγεθος των ηχητικών δεδομένων είναι να μην κωδικοποιούμε ξεχωριστά το κάθε δείγμα δίνοντας του π.χ. 8 bits. Είναι πιο οικονομικό να βρίσκουμε τη διαφορά του κάθε δείγματος με το προηγούμενο και να κωδικοποιούμε αυτήν. Επειδή ο αριθμός των δειγμάτων ανά δευτερόλεπτο είναι μεγάλος, η πιθανότητα δυο διαδοχικά δείγματα να έχουν κοντινές τιμές έντασης είναι μεγάλη, άρα έχουμε να κωδικοποιήσουμε μικρές διαφορές και δεσμεύουμε λίγα bits, συνήθως 2-3 αντί των 8. Η κωδικοποίηση αυτή λέγεται διαφορική παλμοκωδική (Differential Pulse Code Modulation). Μια παραλλαγή της μεθόδου αυτής που επιτρέπει την ύπαρξη μεγαλύτερων διαφορών λέγεται προσαρμοστική-διαφορική παλμοκωδική (Adaptive Differential Pulse Code Modulation) και χρησιμοποιείται στα αρχεία .wav των Windows.

Όλες αυτές οι τεχνικές αντιμετωπίζουν το ηχητικό σήμα σαν μαθηματική κυματομορφή, αγνοώντας τις ιδιαιτερότητες του ανθρώπινου αυτιού. Έτσι, όλες οι συχνότητες αντιμετωπίζονται ισοδύναμα. Στην πράξη όμως το ανθρώπινο αυτί είναι περισσότερο ευαίσθητο σε κάποια ζώνη συχνοτήτων από κάποια άλλη. Νέες τεχνικές κωδικοποίησης που βασίζονται σε αυτή την απλή παρατήρηση, επιτυγχάνουν πολύ μεγαλύτερα ποσοστά συμπίεσης, χωρίς να υπάρχει ιδιαίτερη απώλεια σε ποιότητα ήχου. Το πρότυπο MPEG-3 (Motion Picture Expert Group) πετυχαίνει συμπίεση έως και 12 φορές ως προς την αρχική μορφή χωρίς εμφανείς απώλειες στην ποιότητα.

Τα αρχεία .mp3 είναι κωδικοποιημένα με το πρότυπο αυτό. Αν όμως δεχθούμε συμβιβασμούς στην ακουστική ποιότητα, π.χ. για μονοφωνικό ήχο ποιότητας τηλεφωνικής μετάδοσης, τότε επιτυγχάνεται συμπίεση μέχρι και 100 φορές.

Άλλα πρότυπα

Παρά το γεγονός της η ριζική και ραγδαία εξέλιξη του MP3 το οποίο έχει γνωρίσει ευρεία αποδοχή μεταξύ χρηστών, εντούτοις, επικρατεί πληθώρα διαφορετικών τεχνολογιών που ολοένα και εξελίσσονται με πρωταρχικό στόχο την καλύτερη συμπίεση ψηφιακού ήχου. Τέτοιες προσπάθειες, έχουν υλοποιηθεί από την εταιρία Microsoft με το δικό της γνωστό και διαδεδομένο πρότυπο Windows Media Audio (WMA), το AAC (Advanced Audio Coding), το MP3 Pro, το OGG και το AC-3 Dolby Digital.

WMA (Windows Media Audio)

Η μεταφορά, η αποθήκευση και η χρήση ακουστικού υλικού με βάση την μορφή απωλεστικών συμπιεσμένων ηχητικών αρχείων μέσω υπολογιστή, ωθεί τον μεγαλύτερο κατασκευαστή λειτουργικών συστημάτων μα μην μείνει έξω από το παιχνίδι. Το πρότυπο Windows Media Audio (WMA) της εταιρία Microsoft, προσφέρει όμοιες δυνατότητες με το MP3, με άριστη ποιότητα τόσο αναπαραγωγής όσο και μεγαλύτερη συμπίεση (64 kbps).

Πιο συγκεκριμένα, το WMA αποτελεί ένα σύστημα κωδικοποίησης / αποκωδικοποίησης ήχου, επιτρέποντας την συμπίεση ψηφιακών δεδομένων ήχου στο 1/20 του αρχικού τους όγκου και την εγγραφή τους σε ένα μόνο δίσκο CD με επακόλουθο τα τραγούδια που είναι προστατευμένα να μην μπορούν να μεταδοθούν ελεύθερα.

Συμπερασματικά, γι' αυτό ακριβώς το λόγω ο μεγαλύτερος αριθμός δισκογραφικών εταιριών χρησιμοποιεί στα πλαίσια υλοποίησης του έργου τους το πρότυπο αυτό.

MP3 Pro

Τον Ιανουάριο του 2001 στη CES, παρουσιάστηκε από την Coding Technologies η τεχνολογία Mp3 Pro, μια βελτιωμένη έκδοση του Mp3 με δυνατότητα να προσφέρει όμοια ποιότητα στο μισό μέγεθος των αρχείων, γεγονός που υλοποιείται με μεγαλύτερη συμπίεση δεδομένων. Συγκεκριμένα, η συμπίεση στα 64kbps και 96kbps, προσφέρει τη ίδια απόδοση ήχου με τα 128kbps και 192kbps του Mp3.

MP3 Surround

Το 2004 το Ινστιτούτο Fraunhofer IIS παρουσίασε μία πολυκαναλική έκδοση του MP3, το MP3 Surround το οποίο βασίζεται στην τεχνολογία Binaural Cue Coding της Agere.

Η τεχνολογία αυτή, επιτρέπει την μείξη σημάτων από πολλά κανάλια σε δύο, με στόχο την δημιουργία ενός σήματος συμβατού με τον απλό MP3 codec, ενώ κωδικοποιεί μία σειρά από παραμέτρους που περιγράφουν πλήρως το ηχητικό πεδίο

surround. Τέτοιες παράμετροι είναι, οι χρονικές διαφορές μεταξύ των καναλιών, οι διαφορές στάθμης μεταξύ των καναλιών και η συσχέτιση μεταξύ των καναλιών.

AAC (Advance Audio Coding)

Το πρότυπο ACC αναπτύχθηκε για πρώτη φορά το 1997 από το Ινστιτούτο Fraunhofer IIS και χρησιμοποιεί όπως και το MP3 το ψυχοακουστικό μοντέλο επικάλυψης, με σκοπό να καλύψει τα προβλήματα που υπήρχαν σχετικά με την ποιότητα των αρχείων MP3 στους μικρούς αριθμούς μετάδοσης των δεδομένων. Το πρότυπο ACC, ως ένα πραγματικό αριστούργημα κωδικοποίησης έχει την ικανότητα να αποδίδει εξαιρετικά υψηλή ποιότητα ήχου σε bitrate 64Kb/sec, επιτρέποντας τόσο την κωδικοποίηση 48 καναλιών ήχου και 16 καναλιών χαμηλής συχνότητας για εφέ όσο και την υποστήριξη πολλών γλωσσών ταυτόχρονα.

