Μελέτη Περιπτώσεων στη Λήψη Αποφάσεων

Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο
ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΕΚΠΑΙΔΕΥΣΗ ΚΑΙ ΔΙΑ ΒΙΟΥ ΜΑΘΗΣΗ
επένδυση στην κοινωνία της γνώσης
ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ ΚΑΙ ΘΡΗΣΚΕΥΜΑΤΩΝ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ
ΕΣΠΑ 2007-2013
πρόγραμμα για την ανάπτυξη
ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ
Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης

Σημείωμα Αδειοδότησης

Χρηματοδότηση

- Το παρόν εκπαιδευτικό υλικό έχει αναπτυχθεί στο πλαίσιο του εκπαιδευτικού έργου του διδάσκοντα.
- Το έργο «**Ανοικτά Ακαδημαϊκά Μαθήματα στο Πανεπιστήμιο Πατρών**» έχει χρηματοδοτήσει μόνο την αναδιαμόρφωση του εκπαιδευτικού υλικού.
- Το έργο υλοποιείται στο πλαίσιο του Επιχειρησιακού Προγράμματος «Εκπαίδευση και Δια Βίου Μάθηση» και συγχρηματοδοτείται από την Ευρωπαϊκή Ένωση (Ευρωπαϊκό Κοινωνικό Ταμείο) και από εθνικούς πόρους.

Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο
ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΕΚΠΑΙΔΕΥΣΗ ΚΑΙ ΔΙΑ ΒΙΟΥ ΜΑΘΗΣΗ
επένδυση στην κοινωνία της γνώσης
ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ ΚΑΙ ΘΡΗΣΚΕΥΜΑΤΩΝ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ
ΕΣΠΑ 2007-2013
πρόγραμμα για την ανάπτυξη
ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ
Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης

Συμπίεση Πολυμεσικών Δεδομένων

Εισαγωγή στο πρόβλημα και επιλεγμένες εφαρμογές

Κώστας Μπερμπερίδης

Εργαστήριο Σημάτων & Τηλεπικοινωνιών
Τμήμα Μηχανικών Η/Υ & Πληροφορικής

ΔΠΜΣ ΜΥΑ, 3 Απριλίου 2015

Ψηφιακή Αναπαράσταση

- Συμπίεση για Αποθήκευση ή/και Μετάδοση
- Η ψηφιακή αναπαράσταση των δεδομένων έχει σαφή πλεονεκτήματα:
  - μεγαλύτερη ανοσία σε «θόρυβο»
  - ευελιξία στην ανταλλαγή εύρους ζώνης και ισχύος
  - τεχνικές κρυπτογράφησης και προστασίας
  - ευκολία υλοποίησης σε υλικό (VLSI)
- Αν τα δεδομένα είναι Αναλογικά, θα πρέπει να μετατραπούν σε Ψηφιακά (πώς ;)
- Ερώτηση: Χάνεται πληροφορία κατά τη μετατροπή A/D;
- Τα πολυμεσικά δεδομένα είναι κατά βάση αναλογικά.

Θεωρία Πληροφορίας

- Βασικός στόχος της Συμπίεσης:
  - η αποδοτική αναπαράσταση των δεδομένων που παράγει μια πηγή πληροφορίας
- Το πρόβλημα της συμπίεσης είναι επίσης γνωστό ως κωδικοποίηση πηγής (source coding)
- Το αντικείμενο της συμπίεσης δεδομένων ανήκει στον ευρύτερο χώρο της θεωρίας πληροφορίας (information theory)

Κωδικοποίηση Πηγής

- Στόχος: η αποδοτική αναπαράσταση/κωδικοποίηση/συμπίεση της πληροφορίας/σήματος/εξόδου μιας πηγής

- Ερωτήματα που προκύπτουν:
  - πώς ορίζεται η πληροφορία μιας πηγής;
  - πότε μια πηγή εξάγει περισσότερη πληροφορία;
  - μπορώ να τη μετρήσω μαθηματικά;
  - τι παθαίνει η πληροφορία μιας πηγής όταν εφαρμόζω κάποια επεξεργασία (π.χ. μετατροπή A/D);
  - πόσο πολύ μπορώ να συμπιέσω τα δεδομένα μιας πηγής;

- Απαντήσεις
  - δίνονται μέσα από τη θεωρία πληροφορίας και τις τεχνικές κωδικοποίησης πηγής

Πηγές Πληροφορίας

- Η έξοδος της πηγής είναι
  - κάτι τυχαίο και άγνωστο, μια τυχαία διαδικασία (αν είναι κάτι σταθερό ή γενικότερα ντετερμινιστικό, δεν υπάρχει λόγος να το μεταδώσουμε ή να το αποθηκεύσουμε)
- Παραδείγματα πηγών πληροφορίας:
  - ακολουθία από bits
  - χαρακτήρες ASCII
  - ήχος (ομιλία)
  - εικόνα
  - video
- Διάκριση ως προς το χρόνο:
  - συνεχούς χρόνου (π.χ. αναλογικό ηχητικό σήμα)
  - διακριτού χρόνου (δειγματοληπτημένο σήμα ή σύμβολα)
- Διάκριση ως προς τις δυνατές τιμές (αλφάβητο):
  - συνεχείς τιμές (π.χ. αναλογικό σήμα)
  - διακριτές τιμές (π.χ. ASCII)

Πηγές Πληροφορίας (2)

- Αν η πηγή είναι αναλογική τότε απαιτείται: Μετατροπή της πηγής από συνεχούς σε διακριτού χρόνου
  - δειγματοληψία
  - το σήμα πρέπει να έχει πεπερασμένο εύρος ζώνης
  - αν είναι κατωπερατό με μέγιστη συχνότητα f_{max}, τότε η συνθήκη Nyquist μας λέει ότι αρκεί να το δειγματοληπτήσω με

$$\boxed{f_s \geq 2 f_{\max}}$$

  - και μπορώ να ανακατασκευάσω το αναλογικό σήμα από τα δείγματα του χωρίς απώλειες
- Οι πηγές που μας ενδιαφέρουν,
  - έχουν περιορισμένο εύρος ζώνης
  - ή μπορούμε να το περιορίσουμε εμείς με φιλτράρισμα
  - Περίπτωση στοχαστικών σημάτων: ανακατασκευή με μηδενικό MSE
- Συμπέρασμα: αρκεί να μελετήσω τις πηγές διακριτού χρόνου
- * Compressed Sampling

Πληροφορία

- Η θεωρητική ανάλυση παρουσιάζεται για πηγές πληροφορίας με διακριτό αλφάβητο λόγω ευκολίας
  (μπορεί να γενικευτεί σε πηγές με συνεχές αλφάβητο)

- Αλφάβητο Διακριτής Πηγής:

$$\boxed{\Phi = \left\{ s_1, s_2 \dots s_N \right\}}$$

- Παράδειγμα: ο καιρός στην Ελλάδα κάθε 15 Αυγούστου
  – s_1: χιόνι
  – s_2: βροχή
  – s_3: λιακάδα

- Πότε δίνεται περισσότερη πληροφορία;
  – όταν τυχαίνει το σύμβολο s_1 ή το s_3;
  – με τι σχετίζεται η πληροφορία που φέρει κάθε σύμβολο;

Μέτρο Πληροφορίας (1)

Ιδιότητες του μέτρου της πληροφορίας:

