Κεφάλαιο 1

Μέτρηση πληροφορίας

Η μέτρηση της πληροφορίας ανάγεται στη μέτρηση της αβεβαιότητας για ένα σύνολο ενδεχομένων. Ως πηγή πληροφορίας ορίζεται η διαδικασία παραγωγής ενός συνόλου συμβόλων ή ενδεχομένων :

$$S = \{s_1, s_2, \cdots, s_n\}.$$

Σε κάθε σύμβολο αντιστοιχεί μια πιθανότητα εμφάνισης του αντίστοιχου ενδεχομένου

$$\{p_1, p_2, \cdots, p_n\}.$$

Ολόκληρο το σύνολο συνιστά τη βεβαιότητα με πιθανότητα 1.

Η πληροφορία για κάθε ενδεχόμενο μετρά την αβεβαιότητα εμφάνισής του. Γι΄ αυτό ορίζεται ως μια μη αρνητική συνάρτηση της πιθανότητας πραγματοποίησης ενός ενδεχομένου, ώστε όσο μικρότερη είναι η πιθανότητα του ενδεχομένου τόσο μεγαλύτερη είναι η πληροφορία που κομίζει. Άρα η πληροφορία είναι φθίνουσα συνάρτηση της πιθανότητας και μηδενική για το βέβαιο γεγονός. Επιπλέον η πληροφορία της ταυτόχρονης πραγματοποίησης δύο ανεξάρτητων συμβάντων είναι ίση με το άθροισμα των δύο ξεχωριστών ποσοτήτων πληροφορίας. Η λογαριθμική συνάρτηση ικανοποιεί όλες αυτές τις απαιτήσεις, οπότε για το ενδεχόμενο A, η πληροφορία ορίζεται ως

$$I(A) = f(p_A) = -\log(p_A).$$

Πρόκειται επομένως για τυχαία μεταβλητή, αφού εξαρτάται από τυχαία συμβάντα. Ως μονάδα μέτρησης ορίζεται η ποσότητα πληροφορίας ενός ενδεχομένου με πιθανότητα 1/2. Ως βάση της λογαριθμικής συνάρτησης αρμόζει να ληφθεί το 2, οπότε η μονάδα μέτρησης είναι το bit.

Η ποσότητα πληροφορίας της πηγής, που χαρακτηρίζει το σύνολο των ενδεχομένων, ονομάζεται **εντροπία** και ισούται με την προσδοκητή τιμή της πληροφορίας των ενδεχομένων,

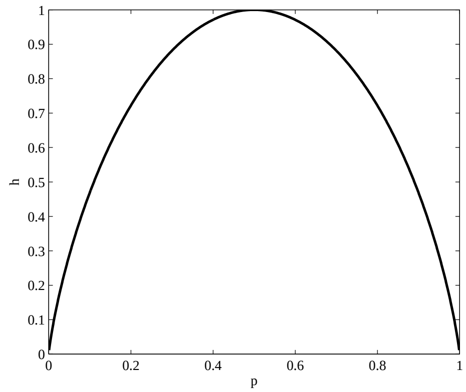
$$H(S) = E\{I(s)\} = -\sum_{i=1}^{n} p_i \log_2 p_i = h(p_1, p_2, \cdots, p_n).$$

Στο Σχήμα 1.1 δίδεται γραφικά η εντροπία για μια πηγή πληροφορίας δύο συμβόλων ως συνάρτηση της πιθανότητας εμφάνισης του ενός ενδεχομένου.

Η εντροπία ορίζεται και στην περίπτωση που το πλήθος των ενδεχομένων είναι άπειρο, αλλά αριθμήσιμο.

Παράδειγμα 1.1. Ας είναι τυχαία μεταβλητή με πιθανότητες

$$\Pr\{X = k\} = 2^{-k}, \quad k = 1, 2, \ldots$$



Σχήμα 1.1: Εντροπία δυαδικής πηγής

Η εντροπία θα είναι

$$H(X) = \sum_{k=1}^{\infty} k2^{-k} = 2. \quad \square$$

Ιδιότητες της συνάρτησης εντροπίας:

1. Η εντροπία είναι συνεχής συνάρτηση των πιθανοτήτων στο διάστημα $(0, 1]$.

2. Η εντροπία είναι συμμετρική συνάρτηση των πιθανοτήτων.

3. Η εντροπία είναι κυρτή συνάρτηση των πιθανοτήτων.

4. Η εντροπία είναι μέγιστη όταν όλα τα ενδεχόμενα είναι εξίσου πιθανά:

$$0 \leqslant h(p_1, p_2, \cdots, p_n) \leqslant \log_2 n.$$

5. $h(p_1, p_2, \cdots, p_n, 0) = h(p_1, p_2, \cdots, p_n)$.

6. $h(\frac{1}{n}, \cdots, \frac{1}{n}) < h(\frac{1}{n+1}, \cdots, \frac{1}{n+1})$

7. $h(\frac{1}{mn}, \cdots, \frac{1}{mn}) = h(\frac{1}{m}, \cdots, \frac{1}{m}) + h(\frac{1}{n}, \cdots, \frac{1}{n})$

8. $h(p_1, \cdots, p_n, q_1, \cdots, q_m) = h(p, q) + ph(\frac{p_1}{p}, \cdots, \frac{p_n}{p}) + qh(\frac{q_1}{q}, \cdots, \frac{q_m}{q})$
   όπου: $p = p_1 + \cdots + p_n, q = q_1 + \cdots + q_m, p + q = 1$.

Η **σχετική εντροπία** ή **απόσταση Kullback** μεταξύ δύο κατανομών πιθανοτήτων ορίζεται ως

$$D(p\|q) = \sum_{i=1}^{n} p_i \log_2 \frac{p_i}{q_i}.$$

Για κάθε τυχαία μεταβλητή ισχύει η **ανισότητα του Jensen**:

$$\text{Αν η συνάρτηση } f \text{ είναι κυρτή, τότε } E\{f(X)\} \leqslant f(E\{X\}).$$

Η απόδειξη γίνεται επαγωγικά. Στην περίπτωση δύο ενδεχομένων με πιθανότητες p_1 και $p_2 = 1 - p_1$ η ανισότητα ισοδυναμεί με

$$p_1 f(x_1) + (1 - p_1) f(x_2) \leqslant f(p_1 x_1 + p_2 x_2),$$

που είναι άμεση συνέπεια του ορισμού της κυρτής συνάρτησης. Ας υποθέσουμε ότι η ανισότητα ισχύει για $k = n - 1$. Θα αποδείξουμε ότι ισχύει για $k = n$.

$$\sum_{i=1}^{n} p_i f(x_i) = \sum_{i=1}^{n-1} p_i f(x_i) + p_n f(x_n) = (1 - p_n) \sum_{i=1}^{n-1} \frac{p_i}{1 - p_n} f(x_i) + p_n f(x_n),$$

$$\sum_{i=1}^{n} p_i f(x_i) \leqslant (1 - p_n) f\left(\sum_{i=1}^{n-1} \frac{p_i}{1 - p_n} x_i\right) + p_n f(x_n) \leqslant f\left(\sum_{i=1}^{n} p_i x_i\right).$$

Από την ανισότητα αυτή προκύπτει η ανισότητα της πληροφορίας

$$D(p\|q) \geqslant 0, \quad \text{με ισότητα, εάν, και μόνο εάν,} \quad p_i = q_i, \forall i.$$

Πράγματι επειδή η λογαριθμική συνάρτηση είναι κυρτή,

$$\sum_{i=1}^{n} p_i \log_2 \frac{q_i}{p_i} \leqslant \log_2 \sum_{i=1}^{n} p_i \frac{q_i}{p_i} = 0.$$

Για δύο πηγές πληροφορίας ορίζεται η από κοινού εντροπία

$$H(S_1, S_2) = -\sum_{i=1}^{m} \sum_{j=1}^{n} \Pr\{s_i, s_j\} \log_2 \Pr\{s_i, s_j\}.$$

Ορίζεται επίσης η δεσμευμένη εντροπία

$$H(S_1|S_2) = -\sum_{i=1}^{m} \sum_{j=1}^{n} \Pr\{s_i, s_j\} \log_2 \Pr\{s_i|s_j\}$$

Ισχύει ο κανόνας της αλυσίδας

$$H(S_1, S_2) = H(S_1) + H(S_2|S_1) = H(S_2) + H(S_1|S_2),$$

και οι ακόλουθες ανισότητες

$$H(S_1|S_2) \leqslant H(S_1), \quad H(S_2|S_1) \leqslant H(S_2),$$

$$H(S_1, S_2) \leqslant H(S_1) + H(S_2).$$

Η αμοιβαία πληροφορία δύο πηγών ισούται με την πρόσθετη πληροφορία της μίας πηγής ως προς την άλλη. Ορίζεται επομένως ως η απόσταση από την περίπτωση της ανεξαρτησίας,

$$I(S_1, S_2) = \sum_{i=1}^{m} \sum_{j=1}^{n} \Pr\{s_i, s_j\} \log_2 \frac{\Pr\{s_i, s_j\}}{\Pr\{s_i\}\Pr\{s_j\}}.$$

Με βάση τις ιδιότητες που αναφέρθηκαν ανωτέρω θα ισχύει

$$
\begin{aligned}
I(S_1, S_2) &= H(S_1) - H(S_1|S_2) \\
I(S_1, S_2) &= H(S_1) + H(S_2) - H(S_1, S_2) \\
I(S_1, S_2) &\geqslant 0
\end{aligned}
$$

Παράδειγμα 1.2. Δίδεται ο πίνακας των από κοινού πιθανοτήτων για ένα ζεύγος τυχαίων μεταβλητών

$$\begin{bmatrix} \frac{1}{3} & \frac{1}{3} \\ 0 & \frac{1}{3} \end{bmatrix}.$$

Η εντροπία τόσο της πηγής S_1, όσο και της πηγής S_2, είναι:

$$H(S_1) = H(S_2) = \frac{1}{3}\log_2 3 + \frac{2}{3}(\log_2 3 - 1) = \log_2 3 - \frac{2}{3} = 0,918 \text{ bits}.$$

Η από κοινού εντροπία είναι:

$$H(S_1, S_2) = \log_2 3 = 1,585 \text{ bits}.$$

Η αμοιβαία πληροφορία είναι:

$$I(S_1, S_2) = \log_2 3 - \frac{4}{3} = 0,252 \text{ bits}. \quad \square$$

Τυπικές ακολουθίες συμβόλων

Η εφαρμογή του νόμου των μεγάλων αριθμών συνεπάγεται την ασυμπτωτικά ομοιόμορφη κατανομή των ακολουθιών από σύμβολα της πηγής

$$-\frac{1}{L}\log_2 p(X_1, \ldots, X_L) \to H(X),$$

όπου οι μεταβλητές $X_1, \ldots, X_L$ είναι ανεξάρτητες με την ίδια κατανομή πιθανοτήτων, και με εντροπία $H(X)$.

