

ΕΡΓΑΣΤΗΡΙΟ 6 ΚΗ ΑΣΚΗΣΗ



Συμπίεση Εικόνων με Χρήση των Κωδικοποιήσεων Δέλτα & CVSD

1 Εισαγωγή

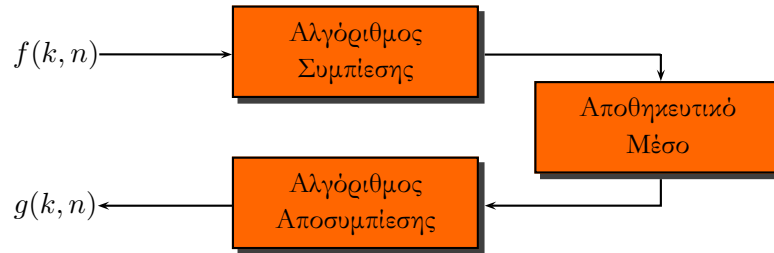
Μια εικόνα είναι ένα σήμα δύο διαστάσεων ή με άλλα λόγια μια συνάρτηση δύο μεταβλητών $f(x, y)$. Η τιμή της συνάρτησης αυτής για ένα συγκεκριμένο ζεύγος τιμών (x_0, y_0) μπορεί να είναι ένα χρώμα, αν πρόκειται για μια έγχρωμη εικόνα, ή μια τιμή φωτεινότητας αν πρόκειται για μια εικόνα αποχρώσεων του γκρι. Στα επόμενα θα θεωρήσουμε μόνο εικόνες αποχρώσεων του γκρι. Ακριβώς όμοια με την περίπτωση όπου ένα αναλογικό σήμα μιας διάστασης (π.χ. ήχος) δειγματοληπτείται με σκοπό την επεξεργασία ή την αποθήκευσή του από ένα ψηφιακό σύστημα, έτσι και μια εικόνα μπορεί να δειγματοληπτηθεί στο χώρο και να προκύψει έτσι μια ψηφιακή εκδοχή της

$$f_d(k, n) = f(k \cdot dx, n \cdot dy) \quad k, n = 0, \pm 1, \pm 2, \dots$$

όπου dx και dy είναι σταθερές οι οποίες αντιστοιχούν στην περίοδο δειγματοληψίας T στην περίπτωση της δειγματοληψίας ενός σήματος μιας διάστασης. Το σημείο (k, n) ονομάζεται *εικονοστοιχείο* (picture element, pixel) της εικόνας.

Για την αποθήκευση των τιμών $f_d(k, n)$ από ένα ψηφιακό σύστημα θα πρέπει να επιλέξουμε μια κατάλληλη αριθμητική αναπαράσταση τους. Η αναπαράσταση αυτή θα έχει πεπερασμένη αριθμητική ακρίβεια. Μια συνήθης αναπαράσταση για τις εικόνες διαβαθμίσεων του γκρι είναι να χρησιμοποιούμε έναν ακέραιο αριθμό με την τιμή μηδέν να αντιστοιχεί στο μαύρο χρώμα και την τιμή 2^k στο λευκό, όπου k είναι το πλήθος των δυαδικών ψηφίων τα οποία έχουμε στη διάθεσή μας για την αποθήκευση της τιμής της φωτεινότητας για κάθε εικονοστοιχείο. Η τιμή k ονομάζεται συνήθως *βάθος χρώματος* της εικόνας. Η αναπαράσταση αυτή της εικόνας ονομάζεται Pulse Code Modulation (PCM) [1].