1. $\boxed{I(s_k)=0 \quad for \quad p(s_k)=1}$

2. $\boxed{I(s_k)\geq 0 \quad for \quad 0\leq p(s_k)\leq 1}$

3. $\boxed{I(s_k)>I(s_l) \quad for \quad p(s_k)<p(s_l)}$

4. $\boxed{I(s_k s_l)=I(s_k)+I(s_l) \quad \varepsilon\alpha\nu \quad s_k, s_l \quad \sigma\tau\alpha\tau\iota\sigma\tau\iota\kappa\acute{\alpha} \quad \alpha\nu\varepsilon\xi\acute{\alpha}\rho\tau\eta\tau\alpha}$

5. Μικρή αλλαγή στην πιθανότητα → μικρή αλλαγή στην πληροφορία (συνεχής συνάρτηση)

Μέτρο Πληροφορίας (2)

- Η Πληροφορία ενός συμβόλου (Information) s_i με πιθανότητα εμφάνισης $p(s_i)$ ορίζεται ως

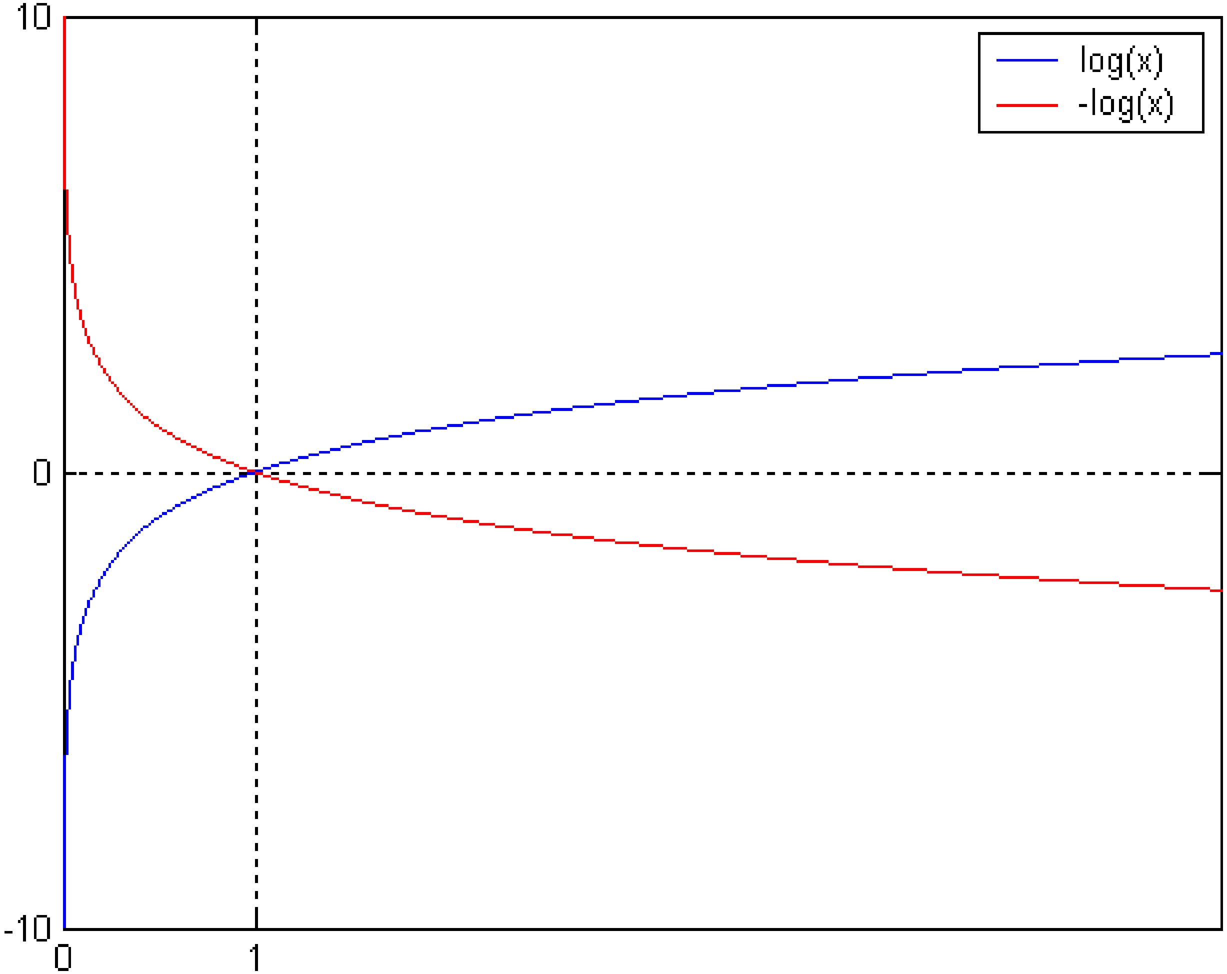
$$\boxed{I(s_i) = \log \frac{1}{p(s_i)} = -\log p(s_i)}$$

- Βάση του λογαρίθμου
  - συνήθως χρησιμοποιείται το 2 με μονάδα μέτρησης bit
  - δεν ταυτίζεται με το bit που γνωρίζουμε ως δυαδικό ψηφίο

- Σύνθετη πηγή:

$$\boxed{I(s_i, a_j) = I(s_i) + I(a_j)}$$

Μέτρο Πληροφορίας (3)



- Παρατηρήσεις:

1. Φθίνουσα
2. Πεδίο ορισμού (που ορίζεται;)
3. Όρια (τι συμβαίνει στα άκρα;)
4. Είναι συνεχής;

Διακριτή Πηγή Χωρίς Μνήμη

- Discrete Memoryless Source (DMS):
  - διακριτού χρόνου
  - διακριτού αλφαβήτου
  - τα σύμβολα στην έξοδό της είναι ανεξάρτητα
  - ακολουθούν την ίδια κατανομή πιθανότητας
- Περιγράφεται πλήρως από:
  - το αλφάβητο $\boxed{\Phi = \{s_1, \cdots, s_N\}}$
  - και τις πιθανότητες εμφάνισης $\boxed{\{p_1, \cdots, p_N\}}$
- Ειδικές Περιπτώσεις:
  - Δυαδική Πηγή Χωρίς Μνήμη: $\boxed{\Phi = \{0,1\}}$ $\boxed{\{p, 1-p\}}$
  - Για $p=0.5$, Δυαδική Συμμετρική Πηγή Χωρίς Μνήμη

Εντροπία

- Η εντροπία μιας DMS ορίζεται ως

$$H(\Phi)=\sum_{i=1}^{N} p_i I(s_i)=-\sum_{i=1}^{N} p_i \log_2 p_i$$

- Φυσική Σημασία:
  – εκφράζει τη μέση αβεβαιότητα που έχω για την πηγή
  – είναι ο μέσος όρος της πληροφορίας των συμβόλων

- Όσο μεγαλύτερη εντροπία έχει μια πηγή,
  – τόσο περισσότερη πληροφορία φέρει, και
  – τόσο περισσότερα bits χρειάζονται για την κωδικοποίησή της

Ορισμός της εντροπίας ενός φυσικού συστήματος στην Στατιστική Μηχανική (εκφράζει τον βαθμό τυχαιότητας) →