Παράλληλα, το ACC διακατέχεται από τρεις διαφορετικές όψεις, την «κύρια», την «χαμηλής πολυπλοκότητα» και την «κλιμακούμενη συχνότητα δειγματοληψίας». Η «κύρια» όψη αναφέρεται σε εφαρμογές που η υπολογιστική ισχύει και εφαρμογές δεν είναι περιορισμένη, η «χαμηλής πολυπλοκότητα» σε εφαρμογές που η ισχύος και η μνήμη βρίσκονται σε μεγάλη ζήτηση, ενώ η τελευταία, είναι φτιαγμένη έτσι ώστε οι αποκωδικοποιητές να έχουν ελάχιστες απαιτήσεις σε μνήμη και ισχύ.

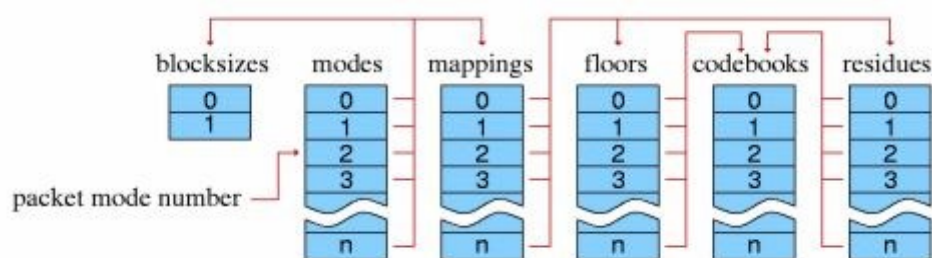
Υποκειμενικά τεστ που έχουν πραγματοποιηθεί με καλά εκπαιδευμένους ακροατές, έδειξαν ότι η συγκεκριμένη κωδικοποίηση προσφέρει καλύτερη ποιότητα ήχου από οποιαδήποτε άλλη κωδικοποίηση ήχου με το μισό μόνο bitrate. Στο σημείο αυτό, οφείλουμε να αναφέρουμε ότι το πρότυπο ACC παρέχει καλύτερη απόδοση από το MP3, ενώ το 2003 η έκδοση του παρουσιάζεται συμβατή με τις προδιαγραφές του Mpeg ούτως ώστε το πρότυπο να αναφέρεται και ως Mpeg-4 ACC.

RA-Real Audio

Το Real Audio ως κλειστό πρότυπο, δημιουργήθηκε και υποστηρίχθηκε από την εταιρία Real Network με σκοπό την αναπαραγωγή ήχων στο Διαδίκτυο χωρίς να προηγείται κατέβασμα των ήχων στο σκληρό δίσκο του υπολογιστή. Το πρότυπο Real Audio, είναι αρκετά δημοφιλές και αυτό εξαιτίας της ελεύθερης διάθεσης του λογισμικού ανάγνωσης των αρχείων ήχου τέτοιου τύπου υποστηρίζοντας, μεγάλη συμπίεση και κατακανόνα χαμηλή ποιότητα ήχου.

OGG Vorbis

Ο codec Ogg Vorbis αναπτύχθηκε γύρω από το πρότυπο αρχείων Ogg και βασίζεται στη open source εφαρμογή απωλεστικής συμπίεσης με την ονομασία Vorbis. Ως προς τον τρόπο κωδικοποίησης, ο τρόπος μοιάζει με αυτό του Mp3 ενώ ταυτόχρονα ο Ogg Vorbis χρησιμοποιεί MDCT για τον μετασχηματισμό του σήματος από το πεδίο του χρόνου στο πεδίο της συχνότητας, καθώς και μία εναλλακτική μέθοδο επεξεργασίας του φάσματος, κατά την οποία κωδικοποιείται το φάσμα βάσης του οποίου η κατανομή είναι σχετικώς ομαλή και με περισσότερα ψηφία το απομένον φάσμα που η δομή και η χρονική εξέλιξη είναι πολύ πιο πολύπλοκη. Συμπερασματικά, η τακτική αυτή σε συνδυασμό με την καλή ποιότητα ήχου που προσφέρει, ωθεί το πρότυπο Ogg σε ένα ανταγωνιστικό παιχνίδι ως προς το WMA και MP3.



Διάγραμμα βαθμίδων του κωδικοποιητή Ogg Vorbis. Η κωδικοποίηση του φάσματος βάσης και του απομένοντος φάσματος.

AC3 Dolby Digital

Ένα από τα πιο διαδεδομένα πρότυπα για τον ψηφιακό πολυκάναλο ήχο είναι το AC3, που εκτός σημαντικού απροόπτου άρχισε να γίνεται το διεθνές πρότυπο για την συμπίεση ηχητικών δεδομένων. Στο ψηφιακό σύστημα ήχου AC3, ο ήχος κωδικοποιείται σε έξι συνολικά κανάλια.

Συγκεκριμένα, υπάρχουν:

(α) τρία κανάλια (αριστερό, κεντρικό, δεξί) που αποσκοπούν στο να φέρουν την βασική ηχητική πληροφορία,

(β) δύο συνοδευτικά κανάλια περιβάλλοντος ήχου και

(γ) ένα κανάλι για τις υπόλοιπες συχνότητες (σύστημα 3/2/.1).

Ως προς τον τρόπο λειτουργίας τους, τα πέντε πρώτα κανάλια διαχειρίζονται συχνότητες ήχου στο διάστημα 3-20000Hz, ενώ το έκτο συχνότητες 3-120Hz. Συμπερασματικά, ο ρυθμός δειγματοληψία είναι 48KHz μεγαλύτερος από το ρυθμό των 44KHz που χρησιμοποιείται στα CDs, ενώ η συμπίεση των ηχητικών δεδομένων ανέρχεται στην αναλογία 10:1.

ΚΕΦΑΛΑΙΟ 6: ΨΗΦΙΑΚΗ ΠΑΡΑΣΤΑΣΗ ΕΙΚΟΝΑΣ

Μία ψηφιακή εικόνα αποτελείται από μικρές κουκίδες που ονομάζονται εικονοστοιχεία (pixels). Στην πιο απλή περίπτωση κάθε εικονοστοιχείο μπορεί να είναι είτε άσπρο είτε μαύρο.

Στο παρακάτω σχήμα βλέπουμε μία ασπρόμαυρη ψηφιακή εικόνα. Η εικόνα του σχήματος αποτελείται από 256 γραμμές και 256 στήλες, περιέχει δηλαδή 65.536 διαφορετικά εικονοστοιχεία. Επειδή η φωτογραφία είναι ασπρόμαυρη, κάθε εικονοστοιχείο κωδικοποιείται με ένα bit: το 0 συμβολίζει το μαύρο χρώμα και το 1 το άσπρο. Έτσι, για την αποθήκευση αυτής της φωτογραφίας στον υπολογιστή χρειαζόμαστε 65.536 bits ή ισοδύναμα 8.192 bytes.