Μία ακολουθία $x_1, \ldots, x_L$ ονομάζεται τυπική, αν

$$2^{-L(H(X)+\epsilon)} \leqslant p(x_1, \ldots, x_L) \leqslant 2^{-L(H(X)-\epsilon)}.$$

Η πιθανότητα του συνόλου των τυπικών ακολουθιών είναι σχεδόν 1. Το πλήθος των τυπικών ακολουθιών είναι περίπου $2^{LH(X)}$.

Συνέπεια των παραπάνω είναι η δυνατότητα συμπίεσης της πηγής πληροφορίας που παριστάνεται από τη μεταβλητή X σε $H(X)$ bits ανά σύμβολο της πηγής πληροφορίας.

Κεφάλαιο 2

Κωδικοποίηση πηγής πληροφορίας

Κωδικοποίηση είναι η αντιστοιχία των συμβόλων μιας πηγής πληροφορίας σε κωδικές λέξεις ενός δοσμένου αλφαβήτου. Ένας κώδικας, εφόσον δεν είναι αμφίσημος, είναι μοναδικά αποκωδικοποιήσιμος, αν δεν υπάρχει ακολουθία συμβόλων που η κωδική της έκφραση να επιδέχεται περισσότερες της μίας ερμηνείες. Ένας κώδικας είναι στιγμιαίος, αν το πέρας μιας κωδικής λέξης προσδιορίζεται μόνο από την ίδια, και όχι με τη βοήθεια των κωδικών λέξεων που ακολουθούν. Αυτό σημαίνει ότι καμία κωδική λέξη ενός στιγμιαίου κώδικα δεν είναι πρόθεμα κάποιας άλλης. Ένας δυαδικός στιγμιαίος κώδικας μπορεί να παρασταθεί με ένα δυαδικό δένδρο, στα φύλλα του οποίου ευρίσκονται οι κωδικές λέξεις. Το δένδρο αυτό μπορεί να χρησιμοποιηθεί για την αποκωδικοποίηση.

Ανισότητα του Kraft

Ικανή και αναγκαία συνθήκη για να είναι ένας δυαδικός κώδικας στιγμιαίος, είναι τα μήκη των κωδικών λέξεων $\{l_1, \ldots, l_N\}$ να ικανοποιούν την ακόλουθη ανισότητα

$$\sum_{n=1}^{N} 2^{-l_n} \leqslant 1$$

Η ίδια συνθήκη είναι ικανή και αναγκαία για να είναι ένας δυαδικός κώδικας μοναδικά αποκωδικοποιήσιμος (McMillan).

Ένας κώδικας ονομάζεται βέλτιστος, εάν το μέσο μήκος κωδικής λέξης,

$$\bar{l} = \sum_{n=1}^{N} p_n l_n,$$

είναι ελάχιστο. Με βάση τις παραπάνω ιδιότητες η εύρεση του βέλτιστου κώδικα μπορεί να περιορισθεί στην κλάση των στιγμιαίων κωδίκων. Μία απλουστευτική επίλυση δίδει την ακόλουθη σχέση ανάμεσα στην πληροφορία ενός συμβόλου και στο μήκος της αντίστοιχης κωδικής λέξης

$$l_n = \lceil -\log_2 p_n \rceil$$

Αν $\bar{l}$ είναι το μέσο μήκος κωδικής λέξης του κώδικα με τα παραπάνω μήκη, και $H(X)$ είναι η εντροπία της πηγής πληροφορίας, τότε

$$H(X) \leqslant \bar{l} < H(X) + 1$$

Κατ΄ επέκταση για την κωδικοποίηση μίας ακολουθίας L συμβόλων της πηγής πληροφορίας ισχύει

$$H(X) \leqslant \bar{l} < H(X) + \frac{1}{L}$$

Αυτό σημαίνει ότι το μέσο μήκος κωδικής λέξης μπορεί να πλησιάσει την εντροπία όσο κοντά είναι επιθυμητό.

Τα παραπάνω μήκη προέκυψαν χωρίς να λαμβάνεται υπόψη ότι τα μήκη είναι ακέραια. Αποδεικνύεται ότι ο βέλτιστος στιγμιαίος κώδικας έχει τις ακόλουθες ιδιότητες :

- Αν $p_m > p_n$, τότε $l_m \leqslant l_n$.
- Οι δύο κωδικές λέξεις μέγιστου μήκους έχουν το ίδιο μήκος.
- Οι δύο κωδικές λέξεις μέγιστου μήκους διαφέρουν μόνο στο τελευταίο ψηφίο.

Ο κώδικας Huffman ικανοποιεί τις παραπάνω ιδιότητες.

2.1 Κωδικοποίηση Huffman

Η κατασκευή του δυαδικού κώδικα Huffman, ή ισοδύναμα του δένδρου Huffman, υλοποιείται με τον ακόλουθο αλγόριθμο:

- Ένωση των δύο συμβόλων με χαμηλότερη πιθανότητα σε ένα κόμβο με πιθανότητα το άθροισμα των δύο. Οι δύο κλάδοι προς τους κόμβους - παιδιά διακρίνονται με τα δυαδικά ψηφία.
- Αντικατάσταση των κόμβων-παιδιά από τον κόμβο-γονέα που δημιουργήθηκε. Εάν η πιθανότητα είναι 1, η διεργασία ολοκληρώθηκε. Άλλως, επαναλαμβάνεται το προηγούμενο βήμα.

Το δένδρο που κατασκευάζεται έχει την αδελφική ιδιότητα.

- Κάθε κόμβος, πλην της ρίζας, έχει αδελφικό κόμβο.
- Όλοι οι κόμβοι του δένδρου μπορούν να διαταχθούν κατά φθίνουσα τάξη πιθανοτήτων, ενώ οι αδελφικοί κόμβοι γειτονεύουν.

Ως αποτέλεσμα των ιδιοτήτων αυτών ικανοποιείται η ισότητα

$$\sum_{n=1}^{N} 2^{-l_n} = 1$$

Πρόκειται για ένα πλήρη στιγμιαίο κώδικα. Ακόμα θα ισχύει

$$H(X) \leqslant \bar{l} < H(X) + 1$$

Λαμβάνοντας υπόψη τη δομή και τις ιδιότητες του δένδρου Huffman είναι δυνατή η προσαρμογή του, όταν οι πιθανότητες των συμβόλων είναι αρχικά άγνωστες και πρέπει να εκτιμηθούν από τα υπό συμπίεση δεδομένα.

Παράδειγμα 2.1. Έστω πηγή που παράγει δύο σύμβολα $\{\alpha, \beta\}$ με πιθανότητες αντίστοιχα $\{3/4, 1/4\}$. Ζητείται η κωδικοποίηση ακολουθιών από τρία σύμβολα της πηγής με την προϋπόθεση της ανεξαρτησίας. Δίδεται κατ' αρχήν ο πίνακας με τις πιθανότητες που υπολογίζονται στα διαδοχικά βήματα του αλγορίθμου. Η αρίθμηση των βημάτων γίνεται με βάση των αριθμό των εναπομεινάντων μηνυμάτων.

Σύμβολο	8	7	6	5	4	3	2	1
$\alpha\alpha\alpha$	27/64	27/64	27/64	27/64	27/64	27/64	27/64	1
$\alpha\alpha\beta$	9/64	9/64	9/64	9/64	18/64	18/64	37/64	
$\alpha\beta\alpha$	9/64	9/64	9/64	9/64				
$\beta\alpha\alpha$	9/64	9/64	9/64	9/64	9/64	19/64		
$\alpha\beta\beta$	3/64	3/64	6/64	10/64	10/64			
$\beta\alpha\beta$	3/64	3/64						
$\beta\beta\alpha$	3/64	4/64	4/64					
$\beta\beta\beta$	1/64							

Ακολουθεί ο πίνακας με τα δυαδικά ψηφία σε κάθε βήμα του αλγορίθμου και οι τελικές κωδικές λέξεις.

Σύμβολο	7	6	5	4	3	2	1	Κωδική λέξη
$\alpha\alpha\alpha$							0	0
$\alpha\alpha\beta$				0		0	1	100
$\alpha\beta\alpha$				1		0	1	101
$\beta\alpha\alpha$					0	1	1	110
$\alpha\beta\beta$		0	1		1	1	1	11110
$\beta\alpha\beta$		1	1		1	1	1	11111
$\beta\beta\alpha$	1		0		1	1	1	11101
$\beta\beta\beta$	0		0		1	1	1	11100

Η εντροπία της αρχικής πηγής πληροφορίας είναι $H(S) = \frac{3}{4}\log_2\frac{4}{3} + \frac{1}{4}\log_2 4 = 0.811$ bits. Το μέσο μήκος κωδικής λέξης ανά σύμβολο της αρχικής πηγής είναι $\bar{l} = 0.823$ bits, πολύ κοντά στην εντροπία.

2.2 Κώδικας Golomb

Ο κώδικας Golomb χρησιμοποιείται για την κωδικοποίηση μη αρνητικών ακεραίων αριθμών με γεωμετρικά φθίνουσα πιθανότητα. Συγκεκριμένα οι πιθανότητες δίδονται ως εξής:

$$P(r) = (1-p)p^r, \qquad r = 0, 1, \ldots$$

Γι' αυτή την κατανομή, που ονομάζεται γεωμετρική, ισχύει:

$$E\{R\} = \frac{p}{1-p} \quad \text{και} \quad H(R) = \frac{h(p, 1-p)}{1-p}$$

Ο ρυθμός μείωσης της πιθανότητας μετράται από την παράμετρο

$$m = \lceil -\frac{1}{\log_2 p} \rceil,$$

που δίδει την απόσταση που χωρίζει δύο ενδεχόμενα με λόγο πιθανοτήτων 1/2. Ορίζονται δύο ακέραιοι αριθμοί, το ακέραιο πηλίκο και το υπόλοιπο της διαίρεσης του r δια m:

$$q = \lfloor \frac{r}{m} \rfloor \quad \text{και} \quad s = r - mq.$$

Η κωδική λέξη για το q συνίσταται από q '1' και ένα '0'. Το s λαμβάνει m διαφορετικές τιμές με την ίδια κατά προσέγγιση πιθανότητα, κι επομένως απαιτούνται $k = \lceil log_2 m \rceil$ δυαδικά ψηφία για την κωδικοποίηση. Αν το m δεν είναι δύναμη του 2 είναι δυνατή μια μικρή επιπλέον συμπίεση. Ο κώδικας που παράγεται με αυτόν τον τρόπο είναι ελεύθερος προθέματος. Αν χρησιμοποιηθούν k δυαδικά ψηφία για το s, το μέσο μήκος κωδικής λέξης θα είναι

$$\bar{l} = k + 1 + \frac{p^m}{1 - p^m} \approx k + 2.$$

Παράδειγμα 2.2. Ας υποθέσουμε $p = 0.84$, οπότε $H(R) = 3.964$ bits. Ευρίσκουμε ότι $m = 4$, κι επομένως απαιτούνται 2 bits για την κωδικοποίηση του s. Άρα για $r = 0, \ldots, 3$ απαιτούνται 3 bits, για $r = 4, \ldots, 7$ απαιτούνται 4 bits κ.ο.κ. Το μέσο μήκος κωδικής λέξης θα είναι 3.976 bits.