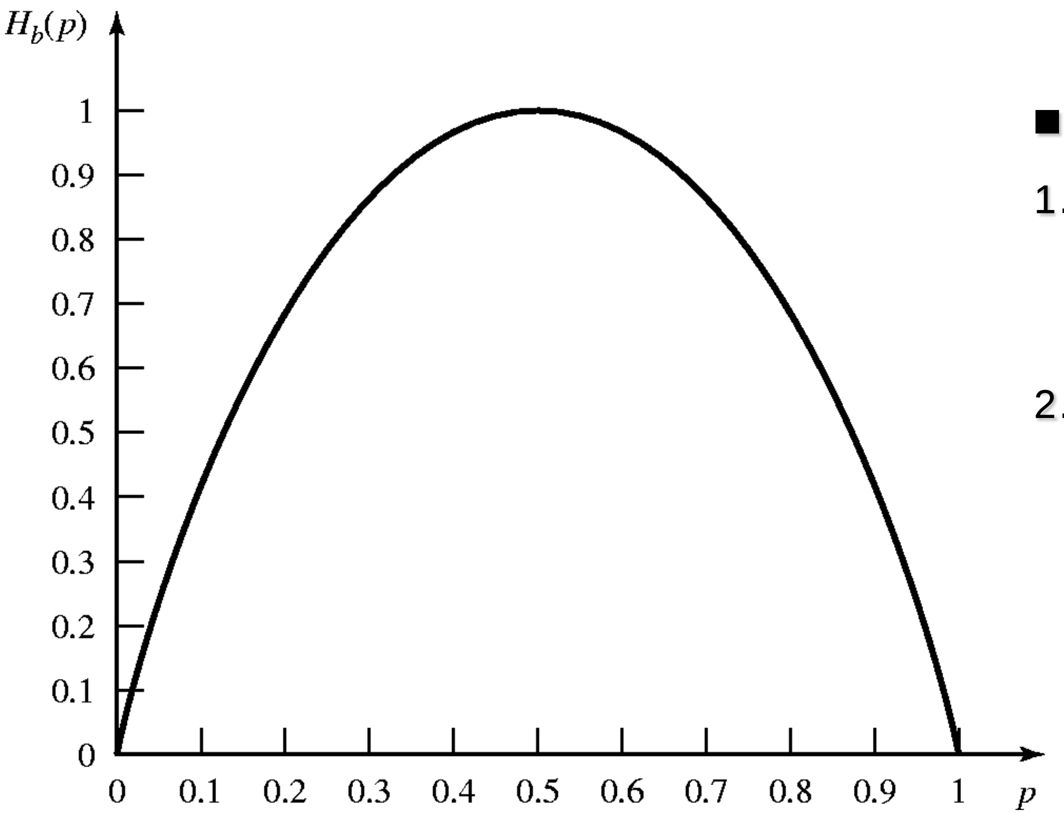
$$J = k \ln \Omega$$

2ος Νόμος Θερμοδυναμικής*: Η εντροπία ενός κλειστού συστήματος που δεν είναι σε ισορροπία τείνει να αυξάνεται.*

Συνάρτηση Δυαδικής Εντροπίας

- Αν έχω δυαδική DMS *Φ={0,1}*, με πιθανότητες εμφάνισης *{p,1-p}*, τότε ορίζεται η συνάρτηση δυαδικής εντροπίας

$$\boxed{H_b(p) = -p\log_2 p - (1-p)\log_2(1-p)}$$



- Παρατηρήσεις:

1. ελαχιστοποιείται όταν *p=0 ή 1*, *H(0)=H(1)=0*
2. μεγιστοποιείται όταν τα σύμβολα είναι ισοπίθανα, *H(0.5)=1*

Εντροπία Ομοιόμορφης Πηγής

- Είδαμε ότι η εντροπία της δυαδικής DMS μεγιστοποιείται για ισοπίθανα σύμβολα

Γενίκευση: Η εντροπία μιας N-αδικής DMS μεγιστοποιείται όταν τα σύμβολά της ακολουθούν ομοιόμορφη κατανομή, δηλαδή $p_i=1/N$ για $i=1,\ldots,N$

- Συμπέρασμα: Η εντροπία φράσσεται ως

$$\boxed{0 \le H(\Phi) \le \log_2 N}$$

  - όπου N το πλήθος του αλφαβήτου
  - και το άνω όριο επιτυγχάνεται για ομοιόμορφη πηγή

Κωδικοποίηση Πηγής

- Στόχος: Η αποδοτική αναπαράσταση μιας Μιαδικής πηγής
- *Κώδικες μεταβλητού μήκους:* Αξιοποιεί τη γνώση των στατιστικών ιδιοτήτων της πηγής
- Λειτουργικές απαιτήσεις:
  - Οι κωδικές λέξεις είναι δυαδικές
  - Ο κώδικας είναι μοναδικά αποκωδικοποιήσιμος
- Μέσο μήκος κώδικα

$$\boxed{\overline{L} = \sum_{i=1}^{N} p(s_i)\, l(s_i)}$$

Θεώρημα Κωδικοποίησης Πηγής

- ή «Το Πρώτο Θεώρημα του Shannon» (1948)
- Χρησιμότητα: πόσο μπορούμε να συμπιέσουμε μια πηγή χωρίς να εισάγουμε σφάλματα;

> Θεώρημα: Έστω πηγή με εντροπία H που κωδικοποιείται ώστε να παρέχει ρυθμό R *(bits/έξοδο πηγής)*.
>
> –Αν $R>H$, η πηγή μπορεί να κωδικοποιηθεί με οσοδήποτε μικρή πιθανότητα σφάλματος
>
> –Αν $R<H$, όσο πολύπλοκος κι αν είναι ο κωδικοποιητής πηγής, η πιθανότητα σφάλματος θα είναι μακριά από το 0

- Σχόλια:
  - Όπου R μπορείτε να θεωρήσετε το μέσο μήκος κώδικα (\bar{L})
  - ο Shannon δίνει την ικανή και αναγκαία συνθήκη
  - όμως δεν προτείνει κάποιο αλγόριθμο/μεθοδολογία για να φτιάξουμε έναν κωδικοποιητή όταν $R>H$
  - $R<H$: *Data compression, Rate-Distortion Theory*

Ειδικές Περιπτώσεις

- Ομοιόμορφη Πηγή:
  - $H(X)=log_2N$
  - δε μπορεί να συμπιεστεί
  - κάθε ακολουθία εξόδου είναι δυνατή (τυπική) και ισοπίθανη
- Πηγές με μνήμη:
  - ο ρυθμός εντροπίας παίζει τον ίδιο ρόλο με την εντροπία για στατικές πηγές
  - ο ρυθμός εντροπίας συγκλίνει γρήγορα στην τελική τιμή
- Παράδειγμα πηγής με μνήμη: αγγλικό κείμενο
  - για $n=1$ (αγνοώντας τη μνήμη), $H(X)=4.03$ *bits/letter*
  - για μπλοκ γραμμάτων (π.χ. $n=10$) συγκλίνει στην τιμή $H(X)=1.3$ *bits/letter*

Μη απωλεστική κωδικοποίηση δεδομένων

- *Προθεματικοί κώδικες*

- *Αλγόριθμος Huffman*

Προθεματικοί κώδικες

- Αλγόριθμοι κωδικοποίησης (συμπίεσης) πηγής
- Επιτυγχάνουν ρυθμούς κωδικοποίησης κοντά στην εντροπία (στο όριο συμπίεσης χωρίς απώλειες)
- Κωδικοποίηση από σταθερό σε μεταβλητό μήκος:
  - είσοδος: μπλοκ συμβόλων σταθερού μήκους (μήκος μπλοκ $\geqq 1$)
  - έξοδος: μπλοκ bits μεταβλητού μήκους (κωδική λέξη)
- Πρόβλημα: Συγχρονισμός
  - πώς μπορώ να βρω τα όρια των μπλοκ στην έξοδο για να γίνει η αποκωδικοποίηση
- Λύση: Προθεματικός
  - Καμμία κωδική λέξη δεν αποτελεί πρόθεμα κάποιας άλλης
    - » μοναδικά αποκωδικοποιήσιμος (κάθε έξοδος αντιστοιχεί σε μοναδική είσοδο)
    - » άμεσος (επιτρέπει απευθείας αποκωδικοποίηση)