Για έγχρωμες εικόνες, κάθε εικονοστοιχείο απαιτεί περισσότερα από 1 bit για την αναπαράσταση του αντίστοιχου χρώματος. Για παράδειγμα, αν έχουμε 16 διαφορετικά χρώματα και κάθε χρώμα εμφανίζει 16 αποχρώσεις, τότε χρειαζόμαστε $16 \times 16 = 256$ διαφορετικές τιμές. Όλες αυτές μπορούμε να τις αναπαραστήσουμε με ένα byte (8 bits). Αν, για παράδειγμα, έχουμε μια εικόνα μεγέθους 640×480 , δηλαδή

307.200 εικονοστοιχεία και για καθένα χρησιμοποιούμε 3 bytes=24bit για την αναπαράσταση του χρώματος, τότε χρειαζόμαστε 921.600 bytes. Τα αρχεία που περιέχουν την εικόνα ασυμπίεστη, ονομάζονται χαρτογραφικά (bitmap, BMP).

Το μέγεθος των ασυμπίεστων εικόνων αυξάνει δραματικά με την αύξηση των διαστάσεων της ή του βάθους χρώματος (δηλ. τον αριθμό των bytes για την αναπαράσταση των διαφορετικών χρωμάτων). Η χρησιμοποίηση της μεθόδου συμπίεσης RLE σε εικόνες που είναι σκίτσα ή σχεδιαγράμματα, κυρίως με αποχρώσεις του γκρι, οδηγεί σε σημαντική μείωση του όγκου τους. Η RLE μέθοδος είναι πολύ αποτελεσματική για εικόνες με συνεχόμενες περιοχές εικονοστοιχείων που έχουν το ίδιο χρώμα, καθώς τα αντικαθιστά με ένα κωδικό και το πλήθος τους. Τα αρχεία αυτά είναι τα PCX αλλά και πολλές μορφές TIFF.

Η συμπίεση LZW εφαρμόζεται σε περισσότερο σύνθετες εικόνες, όπως π.χ. έγχρωμες φωτογραφίες. Εκεί, υπάρχει μεγάλη ποικιλία χρωμάτων, που όμως επαναλαμβάνονται πολλές φορές τα ίδια, άρα, όπως είδαμε στο κεφάλαιο 2.3, η κωδικοποίηση τους με λιγότερα bit φέρνει μεγάλα ποσοστά συμπίεσης. Τέτοια συμπίεση υπάρχει στα αρχεία GIF.

Τέλος, αντίθετα με τις προηγούμενες μεθόδους αναπαράστασης που χρησιμοποιούσαν μη απωλεστικούς αλγόριθμους συμπίεσης, υπάρχει μια ευρύτατα διαδεδομένη μέθοδος για εικόνες, που επιτυγχάνει συμπίεση μέχρι και 100-200 φορές σε ορισμένες περιπτώσεις, με μικρή απώλεια της ποιότητας. Η κωδικοποίηση αυτή λέγεται κωδικοποίηση JPEG (Joint Photographic Expert Group). Βασίζεται στην εφαρμογή μιας σειράς πολύπλοκων μαθηματικών μετασχηματισμών στην εικόνα. Για τη συμπίεση ή την αποσυμπίεση από JPEG χρειάζεται αρκετή υπολογιστική ισχύς, κάτι που παλαιότερα ήταν ιδιαίτερα χρονοβόρα. Σήμερα, που υπάρχουν ισχυροί επεξεργαστές, η κωδικοποίηση JPEG δεν είναι πια πρόβλημα, και έχει το πλεονέκτημα της δυνατότητας επιλογής του ποσοστού συμπίεσης που θέλουμε να επιτύχουμε, με αντίστοιχη επιβάρυνση στην ποιότητα της εικόνας. Συνήθως συμπίεσεις 10 έως και 20 φορές, οδηγούν σε εικόνες με μη ορατή διαφορά από την αρχική.

Συμπύεση Δεδομένων για κινούμενες εικόνων

Γενικές αρχές της συμπύεσης video

Ποιοτική ανοχή

Σε αντίθεση με την αντίληψη ότι η ψηφιακή μετάδοση θα πρέπει να έχει το εύρος ζώνης που χρειάζεται για να μεταδώσει και το πιο απαιτητικό πλαίσιο (frame) μιας σειράς κινούμενων εικόνων, ακόμα και αν αυτό παρουσιάζεται με συχνότητα 1%, τώρα πια κάτω από την εμπορική πίεση που υπάρχει (λόγοι κόστους) θεωρείται λογικό να δεχόμαστε κάποιο ποσοστό παραμόρφωσης σε τέτοιες σπάνιες σκηνές, με αντάλλαγμα να μπορούμε να χρησιμοποιήσουμε τα bit που εξοικονομούμε για την αναβάθμιση του μέσου όρου ανάλυσης του συνόλου των σκηνών.

Έτσι σε αντίθεση με την ηχογράφηση ήχου σε CD, που γίνεται χωρίς καμία συμπύεση, στη συμπύεση video αναζητείται μία χρυσή τομή (sweet spot) ανάμεσα στην ποιότητα και το bandwidth που πολλές φορές βασίζεται σε υποκειμενικές μετρήσεις που γίνονται σε άτομα που θεωρούνται ικανά “δείγματα” πάνω στην εκτίμηση της εικόνα και του ήχου λόγω επαγγελματικής εμπειρίας.

Αν η εμπειρία τους δεν τους επιτρέπει να αντιληφθούν σημαντικές διαφορές στην ποιότητα του κωδικοποιημένου ήχου από τον αρχικό και συγχρόνως θεωρούν ανεκτές τις ατέλειες (artifacts) της εικόνας μετά από κάποιο ποσοστό συμπύεσης, η μέθοδος και το ποσοστό αυτό θεωρούνται αποδεκτά και τυποποιούνται.

Να σημειωθεί εδώ ότι ενώ ο ήχος προσφέρεται για μεγάλα ποσοστά συμπύεσης χωρίς να μπορεί να γίνει αισθητή υποβάθμιση στην ποιότητά του, στον τομέα της εικόνας τα πράγματα είναι λίγο πιο δύσκολα και οι όποιες ατέλειες είναι ορατές, απλά γίνεται προσπάθεια να μην γίνονται ενοχλητικές για το θεατή.

Πλεονασμός (Redundancy)

Η βασική αρχή πάνω στην οποία στηρίζονται όλες οι μέθοδοι ψηφιακής συμπύεσης είναι το γεγονός ότι το σήμα εμπεριέχει ένα ποσοστό πλεονασμού (redundancy). Με τον όρο αυτό εννοούμε την πληροφορία που είτε μπορεί να παραληφθεί, είτε να κωδικοποιηθεί με λιγότερη ακρίβεια, χωρίς αυτό να έχει αξιοσημείωτη επίδραση στο τελικό αποτέλεσμα. Υπάρχουν δύο είδη πλεονασμού:

Στατικός πλεονασμός (Spatial Redundancy):

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

Υποκειμενικός Πλεονασμός (Temporal Redundancy):

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

DCT Coding (Discrete Cosine Transform Coding)

Ο Διακριτός Συνημιτονικός Μετασχηματισμός (Discrete Cosine Transform) είναι μία μέθοδος που βρίσκει μεγάλη εφαρμογή στην ψηφιακή συμπίεση γενικά αλλά και στο MPEG ειδικότερα.