Το μέγεθος το οποίο καταλαμβάνει μια εικόνα σε ένα ψηφιακό μέσο αποθήκευσης μπορεί να είναι ασύμφορα μεγάλο. Για παράδειγμα, μια εικόνα με 1024×768 εικονοστοιχεία και 16 bit βάθος χρώματος θα απαιτούσε αποθηκευτικό χώρο 1.572.864 bytes ή 1,5 MB. Ο χώρος αυτός μπορεί να μειωθεί κατά πολύ αν εφαρμόσουμε κάποια



Σχήμα 1: Διαδικασία συμπίεσης και αποσυμπίεσης εικόνας

τεχνική συμπίεσης των δεδομένων της εικόνας. Οι τεχνικές αυτές μπορούν να είναι είτε απωλεστικές (lossy) είτε μη-απωλεστικές (lossless) ανάλογα με το αν διατηρούν μέρος της πληροφορίας της εικόνας ή όλη την πληροφορία χωρίς υποβάθμιση της ποιότητάς της. Για παράδειγμα, οι εικόνες τύπου JPEG [1] χρησιμοποιούν έναν αλγόριθμο απωλεστικής συμπίεσης ενώ οι εικόνες τύπου GIF χρησιμοποιούν τον μη-απωλεστικό αλγόριθμο Ziv-Lempel [3].

Οι αλγόριθμοι συμπίεσης εικόνων διαφέρουν από τους αλγορίθμους συμπίεσης οποιονδήποτε δεδομένων (π.χ. αρχεία ZIP). Οι δεύτεροι αλγόριθμοι δεν εκμεταλλεύονται το γεγονός πως τα εικονοστοιχεία μιας εικόνας τα οποία βρίσκονται σε κοντινές θέσεις παρουσιάζονται συσχετισμένα, ότι δηλαδή μια εικόνα συνήθως αποτελείται από αριστερές ομοιογενείς περιοχές. Αντίθετα, οι αλγόριθμοι οι οποίοι εξειδικεύονται σε συμπίεση δεδομένων εικόνων προσπαθούν να εκμεταλλευτούν όσο μπορούν αυτήν την ιδιότητα. Αντικείμενο της άσκησης αυτής είναι η υλοποίηση και η αξιολόγηση ορισμένων αλγορίθμων απωλεστικής συμπίεσης εικόνων.

2 Κριτήρια Ποιότητας Συμπίεσης

Ας θεωρήσουμε πως μια εικόνα $f(k, n)$ κωδικοποιείται με χρήση κάποιου αλγορίθμου απωλεστικής συμπίεσης και στη συνέχεια η έξοδος του αλγορίθμου κωδικοποίησης τροφοδοτείται στον αντίστοιχο αλγόριθμο αποκωδικοποίησης δίνοντας έτσι την ανακατασκευασμένη εικόνα $g(k, n)$ όπως παρουσιάζεται στο Σχήμα 1. Εφόσον χρησιμοποιούμε έναν απωλεστικό αλγόριθμο κωδικοποίησης της αρχικής εικόνας, η τελική εικόνα $g(k, n)$ θα παρουσιάζει διαφορές σε σχέση με την αρχική. Για να ποσοτικοποιήσουμε τις διαφορές αυτές και να μετρήσουμε έτσι την υποβάθμιση την οποία εισήγαγε η διαδικασία της κωδικοποίησης θα πρέπει να ορίσουμε κάποιο μέτρο απόστασης ανάμεσα στις δύο εικόνες. Συνηθισμένα μέτρα απόστασης είναι τα ακόλουθα [1]:

Το *κανονικοποιημένο μέσο τετραγωνικό σφάλμα* (Normalized Mean Square Error, NMSE) ανάμεσα σε δύο εικόνες $f(k, n)$ και $g(k, n)$ ορίζεται ως:

$$NMSE = 100 \times \frac{Var[f(k, n) - g(k, n)]}{Var[f(k, n)]} \%$$

όπου $Var[\cdot]$ είναι ο τελεστής υπολογισμού της διασποράς μιας εικόνας.