Μέσο Μήκος Κώδικα

- Έστω DMS με πιθανότητες εμφάνισης *p*(s_i) *και ένας* κωδικοποιητής πηγής που αναθέτει *l*(s_i) bits στο σύμβολο s_i

Εάν ένας κώδικας είναι προθεματικός έχει τις εξής ιδιότητες:

- Kraft-McMillan inequality (αναγκαία και ικανή)

$$\sum_{i=1}^{N} 2^{-l(s_i)} \leq 1$$

- Ικανή συνθήκη για την ύπαρξη προθεματικου κώδικα
- Αναγκαία συνθήκη για την ύπαρξη μοναδικά αποκωδ/μου κώδικα

- Φράγματα στο μέσο μήκος.

Μπορεί να κατασκευαστεί προθεματικός για τον οποίο:

$$H(X) \leq \overline{L} < H(X) + 1$$

Αποδοτικότητα Κώδικα

- Η αποδοτικότητα ενός κώδικα ορίζεται ως

$$\eta = \frac{H(X)}{\overline{L}} \leq 1$$

  – και δείχνει πόσο κοντά βρίσκεται ο κωδικοποιητής στο όριο συμπίεσης της πηγής (εντροπία)

- Ένας κώδικας είναι αποδοτικός, όσο το η πλησιάζει στο 1

N-οστής Τάξης Επέκταση Πηγής

- Ο προθεματικός κώδικας (π.χ. ο αλγόριθμος Huffman) θεωρεί ένα μπλοκ από σύμβολα ως επεκτεταμένη είσοδο
- και τα κωδικοποιεί ως ένα σύνθετο σύμβολο, δηλαδή
  - έστω s_i και s_j
  - θέτει $\sigma_\kappa=(s_i,s_j)$
  - με πιθανότητα εμφάνισης $p(\sigma_\kappa)=p(s_i)p(s_j)$ για πηγή χωρίς μνήμη
- Γενικεύεται σε n-οστή επέκταση της πηγής και ισχύει

$$\boxed{H\left(X^n\right)\leq \overline{L}_n < H\left(X^n\right)+1}$$

- Μέσο μήκος κώδικα της επεκταμένης ακολουθίας πηγής

$$\boxed{\overline{L}=\overline{L}_n/n}$$

N-οστής Τάξης Επέκταση Πηγής (2)

- Για πηγή χωρίς μνήμη, αποδεικνύεται $\boxed{H\left(X^n\right) = nH\left(X\right)}$
- Προκύπτει $\boxed{H\left(X\right) \le \overline{L} < H\left(X\right) + \frac{1}{n}}$
- Συμπέρασμα: Η *n*-οστής τάξης επέκταση μιας πηγής αποφέρει κώδικες που είναι ολοένα και πιο κοντά στο όριο συμπίεσης (εντροπία) της πηγής

$$\boxed{\lim_{n\to\infty} \overline{L} = H\left(X\right)}$$

- Ερώτηση: Γιατί δε χρησιμοποιώ ένα πολύ μεγάλο *n*, ώστε να πετύχω συμπίεση κοντά στο όριο της εντροπίας;

Βήματα Αλγορίθμου Huffman (1)

- Δημιουργία Δυαδικού Δέντρου:

1. Διάταξε τις εισόδους κατά φθίνουσα σειρά πιθανοτήτων
2. Συγχώνευσε τα δύο σύμβολα με τις μικρότερες πιθανότητες και δημιούργησε νέο «σύμβολο»
3. Ανάθεσε στα δύο σύμβολα «0» και «1»
4. Ταξινόμησε εκ νέου τη λίστα των συμβόλων
5. Επανέλαβε τα παραπάνω μέχρι όλα τα σύμβολα συγχωνευτούν σε ένα τελικό σύμβολο

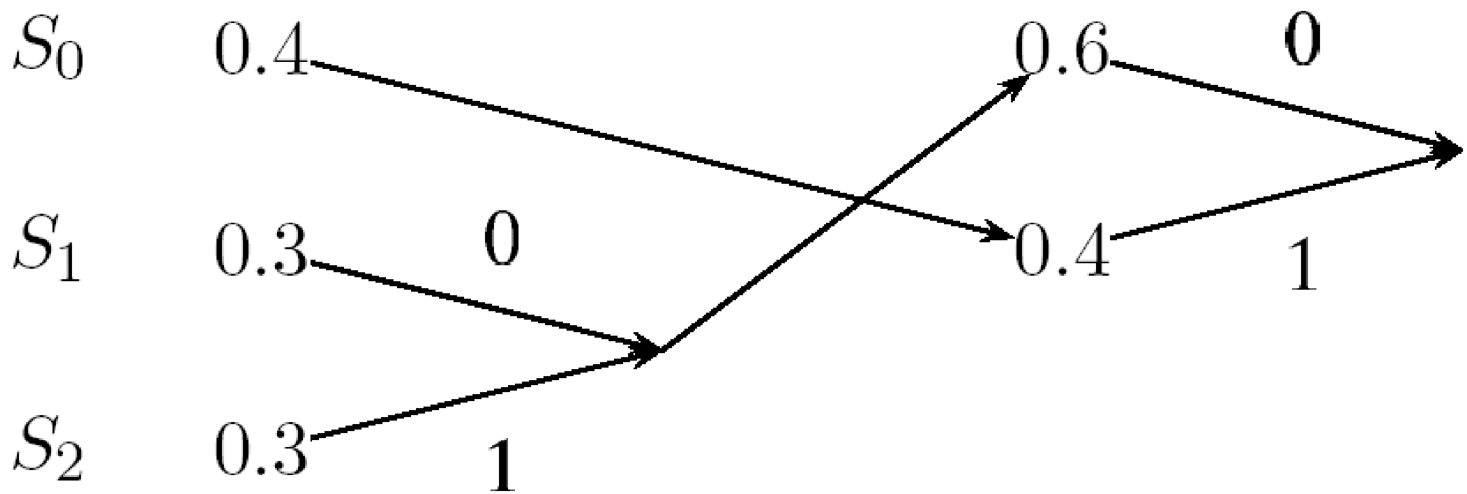
- Δημιουργήθηκε ένα δυαδικό δέντρο:
  - ρίζα: το τελικό σύνθετο σύμβολο
  - φύλλα: τα αρχικά σύμβολα
  - ενδιάμεσοι κόμβοι: σύνθετα σύμβολα

Βήματα Αλγορίθμου (2)

- Ανάθεση Bits σε Σύμβολα Εισόδου

1. Ξεκίνα από τη ρίζα και κινήσου προς ένα φύλλο
2. Η ακολουθία των bits που συναντώνται είναι η ακολουθία κωδικοποίησης
3. Επανέλαβε για όλα τα σύμβολα (φύλλα)