Με το μετασχηματισμό DCT μπορούμε να μεταφέρουμε την πληροφορία που περικλείει η εικόνα από το πεδίο του χώρου στο πεδίο της συχνότητας (αφηρημένο πεδίο), όπου η περιγραφή της μπορεί να γίνει με σημαντικά μικρότερο πλήθος bits, για διάφορους λόγους.

Ο μετασχηματισμός DCT ορίζεται ως εξής :

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

$$(4.1) \quad DCT(i, j) = \frac{1}{\sqrt{2N}} C(i)C(j) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} pixel(x, y) \cos \left[\frac{(2x+1)i\pi}{2N} \right] \cos \left[\frac{(2y+1)j\pi}{2N} \right]$$

όπου $C(x) = 0.7071$, $x = 0$

$$1, x > 0$$

παίρνουμε την τιμή $DCT(i, j)$ που είναι η τιμή του συντελεστή του μετασχηματισμού στο πεδίο της συχνότητας.

Έτσι αντιστοιχίζουμε τις τιμές των pixels στις αντίστοιχες τιμές συντελεστών.

Οι συντελεστές αυτοί μεταφέρουν ο καθένας ένα κομμάτι της αρχικής πληροφορίας (αυτό που αντιστοιχεί στο κομμάτι του φάσματος που περιγράφει).

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

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

$$(4.2) \quad pixel(x, y) = \frac{1}{\sqrt{2N}} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} C(i)C(j) DCT(i, j) \cos \left[\frac{(2x+1)i\pi}{2N} \right] \cos \left[\frac{(2y+1)j\pi}{2N} \right]$$

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

Κβαντοποίηση (Quantization)

Η μέθοδος που μας βοηθάει να απαλλαγούμε από σημαντικό μέρος της πληροφορίας είναι η κβαντοποίηση. Με τον όρο κβαντοποίηση γενικά εννοούμε τη μετατροπή ενός σήματος άπειρων (η πάρα πολλών) τιμών σε ένα σήμα ορισμένων διακριτών τιμών π.χ. η κβαντοποίηση μιας εικόνας που περιέχει εκατομμύρια χρώματα οδηγεί σε μία εικόνα που έχει 256 διαφορετικές τιμές για το χρώμα (πρότυπο JPEG). Με

άλλα λόγια κβαντοποίηση είναι ο περιορισμός των bits με τα οποία περιγράφουμε τα δείγματα του σήματος (προφανώς το 256 έχει πολύ λιγότερα bits από τους τεράστιους αριθμούς με τους οποίους έπρεπε να περιγράψουμε τα δείγματά μας αν δεν γινόταν κβαντοποίηση).

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

Ο αριθμός 45 είναι 101101 έχει δηλαδή 6 bits.

Με 4 bits γίνεται 1011 = 11

Με 3 bits γίνεται 101 = 5 κτλ.

Δηλαδή αν είχαμε διαλέξει να περιγράψουμε το σήμα με 3 bits, τιμές όπως το 45 και το 11 θα έπαιρναν την τιμή 101 (=5). Είναι προφανές ότι η κβαντοποίηση εισάγει σφάλμα ανάλογο με τον αριθμό των bits που απορρίπτονται και κατά συνέπεια ευθύνεται στο μεγαλύτερο βαθμό για την απώλεια πληροφορίας κατά τη συμπίεση (lossy compression) σε αντίθεση με το μετασχηματισμό DCT που είναι μία, σε μεγάλο βαθμό αντιστρεπτή διαδικασία.

Για να περιγράψουμε όσο περισσότερες από τις τιμές του σήματος γίνεται με δοσμένο αριθμό bits, διαιρούμε τις τιμές των δειγμάτων είτε με σταθερές τιμές (uniform quantization) είτε με πίνακες κβαντοποίησης (quantization tables).

Στη συγκεκριμένη περίπτωση, δηλαδή στο πρότυπο MPEG χρησιμοποιείται η δεύτερη μέθοδος και μάλιστα υπάρχει ένας πίνακας για τα πλαίσια που έχουν κωδικοποιηθεί με ενδοπλαισιακή (intra-frame coding) και ένας για αυτά με διαπλαισιακή (inter-frame coding).

Τμηματική Πρόβλεψη Κίνησης (Block Motion Compensation)

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

Για παράδειγμα στο παρακάτω σχήμα έχουμε μία σκακιέρα σε δύο φάσεις που διαφέρουν μεταξύ τους στο ότι κάποια πιόνια έχουν μετακινηθεί.

Εναλλακτικά με το να κωδικοποιήσουμε ανεξάρτητα τα δύο πλαίσια μπορούμε να περιγράψουμε τη δεύτερη εικόνα με το να τη χωρίσουμε σε ίσα τμήματα και να φτιάξουμε ένα πίνακα που να περιέχει τμήματα που έχουν μείνει ίδια και διανύσματα που να δείχνουν τη νέα θέση των τμημάτων που άλλαξαν θέση.

Έτσι αν έχουμε ήδη αποστείλει την πρώτη εικόνα μπορούμε να στείλουμε τη δεύτερη σαν ένα πίνακα 20 διανυσμάτων και ορισμένων σταθερών τμημάτων, που προφανώς έχει πολύ μικρότερο μέγεθος.

Το παραπάνω είναι μία καλή προσέγγιση της μεθόδου αλλά στην πραγματικότητα οι πραγματικές εικόνες δεν θα είναι τόσο όμοιες μεταξύ τους όσο η σκακιέρα.

Θα έχουν κάποια κοινά τμήματα που αλλάζουν θέση από πλαίσιο σε πλαίσιο αλλά θα υπάρχουν και τμήματα που αλλάζουν θέση διατηρώντας το σχήμα τους αλλά μεταβάλλεται το χρώμα τους, καθώς και άλλα που δεν υπάρχουν σε προηγούμενο πλαίσιο αλλά εμφανίζονται σε κάποιο για πρώτη φορά.

Για την κάλυψη αυτών των περιπτώσεων χρησιμοποιείται μία πιο βελτιωμένη εκδοχή της παραπάνω ιδέας (ή για την ακρίβεια διάφορες εκδοχές της παραπάνω ιδέας).

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

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

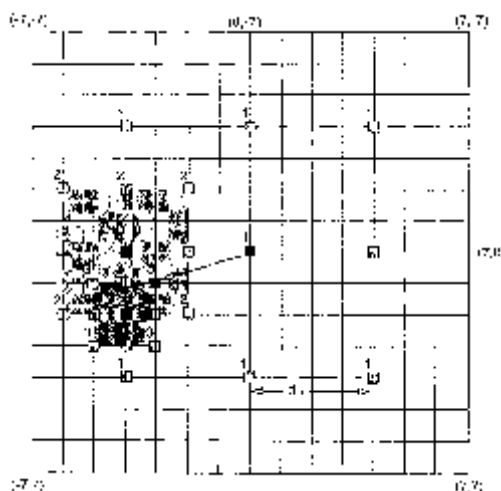
$$(4.3) \quad AE(d_x, d_y) = \sum_{i=0}^{15} \sum_{j=0}^{15} |f(i, j) - g(i - d_x, j - d_y)|$$

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

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

Για να μειωθεί λίγο η υπολογιστική πολυπλοκότητα έχει επινοηθεί η μέθοδος αναζήτησης τριών βημάτων (TSS - Three Step Search). Ο αλγόριθμος υπολογίζει την απόλυτη διαφορά (AE) στο κέντρο και σε οχτώ περιοχές της περιοχής αναζήτησης που είναι 32×32 pixels. Η περιοχή που θα έχει τη μικρότερη απόλυτη διαφορά γίνεται το κέντρο για την επόμενη αναζήτηση, η οποία έτσι έχει το μισό μέγεθος.