Ο λόγος *σήματος προς θόρυβο* όπου ως σήμα θεωρούμε την αρχική εικόνα $f(k, n)$ και ως θόρυβο θεωρούμε τη διαφορά ανάμεσα στην αρχική εικόνα και την ανακατασκευασμένη

εκδοχή της ορίζεται με χρήση του NMSE ως

$$SNR = 10 \log_{10} \left(\frac{NMSE}{100} \right)^{-1} dB$$

Τα δύο προηγούμενα μέτρα της ποιότητας ανακατασκευής μιας συμπίεσμνης εικόνας δεν καταφέρνουν πάντα να αποτυπώσουν την αίσθηση της ποιότητας έτσι όπως την αντιλαμβάνεται το ανθρώπινο οπτικό σύστημα. Ένα άλλο μέτρο το οποίο καταφέρνει καλύτερα να αποτυπώσει την ποιότητα μιας συμπίεσμνης εικόνας ονομάζεται ανεπηρέαστο από αλλαγή κλίμακας κανονικοποιημένο μέσο τετραγωνικό σφάλμα (Scale-Invariant Normalized Mean Square Error, SINMSE) και ορίζεται ως

$$SINMSE = 100 \times \frac{Var[f(k, n) - a \cdot g(k, n)]}{Var[f(k, n)]} \%$$

όπου a είναι εκείνη η σταθερά για την οποία το παραπάνω μέτρο λαμβάνει την ελάχιστη τιμή του, και η οποία δίνεται από τον τύπο

$$a = \frac{E[f(k, n)g(k, n)] - E[f(k, n)]E[g(k, n)]}{Var[g(k, n)]}$$

όπου $E[\cdot]$ είναι ο τελεστής της αναμενόμενης τιμής μιας εικόνας.

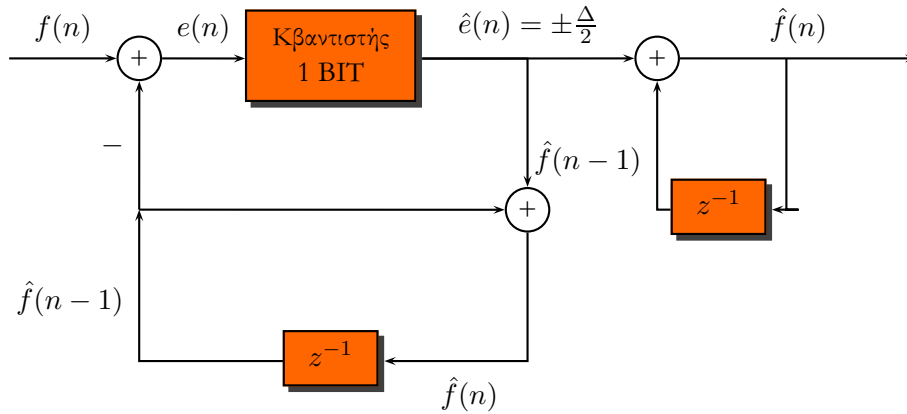
3 Κωδικοποίηση Δέλτα

Στην αναπαράσταση PCM μιας εικόνας διαβαθμίσεων του γκρι για κάθε εικονοστοιχείο αποθηκεύουμε μια κβαντισμένη τιμή της φωτεινότητάς του. Με τον τρόπο αυτό, δεν εκμεταλλευόμαστε καθόλου το γεγονός πως εικονοστοιχεία τα οποία βρίσκονται σε κοντινές θέσεις παρουσιάζονται συσχετισμένα ή με άλλα λόγια οι τιμές της φωτεινότητάς τους είναι παρόμοιες. Ένας τρόπος να εκμεταλλευτούμε την ιδιότητα αυτή είναι να χρησιμοποιήσουμε την κωδικοποίηση δέλτα. Στην κωδικοποίηση δέλτα, κωδικοποιούμε τη διαφορά ανάμεσα στις τιμές φωτεινότητας γειτονικών εικονοστοιχείων της εικόνας χρησιμοποιώντας έναν κβαντιστή του ενός bit ο οποίος προσφέρει 2 επίπεδα. Με τον τρόπο αυτό, αποθηκεύουμε το δυαδικό ψηφίο 0 αν το επόμενο εικονοστοιχείο παρουσιάζει χαμηλότερη φωτεινότητα και το δυαδικό ψηφίο 1 αν η φωτεινότητά του είναι υψηλότερη πάντα σε σχέση με τη φωτεινότητα του τρέχοντος εικονοστοιχείου. Αν και γενικά η πράξη της αφαίρεσης δύο ποσοτήτων οδηγεί σε μια νέα ποσότητα η οποία έχει τη διπλάσια δυναμική περιοχή (και άρα θα θέλαμε περισσότερα δυαδικά ψηφία για να την περιγράψουμε) η συσχέτιση ανάμεσα σε γειτονικά εικονοστοιχεία της εικόνας είναι τέτοια έτσι ώστε η διασπορά του σήματος διαφοράς είναι πολύ μικρή.