Παράδειγμα Huffman



- Προθεματική αντιστοίχηση:
  - s_0: 1
  - s_1: 00
  - s_2: 01
- Μονοσήμαντη και άμεση αποκωδικοποίηση

$$\underbrace{1}_{s_0}\underbrace{01}_{s_2}\underbrace{00}_{s_1}\underbrace{01}_{s_2}\underbrace{1}_{s_0}\underbrace{00}_{s_1}$$

Χαρακτηριστικά Huffman

- Μειονέκτημα:
  - απαιτεί να γνωρίζει εκ των προτέρων τις πιθανότητες εμφάνισης των συμβόλων της πηγής
  - δε μπορεί να χρησιμοποιηθεί σε εφαρμογές πραγματικού χρόνου

- Βέλτιστος:
  - ανάμεσα σε όλους τους προθεματικούς κώδικες (άρα μονοσήμαντα αποκωδικοποιήσιμους και άμεσους)

  πετυχαίνει το ελάχιστο μέσο μήκος κώδικα

- Συμβάσεις:
  - Ο τρόπος ανάθεσης 0 και 1
  - Η ταξινόμηση σε φθίνουσα σειρά (σχετίζεται με τη διασπορά του κώδικα)

Συμπίεση Πολυμεσικών Δεδομένων

Εισαγωγή στο πρόβλημα και επιλεγμένες εφαρμογές

Παράδειγμα: Συμπίεση Εικόνας

ΔΠΜΣ ΜΥΑ,   3 Απριλίου 2015

Εισαγωγή (1)

- Οι τεχνικές συμπίεσης βασίζονται στην απόρριψη της πλεονάζουσας πληροφορίας
- Ανάγκες που καλύπτονται
  - Εξοικονόμηση μνήμης
  - Ελάττωση χρόνου και εύρους ζώνης μετάδοσης
- Ιδιαίτερα μεγάλη η σημασία της συμπίεσης στις 2-D και 3-D εφαρμογές
- Κατηγορίες τεχνικών συμπίεσης
  - Απωλεστικές
  - Μη απωλεστικές

Εισαγωγή (2)

Εκμετάλλευση 3 ειδών πλεονασμού πληροφορίας:

1. Πλεονασμός κωδικοποίησης
   μπορούμε να χρησιμοποιήσουμε διαφορετική αναπαράσταση για gray levels που εμφανίζονται με διαφορετικές πιθανότητες (π.χ. κωδικοποίηση Huffman)

2. Πλεονασμός μεταξύ των pixel (interpixel)
   μπορούμε να εκμεταλλευτούμε τα επαναλαμβανόμενα μοτίβα ή τις μεγάλες ομαλές επιφάνειες (πχ κωδικοποίηση μήκους διαδρομών)

3. Φυσικο-οπτικός πλεονασμός
   εμφανίζεται λόγω του τρόπου με τον οποίο αντιλαμβάνεται ο άνθρωπος την οπτική πληροφορία (π.χ. κβαντισμός)

Εισαγωγή (3)

Τα βασικά στάδια μιας απωλεστικής τεχνικής:

1. Μετασχηματισμός της εικόνας στο κατάλληλο πεδίο

   - Πεδίο εικονοστοιχείων (π.χ., ADPCM)
   - Πεδίο συχνοτήτων (π.χ., DCT, Wavelet)
   - Πεδίο παραμέτρων μοντέλου
     (π.χ., στοχαστικά, γεωμετρικά, fractals)

Εισαγωγή (4)

2. Κβαντισμός των αποτελεσμάτων του 1ου σταδίου

   - ομοιόμορφος
   - ανομοιόμορφος
   - διανυσματικός

3. Προσδιορισμός του λεξικού (codebook) για την αναπαράσταση των εξόδων του κβαντιστή

   (συνήθως αντιμετωπίζεται ως πρόβλημα μη απωλεστικής κωδικοποίησης διακριτής πηγής)

Εισαγωγή (5)

Μη απωλεστικές τεχνικές:

- Κωδικοποίηση Huffman
- Bit plane coding
- Constant Area Coding
- Contour Tracing
- Κωδικοποίηση Μήκους Διαδρομών
- Συμπίεση LZW
- Κωδικοποίηση με Πρόβλεψη
- …

Απωλεστικές τεχνικές:

- Κωδικοποίηση με Πρόβλεψη (και κβαντισμό)
- Κωδικοποίηση με Μετασχηματισμούς
- … και πολλές άλλες

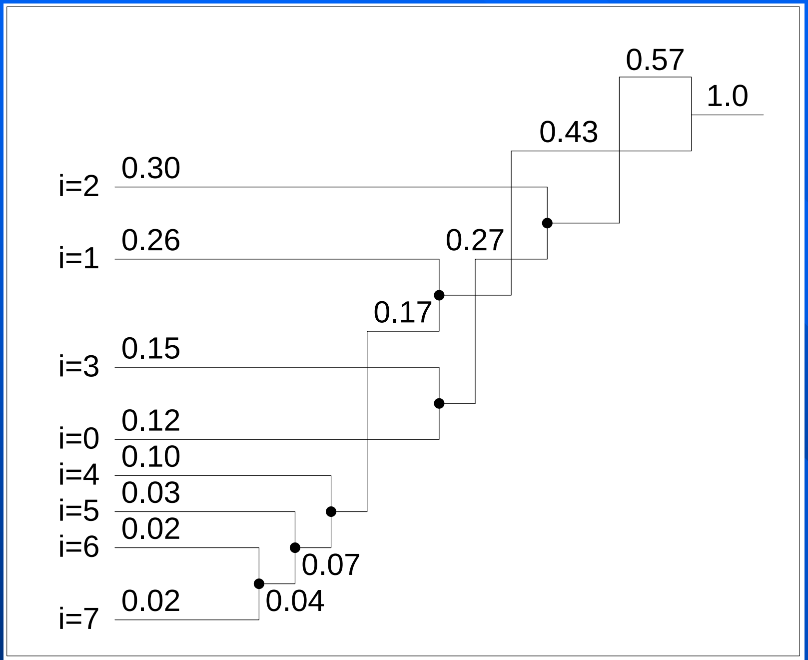
Κωδικοποίηση Huffman (1)

- Έστω εικόνα NxM που αναπαριστάται με B bits/pixel. Μπορούμε να εκτιμήσουμε την *συνάρτηση πυκνότητας πιθανότητας* p(i) κατασκευάζοντας το ιστόγραμμά της.
- <u>Κωδικοποίηση Εντροπίας</u>: αντιστοιχίζουμε μικρές κωδικές λέξεις στα επίπεδα φωτεινότητας που εμφανίζονται με μεγάλη πιθανότητα και μεγάλες κωδικές λέξεις σε αυτά που εμφανίζονται σπανιότερα.
- Τα μήκη των λέξεων επιλέγονται έτσι ώστε το μέσο μήκος να ελαχιστοποιείται. Από την Θεωρία Πληροφορίας ισχύει:

$$H(B) \le \bar{L} \le H(B) + 1$$

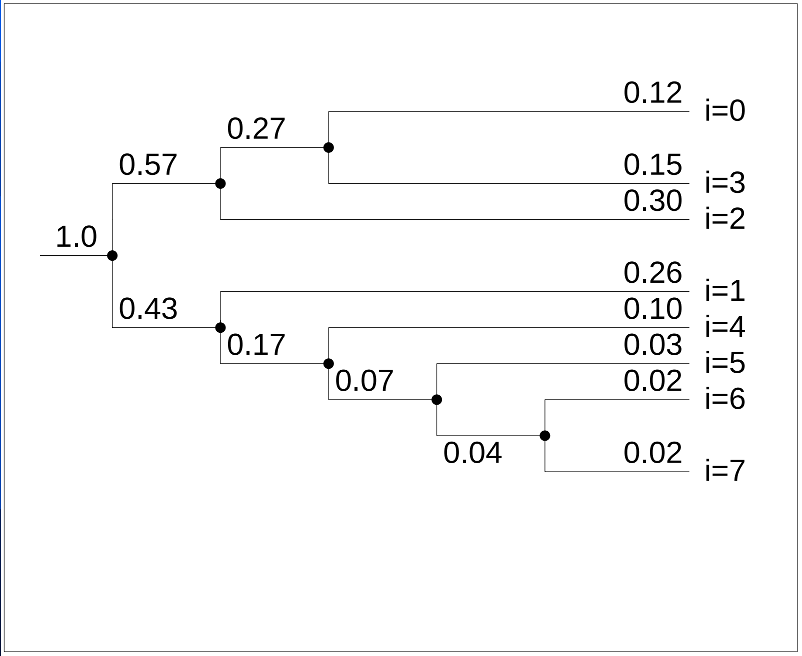
όπου *H(B)* η εντροπία της εικόνας.

Κωδικοποίηση Huffman (2)



← Κατασκευή του δέντρου Huffman

Δημιουργία των κωδικών λέξεων →



Κωδικοποίηση Μήκους Διαδρομών

Έστω $\boxed{x_1 \mid x_2 \mid \quad \ldots \quad \mid x_M}$ μία γραμμή εικόνας και x_i οι τιμές φωτεινότητας των pixels. Μπορεί να θεωρηθεί ότι η γραμμή αποτελείται από k τμήματα με μήκος l_i και φωτεινότητα g_i ,$1 \le i \le k$, και να αναπαρασταθεί με ζεύγη (g_i, l_i) ως εξής:

$$(x_1, \ldots, x_M) \to (g_1, l_1), (g_2, l_2), \ldots, (g_k, l_k)$$

Για δυαδικές εικόνες, μόνο τα μήκη l_i είναι απαραίτητο να κωδικοποιηθούν.

Για επιπλέον συμπίεση, μπορεί να χρησιμοποιηθεί κωδικοποίηση Huffman για τα μήκη.

Παράδειγμα

- Παράδειγμα γραμμής δυαδικής εικόνας και αναπαράστασή της με ζεύγη (g_i, l_i)

$k = 10$

$$
\begin{array}{ll}
l_1 = 3 & l_6 = 1 \\
l_2 = 6 & l_7 = 1 \\
l_3 = 1 & l_8 = 3 \\
l_4 = 1 & l_9 = 4 \\
l_5 = 1 & l_{10} = 8
\end{array}
$$

$$C = \frac{1 - p^M}{m(1-p)}$$

$$(1,3),(0,6),(1,1),(0,1),(1,1),(0,1),(1,1),(0,3),(1,4),(0,8)$$

$$\rightarrow (3,6,1,1,1,1,1,3,4,8)$$

Τροποποιημένη Κωδικοποίηση Read

- Είναι μία επέκταση της κωδικοποίησης μήκους διαδρομών στις δύο διαστάσεις.
- Εκμεταλλευόμαστε την συσχέτιση μεταξύ των διαδοχικών γραμμών της εικόνας. Κάθε καινούρια γραμμή κωδικοποιείται με βάση την προηγούμενή της.
- Η READ χρησιμοποιεί 3 στοιχεία μετάβασης στην τρέχουσα γραμμή, (a_0, a_1, a_2) και 2 στην γραμμή αναφοράς, (b_1, b_2) .

Άλλες μη-απωλεστικές τεχνικές

- Bit plane coding

Αποσύνθεση σε L δυαδικές εικόνες και διαφορετική κωδικοποίηση ανάλογα με τη «βαρύτητά» τους

- Constant Area Coding

Υποδιαίρεση σε υποεικόνες και προθεματική κωδικοποίηση των υποεικόνων ανάλογα με το αν είναι Λευκές, Μαύρες ή Μικτές (0, 10, 11). Στην τελευταία περίπτωση αποστέλλεται όλη η υποεικόνα

- Contour Tracing

Βρίσκεται το σύνολο των «ισουψών» καμπυλών της εικόνας οι οποίες κωδικοποιούνται με αποδοτικό τρόπο

Κωδικοποίηση με Πρόβλεψη (1)

- Συνήθως τα δεδομένα της εικόνας σε μία γειτονιά είναι συσχετισμένα μεταξύ τους, οπότε είναι δυνατόν να «προβλέψουμε» την τιμή ενός pixel με βάση αυτά που βρίσκονται στην γειτονιά του.
- Για κάθε pixel (x,y) αποθηκεύουμε τη «νέα» πληροφορία που περιέχει.
- «Νέα» πληροφορία είναι η διαφορά της προβλεφθείσας από την πραγματική τιμή του pixel.

Κωδικοποίηση με Πρόβλεψη (2)

Για κάθε διαδοχικό pixel f_n της εικόνας, η διάταξη πρόβλεψης παράγει την αναμενόμενη τιμή βασισμένη στις προηγούμενες εισόδους. Η τιμή αυτή στρογγυλοποιείται στον κοντινότερο ακέραιο. Το λάθος πρόβλεψης τότε είναι

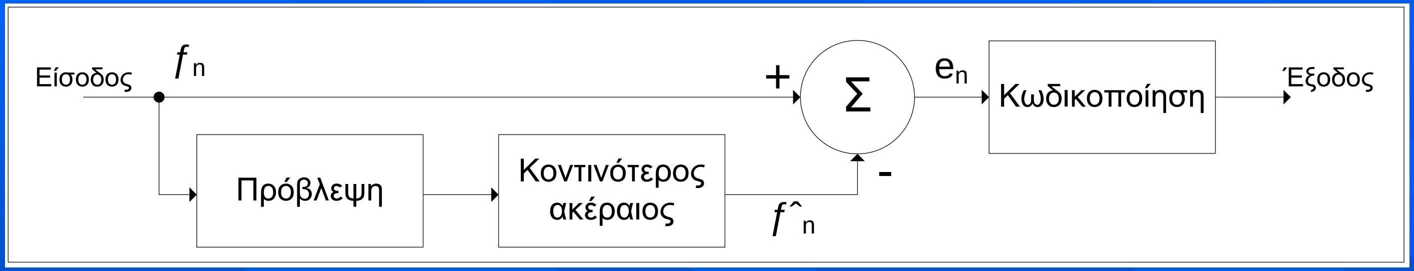
$$e_n = f_n - \hat{f}_n$$

Η αρχική τιμή μπορεί να ανακατασκευαστεί ως

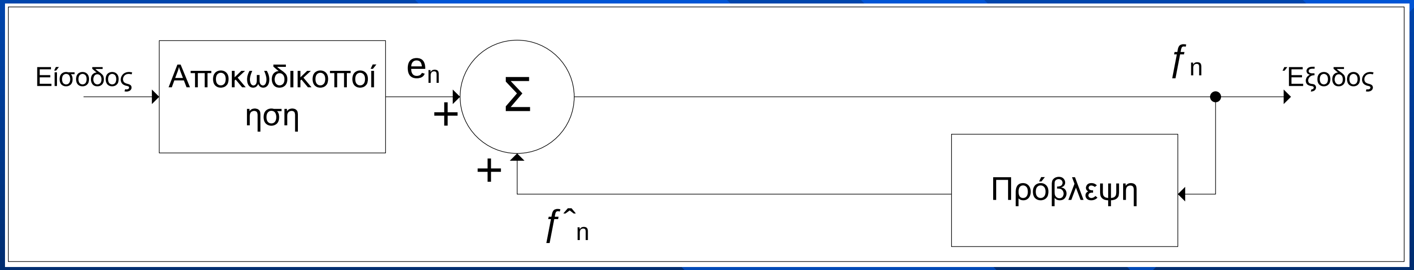
$$f_n = e_n + \hat{f}_n$$

όπου $\hat{f}_n$ είναι μία συνάρτηση πρόβλεψης: $\hat{f}_n = \text{round}\left[\sum_{i=1}^{m} a_i f_{n-i}\right]$