Η διαδικασία αυτή επαναλαμβάνεται τρεις φορές, όπως φαίνεται στο παρακάτω σχήμα (η περιοχή του κάθε βήματος περικλείεται από σημεία που φέρουν την αντίστοιχη αρίθμηση).



Ένας από τους πλέον κατάλληλους τρόπους συμπίεσης των παραπάνω εικόνων είναι η τεχνική MPEG (Moving Pictures Experts Group) η οποία χρησιμοποιείται για την συμπίεση κινούμενων εικόνων. Η τεχνική αυτή παρουσιάζει το πλεονέκτημα ότι χρησιμοποιεί την πλέον διαδεδομένη μορφή αποτελεσματικής συμπίεσης ακίνητης

εικόνας JPEG για την κωδικοποίηση ορισμένων εικόνων της χρονοσειράς, και προβλέπει βάση αυτών και με χρήση αλγορίθμων βέλτιστου ταιριάσματος περιοχές στις υπόλοιπες εικόνες που παρουσιάζουν πολύ μικρές διαφορές.

Εξέλιξη Προτύπου MPEG

Η δημιουργία και εξέλιξη του προτύπου MPEG ακολούθησε τα παρακάτω στάδια ή όπως αναφέρονται στη βιβλιογραφία επίπεδα (layers):

MPEG-1: Σχεδιάστηκε για χρήση με ρυθμούς μέχρι 1.5 Mbit/sec και είναι ένα πολύ δημοφιλές πρότυπο στο internet για κωδικοποίηση εικόνας και ήχου. Το 3ο επίπεδο του, είναι γνωστό και ως MP3 (το γνωστό μας πρότυπο για κωδικοποίηση ήχου).

MPEG-2: Σχεδιάστηκε για χρήση με ρυθμούς από 1.5 ως 15 Mbit/sec. Είναι το πρότυπο πάνω στο οποίο βασίζονται οι κωδικοποιητές της ψηφιακής τηλεόρασης Αλγόριθμοι Συμπίεσης Δεδομένων Ολοκληρωτικής Φωτογραφίας (Digital Television set top boxes) και η κωδικοποίηση DVD. Το MPEG-2 είναι η 3.

MPEG-4: Σχεδιάστηκε για κωδικοποίηση εφαρμογών πολυμέσων και διαδικτύου. Βασίζεται σε αντικειμενοστραφή κωδικοποίηση. Τα αντικείμενα (objects) και το υπόβαθρο (background) διαχωρίζονται και κωδικοποιούνται ξεχωριστά. Αυτή η μέθοδος επιτρέπει τη μετάδοση των video σε διαφορετικούς ρυθμούς ανάλογα με το σύστημα που χρησιμοποιούμε.

Γενική Περιγραφή Προτύπου MPEG-1

Η κωδικοποίηση MPEG μειώνει δραματικά το χώρο που απαιτείται για να εγγράψουμε ακολουθίες κινούμενων εικόνων εξαλείφοντας πλεοναστική και μη απαραίτητη πληροφορία από τα αρχικά δεδομένα. Λιγότερα bits σημαίνει πως οι κινούμενες εικόνες μπορούν να μεταφερθούν πολύ γρήγορα, τόσο γρήγορα για την ακρίβεια, που αργά και φτηνά δίκτυα μπορούν να τις μεταφέρουν προσφέροντας ένα

τεράστιο φάσμα νέων δυνατοτήτων. Δύο γειτονικά πλαίσια σε μία ακολουθία κινούμενων εικόνων είναι συνήθως πολύ όμοια. Συχνά, η μόνη διαφορά τους είναι πως κάποια κομμάτια της εικόνας έχουν ελαφρώς μετακινηθεί ανάμεσα στα πλαίσια.

Η κωδικοποίηση MPEG εκμεταλλεύεται αυτή τη χρονική συσχέτιση, διασπώντας κάθε νέο πλαίσιο σε μικρότερα κομμάτια και ψάχνει στο προηγούμενο πλαίσιο να βρει από ποιο σημείο ήρθαν αυτά τα κομμάτια κάνοντας ανίχνευση κίνησης (motion compensation). Αν το περιεχόμενο αυτού του frame είχε σταλεί σχεδόν το ίδιο με το προηγούμενο πλαίσιο, δεν υπάρχει χρησιμότητα στο να το στείλουμε ξανά.

Απλά, στέλνουμε τις οδηγίες για να μετακινήσουμε τα κομμάτια από το προηγούμενο πλαίσιο στις θέσεις που έχουν στο τρέχον. Επιπρόσθετα, μέσα σε ένα πλαίσιο, πολλές περιοχές όπως ο ουρανός ή οι τοίχοι έχουν ακριβώς το ίδιο χρώμα.

Το MPEG εκμεταλλεύεται αυτή τη χωρική συσχέτιση διασπώντας τις εικόνες σε κατάλληλα κομμάτια και μειώνοντας αυτά τα κομμάτια σε ένα μόνο χρώμα. Αν κάποιες περιοχές από pixels έχουν περίπου το ίδιο χρώμα, γιατί να στέλνουμε το ίδιο χρώμα συνέχεια; Απλά μπορούμε να στείλουμε το χρώμα όλης της περιοχής με τη μια. Επίσης, το ανθρώπινο μάτι συγχωρεί την προσέγγιση και την εξάλειψη λεπτομερειών. Αυτό μπορούμε να το εκμεταλλευτούμε καθώς στις εικόνες υπάρχει πολύ περισσότερη λεπτομέρεια από ότι χρήσιμη πληροφορία.

Η κωδικοποίηση MPEG προσεγγίζει την ένταση της λεπτομέρειας με μερικές σκιές γλιτώνοντας πολλά bits σε σχέση με την αυθεντική απεικόνιση. Επίσης, το μάτι αντιλαμβάνεται λιγότερες μεταβολές χρώματος ανά εκατοστό από μεταβολές φωτεινότητας. Το MPEG εκμεταλλεύεται αυτή την ιδιότητα απαλείφοντας ασήμαντες χρωματικές λεπτομέρειες που το ανθρώπινο μάτι ούτως ή άλλως δεν μπορεί να αντιληφθεί.