2.3 Κωδικοποίηση αριθμού επαναλήψεων

Η κωδικοποίηση του αριθμού των επαναλήψεων ενός συμβόλου χρησιμοποιείται κύρια για πηγές πληροφορίας με δύο σύμβολα συσχετισμένα ή μη. Σε περίπτωση ανεξαρτησίας προκύπτει η ακόλουθη κατανομή πιθανοτήτων:

$$\{1 - p, (1 - p)p, \ldots, (1 - p)p^m, \ldots, (1 - p)p^{M-1}, p^M\}.$$

Για την κατανομή αυτή μπορεί να κατασκευασθεί κώδικας Huffman.
Παράδειγμα 2.3. Ας θεωρήσουμε πάλι την πηγή που παράγει δύο σύμβολα $\{\alpha, \beta\}$ με πιθανότητες αντίστοιχα $\{3/4, 1/4\}$. Με ανεξάρτητες επαναλήψεις, μπορούμε να θεωρήσουμε τα μηνύματα $\{\beta, \alpha\beta, \alpha\alpha\beta, \alpha\alpha\alpha\beta, \alpha\alpha\alpha\alpha\}$ με πιθανότητες $\{1/4, 3/16, 9/64, 27/256, 81/256\}$. Ο κώδικας Huffman θα είναι αντίστοιχα: {10, 00, 011, 010, 11}. Το μέσο μήκος κωδικής λέξης ανά σύμβολο της αρχικής πηγής πληροφορίας είναι 0.89 bits.

2.4 Αριθμητική κωδικοποίηση

Για να επιτευχθεί υψηλός βαθμός συμπίεσης κωδικοποιούνται ακολουθίες από σύμβολα της πηγής πληροφορίας. Κάθε ακολουθία συμβόλων της πηγής πληροφορίας παρίσταται από ένα διάστημα πραγματικών αριθμών μεταξύ 0 και 1. Η κωδική λέξη για κάθε ακολουθία συμβόλων υπολογίζεται με αριθμητικές πράξεις. Όσο πιο πιθανή είναι μια ακολουθία συμβόλων, τόσο μεγαλύτερο είναι το διάστημα που της αντιστοιχεί, για το οποίο λιγότερα bits απαιτούνται για να κωδικοποιηθεί. Κάθε διάστημα κωδικοποιείται χρησιμοποιώντας έναν αριθμό που ανήκει σ' αυτό. Κάθε αριθμός F, τέτοιος ώστε, $0 \leqslant F < 1$, μπορεί να παρασταθεί από μία ακολουθία δυαδικών ψηφίων $\{b_1, \ldots, b_i, \ldots\}$, ως εξής

$$F = \sum_{i=1}^{\infty} b_i 2^{-i}$$

Αν η ακολουθία συμβόλων προς παράσταση εμφανίζεται με πιθανότητα p, για τη δυαδική παράστασή της αρκούν σε κάθε περίπτωση $\lceil -\log_2 p \rceil + 1$ ψηφία, ώστε να προσδιορίζεται με μοναδικό τρόπο το διάστημα που ανήκει. Πράγματι αρκεί να παρασταθεί με πεπερασμένο αριθμό δυαδικών ψηφίων ο πραγματικός αριθμός

$$F + \frac{p}{2},$$

με τρόπο ώστε $F < \overline{F} < F + p$.

Δίδεται στη συνέχεια ένας αριθμητικός κωδικοποιητής/αποκωδικοποιητής, για μία μεταβλητή X, με πιθανοτητες $\{p_1, p_2, \ldots, p_n\}$. Ας είναι q οι αθροιστικές πιθανότητες

$$q(x_i) = \sum_{j=1}^{i-1} p_j$$

Για την κωδικοποίηση μιας ακολουθίας από L σύμβολα μπορεί να χρησιμοποιηθεί ο ακόλουθος αλγόριθμος, όπου F δηλώνουν τα κάτω όρια των διαστημάτων και S τις αντίστοιχες πιθανότητες.

1. $F_0 = 0$, $S_0 = 1$
2. Για $i = 1, \ldots, L$
$F_i = F_{i-1} + S_{i-1} q(x_i)$
$S_i = S_{i-1} p(x_i)$
3. $K = \lceil \log_2 \frac{1}{S_L} \rceil + 1$
4. Παράσταση του $F = F_L + 2^{-K}$ χρησιμοποιώντας K bits
5. Επιστροφή στο βήμα 1, για το επόμενο μήνυμα

Ο αποκωδικοποιητής χρησιμοποιεί τον ακόλουθο αλγόριθμο

1. $G_0 = F$, $S_0 = 1$
2. Για $i = 1, \ldots, L$
Εύρεση του μέγιστου ακέραιου ώστε
$G_{i-1} \geqslant S_{i-1} q(s_j)$
Τότε $x_i = s_j$
$G_i = G_{i-1} - S_{i-1} q(x_i)$
$S_i = S_{i-1} p(x_i)$
3. $K = \lceil \log_2 \frac{1}{S_L} \rceil + 1$
4. Επιστροφή στο Βήμα 1, για το επόμενο μήνυμα, K bits μετά την αρχή του παρόντος μηνύματος

Επειδή η ακρίβεια των υπολογισμών και της παράστασης των αριθμών είναι πεπερασμένη, απαιτείται να ληφθεί πρόνοια, ώστε να αποφευχθούν σφάλματα λόγω αριθμητικών υπολογισμών.

2.5 Μέθοδοι με λεξικά

Σε πολλές περιπτώσεις το μοντέλο πιθανοτήτων της πηγής είναι άγνωστο, ενώ μπορεί επιπλέον να είναι δύσκολη και πολύπλοκη η μοντελοποίησή της. Σε αυτές τις περιπτώσεις είναι επιθυμητό να προσδιορίζεται έμμεσα το μοντέλο και να αξιοποιείται για την κωδικοποίηση. Χρησιμοποιούνται γι' αυτό λέξεις από σύμβολα της πηγής που προκύπτουν με ανάλυση που λαμβάνει υπόψη την περίοδο εμφάνισής τους.

Ας είναι $X_1^l = x_1 x_2 \ldots x_l$ μία ακολουθία από l σύμβολα της πηγής πληροφορίας. Ας είναι $N_l(X)$ ο μικρότερος ακέραιος, ώστε $X_1^l = X_{1-N}^{l-N} (N \geqslant 1)$, δηλαδή είναι το διάστημα που μεσολαβεί για την επανεμφάνιση της ακολουθίας των συμβόλων. Αποδεικνύεται ότι η προσδοκητή τιμή του διαστήματος επανεμφάνισης είναι

$$E\{N_l(X)|X_1^l = z\} = \frac{1}{\Pr\{z\}}.$$

Η συμπίεση βασίζεται στο ασυμπτωτικό θεώρημα του διαστήματος επανάληψης μίας ακολουθίας συμβόλων

$$\lim_{l\to\infty} \frac{\log_2 N_l(X)}{l} = H(X).$$

Ισοδύναμα μπορεί να αναζητηθεί η μέγιστου μήκους ακολουθία συμβόλων της πηγής πληροφορίας, ώστε ένα αντίγραφο να αρχίζει στα n προηγούμενα σύμβολα, δηλαδή στο διάστημα $[-n+1, 0]$. Τα n προηγούμενα σύμβολα συνιστούν μία μνήμη και αναμένεται το μέγιστο μήκος αντιστοίχησης να αυξάνει με το μέγεθος της μνήμης. Αποδεικνύεται ότι

$$\lim_{n\to\infty} \frac{\log_2 n}{L_n(X)} = H(X).$$

2.5.1 Αλγόριθμος Lempel-Ziv' 77

Ο πρώτος αλγόριθμος χρησιμοποιεί κυλιόμενη μνήμη σταθερού μεγέθους, που αποτελείται από τα W πιο πρόσφατα γνωστά σύμβολα. Αρχικά, εφόσον είναι άγνωστη η πηγή, θεωρείται ένα παράθυρο μήκους W που κωδικοποιείται χωρίς συμπίεση. Αναζητείται στη συνέχεια η πλησιέστερη μέγιστου μήκους ακολουθία X_{W+1}^{W+L} για την οποία ευρίσκεται ταίριασμα με αφετηρία στο παράθυρο που έχει απομνημονευθεί.

$$X_{W+1}^{W+L} = X_{W-m}^{W-m+L-1}, m \in \{0, \ldots, W-1\} \tag{2.1}$$

Ας είναι για παράδειγμα μία πηγή με δύο σύμβολα $\{\alpha, \beta\}$, κι έστω ότι παράγεται η ακολουθία

$$\beta\beta\beta\alpha\alpha\alpha\alpha\alpha\alpha\alpha\alpha\alpha\alpha\beta\ldots$$

Αν είναι $W = 4$, τότε θα είναι $L = 9$ και $m = 0$.

Απαιτείται ένας αποτελεσματικός τρόπος κωδικοποίησης του μήκους L.

Κωδικοποίηση Ορίζεται κατ' αρχήν η δυαδική παράσταση του L, $b(L)$. Απαιτείται επιπλέον το μήκος αυτής της παράστασης, $|b(L)|$. Ο κώδικας έχει τη μορφή

$$c(L) = e(|b(L)|)b(L)$$

όπου $e(k) = 0^{|b(k)|-1}1b(k)$. Ας είναι για παράδειγμα $L = 7$. Τότε $b(L) = 111$ και $|b(L)| = 3$, οπότε $c(7) = 0111111$.

Χρησιμοποιείται η δυαδική παράσταση του m, με $\lceil \log_2 W \rceil$, αν είναι συμφερότερη από την παράσταση ασυμπίεστης της ακολουθίας των L συμβόλων, αλλιώς χρησιμοποιούνται $L\lceil \log_2 |S| \rceil$ δυαδικά ψηφία.

Αποκωδικοποίηση Αποκωδικοποεείται πρώτα ο κώδικας του L. Αν $\lceil \log_2 W \rceil < L\lceil \log_2 |S| \rceil$, τότε αποκωδικοποιείται το m και εφαρμόζεται το ταίριασμα της εξίσωσης (2.1). Αλλιώς, αποκωδικοποεείται άμεσα η ακολουθία των L συμβόλων.

Μία παραλλαγή του αλγορίθμου αυτού γνωστή με το όνομα Deflate χρησιμοποιείται στη μορφή gzip για τη συμπίεση αρχείων και στη μορφή PNG (Portable Network Graphics) για τη συμπίεση εικόνων.

2.5.2 Αλγόριθμος Lempel-Ziv-Welch

Στο δεύτερο αλγόριθμο κατασκευάζεται ένα λεξικό, όπου εισάγονται όσο γίνεται μεγαλύτερες σε μήκος λέξεις που απαντώνται συχνά. Το μέγεθος του λεξικού είναι πάντοτε ορισμένο και η κωδικοποίηση γίνεται με κωδικές λέξεις ορισμένου μήκους που αντιστοιχούν στους δείκτες των λέξεων του λεξικού. Ακολουθεί η περιγραφή του αλγορίθμου.