Κωδικοποίηση με Πρόβλεψη (3)



Διάταξη συμπίεσης



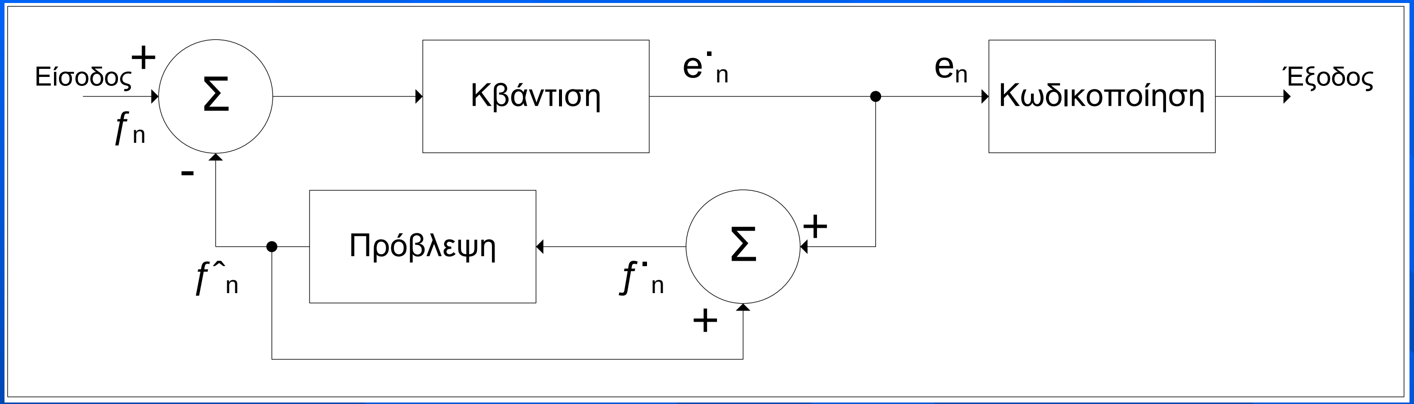
Διάταξη ανακατασκευής

Κωδικοποίηση με Πρόβλεψη (4)

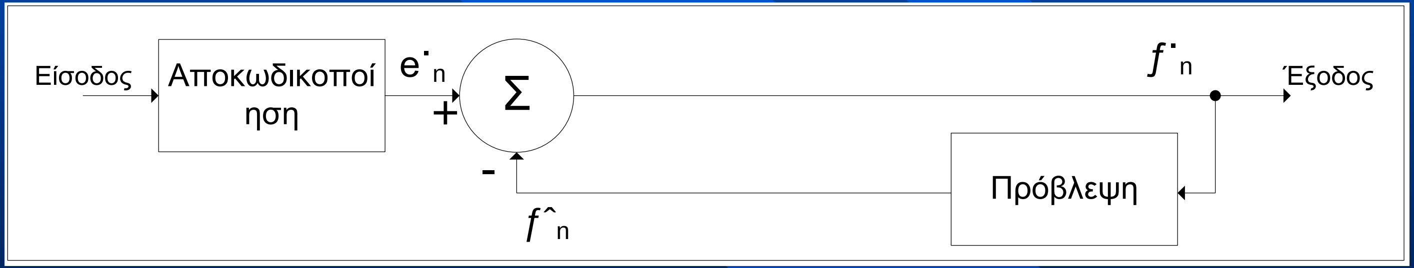
Για καλύτερη συμπίεση χρησιμοποιούμε κβαντισμό για το e_n . Σε αυτή την περίπτωση το λάθος πρόβλεψης κβαντίζεται σε ένα περιορισμένο εύρος τιμών.

Η διαδικασία της κβάντισης είναι μη αντιστρέψιμη και απωλεστική (χάνεται πληροφορία).

Κωδικοποίηση με Πρόβλεψη (5)



Διάταξη συμπίεσης



Διάταξη ανακατασκευής

Κωδικοποίηση με Μετασχηματισμούς (1)

Πρόκειται για απωλεστική μέθοδο συμπίεσης.

Επιλέγουμε και κωδικοποιούμε τους συντελεστές του μετασχηματισμού οι οποίοι περιέχουν την μεγάλη ενέργεια, και απορρίπτουμε τους υπόλοιπους. Η εικόνα ανακτάται (με απώλειες) με την αντιστροφή του μετασχηματισμού.

$$F = T[f] \quad \rightarrow \quad \widehat{f} = T^{-1}[\widehat{F}]$$

Κωδικοποίηση με Μετασχηματισμούς (2)

Βήματα

1. Επιλογή μετασχηματισμού (DFT, DCT, WHT, DST, Wavelet)
2. Επιλογή διαστάσεων των blocks (8x8, 16x16)
3. Επιλογή του αριθμού των bits για την κβάντιση των συντελεστών
4. Σχεδιασμός βέλτιστου κβαντιστή

Για επιπλέον συμπίεση, οι κβαντισμένοι συντελεστές μπορούν να κωδικοποιηθούν με Huffman.

Ο Μετασχηματισμός DCT

1-D DCT & IDCT

$$C(k) = \begin{cases} 2\sum_{n=0}^{N-1} x(n)\cos\frac{k\pi}{2N}(2n+1), & 0 \le k \le N-1 \\ 0, & \alpha\lambda\lambda o\acute{\upsilon} \end{cases}$$

$$x(n) = \begin{cases} \frac{1}{N}\sum_{k=0}^{N-1} w(k)C(k)\cos\frac{k\pi}{2N}(2n+1), & 0 \le n \le N-1 \\ 0, & \alpha\lambda\lambda o\acute{\upsilon} \end{cases}$$

όπου $w(k) = \begin{cases} 1/2, & k = 0 \\ 1, & 1 \le k \le N-1 \end{cases}$

Υπολογισμός του DCT μέσω DFT

DCT

1. $y(n) = x(n) + x(2N-1-n)\,, \quad 0 \le n \le 2N-1$
2. $Y(k) = DFT\{y(n)\}\,, \quad 0 \le k \le 2N-1$
3. $C(k) = \begin{cases} e^{-j\frac{k\pi}{2N}}Y(k)\,, & 0 \le k \le N-1 \\ 0\,, & \alpha\lambda\lambda o\acute{\upsilon} \end{cases}$

IDCT

1. $Y(k) = \begin{cases} e^{j\frac{k\pi}{2N}}C(k)\,, & 0 \le k \le N-1 \\ 0\,, & k = N \\ -e^{j\frac{k\pi}{2N}}C(2N-k)\,, & N+1 \le k \le 2N-1 \end{cases}$
2. $y(n) = IDFT\{Y(k)\}\,, \quad 0 \le n \le 2N-1$
3. $x(n) = \begin{cases} y(n)\,, & 0 \le n \le N-1 \\ 0\,, & \alpha\lambda\lambda o\acute{\upsilon} \end{cases}$

