



ΒΕΣ 04:
Συμπίεση και Μετάδοση Πολυμέσων
Συμπίεση Βίντεο



Περιεχόμενα

- Αρχές Συμπίεσης
 - Τύποι πλαισίων
 - Εκτίμηση και αντιστάθμιση κίνησης
 - Θέματα υλοποίησης
- H261
- H263
- MPEG
 - MPEG-1
 - MPEG-2
 - MPEG-4



Βιβλιογραφία

- Καγιάφας [2000]: Το Πρότυπο Συμπίεσης MPEG [link], και Proprietary Compression [link]
- Halsall [2001]: Chapter 4, pp. 193-235
- Chapman [2004]: Chapter 7

Αρχές συμπίεσης βίντεο

- Στη συμπίεση video αναζητείται μία χρυσή τομή (sweet spot) ανάμεσα στην ποιότητα και το εύρος ζώνης (bandwidth)
- Πολλές φορές βασίζεται σε υποκειμενικές μετρήσεις που γίνονται σε άτομα που θεωρούνται ικανά "δείγματα" πάνω στην εκτίμηση της κινούμενης εικόνας.
- Αν η εμπειρία τους δεν τους επιτρέπει να αντιληφθούν σημαντικές διαφορές στην ποιότητα του κωδικοποιημένου ήχου από τον αρχικό και συγχρόνως θεωρούν ανεκτές τις ατέλειες (artifacts) της εικόνας μετά από κάποιο ποσοστό συμπίεσης, η μέθοδος και το ποσοστό αυτό θεωρούνται αποδεκτά και τυποποιούνται.
- Να σημειωθεί εδώ ότι ενώ ο ήχος προσφέρεται για μεγάλα ποσοστά συμπίεσης χωρίς να μπορεί να γίνει αισθητή υποβάθμιση στην ποιότητά του, στον τομέα της εικόνας τα πράγματα είναι λίγο πιο δύσκολα και οι όποιες ατέλειες είναι ορατές, απλά γίνεται προσπάθεια να μην γίνονται ενοχλητικές για το θεατή.
- Η βασική αρχή πάνω στην οποία στηρίζονται όλες οι μέθοδοι ψηφιακής συμπίεσης είναι το γεγονός ότι το σήμα περιέχει 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.

Αρχές συμπίεσης βίντεο (II)

- Υπάρχουν δύο είδη πλεονασμού σε ένα σήμα βίντεο:
- Χωρικός πλεονασμός (Spatial 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.
 - Ο υποκειμενικός πλεονασμός και οι ιδιαιτερότητες της ανθρώπινης όρασης έχουν αξιοποιηθεί στις τεχνικές συμπίεσης ακίνητων εικόνων (π.χ. πρότυπο JPEG).
- Χρονικός Πλεονασμός (Temporal Redundancy):
 - Οι τιμές φωτεινότητας και χρώματος μπορούν σε κάθε δεδομένη χρονική στιγμή να προϋπολογιστούν από τις προηγούμενες τιμές τους, γιατί οι τιμές δειγμάτων της εικόνας σε μεγάλο ποσοστό σχετίζονται μεταξύ τους.
 - Για παράδειγμα στα δελτία ειδήσεων ένα μεγάλο κομμάτι της εικόνας (φόντο) παραμένει αμετάβλητο και μόνο το κομμάτι της εικόνας που καταλαμβάνει ο παρουσιαστής μεταβάλλεται ελαφρά (κινήσεις των χεριών, 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.