Κωδικοποίηση (συμπίεση)

Εκκίνηση Αρχικό λεξικό

- Λεξικό αποτελούμενο από όλα τα γράμματα
- Πρώτη λέξη: πρώτο γράμμα

Βήμα : Ενημέρωση του λεξικού και κωδικοποίηση

- Αν δεν υπάρχει άλλο γράμμα
  - κωδικοποίηση (ω)
  - τέλος
- Αν $\omega\alpha$ υπάρχει στο λεξικό
  - νέα λέξη: $\omega\alpha$
  - επανάληψη του βήματος
- Αν $\omega\alpha$ δεν υπάρχει στο λεξικό
  - κωδικοποίηση ω
  - καταχώρηση της $\omega\alpha$ στο λεξικό
  - νέα λέξη: α
  - επανάληψη του βήματος

Αποκωδικοποίηση (αποσυμπίεση)

Εκκίνηση Μεταβλητές: Code, OldCode, InCode, FinChar

- Πρώτος κώδικας στο Code και στο OldCode
- Αποκωδικοποίηση(Code) στο πρώτο γράμμα (α)
- α στο FinChar

Επόμενος κώδικας στο Code και στο InCode

- Αν δεν υπάρχει άλλος κώδικας: τέλος
- Αν ο Code δεν έχει ορισθεί
  - FinChar στην έξοδο
  - OldCode στο Code
  - κώδικας(OldCode, FinChar) στο InCode

Επόμενο γράμμα Αποκωδικοποίηση

- Αν Code = κώδικας($\omega\alpha$)
  - α στη στήλη
  - κώδικας(ω) στο Code
  - Στο Επόμενο γράμμα
- Αν Code = κώδικας(α)
  - α στην έξοδο και στο FinChar
  - Όσο η στήλη δεν είναι κενή, κορυφή της στήλης στην έξοδο και ανέβασμα της στήλης
  - (OldCode, α) στο λεξικό
  - InCode στο OldCode
  - Στον Επόμενο κώδικα

Ο αλγόριθμος Lempel-Ziv-Welch χρησιμοποιείται για τη συμπίεση εικόνων στη μορφή GIF Το λεξικό μεγαλώνει δυναμικά αρχίζοντας από μέγεθος 2^{b+1}, όπου b είναι ο αριθμός των bits ανά εικονοστοιχείο. Το μέγεθος του λεξικού μπορεί να φθάσει τις 4096 λέξεις, ενώ ο αριθμός των bits στα αρχικά δεδομένα είναι το πολύ 8. Προβλέπεται η δυνατότητα επανεκκίνησης του λεξικού.

Κεφάλαιο 3

Συνεχής πηγή πληροφορίας

3.1 Διαφορική εντροπία

Η διαφορική εντροπία χαρακτηρίζει συνεχείς πηγές κι επομένως μπορεί να χρησιμεύσει σε μετρήσεις πληροφορίας για σήματα που λαμβάνουν συνεχείς τιμές. Η διαφορική εντροπία ωστόσο δε μετρά άμεσα μια ποσότητα πληροφορίας. Ας είναι X μια τυχαία μεταβλητή με συνάρτηση πυκνότητας πιθανότητας $p_X(x)$. Η διαφορική εντροπία ορίζεται ως

$$h(X) = \int p_X(x) \log \frac{1}{p_X(x)} dx,$$

όπου το ολοκλήρωμα ορίζεται στο σύνολο των τιμών της μεταβλητής, και εφόσον βεβαίως υπάρχει.

Στην περίπτωση μιας ομοιόμορφης κατανομής στο διάστημα $[0, a]$ η διαφορική εντροπία είναι

$$h(X) = \log a.$$

Στην περίπτωση μιας κατανομής Gauss με διασπορά σ^2 η διαφορική εντροπία είναι, σε φυσικούς λογάριθμους,

$$h(X) = \frac{1}{2} \ln 2\pi e \sigma^2.$$

Αποδεικνύεται εύκολα ότι η μετακίνηση της μέσης τιμής της μεταβλητής δεν επηρεάζει τη διαφορική εντροπία. Ενώ η αλλαγή κλίμακας κατά α συνεπάγεται αύξηση κατά $\log|\alpha|$.

Η από κοινού διαφορική εντροπία ορίζεται κατά παρόμοιο τρόπο

$$h(X, Y) = \int \int p_{X,Y}(x, y) \log \frac{1}{p_{X,Y}(x, y)} dxdy,$$

όπως και η δεσμευμένη διαφορική εντροπία

$$h(X|Y) = \int \int p_{X,Y}(x, y) \log \frac{1}{p_{X|Y}(x|y)} dxdy,$$

$$h(Y|X) = \int \int p_{X,Y}(x, y) \log \frac{1}{p_{Y|X}(y|x)} dxdy.$$

Ισχύει ο κανόνας της αλυσίδας

$$h(X,Y) = h(X) + h(Y|X) = h(Y) + h(X|Y).$$

Η σχετική εντροπία ή απόσταση Kullback μεταξύ δύο κατανομών πιθανοτήτων ορίζεται ως

$$D(p\|q) = \int p(x) \log \frac{p(x)}{q(x)} dx.$$

Αποδεικνύεται ότι

$$D(p\|q) \geqslant 0.$$

Χρησιμοποιώντας την ανισότητα αυτή αποδεικνύεται ότι για δοσμένη διασπορά η κατανομή Gauss μεγιστοποιεί τη διαφορική εντροπία. Πράγματι μπορούμε να γράψουμε

$$\int_{-\infty}^{\infty} p(x) \ln p(x) \sigma\sqrt{2\pi} e^{\frac{(x-\mu)^2}{2\sigma^2}} dx \geqslant 0 \Leftrightarrow -h(X) + \ln(\sigma\sqrt{2\pi}) + \int_{-\infty}^{\infty} \frac{(x-\mu)^2}{2\sigma^2} p(x) dx \geqslant 0 \Leftrightarrow$$

$$h(X) \leqslant \frac{1}{2} \ln 2\pi e \sigma^2.$$

Ορίζεται η αμοιβαία πληροφορία

$$I(X,Y) = \int\int p_{X,Y}(x,y) \log \frac{p_{X,Y}(x,y)}{p_X(x) p_Y(y)} dxdy.$$

Η αμοιβαία πληροφορία συνδέεται με τις ποσότητες διαφορικής εντροπίας ως ακολούθως

$$I(X,Y) = h(X) - h(X|Y) = h(Y) - h(Y|X) = h(X) + h(Y) - h(X,Y).$$

Επειδή η αμοιβαία πληροφορία είναι απόσταση μεταξύ δύο κατανομών

$$I(X,Y) \geqslant 0.$$

Η αμοιβαία πληροφορία μετρά το βαθμό εξάρτησης μεταξύ των μεταβλητών X, Y και μηδενίζεται όταν είναι ανεξάρτητες.

Εάν X, Y είναι δύο μεταβλητές που ακολουθούν την κατανομή Gauss, με συντελεστή συσχέτισης ϱ, τότε

$$I(X,Y) = \frac{1}{2} \log \frac{1}{1-\varrho^2}.$$

Εάν X, Z είναι δύο ανεξάρτητες μεταβλητές που ακολουθούν την κατανομή Gauss, με διασπορές σ_X^2, σ_Z^2 αντίστοιχα, και $Y = X + Z$, τότε

$$I(X,Y) = \frac{1}{2} \log \left(1 + \frac{\sigma_X^2}{\sigma_Z^2}\right).$$

3.2 Ποσότητα πληροφορίας και προσέγγιση

Ας είναι $\hat{X}$ μια προσέγγιση της μεταβλητής X. Ορίζεται ένα μέτρο απόστασης ή παραμόρφωσης,

$$d : \mathcal{X} \times \hat{\mathcal{X}} \to \mathcal{R}^{+}.$$

Για δυαδικές τιμές μπορεί να είναι η απόσταση Hamming,

$$d(x, \hat{x}) = \begin{cases} 0 & x = \hat{x} \\ 1 & x \neq \hat{x} \end{cases}$$

Για οποιεσδήποτε πραγματικές τιμές μπορεί να ορισθεί η τετραγωνική απόκλιση,

$$d(x, \hat{x}) = (x - \hat{x})^2.$$

Αν απαιτήσουμε η μέση παραμόρφωση να μην υπερβαίνει κάποιο φράγμα,

$$\int \int p(x, \hat{x}) d(x, \hat{x}) dx d\hat{x} \leqslant D,$$

τότε η αμοιβαία πληροφορία $I(X, \hat{X})$ έχει κάτω φράγμα μια ποσότητα πληροφορίας,

$$R(D) = \min I(X, \hat{X}),$$

όπου το ελάχιστο λαμβάνεται μεταξύ όλων των δεσμευμένων κατανομών $p(\hat{x}|x)$. Επομένως αυτή η ποσότητα πληροφορίας είναι η ελάχιστη που απαιτείται, ώστε το σφάλμα της προσέγγισης να μην υπερβαίνει το φράγμα D.

Για μια δυαδική πηγή με πιθανότητα p και απόσταση Hamming, η ελάχιστη ποσότητα πληροφορίας είναι:

$$R(D) = \begin{cases} H(p) - H(D), & 0 \leqslant D \leqslant \min(p, 1-p) \\ 0, & D > \min(p, 1-p) \end{cases}$$

Ας υποθέσουμε κατ΄ αρχήν για τις ανάγκες της παρουσίασης ότι οι δύο δυνατές τιμές της μεταβλητής X είναι $\{0, 1\}$ με αντίστοιχες τιμές πιθανότητας $\{1 - p, p\}$. Η μέση παραμόρφωση θα είναι $D =$

Για την προσέγγιση μιας μεταβλητής Gauss με διασπορά σ^2 και τετραγωνική απόσταση απαιτείται ποσότητα πληροφορίας τουλάχιστον

$$R(D) = \begin{cases} \frac{1}{2} \log \frac{\sigma^2}{D}, & 0 \leqslant D \leqslant \sigma^2 \\ 0, & D > \sigma^2 \end{cases}$$

Κεφάλαιο 4

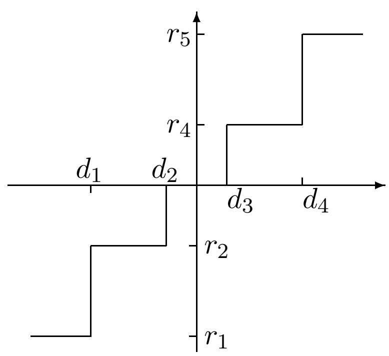
Κβαντισμός

4.1 Βαθμωτός κβαντιστής

Ένας κβαντιστής (Σχήμα 4.1) αντιστοιχεί σε μία τιμή μιας συνεχούς μεταβλητής X μία από τις διακριτές τιμές ενός πεπερασμένου συνόλου $\{r_1, r_2, \ldots, r_N\}$ αντιπροσωπευτικών τιμών. Για τον προσδιορισμό του κβαντιστή απαιτούνται το σύνολο των αντιπροσωπευτικών τιμών και το σύνολο των σημείων απόφασης $\{d_1, d_2, \ldots, d_{N-1}\}$ που χωρίζουν τον άξονα των πραγματικών αριθμών στα ακόλουθα διαστήματα

$$\mathbb{R} = (-\infty, d_1] \cup (d_1, d_2] \cup \cdots \cup (d_{N-1}, \infty) \tag{4.1}$$

Για το σχεδιασμό ενός κβαντιστή παρουσιάζεται στη συνέχεια μία στατιστική μέθοδος προσ-



Σχήμα 4.1: Ένας συμμετρικός κβαντιστής

διορισμού των αντιπροσωπευτικών τιμών και των τιμών απόφασης.