- Ο μετασχηματισμός 2-D DCT μπορεί να υπολογιστεί μέσω του 1-D DCT με τη μέθοδο γραμμών-στηλών
- Ερμηνεία της μεγάλης ενεργειακής συγκέντρωσης

Κωδικοποίηση των συντελεστών του DCT

Μετασχηματισμός DCT

8	7	6	4	3	2	1	0
7	6	5	4	3	2	1	0
6	5	4	3	3	1	1	0
4	4	3	3	2	1	0	0
3	3	3	2	1	1	0	0
2	2	1	1	1	0	0	0
1	1	1	0	0	0	0	0
0	0	0	0	0	0	0	0

1	1	1	1	1	0	0	0
1	1	1	1	0	0	0	0
1	1	1	0	0	0	0	0
1	1	0	0	0	0	0	0
1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

1	1	0	1	1	0	0	0
1	1	1	1	0	0	0	0
1	1	0	0	0	0	0	0
1	0	0	1	0	0	0	0
0	1	0	0	0	0	0	0
0	1	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

- Αντιστοίχιση bits κβαντισμού σε block 8x8
- ο αριθμός bits ανά συντελεστή είναι ανάλογος της μεταβλητότητάς του

Διαδικασία συμπίεσης με DCT

Συμπίεση:

1. Διαίρεση σε block
2. Μετασχηματισμός DCT
3. Επιλογή συντελεστών (μέθοδοι: ζώνης, κατωφλίου)
4. Κβάντιση και αποθήκευση συντελεστών

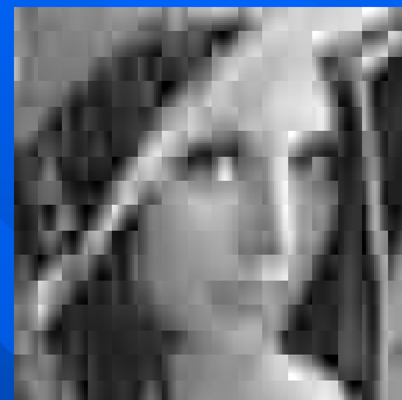
Αποσυμπίεση:

1. Αντιστροφή κβάντισης με το βήμα κβαντισμού
2. Ανακατασκευή block
3. Μετασχηματισμός IDCT
4. Επανένωση των block

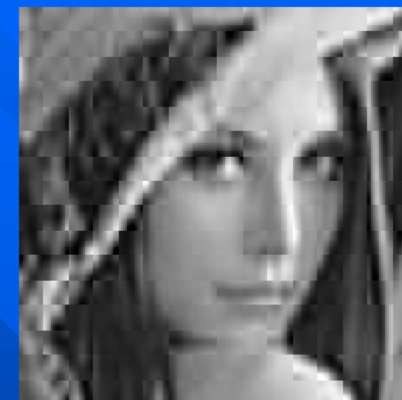
Παραδείγματα συμπίεσης (1)

Αρχική

126:1



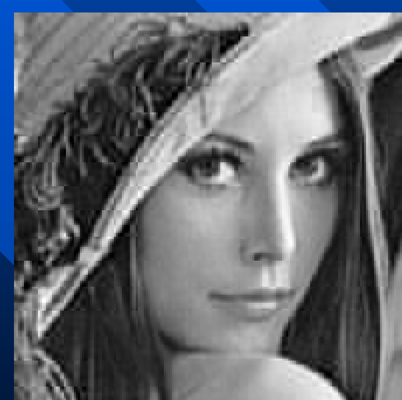
51:1



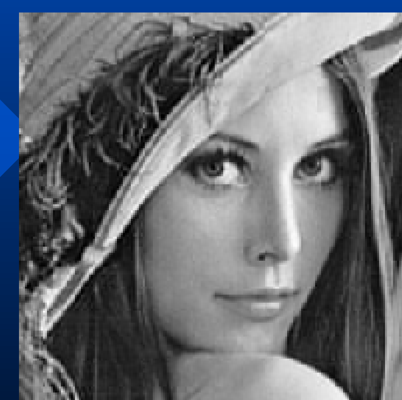
28:1

15:1

8:1

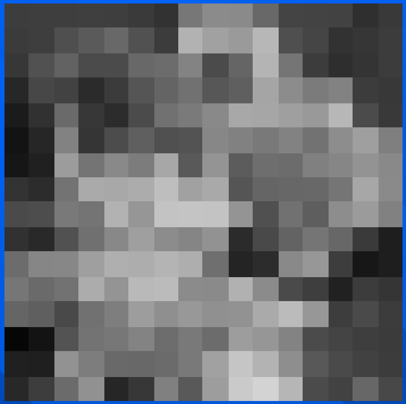


4:1

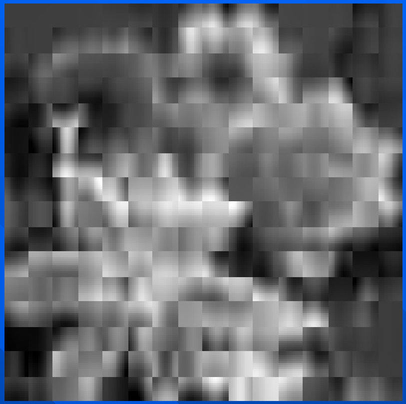


Παραδείγματα συμπίεσης (2)

Αρχική



126:1



51:1

28:1

15:1

8:1

4:1

Τέλος Ενότητας

Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο
ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΕΚΠΑΙΔΕΥΣΗ ΚΑΙ ΔΙΑ ΒΙΟΥ ΜΑΘΗΣΗ
επένδυση στην κοινωνία της γνώσης
ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ ΚΑΙ ΘΡΗΣΚΕΥΜΑΤΩΝ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ
ΕΣΠΑ 2007-2013
πρόγραμμα για την ανάπτυξη
ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ
Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης

Σημείωμα Ιστορικού Εκδόσεων Έργου

Το παρόν έργο αποτελεί την έκδοση 1.0.

Σημείωμα Αναφοράς

Copyright Πανεπιστήμιο Πατρών, Κώστας Μπερμπερίδης «Μελέτη Περιπτώσεων στη Λήψη Αποφάσεων: Συμπίεση Πολυμεσικών Δεδομένων». Έκδοση: 1.0. Πάτρα 2015. Διαθέσιμο από τη δικτυακή διεύθυνση:

https://eclass.upatras.gr/courses/MATH959/

Σημείωμα Αδειοδότησης

Σημείωμα Χρήσης Έργων Τρίτων

Το Έργο αυτό κάνει χρήση των ακόλουθων έργων:

Διαφάνειες: 54- 55

http://students.ceid.upatras.gr/~mprokala/techarticles/DigitalImageProcessing/imageRestoration.htm

Διατήρηση Σημειωμάτων

Οποιαδήποτε αναπαραγωγή ή διασκευή του υλικού θα πρέπει να συμπεριλαμβάνει:

- το Σημείωμα Αναφοράς
- το Σημείωμα Αδειοδότησης
- τη δήλωση Διατήρησης Σημειωμάτων
- το Σημείωμα Χρήσης Έργων Τρίτων (εφόσον υπάρχει) μαζί με τους συνοδευόμενους υπερσυνδέσμους.