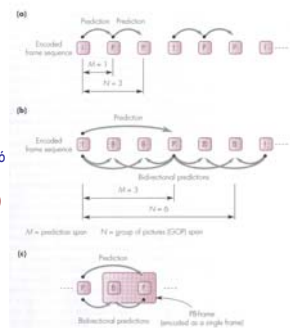
Αρχές συμπίεσης βίντεο (III)

- Χωρική συμπίεση (spatial / intra-frame compression)
 - Κάθε πλαίσιο συμπίεζεται αυτόνομα ως χαρτογραφική εικόνα (μορφή συμπίεσης JPEG)
 - Η συμπίεση μπορεί να είναι με ή χωρίς απώλειες (lossy or lossless). Η συμπίεση με απώλειες οδηγεί σε απώλεια ποιότητας όταν έχουμε διαδοχικές συμπίεσεις και αποσυμπίεσεις
- Χρονική συμπίεση (temporal / inter-frame compression)
 - Η συμπίεση εφαρμόζεται σε ακολουθίες από πλαίσια με κωδικοποίηση των διαφορών ανάμεσα σε αυτά
 - Πλαίσιο κλειδιά (key frame) κωδικοποιούνται μόνο χωρικά
 - Σε αρκετές περιπτώσεις η απόσταση ανάμεσα σε πλαίσια κλειδιά είναι σταθερή (π.χ. Κάθε 12 πλαίσια)
 - Η διαφορές ανάμεσα στα πλαίσια κωδικοποιούνται χωρικά (λογική JPEG)
- Σε όλες τις περιπτώσεις έχουμε κάποια μορφή συμπίεσης εξαιτίας της δειγματοληψίας φωτεινότητας /χρωματικών καναλιών (4:2:0, 4:1:1 κλπ)

Αρχές Συμπίεσης =>

Τύποι πλαισίων

- **I-πλαίσια (intra-coded frames)**
 - Αυτόνομα κωδικοποιημένα frame (κωδικοποιημένα ως εικόνες)
- **P-πλαίσια (predictive frames)**
 - Κωδικοποιημένα με βάση προβλεπτική κωδικοποίηση από άλλα P ή I frames
- **B-πλαίσια (inter-coded frames)**
 - Παρεμβλλόμενα (Interpolation) frames ανάμεσα σε P και I frames



Αρχές Συμπίεσης => Τύποι πλαισίων =>

I (Intra frames)

- Το είδος αυτό των πλαισίων κάνει χρήση του intra frame-coding. Τα πλαίσια τύπου I είναι τα μόνα που είναι κωδικοποιημένα στο σύνολό τους και η αποκωδικοποίηση μπορεί να γίνει χωρίς αναφορά σε κάποιο άλλο.
- Αποτελούν σημεία αναφοράς κατά την τυχαία προσέλαση ενός σήματος. Επειδή η παρουσία τους είναι απαραίτητη σε σημείο χρονικής αναφοράς και για να αποφευχθεί η διάδοση των σφαλμάτων που δημιουργούν τα P πλαίσια επιβάλλεται να μεταδίδονται ανά τακτά χρονικά πλαίσια. Έτσι υπάρχει ένα I πλαίσιο τουλάχιστον κάθε 15 πλαίσια (δηλαδή δύο φορές το δευτερόλεπτο αν η συχνότητα είναι 30 Hz).
- Η διαδικασία της κωδικοποίησης ενός I πλαισίου βασίζεται στη λογική του προτύπου JPEG.
- Η εικόνα χωρίζεται σε macroblocks και για κάθε block ξεχωριστά εφαρμόζεται DCT, Κβαντοποίηση, Zig-Zag Scanning, Run-Length-Encoding και Huffman Encoding.

Αρχές Συμπίεσης => Τύποι πλαισίων =>

P (Predicted frames)