Ας είναι X μία τυχαία μεταβλητή, της οποίας η συνάρτηση πυκνότητας πιθανότητας είναι $p(x)$. Ζητείται ο προσδιορισμός τόσο των επιπέδων κβαντισμού, όσο και των σημείων απόφασης, για δοσμένο αριθμό N διαστημάτων και επιπέδων κβαντισμού, με τρόπο ώστε το μέσο τετραγωνικό σφάλμα να ελαχιστοποιείται

$$D = E\{(X - Q(X))^2\} = \sum_{i=1}^{N} \int_{d_{i-1}}^{d_i} (x - r_i)^2 p(x) dx, \tag{4.2}$$

όπου $Q(.)$ είναι ο τελεστής του κβαντισμού, και $d_0 = -\infty$, $d_N = \infty$.

Οι Lloyd και Max έδωσαν τις αναγκαίες συνθήκες για την ελαχιστοποίηση του D, που συνίστανται σε δύο συστήματα εξισώσεων. Το πρώτο σύστημα δίδει την καλύτερη αντιπροσώπευση, με δοσμένα τα διαστήματα

$$r_i = E\{X|X \in (d_{i-1}, d_i)\} = \frac{\int_{d_{i-1}}^{d_i} xp(x)dx}{\int_{d_{i-1}}^{d_i} p(x)dx}; i = 1, \ldots, N \tag{4.3}$$

Το δεύτερο σύστημα δίδει την καλύτερη διαμέριση του άξονα των πραγματικών αριθμών με δοσμένα τα επίπεδα αντιπροσώπευσης

$$d_i = \frac{r_i + r_{i+1}}{2}; i = 1, \ldots, N-1 \tag{4.4}$$

Στη γενική περίπτωση το παραπάνω σύστημα εξισώσεων (4.3) και (4.4) μπορεί να επιλυθεί μόνο με αριθμητικές επαναληπτικές μεθόδους. Μετά από κάποιες αρχικές τιμές για τα επίπεδα κβαντισμού, χρησιμοποιούνται διαδοχικά, και επαναληπτικά μέχρι τη σύγκλιση, οι εξισώσεις (4.4) και (4.3). Πρέπει να σημειωθεί ότι η επαναληπτική επίλυση του συστήματος των εξισώσεων δεν εξασφαλίζει πάντοτε την ελαχιστοποίηση του μέσου τετραγωνικού σφάλματος. Μπορεί η σύγκλιση να οδηγήσει σ' ένα τοπικό ελάχιστο, όταν οι παραπάνω συνθήκες είναι μόνο αναγκαίες, αλλά όχι και ικανές για την ελαχιστοποίηση του D.

Ο Lloyd-Max κβαντιστής έχει τις ακόλουθες ιδιότητες.

- Η μέση κβαντισμένη τιμή είναι ίση με τη μεση τιμή της μεταβλητής

$$E\{Q(X)\} = E\{X\}$$

- Η κβαντισμένη τιμή και το σφάλμα κβαντισμού είναι ασυσχέτιστες μεταβλητές

$$E\{(X - Q(X))Q(X)\} = 0$$

- Η διασπορά του σφάλματος κβαντισμού είναι ίση με τη διαφορά της διασποράς της μεταβλητής μείον τη διασπορά των κβαντισμένων τιμών

$$E\{(X - Q(X))^2\} = \int_{-\infty}^{\infty} x^2 p(x)dx - \sum_{i=1}^{N} P_i r_i^2$$

  όπου $P_i = \int_{d_{i-1}}^{d_i} p(x)dx$.

- Ασυμπτωτικά, δηλαδή για μεγάλο N, η διασπορά του σφάλματος κβαντισμού μειώνεται όπως το $1/N^2$,

$$D \approx \frac{\left(\int_{-\infty}^{\infty} (p(x))^{1/3} dx\right)^3}{12N^2}.$$

Αν η τυχαία μεταβλητή X κατανέμεται ομοιόμορφα, τα διαστήματα κβαντισμού είναι ισομήκη, και τα επίπεδα κβαντισμού κατανέμονται ομοιόμορφα. Ο ομοιόμορφος κβαντιστής προσδιορίζεται με τη χρήση μίας μόνο παραμέτρου, το βήμα του κβαντισμού Δ. Η διασπορά του σφάλματος κβαντισμού στην περίπτωση της ομοιόμορφης κατανομής είναι

$$D = \frac{\sigma^2}{N^2}.$$

Λόγω της απλότητάς του ο ομοιόμορφος κβαντιστής θα μπορούσε να είναι χρήσιμος και στην περίπτωση μεταβλητών που δεν κατανέμονται ομοιόμορφα. Αν μία μεταβλητή κατανέμεται στο σύνολο των πραγματικών αριθμών, υπάρχει ένα βέλτιστο βήμα κβαντισμού που ελαχιστοποιεί το μέσο τετραγωνικό σφάλμα. Παρουσιάζεται στη συνέχεια το πρόβλημα βελτιστοποίησης στην περίπτωση μίας μεταβλητής με συμμετρική κατανομή, όπου $p(x) = p(-x)$, με συνέπεια ο κβαντιστής να είναι συμμετρικός, κι όταν επιπλέον ο αριθμός των επιπέδων κβαντισμού είναι άρτιος.

Ας είναι Δ το βήμα του κβαντιστή και $2N$ ο αριθμός των επιπέδων κβαντισμού. Θα έχουμε

$$\left.\begin{array}{rcll} r_i & = & (2i - \mathrm{sgn}(i))\,\frac{\Delta}{2} & , \quad i = \pm 1, \ldots, \pm N \\ d_i & = & i\Delta & , \quad i = 0, \pm 1, \ldots, \pm(N-1) \end{array}\right\} \tag{4.5}$$

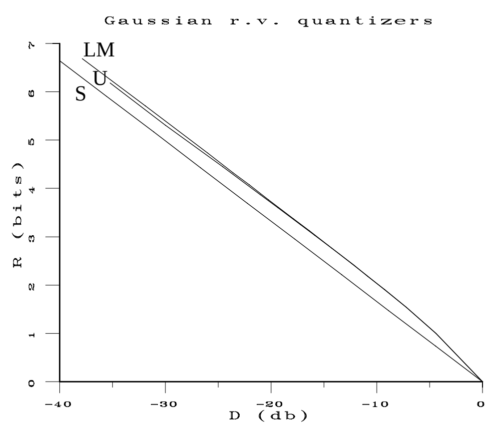
Το μέσο τετραγωνικό σφάλμα είναι

$$D = 2\sum_{i=1}^{N-1} \int_{d_{i-1}}^{d_i} \left(x - i\frac{\Delta}{2}\right)^2 p(x)dx + 2\int_{d_{N-1}}^{\infty} \left(x - N\frac{\Delta}{2}\right)^2 p(x)dx \tag{4.6}$$

Η λύση ως προς Δ ευρίσκεται με αριθμητικές μεθόδους.

Στο Σχήμα 4.2 δίδεται για μία τυχαία μεταβλητή που ακολουθεί την κατανομή Gauss (με διασπορά $\sigma^2 = 1$) η απαιτούμενη ποσότητα πληροφορίας σα συνάρτηση της παραμόρφωσης για τον κβαντιστή Max-Lloyd (LM), για τον ομοιόμορφο κβαντιστή (U), και για τον ιδανικό κβαντιστή Shannon (S), για τον οποίο ισχύει

$$R = \max\left(0, \frac{1}{2}\log_2 \frac{\sigma^2}{D}\right) \tag{4.7}$$



Σχήμα 4.2: Ποσότητα πληροφορίας σε σχέση με την παραμόρφωση για μια μεταβλητή Gauss

4.2 Διανυσματικός κβαντιστής

Σ' αυτή την περίπτωση ο κβαντισμός αφορά N μεταβλητές από κοινού, που παριστάνονται από ένα διάνυσμα N διαστάσεων. Ο κβαντιστής διαθέτει ένα σύνολο από αντιπροσωπευτικά διανύσματα $\{r_k : k = 1, \ldots, K\}$ που ονομάζεται λεξικό. Για κάθε διάνυσμα προς κβαντισμό επιλέγεται το πλησιέστερο σ' αυτό αντιπροσωπευτικό διάνυσμα, και κωδικοποιείται ο δείκτης αυτού του διανύσματος. Ο αποσυμπιεστής διαθέτει το λεξικό και με την αποκωδικοποίηση προσδιορίζει το αντίστοιχο διάνυσμα από το λεξικό, το οποίο και αποκαθιστά.

Το λεξικό κατασκευάζεται από ένα σύνολο διανυσμάτων εκμάθησης $\{x_i : i = 1, \ldots, N_t\}$ και βασίζεται σ' ένα κριτήριο ελάχιστης τετραγωνικής παραμόρφωσης

$$D = \sum_{k=1}^{K} \sum_{x_i \in \mathcal{C}_k} \|x_i - r_k\|^2 \tag{4.8}$$

για K κλάσεις $\mathcal{C}_k$.

Ο βέλτιστος διανυσματικός κβαντιστής ικανοποιεί δύο αναγκαίες συνθήκες για την ελαχιστοποίηση της D. Για δοσμένη κλάση ο καλύτερος αντιπρόσωπος είναι το κέντρο βάρους

$$r_k = \frac{1}{\mathrm{card}[\mathcal{C}_k]} \sum_{x_i \in \mathcal{C}_k} x_i \tag{4.9}$$

Για δοσμένο λεξικό η καλύτερη τιμή κβαντισμού ενός διανύσματος x συνίσταται στην επιλογή του πλησιέστερου αντιπρόσωπου

$$\|x - r_k\| < \|x - r_l\| \quad \forall l \neq k \quad \Rightarrow \quad x \in \mathcal{C}_k \tag{4.10}$$

Η χρησιμοποίηση των δύο αυτών αναγκαίων συνθηκών δίδει έναν επαναληπτικό αλγόριθμο κατασκευής ενός λεξικού από ένα σύνολο διανυσμάτων εκμάθησης.

- Αρχικό βήμα: Αρχικό λεξικό, $i = 1$, και αρχική μεγάλη τιμή για την παραμόρφωση $D^{(0)}$
- Βήμα 1: Εύρεση των κλάσεων (Εξίσωση (4.10))
- Βήμα 2: Υπολογισμός της παραμόρφωσης $D^{(i)}$
- Βήμα 3: Έλεγχος σύγκλισης

  $$\text{Αν} \quad \frac{D^{(i-1)} - D^{(i)}}{D^{(i-1)}} \leqslant \epsilon, \quad \text{τέλος}$$

  Διαφορετικά, προσαύξηση του i, και συνέχιση των επαναλήψεων
- Βήμα 4: Εύρεση του καλύτερου αντιπρόσωπου για κάθε κλάση (Εξίσωση (4.9)), κι επιστροφή στο Βήμα 1.