Στην κωδικοποίηση δέλτα, θεωρούμε αρχικά πως η εικόνα προς κωδικοποίηση $f(k, n)$ μετατρέπεται σε ένα ισοδύναμο μονοδιάστατο σήμα, απλά με το να θεωρήσουμε ένα σταθερό τρόπο σάρωσης σύμφωνα με τον οποίο διατρέχουμε τα εικονοστοιχεία της εικόνας

$$f(k, n) \quad k = 1..K, n = 1..N \rightarrow f(i) \quad i = 1..K \cdot N$$

Ορισμένοι τρόποι με τους οποίους μπορούμε να σαρώσουμε τα εικονοστοιχεία μιας εικόνας και να μετασχηματίσουμε έτσι το δύο διαστάσεων σήμα $f(k, n)$ στο μονοδιάστατο σήμα $f(n)$ παρουσιάζονται στο Σχήμα 2. Είναι φανερό ότι το πόσο καλά



Σχήμα 3: Δομή του κωδικοποιητή και του αποκωδικοποιητή δέλτα

πλήρως το σήμα $\hat{e}(n)$.

Στο δέκτη, το σήμα $\hat{e}(n)$ προστίθεται στο σήμα $\hat{f}(n-1)$ έτσι ώστε να υπολογιστεί το σήμα $\hat{f}(n)$. Παρατηρούμε από το Σχήμα 3 πως με τον ίδιο τρόπο και ο δέκτης κατασκευάζει το σήμα $\hat{f}(n)$ γιατί θα το χρησιμοποιήσει στην κωδικοποίηση του $f(n+1)$. Οι εξισώσεις οι οποίες περιγράφουν τη λειτουργία ενός συστήματος κωδικοποίησης δέλτα είναι οι ακόλουθες

$$e(n) = f(n) - \hat{f}(n-1)$$

$$\hat{e}(n) = \begin{cases} \frac{\Delta}{2} & e(n) > 0 \\ -\frac{\Delta}{2} & e(n) \leq 0 \end{cases}$$

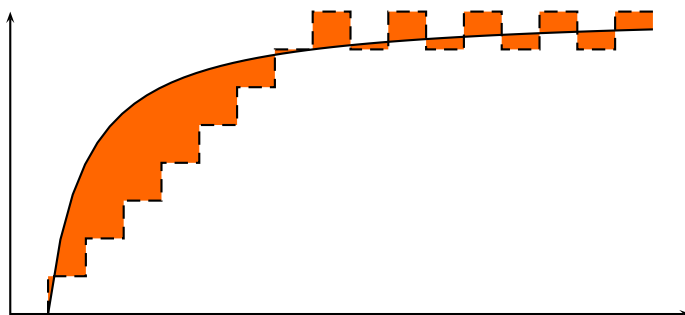
$$\hat{f}(n) = \hat{f}(n-1) + \hat{e}(n)$$