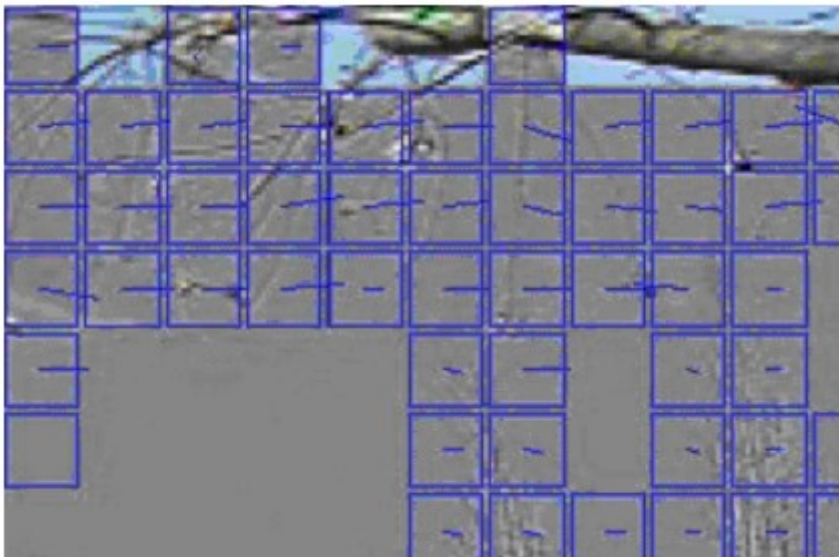
Η εικόνα που ακολουθεί αναπαριστά μια εικόνα η οποία έχει διασπαστεί σε περίπου 300 τετράγωνα χρησιμοποιώντας τις ανωτέρω τεχνικές συμπίεσης. Αυτή η εικόνα αποτελεί μέρος μιας ακολουθίας video που αναπαριστά το πέρασμα μπροστά από μια σειρά σπιτιών που έχουν ,μπροστά τους πολύχρωμα λουλούδια.

Οι γκρι περιοχές υποδηλώνουν περιοχές που είχαν βρεθεί στο προηγούμενο πλαίσιο και έχουν μετακινηθεί στις τωρινές τους θέσεις.



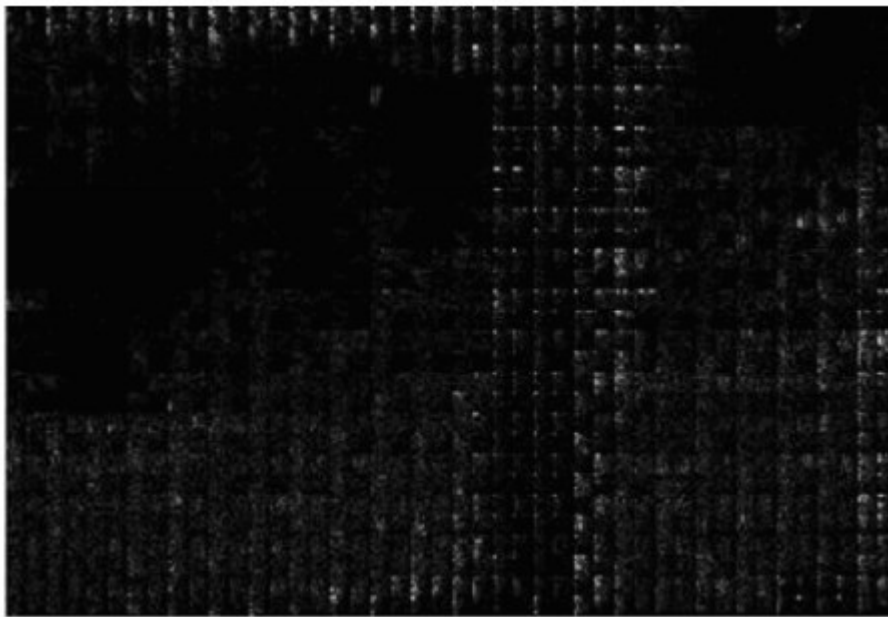
Στην ακόλουθη εικόνα έχουν περιγραφεί με μπλε χρώμα κάποια κομμάτια του πλαισίου και με τις μπλε γραμμές (που καλούνται motion vectors) φαίνεται η ακριβής θέση της συγκεκριμένης περιοχής στο προηγούμενο πλαίσιο.

Το διάνυσμα κίνησης -motion vector - (που μπορεί να χρειάζεται 4 bits να κωδικοποιηθεί) κωδικοποιείται στη ροή του MPEG αντί ολόκληρης της ομάδας εικονοστοιχείων (2048 bits).



Η αυθεντική εικόνα η οποία είχε πάρα πολύ λεπτομέρεια και χρώματα, έχει μετασχηματιστεί στην εικόνα που είδαμε στο πρώτο βήμα της κωδικοποίησης (χρονική συσχέτιση).

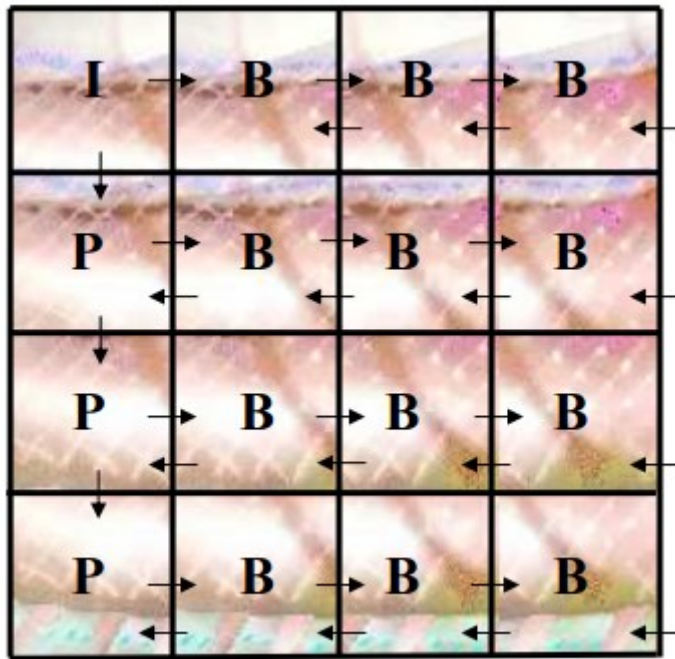
Είναι σημαντικό το γεγονός πως η καινούρια εικόνα έχει σχεδόν ένα χρώμα (γκρι) σε συνδυασμό με κάποιες λεπτομέρειες. Η ακόλουθη εικόνα αναπαριστά την αρχική μετά την προετοιμασία για το δεύτερο βήμα της κωδικοποίησης (χωρική συσχέτιση και απαλοιφή λεπτομερειών).



Η παραπάνω εικόνα είναι ένα μέτρο της λεπτομέρειας που απομένει στην εικόνα μετά το πρώτο βήμα. Οι φωτεινότερες περιοχές αναπαριστούν τις μη κωδικοποιημένες λεπτομέρειες που απομένουν στην εικόνα. Η τελική απεικόνιση είναι σχεδόν μαύρη, δηλαδή αποτελείται από μηδενικά ή μικρούς αριθμούς. Αυτοί οι μικροί αριθμοί απαιτούν μικρό αριθμό από bits για να αποθηκευτούν και μεταφερθούν. Οι αρχικά πολύχρωμη και πολύπλοκη ακολουθία εικόνων μετατράπηκε σε μια απεικόνιση εύκολο να αποθηκευτεί και να μεταδοθεί.