Το υπολογιστικό κόστος για την εύρεση του πλησιέστερου αντιπρόσωπου από το λεξικό είναι $O(NK)$. Το κόστος αποθήκευσης του λεξικού είναι NK. Επειδή το υπολογιστικό κόστος είναι υψηλό, κι ενδεχόμενα απαγορευτικό για μετάδοση σε πραγματικό χρόνο, μπορεί να επιβληθεί μία δομή δένδρου στο λεξικό. Αν το δένδρο περιλαμβάνει ν κλάδους, τότε το υπολογιστικό κόστος περιορίζεται σε $O(N\nu \log_\nu K)$, ενώ το κόστος αποθήκευσης του λεξικού αυξάνει σε $N(K-1)\frac{\nu}{\nu-1}$. Αν για παράδειγμα $\nu = 2$, το υπολογιστικό κόστος είναι $O(2N \log_2 K)$ και το κόστος αποθήκευσης $2N(K-1)$. Η κατασκευή του λεξικού γίνεται με τη χρήση του παραπάνω αλγόριθμου για ν κλάσεις και για όλους τους ενδιάμεσους κόμβους του δένδρου από τη ρίζα προς τα φύλλα.

Κεφάλαιο 5

Συμπίεση σημάτων

Η ανάγκη για συμπίεση των σημάτων προκύπτει από το μεγάλο όγκο των δεδομένων προς μετάδοση ή αποθήκευση. Ζητούμενο της συμπίεσης είναι η οικονομικότερη μετάδοση ή αποθήκευση των σημάτων. Η συμπίεση των σημάτων βασίζεται στον περιορισμό της περιττής πληροφορίας, χωρίς αντιληπτή αλλοίωση των δεδομένων του σήματος, ή έστω ως τα όρια αλλοίωσης ανεκτής από το ανθρώπινο σύστημα αντίληψης του σήματος. Κριτήρια επίδοσης ενός συστήματος συμπίεσης είναι: ο βαθμός συμπίεσης, η ποιότητα των σημάτων που μπορούν να ανασυντεθούν από τη συμπιεσμένη ποσότητα πληροφορίας, η πολυπλοκότητα του συμπιεστή/αποσυμπιεστή, και, στην περίπτωση ύπαρξης λαθών μεταξύ συμπιεστή και αποσυμπιεστή, η αντοχή του συστήματος συμπίεσης σ' αυτά τα λάθη.

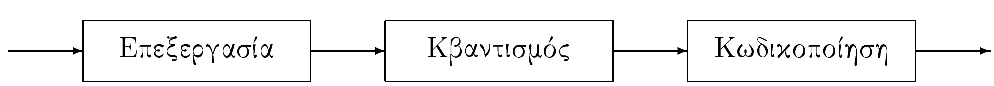
Ο βαθμός συμπίεσης ορίζεται ως ο λόγος της ποσότητας πληροφορίας πριν και μετά τη συμπίεση. Καταλληλότερος κριτής της ποιότητας των συμπιεσμένων σημάτων είναι ο δέκτης και χρήστης των σημάτων. Ωστόσο για τη σχεδίαση συστημάτων συμπίεσης σημάτων μπορούν να χρησιμοποιηθούν στατιστικά κριτήρια μέτρησης της επίδοσής τους, όπως η σηματοθορυβική σχέση σε λογαριθμική κλίμακα (decibel)

$$\text{SNR } = 10\log_{10}\frac{\sum_{n=1}^{N} s^2(n)}{\sum_{n=1}^{N}(s(n)-\bar{s}(n))^2},$$

που επίσης μπορεί να ορισθεί με βάση τη μέγιστη τιμή του σήματος

$$\text{PSNR } = 10\log_{10}\frac{N(s_{\max}-s_{\min})^2}{\sum_{n=1}^{N}(s(n)-\bar{s}(n))^2}.$$

Κατά γενικό κανόνα ένας συμπιεστής συνίσταται από τρία μέρη: στο πρώτο μέρος το σήμα υφίσταται μία επεξεργασία, ανάλυση ή μετασχηματισμό, χωρίς απώλεια πληροφορίας, στο δεύτερο μέρος το αποτέλεσμα κβαντίζεται, με απώλεια πληροφορίας, και στο τρίτο μέρος τα κβαντισμένα μεγέθη κωδικοποιούνται (Σχήμα 5.1). Συμμετρικά ο αποσυμπιεστής αποκωδικοποιεί, αποκαθιστά



Σχήμα 5.1: Τυπικό σχήμα συμπίεσης σημάτων

την αντιστοιχία με τα κβαντισμένα μεγέθη και ανασυνθέτει το σήμα.

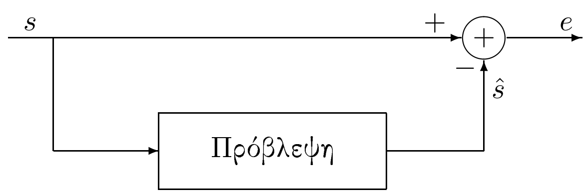
Η χρήση ενός εντροπικού κωδικοποιητή επιτρέπει τη μέγιστη δυνατή συμπίεση στην έξοδο του κβαντιστή χωρίς απώλεια πληροφορίας. Ο κβαντισμός οδηγεί σε σημαντική συμπίεση, εισάγοντας παραμόρφωση, που επιδιώκεται να περιορίζεται σε όρια μη αντιληπτά. Και τα δύο αυτά θέματα καλύφθηκαν σε προηγούμενα κεφάλαια. Στη συνέχεια θα αναφερθούμε κυρίως στην επεξεργασία των σημάτων με σκοπό τη συμπίεση.

5.1 Συμπίεση με πρόβλεψη

Η βασική αρχή ενός συστήματος συμπίεσης με πρόβλεψη είναι 'ό,τι είναι προβλέψιμο, πλεονάζει'. Στο Σχήμα 5.2 δίδεται ένα σύστημα εξαγωγής της πλεονάζουσας πληροφορίας με πρόβλεψη. Το σύστημα αυτό χαρακτηρίζεται από το κέρδος της πρόβλεψης

$$G_p = \frac{\sigma_s^2}{\sigma_e^2}. \tag{5.1}$$

Για δοσμένο ρυθμό πληροφορίας, και εάν χρησιμοποιηθεί κβαντισμός, που συνεπάγεται παρα-



Σχήμα 5.2: Εξαγωγή του σφάλματος της πρόβλεψης

μόρφωση, η πρόβλεψη οδηγεί σε μείωση της παραμόρφωσης πρακτικά ίση με το κέρδος της πρόβλεψης

$$\frac{D_s}{D_e} = G_p.$$

Ενώ για δοσμένη παραμόρφωση, ο αριθμός bits ανά δείγμα του σήματος, μειώνεται

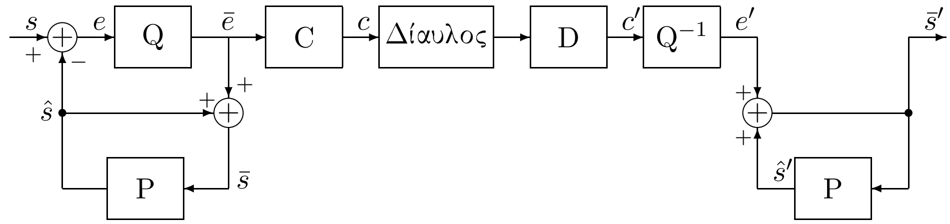
$$B_s - B_e = \frac{1}{2}\log_2 G_p.$$

Το σύστημα συμπίεσης που περιλαμβάνει και κβαντιστή δίδεται στο Σχήμα 5.3. Απ' αυτό το σχήμα προκύπτει ότι η παραμόρφωση είναι ίση με το σφάλμα του κβαντισμού

$$s - \bar{s} = e - \bar{e}.$$

Κατά κανόνα για την πρόβλεψη χρησιμοποιείται ένα γραμμικό αναδρομικό φίλτρο. Ένας μικρός αριθμός συντελεστών είναι αρκετός για τον προσδιορισμό του φίλτρου πρόβλεψης, που χρησιμοποείται ως ακολούθως

$$\hat{s}(n) = \sum_{k=1}^{K} a(k)\bar{s}(n-k). \tag{5.2}$$



Σχήμα 5.3: Συμπιεστής/αποσυμπιεστής με πρόβλεψη (P: πρόβλεψη, Q: κβαντισμός, C: κωδικοποίηση, D: αποκωδικοποίηση, Q^{-1}: αντίστροφος κβαντισμός).

Οι συντελεστές του φίλτρου μπορούν να εκτιμηθούν μέσω της συνάρτησης συμμεταβλητότητας του σήματος.

Ας θεωρήσουμε ένα γραμμικό μοντέλο πρόβλεψης για ένα σήμα με μηδενική μέση τιμή

$$\hat{s}(n) = \sum_{k=1}^{K} a(k)s(n-k). \tag{5.3}$$

Οι συντελεστές μπορούν να προσδιορισθούν χρησιμοποιώντας το κριτήριο ελάχιστου τετραγωνικού σφάλματος

$$E\{\left(s(n) - \sum_{k=1}^{K} a(k)s(n-k)\right)^2\}.$$

Η ελαχιστοποίηση του παραπάνω σφάλματος δίδει τις ακόλουθες εξισώσεις

$$E\{\left(s(n) - \sum_{k=1}^{K} a(k)s(n-k)\right)s(n-l)\} = 0,$$

για $l = 1, \ldots, K$, που γράφονται ως ακολούθως

$$\sum_{k=1}^{K} \gamma(l-k)a(k) = \gamma(l). \tag{5.4}$$

Οι τιμές της συνάρτησης $\gamma(m)$ μπορούν να προσδιορισθούν είτε εμπειρικά είτε μέσω ενός μοντέλου. Το απλούστερο μοντέλο είναι πρώτης τάξης

$$\gamma(m) = \sigma^2 \rho^{|m|}.$$

Η λύση του παραπάνω συστήματος σ' αυτήν την περίπτωση δίδει μόνο ένα συντελεστή για το φίλτρο πρόβλεψης,

$$a(1) = \rho.$$

Το κέρδος της πρόβλεψης είναι τόσο μεγαλύτερο, όσο ο συντελεστής συσχέτισης είναι υψηλότερος

$$G_p = \frac{1}{1-\rho^2}.$$

Ο κβαντιστής μπορεί να σχεδιασθεί με βάση τη μέθοδο Max-Lloyd. Ωστόσο το σφάλμα της πρόβλεψης διαφέρει σημαντικά ανάλογα με τα τοπικά χαρακτηριστικά του σήματος. Οπως είναι αναμενόμενο, το σφάλμα της πρόβλεψης είναι ισχυρότερο κοντά στις απότομες μεταβολές των τιμών του σήματος, κι ασθενέστερο σε ομοιογενείς περιοχές. Αν επιπλέον ληφθεί υπόψη ότι το ανθρώπινο σύστημα αντίληψης είναι γενικά πιο ευαίσθητο σε παραμορφώσεις της έντασης σε ομοιογενείς περιοχές, είναι προτιμότερο ο κβαντιστής να προσαρμόζεται τοπικά στο σφάλμα πρόβλεψης του σήματος, και το βήμα του κβαντισμού να είναι μικρότερο στις ομοιογενείς περιοχές, και μεγαλύτερο κοντά στις μεταβάσεις.