- Τα πλαίσια τύπου P είναι βασισμένα σε ένα προηγούμενο I ή P πλαίσιο.
- Με τη βοήθεια του motion compensation προβλέπουν τη νέα θέση όπου των macroblocks έχουν απλά μετακινηθεί και κωδικοποιούν τον αριθμό του macroblock, ένα διάνυσμα κίνησης και τη διαφορά (σφάλμα πρόβλεψης).
- Με τη σειρά τους μπορούν να αποτελέσουν και αυτά σημείο αναφοράς για επόμενα πλαίσια και αυτός είναι και ο λόγος που συμβάλλουν στην εισαγωγή και διάδοση σφαλμάτων, αφού η διαδικασία της πρόβλεψης κίνησης δεν μπορεί να είναι 100% ακριβής.
- Δεν έχουν το μέγεθος των I πλαισίων γιατί δεν έχουν περιγραφεί με την ίδια ακρίβεια, δηλαδή παρουσιάζουν μεγαλύτερο ποσοστό συμπίεσης.
- Η διαδικασία λοιπόν της κωδικοποίησης τους περιλαμβάνει:
 - σύγκριση macroblocks του τρέχοντος πλαισίου με αυτά του προηγούμενου και εύρεση της βέλτιστης ομοιότητας (motion estimation)
 - δημιουργία motion vector, και υπολογισμός των διαφορών πρόβλεψης (motion compensation)
 - μετασχηματισμός DCT των διαφορών σε κάθε block του νέου macroblock, Κβαντοποίηση, Run-Length-Encoding και το αποτέλεσμα κωδικοποιείται με κωδικοποίηση Huffman, όπως και στα I πλαίσια.

Αρχές Συμπίεσης => Τύποι πλαισίων =>

B (Bi-directional frames)

- Τα πλαίσια τύπου B είναι πλαίσια που δημιουργούνται λαμβάνοντας το μέσο όρο σε επίπεδο macroblock ενός προηγούμενου και ενός επόμενου πλαισίου I και P (ένα από το κάθε είδος).
- Δε συντελούν τόσο πολύ στη διάδοση των σφαλμάτων γιατί δεν χρησιμοποιούνται ως σημεία αναφοράς και επιπλέον μειώνουν σημαντικά το σφάλμα παίρνοντας το μέσο όρο από δύο πλαίσια.
- Μπορούμε να πούμε ότι ο 'κύκλος της ζωής' τους περιορίζεται μόνο σε αυτά και δεν επεκτείνεται με το να κληροδοτούν πληροφορίες σε άλλα πλαίσια, κάτι που πολλές φορές σε συνδυασμό και με την υπολογιστική πολυπλοκότητα που απαιτούν για την κωδικοποίηση και αποκωδικοποίηση τα κάνει μη επιθυμητά από τους κατασκευαστές.
- Η διαδικασία της κωδικοποίησης περιλαμβάνει συνδυασμό των αντίστοιχων macroblocks που παρουσιάζουν μικρές διαφορές με τα αντίστοιχα των πλαισίων αναφοράς (προηγούμενα και επόμενα) δηλαδή αφαίρεση του μέσου όρου των άλλων δύο από το τρέχον πλαίσιο, συνδυασμό των διανυσμάτων κίνησης των πλαισίων αναφοράς (που συνδυάζονται όπως και τα αντίστοιχα macroblocks, δηλαδή λαμβάνεται ο μέσος όρος τους) και στη συνέχεια την ίδια διαδικασία με τα I και P πλαίσια για την κωδικοποίηση του macroblock που προκύπτει.