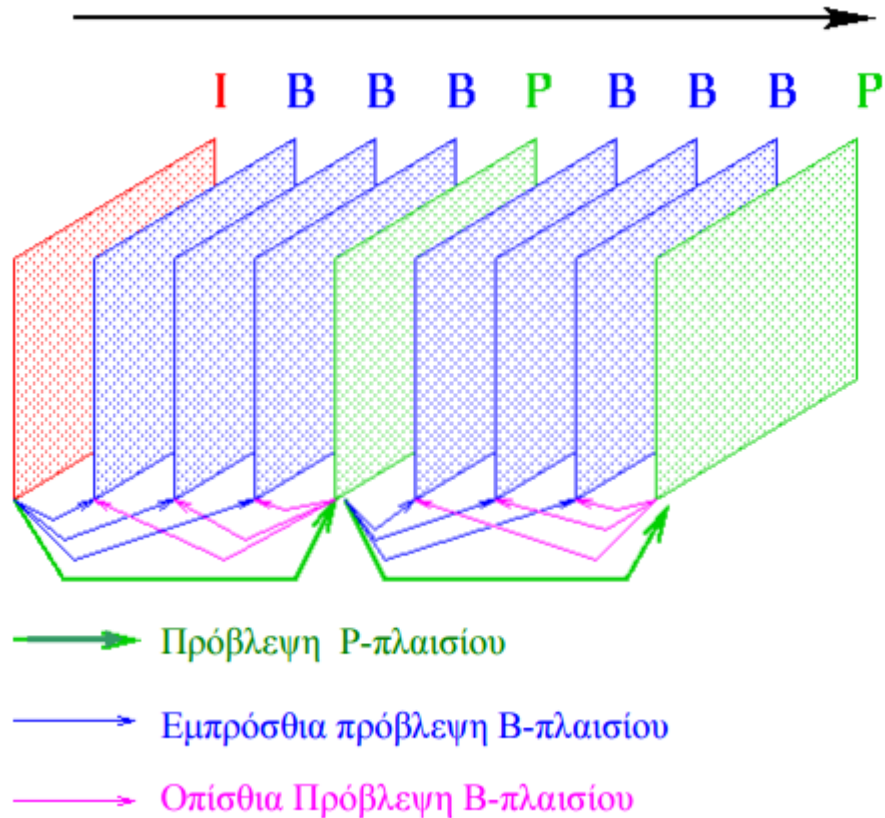
Αναλυτική περιγραφή λειτουργίας MPEG

Για να επιτύχουμε την κωδικοποίηση MPEG χωρίζουμε τα πλαίσια σε 3 τύπους, στα I, P, B πλαίσια. Τα I frames χρησιμοποιούνται ως frames αναφοράς για την πρόβλεψη μελλοντικών frames.



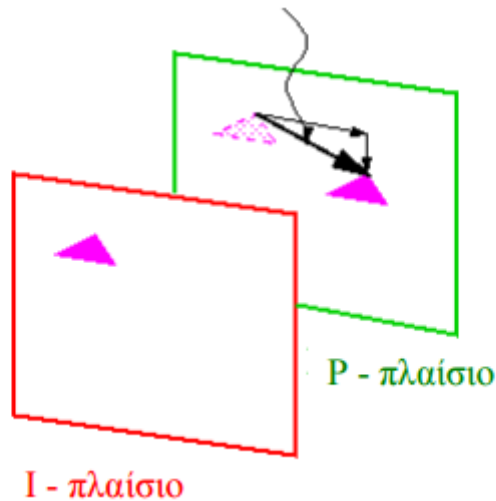
Τα I πλαίσια, τα οποία εμφανίζονται με ρυθμό 1 στα 10 ή στα 15, περιέχουν μόνο την πληροφορία που έχουν τα ίδια. Τα P πλαίσια μπορούν να προβλεφθούν από την πληροφορία που βρίσκεται στο κοντινότερο προηγούμενο I ή P πλαίσιο. Τα B πλαίσια, πλαίσια διπλής κατεύθυνσης, κωδικοποιούνται χρησιμοποιώντας δεδομένα από τα προηγούμενα I ή P πλαίσια καθώς και από τα επόμενα. Για παράδειγμα, στην επόμενη ακολουθία κινούμενων εικόνων έχουμε ονοματίσει τα I, P και B πλαίσια. Έτσι, για κάποια ακολουθία εικόνων θα πάρουμε το εξής σχήμα κωδικοποίησης:

Διάταξη πλαισίων του MPEG

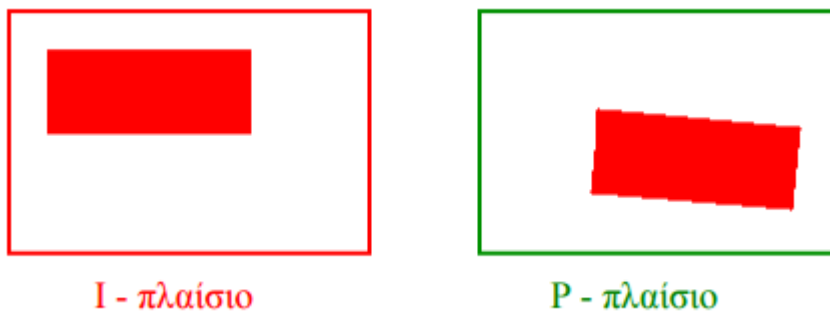


Η ανίχνευση κίνησης (motion compensation) που αναφέρθηκε παραπάνω και αποτελεί τη βασική τεχνική κωδικοποίησης του MPEG είναι η εξής: Ας φανταστούμε ένα I-πλαίσιο που δείχνει ένα τρίγωνο σε λευκό υπόβαθρο. Ένα ακόλουθο P-πλαίσιο δείχνει το ίδιο τρίγωνο αλλά σε άλλη θέση.

Η πρόβλεψη και η ανίχνευση κίνησης σημαίνουν πως υπολογίζουμε ένα διάνυσμα κίνησης (motion vector) το οποίο δηλώνει πώς να μετακινήσουμε το τρίγωνο του I-πλαisiού έτσι ώστε να συμπίπτει με το τρίγωνο του I-πλαisiού. Αυτή το διάνυσμα κίνησης είναι κομμάτι της ροής του MPEG και διαιρείται σε μία οριζόντια και μία κάθετη παράμετρο. Αυτές οι συνιστώσες μπορούν να απέχουν θετικές και αρνητικές τιμές. Η θετική τιμή σημαίνει πως πρέπει να κινηθούμε προς τα δεξιά ή προς τα κάτω αντίστοιχα. Η αρνητική τιμή σημαίνει πως πρέπει να κινηθούμε προς τα αριστερά ή προς τα πάνω αντίστοιχα.



Τα κομμάτια του διανύσματος κίνησης είναι σε ένα εύρος τιμών από -64 έως και +63. Δεδομένου αυτού του γεγονότος, η προαναφερθείσα περιοχή μπορεί να βρεθεί μέχρι και 64*64 pixels μακριά. Όμως, αυτό το μοντέλο, προϋποθέτει ότι κάθε αλλαγή μεταξύ των πλαισίων μπορεί να εκφραστεί σαν μια απλή μετατόπιση των pixels. Αλλά αυτή η εικόνα που παρουσιάζεται στην παρακάτω, αποδεικνύει ότι αυτή η σκέψη δεν είναι σωστή. Το κόκκινο ορθογώνιο της εικόνας, είναι μετατοπισμένο και περιστραμμένο κατά 5 μοίρες προς τα δεξιά. Έτσι, μια απλή μετατόπιση του κόκκινου ορθογώνιου θα προκαλέσει ένα λάθος πρόβλεψης. Γι' αυτόν το λόγο, η ροή του MPEG, περιέχει έναν πίνακα για να αντισταθμίσει αυτό ακριβώς το λάθος πρόβλεψης.