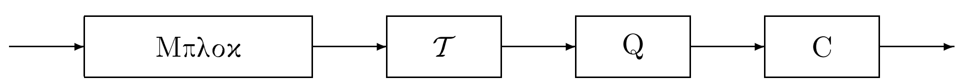
Χωρίς κβαντισμό πρόβλεψη χρησιμοποιείται στο σύστημα κωδικοποίησης JPEG για συμπίεση εικόνων χωρίς απώλειες. Πρόβλεψη από εικόνα σε εικόνα μετά και από εκτίμηση της κίνησης χρησιμοποιείται στα συστήματα κωδικοποίησης βίντεο MPEG.

5.2 Συμπίεση με χρήση ορθομοναδιαίου μετασχηματισμού

Η χρήση ενός ορθομοναδιαίου μετασχηματισμού επιτυγχάνει τη συμπίεση, χάρη στη σημαντική αποσυσχέτιση των δεδομένων και στη δυνατότητα ο κβαντισμός να σχεδιασθεί για ένα ολόκληρο διάνυσμα ή μπλοκ τιμών του σήματος. Τυπική διάσταση των μπλοκ που χρησιμοποιούνται για το μετασχηματισμό για συμπίεση εικόνων είναι 8×8. Μεταξύ των μετασχηματισμών ο πλέον χρησιμοποιούμενος με σκοπό τη συμπίεση είναι ο διακριτός μετασχηματισμός συνημιτόνου (βλέπε σύστημα JPEG). Το κέρδος που επιτυγχάνεται με τη χρήση μετασχηματισμού, αν σ^2 είναι η διασπορά του αρχικού σήματος και σ_n^2 είναι η διασπορά του συντελεστή του μετασχηματισμού στη θέση n, είναι

$$G_T = \frac{\sigma^2}{\left(\prod_{n=1}^{N} \sigma_n^2\right)^{\frac{1}{N}}} \tag{5.5}$$

πάντοτε μεγαλύτερο της μονάδας, αφού $\sum_{n=1}^{N} \sigma_n^2 = N\sigma^2$.



Σχήμα 5.4: Συμπίεση με χρήση μετασχηματισμού ($\mathcal{T}$: ορθομοναδιαίος μετασχηματισμός, Q: κβαντισμός, C: κωδικοποίηση).

Μπορεί κανείς να διακρίνει δύο τεχνικές για τον κβαντισμό και την κωδικοποίηση. Κατά την πρώτη χρησιμοποιείται διαφορετικός κβαντιστής για κάθε συντελεστή του μετασχηματισμού, και σταθερός αριθμός bits για κάθε επίπεδο κβαντισμού. Κατά τη δεύτερη κατά κανόνα χρησιμοποιείται ο ίδιος ομοιόμορφος κβαντιστής για όλους τους συντελεστές του μετασχηματισμού και μεταβλητός αριθμός bits ανάλογα με το επίπεδο κβαντισμού. Οι δύο αυτές τεχνικές παρουσιάζονται στη συνέχεια.

5.2.1 Σταθερός ρυθμός πληροφορίας

Έστω ότι ο ρυθμός μετάδοσης της πληροφορίας είναι σταθερός και καθορισμένος

$$R = \frac{1}{N}\sum_{n=01}^{N} B_n, \tag{5.6}$$

όπου B_n είναι ο αριθμός των bits που χρησιμοποιούνται για το συντελεστή του μετασχηματισμού στη θέση n. Ζητείται η βέλτιστη κατανομή των bits ανά συντελεστή του μετασχηματισμού, με την έννοια της ελαχιστοποίησης της μέσης παραμόρφωσης λόγω κβαντισμού

$$D = \frac{1}{N}\sum_{n=1}^{N} \sigma_n^2 f(B_n) \tag{5.7}$$

Η συνάρτηση $f(.)$ χαρακτηρίζει τον κβαντιστή. Για παράδειγμα για τον ιδανικό κβαντιστή του Shannon είναι

$$f(x) = 2^{-2x}. \tag{5.8}$$

Η λύση του παραπάνω προβλήματος είναι γνωστή. Αν $h(.)$ είναι η αντίστροφη συνάρτηση της $f'(.)$, τότε ο αριθμός bits είναι

$$B_n = \begin{cases} h(\frac{\theta f'(0)}{\sigma_n^2}) & 0 < \theta < \sigma_n^2 \\ 0 & \theta \geqslant \sigma_n^2 \end{cases}, \tag{5.9}$$

όπου θ είναι η λύση της εξίσωσης

$$R = \frac{1}{N}\sum_{\sigma_n^2 \geqslant \theta} h(\frac{\theta f'(0)}{\sigma_n^2}). \tag{5.10}$$

Για την περίπτωση του κβαντιστή του Shannon (5.8), το αποτέλεσμα είναι

$$B_n = \max(0, \frac{1}{2}\log_2\frac{\sigma_n^2}{\theta}),$$

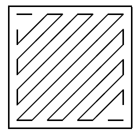
οπότε

$$D = \frac{1}{N}\sum_{n=1}^{N}\min(\theta, \sigma_n^2).$$

Η Εξίσωση (5.9) δεν εξασφαλίζει ακέραιο αριθμό bits. Απαιτείται επομένως η στρογγύλευση του αποτελέσματος, για να έχουμε ακέραιο αριθμό bits ανά συντελεστή του μετασχηματισμού, και το δοσμένο ρυθμό μετάδοσης της πληροφορίας.

5.2.2 Μεταβλητός και προσαρμοζόμενος ρυθμός πληροφορίας

Η τεχνική αυτή χρησιμοποιείται για τη συμπίεση εικόνων, γι' αυτό η παρουσίαση θα γίνει για δισδιάστατα μπλοκ. Η χρήση μεταβλητού αριθμού bits ανά μπλοκ δίδει τη δυνατότητα προσαρμογής του ρυθμού μετάδοσης της πληροφορίας στο περιεχόμενο του κάθε μπλοκ. Προς τούτο ο κβαντιστής είναι ομοιόμορφος και συμμετρικός με περιττό αριθμό επιπέδων κβαντισμού. Παρατηρείται ότι ένας μεγάλος αριθμός συντελεστών του μετασχηματισμού αντιστοιχεί στο ενδιάμεσο επίπεδο κβαντισμού με τιμή '0'. Κατά κύριο λόγο οι συντελεστές αυτοί αντιστοιχούν σε υψηλούς δείκτες, που με τη σειρά τους αντιστοιχούν σε υψηλές συχνότητες για τους χρησιμοποιούμενους μετασχηματισμούς. Αρκεί να δοθεί το ορθογώνιο παραλληλόγραμμο που περιλαμβάνει όλους τους μη 'μηδενικούσ' συντελεστές, για να καθορισθεί η τιμή όλων των υπόλοιπων, ή να χρησιμοποιηθεί μία αντι-διαγώνια σάρωση του μπλοκ, εφόσον είναι δισδιάστατο, που να διέρχεται πρώτα από τους συντελεστές με χαμηλούς δείκτες, και να δοθεί η διεύθυνση του τελευταίου μη 'μηδενικού' συντελεστή (Σχήμα 5.5). Επιπλέον στη ζώνη που περιλαμβάνει τους



Σχήμα 5.5: Αντι-διαγώνια σάρωση του μπλοκ

μη 'μηδενικούσ' συντελεστές μπορεί να υπάρχει ακόμη ικανός αριθμός 'μηδενικών' συντελεστών. Ο βαθμός συμπίεσης είναι υψηλότερος, εάν κωδικοποιηθούν τα διαστήματα 'μηδενικών' συντελεστών, αντί να κωδικοποιηθεί κάθε τιμή ξεχωριστά.

	0	1	2	3	4	5	6	7
0	16	11	10	16	24	40	51	61
1	12	12	14	19	26	58	60	55
2	14	13	16	24	40	57	69	56
3	14	17	22	29	51	87	80	62
4	18	22	37	56	68	109	103	77
5	24	35	55	64	81	104	113	92
6	49	64	78	87	103	121	120	101
7	72	92	95	98	112	100	103	99

Table 5.1: Πίνακας κβαντισμού για ένα μπλοκ συντελεστών DCT 8×8 συνιστώμενος στο σύστημα JPEG

Η χρήση μεταβλητού αριθμού βιτς ανάλογα με το περιεχόμενο του κάθε μπλοκ αξιοποιεί επίσης τη χρήση ενός πίνακα κβαντισμού, που λαμβάνει υπόψη, στην περίπτωση των εικόνων, την απόκριση του ανθρώπινου συστήματος όρασης ανάλογα με τη συχνότητα, δηλαδή με τη θέση του συντελεστή. Ένα παράδειγμα τέτοιου πίνακα με διαφορετικό βήμα κβαντισμού ανάλογα με τη θέση του συντελεστή δίδεται στον Πίνακα 5.1.

Η μέθοδος αυτή κωδικοποίησης με μεταβλητό αριθμό bits ανά μπλοκ χρησιμοποιείται στο σύστημα συμπίεσης JPEG. Ένα παράδειγμα αποτελεσμάτων δίδεται στο Σχήμα 5.6.

5.3 Συμπίεση με ανάλυση σε ζώνες συχνοτήτων

Σ' αυτή τη μέθοδο το σήμα διαχωρίζεται σε συνιστώσες σήματος που η καθεμία αντιστοιχεί σε μία διαφορετική ζώνη συχνοτήτων.