Αρχές Συμπίεσης =>

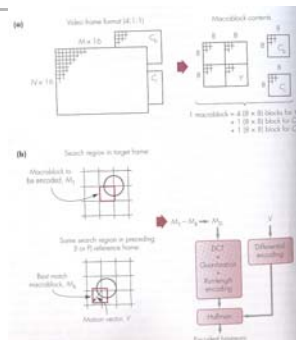
Εκτίμηση και αντιστάθμιση κίνησης

- Ένας τρόπος με τον οποίο μπορούμε να περιγράψουμε πιο αποτελεσματικά εικόνες με κίνηση είναι η τμηματική πρόβλεψη κίνησης. Με τη βοήθεια αυτής της μεθόδου μπορούμε να εκτελέσουμε τη διαπλασιακή (inter-frame) κωδικοποίηση και να περιγράψουμε την αλληλουχία εικόνων ως σειρά ομοιοτήτων και διαφορών.
- 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 (εκτίμηση κίνησης). Έτσι αν έχουμε ήδη αποστείλει την πρώτη εικόνα μπορούμε να στείλουμε τη δεύτερη σαν ένα πίνακα διανυσμάτων και ορισμένων σταθερών τμημάτων, που προφανώς έχει πολύ μικρότερο μέγεθος.
- Το παραπάνω είναι μία καλή προσέγγιση της μεθόδου αλλά στην πραγματικότητα οι πραγματικές εικόνες δεν είναι τόσο όμοιες μεταξύ τους. Οσα έχουν κάποια κοινά τμήματα που αλλάζουν θέση από πλαίσιο σε πλαίσιο αλλά θα υπάρχουν και τμήματα που αλλάζουν θέση διατηρώντας το σχήμα τους αλλά μεταβάλλεται το χρώμα τους, καθώς και άλλα που δεν υπάρχουν σε προηγούμενο πλαίσιο αλλά εμφανίζονται σε κάποιο για πρώτη φορά.
- Για την κάλυψη αυτών των περιπτώσεων κωδικοποιείται το σφάλμα πρόβλεψης (αντιστάθμιση κίνησης).

Αρχές Συμπίεσης =>

Εκτίμηση και αντιστάθμιση κίνησης (II)

- Format 4:1:1 => τα macroblock αποτελούνται από 4 8x8 block για το κανάλι φωτεινότητας και από 1 8x8 block για τα κανάλια χρώματος (Cr, Cb)
- Οι διαφορές πρόβλεψης κωδικοποιούνται με τη λογική JPEG ενώ το διάνυσμα κίνησης κωδικοποιείται με διαφορετική κωδικοποίηση



H263

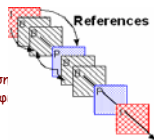
- Βασίζεται στο H261 αλλά έχει σχεδιαστεί για μετάδοση μέσω του πρωτοκόλλου IP
- Χρήση πλαισίων I, P αλλά και B για επίτευξη μεγαλύτερης συμπίεσης
- Digitization Format= QCIF ελαφρά τροποποιημένο:
 - Y=>176x144 (αντί 180x144) ώστε να έχουμε ακέραιο αριθμό από macroblock (11x9),
 - Cr=>88x72, Cb=>88x72,
 - 15 ή 7.5 fps (frame per second)
- Συμμετρική πολυπλοκότητα (ίδιος χρόνος συμπίεσης και αποσυμπίεσης)
- Παράδειγμα: Λογισμικό τηλεδιάσκεψης (π.χ. NetMeeting)

MPEG

- Τα πρότυπα δημιουργήθηκαν από τον οργανισμό τυποποίησης ISO και από μια ομάδα ερευνητών (Moving Picture Experts Group)
- Η σειρά προτύπων MPEG περιλαμβάνει:
 - MPEG-1 (στόχος η διανομή σε CD-ROM οπότε ο μέγιστος ρυθμός μετάδοσης είναι περίπου 1.41Mbps)
 - MPEG-2 (στόχος η χρήση σε DVD και ψηφιακή τηλεοπτική μετάδοση - broadcast)
 - MPEG-4 (στόχος η χρήση σε πολυμεσικές εφαρμογές χαμηλού εύρους ζώνης - bitrate multimedia)
- Στα πρότυπα MPEG δεν είναι υποχρεωτική η χρήση όλων των προδιαγραφών
 - Τα profiles ορίζουν υποσύνολα χαρακτηριστικών του προτύπου που μπορούν να χρησιμοποιηθούν
 - Τα Levels ορίζουν τις τιμές παραμέτρων όπως το μέγεθος πλαισίου αλλά και ο επιθυμητός ρυθμός μετάδοσης (data rate)
 - Κάθε profile μπορεί να υλοποιηθεί σε ένα ή περισσότερα επίπεδα
 - Συμβολισμός: profile@level, π.χ. MP@ML