Σύγκριση επιπέδων MPEG

Τα MPEG1 και MPEG2 είναι πρότυπα που επικεντρώνονται στη συμπίεση και αποκωδικοποίηση ροών από video (video streams). Το MPEG1 σχεδιάστηκε για να παρέχει κωδικοποίηση σε μέσα όπως Video CD και CDROM τα οποία έχουν ρυθμό αναπαραγωγής γύρω στα 1,2 Mbit/s.

Το MPEG2 σχεδιάστηκε για να προσφέρει υψηλότερη ποιότητα σε εφαρμογές μετάδοσης, επικεντρώνοντας κυρίως στην ψηφιακή τηλεόραση.

Η βασική διαφορά ανάμεσα στο MPEG4 και στα MPEG1 και MPEG2 είναι ο τρόπος που το MPEG4 σχετίζεται με το επίπεδο εφαρμογής. Το MPEG4 καθορίζει το περιεχόμενο που πρέπει να μεταφερθεί μέσω ενός δικτύου ως ένα σύνολο από αντικείμενα και σκηνές. Στο MPEG4 κάθε στοιχείο της εικόνας θεωρείται ως ένα αντικείμενο, Κάθε αντικείμενο έχει χρονικά και χωρικά χαρακτηριστικά που καθορίζουν τη συμπεριφορά του και τη θέση του στη σκηνή του video.

ΣΥΜΠΕΡΑΣΜΑΤΑ

Με τον όρο συμπίεση δεδομένων (data compression) εννοούμε τη μετατροπή ενός ψηφιακού αρχείου σε μικρότερο αρχείο (που περιέχει μικρότερο αριθμό μπιτ) με τρόπο ώστε να είναι δυνατή η επαναμετατροπή του συμπιεσμένου αρχείου στο αρχικό. Για το σκοπό αυτό χρησιμοποιούνται πολλές μέθοδοι, οι οποίες χωρίζονται σε δύο μεγάλες κατηγορίες: τις μη απωλεστικές και τις απωλεστικές.

Αν θεωρήσουμε τα δεδομένα μας σαν μια ακολουθία τότε η συμπίεση ορίζεται όπως παρακάτω:

Συμπίεση (compression) μιας ακολουθίας δεδομένων, ονομάζουμε την ελάττωση του μεγέθους της ακολουθίας, ώστε να χρειάζεται λιγότερος χώρος για την αποθήκευση ή τη μετάδοσή της. Η διαδικασία της συμπίεσης εφαρμόζεται συστηματικά στα υπολογιστικά συστήματα που χρησιμοποιούν και επεξεργάζονται μεγάλο όγκο ψηφιακών δεδομένων. Υπάρχουν πολλοί αλγόριθμοι για να καταφέρουμε να μειώσουμε το μέγεθος μίας ακολουθίας. Ένας αλγόριθμος συμπίεσης μπορεί να επιφέρει μεγάλη ελάττωση του μήκους σε ακολουθίες δεδομένων με κάποιο ιδιαίτερο χαρακτηριστικό, ενώ ο ίδιος να είναι αναποτελεσματικός για άλλες ακολουθίες, αφήνοντάς τις στο ίδιο μήκος ή ακόμα και μεγαλώνοντάς το σε μερικές περιπτώσεις.

Όταν έχουμε την ακολουθία των δεδομένων σε συμπιεσμένη μορφή, πρέπει να εφαρμοστεί η αντίστροφη διαδικασία της *αποσυμπίεσης* (decompression, extraction) προκειμένου τα δεδομένα να μπορούν να χρησιμοποιηθούν και πάλι. Η διαδικασία αυτή της αποσυμπίεσης των δεδομένων καθορίζει και τις κατηγορίες των μεθόδων συμπίεσης. Έτσι, υπάρχουν δύο κατηγορίες αλγορίθμων συμπίεσης, οι *απωλεστικοί* (lossy) και οι μη *απωλεστικοί* (lossless) αλγόριθμοι.

Στους απωλεστικούς αλγορίθμους, όταν γίνει η συμπίεση και μετά ακολουθήσει αποσυμπίεση των δεδομένων, η τελική ακολουθία των δεδομένων διαφέρει από την αρχική. Αντίθετα στους μη απωλεστικούς αλγορίθμους, η διαδικασία συμπίεσης και αποσυμπίεσης επαναφέρει την αρχική ακολουθία. Αν πρέπει να μεταφερθούν δεδομένα

με απόλυτη ακρίβεια, πιστότητα, χωρίς να αλλοιωθεί το περιεχόμενό τους, πρέπει να εφαρμοστεί μια μη απωλεστική μέθοδος συμπίεσης.

Στην εργασία αυτή παρουσιάστηκαν μια σειρά από αλγόριθμους και διαδικασίες που εφαρμόζονται για την συμπίεση διαφόρων τύπων δεδομένων όπως ο ήχος και η εικόνα.

Ιδιαίτερη έμφαση δόθηκε στους αλγόριθμους MPEG αφού είναι μια πολύ αποδοτική λύση για την συμπίεση δεδομένων που εφαρμόζεται πλέον σε όλα τα επίπεδα μεταφοράς εικόνας και βίντεο.

BIBΛΙΟΓΡΑΦΙΑ

Βιβλία:

- Behrouz A. Forouzan. «Εισαγωγή στην Επιστήμη των Υπολογιστών», Επιμέλεια Γιώργος Στεφανίδης, Αλέξανδρος Χατζηγεωργίου, Εκδόσεις Κλειδάριθμος.
- Mohammed Ghanbari, Standard Codecs: Image Compression to Advanced Video Coding, Institution of Electrical Engineers (IEE) 2003.
- M. Rabbani and P.W. Jones, Digital image compression techniques, SPIE Optical Engineering Press, 1991.

Δημοσιεύσεις:

- T. Painter and A. Spanias, Perceptual coding of digital audio, Proceedings of the IEEE, pp. 451-513, April 2000.
- P. Noll, MPEG digital audio coding, IEEE Signal Processing Magazine, pp.59-81, Sept. 1997.

URLs:

- <https://avraammauridis.wordpress.com/2011/03/02/%CF%83%CF%85%CE%BC%CF%80%CE%AF%CE%B5%CF%83%CE%B7-%CE%B4%CE%B5%CE%B4%CE%BF%CE%BC%CE%AD%CE%BD%CF%89%CE%BD-data-compression/>
- http://gravitonio.blogspot.gr/2011/09/blog-post_23.html
- http://nefeli.lib.teicrete.gr/browse/sefe/mta/2009/VlahouEleniMaria/attached-document-1263382994-171555-3795/Vlaxou_Eleni_Maria.pdf