Σχήμα 5.6: Αποτελέσματα συμπίεσης με το σύστημα JPEG. Αριστερά: η αρχική εικόνα, Κέντρο: συμπιεσμένη εικόνα με συντελεστή συμπίεσης 9.5, Δεξιά: συμπιεσμένη εικόνα με συντελεστή συμπίεσης 18.5

Ας ξεκινήσουμε με κάποιους χρήσιμους ορισμούς. Ονομάζεται αποδεκατισμός τάξης p (όπου $p > 1$ φυσικός αριθμός) η περιοδική δειγματοληψία ενός από κάθε p δείγματα ενός σήματος. Το αποδεκατισμένο σήμα που προκύπτει από το σήμα $x(n)$ είναι

$$y(n) = x(pn). \tag{5.11}$$

Ας ονομάσουμε $X(z)$ (αντίστοιχα $Y(z)$) το μετασχηματισμό Z του $x(n)$ (αντίστοιχα $y(n)$). Η σχέση μεταξύ των δύο μετασχηματισμών Z προκύπτει ως εξής

$$\begin{aligned} Y(z) &= \sum_{n=-\infty}^{\infty} x(pn) z^{-n} = \sum_{n=-\infty}^{\infty} \frac{z^{-n}}{2\pi j} \oint X(\zeta)\zeta^{pn-1} d\zeta \\ &= \frac{1}{2\pi j} \oint X(\zeta) \sum_{n=-\infty}^{\infty} \left(\frac{z}{\zeta^p}\right)^n \frac{d\zeta}{\zeta} = \frac{1}{2\pi j} \oint X(\zeta) \frac{\zeta^{p-1}}{\zeta^p - z} d\zeta \\ &= \frac{1}{2\pi j} \oint X(\zeta) \frac{1}{p} \sum_{k=0}^{p-1} \frac{1}{\zeta - (ze^{-j2\pi k})^{1/p}} d\zeta \\ &= \frac{1}{p} \sum_{k=0}^{p-1} X(z^{1/p} e^{-\frac{j2\pi k}{p}}) \end{aligned} \tag{5.12}$$

Απ΄ αυτή τη σχέση μπορεί να εξαχθεί επίσης η σχέση μεταξύ των δύο μετασχηματισμών Fourier

$$\mathcal{Y}(u) = \frac{1}{p} \sum_{k=0}^{p-1} \mathcal{X}\left(\frac{u-k}{p}\right). \tag{5.13}$$

Είναι φανερό ότι με τον αποδεκατισμό δημιουργείται ένα πρόβλημα ψευδώνυμων συχνοτήτων. Το φαινόμενο αυτό δεν είναι επιθυμητό, αντίθετα επιζητείται ο περιορισμός του αποδεκατισμένου σήματος σε μια ζώνη συχνοτήτων εύρους $1/p$. Γι΄ αυτό απαιτείται η χρήση ενός βαθυπερατού φίλτρου πριν τον αποδεκατισμό. Το ιδανικό φίλτρο, εάν υπήρχε, θα είχε την ακόλουθη απόκριση στις συχνότητες

$$\mathcal{H}(u) = \begin{cases} 1 & |u| \leqslant \frac{1}{2p} \\ 0 & \frac{1}{2p} < |u| \leqslant \frac{1}{2} \end{cases} \tag{5.14}$$

Στην πράξη για ένα οποιοδήποτε φίλτρο και μετά τον αποδεκατισμό θα ομιλούμε για προσέγγιση του σήματος σε κλίμακα $1/p$ και θα έχουμε τις ακόλουθες σχέσεις

$$\begin{aligned} x_a(n:p) &= \sum_{k=-\infty}^{\infty} h(pn-k)x(k) && (5.15)\\ X_a(z:p) &= \frac{1}{p}\sum_{k=0}^{p-1} H(z^{1/p}e^{-\frac{j2\pi k}{p}})X(z^{1/p}e^{-\frac{j2\pi k}{p}}) && (5.16)\\ \mathcal{X}_a(z:p) &= \frac{1}{p}\sum_{k=0}^{p-1} \mathcal{H}(\frac{u-k}{p})\mathcal{X}(\frac{u-k}{p}) && (5.17) \end{aligned}$$

Στην αντίθετη κατεύθυνση σε σχέση με τον αποδεκατισμό ευρίσκεται η παρεμβολή δειγμάτων με μηδενική τιμή για την επέκταση του σήματος κατά ένα παράγοντα p. Το σήμα που προκύπτει είναι το ακόλουθο

$$y(n) = \begin{cases} x(\frac{n}{p}) & n\%p = 0 \\ 0 & n\%p \neq 0 \end{cases} \tag{5.18}$$

όπου $n\%p$ δηλώνει το υπόλοιπο της διαίρεσης του n δια p. Αποδεικνύεται εύκολα ότι μεταξύ των μετασχηματισμών Z και Fourier ισχύουν οι ακόλουθες σχέσεις

$$Y(z) = X(z^p) \tag{5.19}$$

$$\mathcal{Y}(u) = \mathcal{X}(pu) \tag{5.20}$$

Η σταθμισμένη παρεμβολή $p-1$ τιμών μεταξύ δύο τιμών του αρχικού σήματος $x(n)$ προκύπτει με τη χρήση ενός βαθυπερατού φίλτρου μετά την παρεμβολή των μηδενικών τιμών. Οπότε θα έχουμε

$$x_i(n:p) = \sum_{k=-\infty}^{\infty} h(n-k)y(k) = \sum_{l=-\infty}^{\infty} h(n-pl)x(l) \tag{5.21}$$

Μεταξύ των μετασχηματισμών Z και Fourier ισχύουν οι ακόλουθες σχέσεις

$$X_i(z:p) = H(z)X(z^p) \tag{5.22}$$

$$\mathcal{X}_i(u:p) = \mathcal{H}(u)\mathcal{X}(pu). \tag{5.23}$$

Στη συνέχεια θα περιορισθούμε μόνο στην περίπτωση της ανάλυσης σε δύο ζώνες συχνοτήτων ($p = 2$) με βάση τους παραπάνω ορισμούς. Η προσέγγιση του σήματος σε κλίμακα $1/2$ αντιστοιχεί στη ζώνη χαμηλών συχνοτήτων. Συνήθως το φίλτρο $h(n)$ επιλέγεται ώστε να ικανοποιεί τη συνθήκη

$$\sum_{n=-\infty}^{\infty} h(2n) = \sum_{n=-\infty}^{\infty} h(2n+1).$$

Με παρόμοιο τρόπο και στην ίδια κλίμακα μπορούν να εξαχθούν οι λεπτομέρειες του σήματος που αντιστοιχούν στις υψηλές συχνότητες

$$x_d(n) = \sum_{k=-\infty}^{\infty} g(2n-k)x(k) \tag{5.24}$$

Οπότε θα έχουμε, χρησιμοποιώντας την Εξίσωση (5.16)

$$X_d(z) = \frac{1}{2}\left(G(z^{1/2})X(z^{1/2}) + G(-z^{1/2})X(-z^{1/2})\right) \tag{5.25}$$

Το φίλτρο για την εξαγωγή των υψηλών συχνοτήτων επιλέγεται με βάση το βαθυπερατό φίλτρο που χρησιμοποιείται για την προσέγγιση. Συνήθης επιλογή είναι η ακόλουθη

$$\begin{aligned} \mathcal{G}(u) &= e^{-j2\pi u}\mathcal{H}^*(u-\frac{1}{2}) && (5.26)\\ G(z) &= z^{-1}H(-\frac{1}{z}) && (5.27)\\ g(n) &= (-1)^{1-n}h(1-n) && (5.28) \end{aligned}$$

Από την προσέγγιση και από τις λεπτομέρειες του σήματος, ακόμα κι αν το φίλτρο h δεν είναι ιδανικό μπορεί να αποκατασταθεί τέλεια το αρχικό σήμα αρκεί να χρησιμοποιηθεί σταθμισμένη παρεμβολή με το φίλτρο $H(\frac{1}{z})$ για τις χαμηλές συχνότητες και με το $G(\frac{1}{z})$ για τις υψηλές συχνότητες, και τα δύο προκύπτοντα σήματα να προστεθούν. Ας είναι $\hat{X}(z)$ ο μετασχηματισμός Z που προκύπτει μ' αυτό τον τρόπο. Θα έχουμε

$$\begin{aligned} \hat{X}(z) &= \frac{1}{2}\left(H(z)X(z) + H(-z)X(-z)\right)H(\frac{1}{z}) + \frac{1}{2}\left(G(z)X(z) + G(-z)X(-z)\right)G(\frac{1}{z}) \\ &= \frac{1}{2}\left(H(z)H(\frac{1}{z}) + H(-z)H(-\frac{1}{z})\right)X(z) + \frac{1}{2}\left(H(-z)H(\frac{1}{z}) - H(\frac{1}{z})H(-z)\right)X(-z) \end{aligned}$$

Αν επομένως $H(z)H(\frac{1}{z}) + H(-z)H(-\frac{1}{z}) = 2$, τότε $\hat{X}(z) = X(z)$. Η παραπάνω συνθήκη είναι ισοδύναμη με

$$|\mathcal{H}(u)|^2 + |\mathcal{H}(u-\frac{1}{2})|^2 = 2. \tag{5.29}$$

Επιπλέον μεταξύ των φίλτρων H και G απαιτείται μία σχέση ορθογωνιότητας

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} \mathcal{H}(u)\mathcal{G}^*(u)du = 0.$$

Φίλτρα που επαληθεύουν αυτές τις σχέσεις ανήκουν στην κατηγορία των συζυγών ορθογώνιων φίλτρων.

Ένα ζεύγος φίλτρων που ικανοποιεί όλες τις παραπάνω απαιτήσεις είναι το ακόλουθο

$$H(z) = \frac{1}{\sqrt{2}}(1+z^{-1}) \quad \text{και} \quad G(z) = \frac{1}{\sqrt{2}}(-1+z^{-1})$$

που όμως έχει πτωχές επιδόσεις ως προς την ανάλυση σε ζώνες συχνοτήτων. Κατά κανόνα χρησιμοποιούνται φίλτρα με πεπερασμένη, αλλά μεγαλύτερη σε έκταση κρουστική απόκριση, που ικανοποιούν τις παραπάνω απαιτήσεις. Τέτοια παραδείγματα φίλτρων δίδονται στον Πίνακα 5.2.

Η ανάλυση μπορεί να επαναληφθεί για οποιαδήποτε από τις ζώνες συχνοτήτων, αλλά κυρίως έχει ενδιαφέρον η περαιτέρω ανάλυση της ζώνης χαμηλών συχνοτήτων για προσέγγιση σε μικρότερη κλίμακα. Συνολικά το σήμα μπορεί να αναλυθεί σε K σήματα s_k, με διαφορετική διάρκεια το

n	$h_4(n)$	$h_6(n)$	$h_8(n)$	$h_{10}(n)$	$h_{12}(n)$
0	.483	.333	.230	.160	.112
1	.837	.807	.715	.604	.495
2	.224	.460	.631	.724	.751
3	-.129	-.135	-.028	.138	.315
4		-.085	-.187	-.242	-.226
5		.035	.031	-.032	-.130
6			.033	.078	.098
7			-.011	-.006	.028
8				-.013	-.032
9				.003	.001
10					.005
11					-.001

Table 5.2: Φίλτρα προσέγγισης σε κλίμακα 1/2.

καθένα στη γενική περίπτωση, διάρκεια που προσδιορίζεται από το συντελεστή υπο-δειγματοληψίας, V_k. Θα ισχύει η ακόλουθη σχέση

$$\sum_{k=1}^{K} \frac{1}{V_k} = 1.$$

Το κέρδος που προκύπτει για τη συμπίεση, αν σ_k^2 είναι η διασπορά του σήματος s_k, είναι

$$G_s = \frac{\sigma^2}{\prod_{k=1}^{K} (\sigma_k^2)^{\frac{1}{V_k}}} \tag{5.30}$$

Για την επίτευξη αυτού του κέρδους το σήμα κάθε ζώνης, ανάλογα και με τα στατιστικά χαρακτηριστικά του, είτε κβαντίζεται βαθμωτά ή διανυσματικά, είτε μεσολαβεί ένα φίλτρο πρόβλεψης ή κάποιος μετασχηματισμός πριν τον κβαντισμό.