MPEG (II)

- Η σειρά με την οποία τα διάφορα πλαίσια προβάλλονται δεν είναι κατ' ανάγκη η ίδια με αυτή που αποκωδικοποιούνται
- Π.χ., αν ένα Πλαίσιο-P πρέπει να προβληθεί μετά από ένα Πλαίσιο-B και το πρώτο χρησιμοποιείται για την κωδικοποίηση του δευτέρου, τότε η σειρά αποκωδικοποίησης είναι αντίστροφη από αυτή της προβολής



Σειρά προβολής

	B	B	I	B	B	P	B	B	P	B	B	P
#	0	1	2	3	4	5	6	7	8	9	10	11

Σειρά αποκωδικοποίησης

	I	B	B	P	B	B	P	B	B	P	B	B
#	2	0	1	5	3	4	8	6	7	11	9	10

MPEG =>



MPEG-1

- Πρότυπο για αποθήκευση βίντεο ποιότητας VHS σε CD-ROM
- Bit rate: 1.2 Mbps για το βίντεο και 250 kbps για τον ήχο
- Χωρική συμπίεση: Λογική JPEG
- Χρονική συμπίεση βασισμένη στην αντιστάθμιση κίνησης (motion compensation)
 - Διάγραμμα μετατόπισης σε σχέση με το προηγούμενο I ή P frame και κωδικοποίηση της διαφοράς ανάμεσα στη πρόβλεψη (macroblock στο προηγούμενο frame) και τα πραγματικά δεδομένα (τρέχον macroblock)
- I-frames: αμιγής JPEG κωδικοποίηση
 - Δεν υπάρχει εξάρτηση από προηγούμενα frames
- P-frames: κωδικοποίηση της διαφοράς πρόβλεψης
 - Η πρόβλεψη βασίζεται σε macroblocks από προηγούμενα I ή P frames
- B-frames: αμφίδρομη πρόβλεψη ('bi-directionally predictive') και κωδικοποίηση της διαφοράς πρόβλεψης από τα πραγματικά δεδομένα
 - Η πρόβλεψη βασίζεται σε macroblocks από το αμέσως προηγούμενο I ή P frame σε συνδυασμό με το αμέσως επόμενο I ή P frame.

MPEG =>



MPEG-1 (II)

- Group of Pictures (GOP)
 - Επαναλαμβανόμενες ακολουθίες από I-, P- και B-frames
 - Αρχίζουν πάντα με I- frame και καταλήγουν σε B ή P frame μετά από τα οποία ακολουθεί υποχρεωτικά I- frame (GOP=απόσταση ανάμεσα σε δύο διαδοχικά I-frames)
 - Σειρά προβολής (Display order): Η σειρά με την οποία τα frames προβάλλονται τελικά για να σχηματιστεί το βίντεο
 - Σειρά μετάδοσης και αποκωδικοποίησης (Bitstream order): Αναδιάταξη των πλαίσιων ώστε κάθε P- ή B-frame να μεταδίδεται μετά από frames από τα οποία εξαρτάται η ανακατασκευή του (frames από τα οποία έχουν προβλεφθεί τα περιεχόμενα του)
- Το MPEG-1 βασίζεται στο Source Input Format (SIF) (ποιότητα VHS)
 - 4:1:1 digitization format
 - 352x288 pixels
 - 25 frames per second
 - data rate ~ 1.5Mbits per second
- Μπορεί να χρησιμοποιηθεί και για μεγαλύτερα frames αλλά δεν μπορεί να χειριστεί interlaced video streams

MPEG => MPEG-1 =>



Άσκηση

- Ένα σύστημα κωδικοποίησης βίντεο MPEG-1, χρησιμοποιεί την ακολουθία πλαισίων:
 - I B B P B B P B B P B B I
- 1. Πόσο είναι το prediction span και το GOP;
- 2. Ποια είναι η σωστή σειρά αποκωδικοποίησης;
- 3. Αν υποθέσουμε ότι ο βαθμός συμπίεσης για τα I πλαίσια είναι 10:1 για τα P πλαίσια είναι 20:1 και για τα B πλαίσια είναι 50:1 να υπολογιστεί ο μέσος βαθμός συμπίεσης
- 4. Με τα δεδομένα του ερωτήματος 3 να υπολογιστεί το bitrate της ακολουθίας MPEG-1 στις περιπτώσεις NTSC και PAL

MPEG =>

MPEG-2

- Πρότυπο για αποθήκευση βίντεο ποιότητας DV (αλλά και HDTV) σε DVD-ROM
- Χρησιμοποιείται επίσης για τηλεοπτική μετάδοση ψηφιακού βίντεο:
 - Δορυφορικά συνδρομητικά κανάλια (π.χ. Nova)
- Bit rate: 3-15 Mbps
- MPEG-2 Main Profile at Main Level (MP@ML) used for DVD video
 - CCIR 601 scanning
 - 4:2:0 chrominance sub-sampling
 - Frame size: 720x576, 25 frame per second για PAL ή 720x576, 30 frame per second για NTSC
 - Αντί για progressive scan όπως το MPEG-1 χρησιμοποιείται interlaced scan
 - 15 Mbits per second
 - Αποτελεί την πλέον σύνθετη και επεξεργασμένη αναπαράσταση του προτύπου MPEG-2

MPEG =>

MPEG-4

- Σχεδιάστηκε για υποστήριξη πολυμεσικών εφαρμογών με ρυθμούς μετάδοσης από 10kbps έως και μεγαλύτερους από 1.8Mbps
- Εφαρμογές:
 - Από κινητά τηλέφωνα έως τηλεόραση υψηλής ευκρίνειας (HDTV)
- Δημοφιλείς codecs (αλγόριθμοι συμπίεσης-αποσυμπίεσης) για χρήση στο Διαδίκτυο:
 - QuickTime, RealMedia και DivX
- Το πρότυπο υποστηρίζει τη δημιουργία του data stream από μια ποικιλία αντικειμένων /δομών
 - video, ακίνητες εικόνες, ακίνητα και κινούμενα γραφικά animation, 3-D μοντέλα ...
- Στα πιο σύνθετα profiles κάθε σκηνή συντίθεται από τυχαία βιντεοαντικείμενα (VOP=Video Object Plane) τα οποία έχουν τυχαίο σχήμα και μπορούν να μεταδοθούν ανεξάρτητα. Η σκηνή συντίθεται στο δέκτη ο οποίος συνδυάζει τα data streams των επιμέρους αντικειμένων
- Στα profile SP (Simple Profile) και ASP (Advanced Simple Profile) περιορίζομασε σε ορθογώνια βιντεοαντικείμενα τα οποία συνήθως είναι τα ίδια τα πλαίσια (frames) ολόκληρα.

MPEG =>

MPEG-4 (II)

- Η συμπίεση στο MPEG-4 αποτελεί βελτίωση της συμπίεσης MPEG-1:
 - Τα I-frames συμπίεζονται όπως στο MPEG-1
 - Βελτίωση στην αντιστάθμιση κίνησης (motion compensation) οδηγεί σε καλύτερη ποιότητα σε σχέση με το MPEG-1 για το ίδιο bit rate
- Simple Profile
 - Χρησιμοποιούνται μόνο I και P-frames
 - Αποτελεσματική αποσυμπίεση κατάλληλη για χρήση σε υλικό περιορισμένων δυνατοτήτων (π.χ. PDAs)
- Advanced Simple Profile (ASP)
 - Χρησιμοποιούνται I, B και P-frames
 - Ολική αντιστάθμιση κίνησης (Global Motion Compensation) στην οποία μοντελοποιούνται κάποιες συγκεκριμένες κινήσεις της κάμερας (π.χ. Pan, zoom).