και ο θόρυβος εξαιτίας της κβάντισης θα δίνεται έτσι από τη σχέση

$$e_Q(n) = \hat{f}(n) - f(n) = \hat{e}(n) - e(n)$$

Παρατηρούμε πως σε ένα σύστημα κωδικοποίησης δέλτα, το σήμα σφάλματος $e(n)$ παράγεται ως η διαφορά ανάμεσα στο σήμα $f(n)$ και το σήμα $\hat{f}(n-1)$. Θα περίμενε κάποιος πως αν χρησιμοποιούσαμε το σήμα $f(n-1)$ αντί του σήματος $\hat{f}(n-1)$ για να δημιουργήσουμε το σήμα σφάλματος τότε αφενός δεν θα χρειαζόταν ο κωδικοποιητής να υπολογίζει το σήμα $\hat{f}(n-1)$ και αφετέρου το σήμα σφάλματος θα έπαιρνε κατά απόλυτο τιμή μικρότερες τιμές και επομένως θα ήταν σωστότερο. Ωστόσο, στην περίπτωση αυτή, ο αποκωδικοποιητής ο οποίος δεν μπορεί να έχει στη διάθεσή του το σήμα $f(n-1)$ θα ακολουθούσε μια διαδικασία ανακατασκευής του σήματος στην οποία το σφάλμα θα συσσωρευόταν [1]. Αντίθετα, αν ο κωδικοποιητής χρησιμοποιεί το σήμα $\hat{f}(n-1)$, τότε η διαδικασία αποκωδικοποίησης ακολουθεί ακριβώς την ίδια πορεία που ακολούθησε και η διαδικασία κωδικοποίησης και έτσι το σφάλμα ανακατασκευής δεν συσσωρεύεται [1].

Μια σημαντική παράμετρος του συστήματος κωδικοποίησης / αποκωδικοποίησης δέλτα είναι η σταθερά Δ . Ας θεωρήσουμε την ανακατασκευή ενός σήματος όπως παρουσιάζεται στο Σχήμα 4. Από το Σχήμα αυτό παρατηρούμε πως όταν το σήμα



Σχήμα 4: Υπερφόρτωση κλίσης και κοκκώδης θόρυβος

$f(n)$ (συνεχής γραμμή) μεταβάλλεται απότομα, το ανακατασκευασμένο σήμα $\hat{f}(n)$ (διακεκομμένη γραμμή) αργεί αρκετά να το ακολουθήσει και το φτάνει μετά από αρκετές χρονικές στιγμές. Το φαινόμενο αυτό ονομάζεται *υπερφόρτωση κλίσης* και υπαγορεύει τη χρήση μιας σταθεράς Δ αρκετά μεγάλης έτσι ώστε ο κωδικοποιητής να μπορεί να παρακολουθεί τις απότομες μεταβάσεις του σήματος. Αντίθετα, στις περιοχές όπου το σήμα $f(n)$ είναι σχετικά αμετάβλητο, παρατηρούμε πως το σήμα $\hat{f}(n)$ "ταλαντώνεται" ανάμεσα σε 2 στάθμες που περιβάλλουν το σήμα $f(n)$. Το φαινόμενο αυτό εισάγει ένα θόρυβο ο οποίος ονομάζεται *κοκκώδης θόρυβος* και υπαγορεύει τη χρήση μιας μικρής σταθεράς Δ . Παρατηρούμε πως οι δύο υποβαθμίσεις αυτές τις οποίες εισάγει η κωδικοποίηση δέλτα έχουν αντικρουόμενες απαιτήσεις σχετικά με τη σταθερά Δ ¹

3.1 Περισσότερα δυαδικά ψηφία ανά εικονοστοιχείο

Προκειμένου να πετύχουμε καλή ποιότητα ανακατασκευής χρησιμοποιώντας κωδικοποίηση δέλτα, συνήθως απαιτείται να χρησιμοποιήσουμε περισσότερα του ενός δυαδικά ψηφία για κάθε ένα από τα εικονοστοιχεία της εικόνας. Κάτι τέτοιο επιτυγχάνεται με το να υπερδειγματοληψήσουμε το αρχικό αναλογικό δυδιάστατο σήμα $f(x, y)$ χρησιμοποιώντας μικρότερες τιμές για τις σταθερές δειγματοληψίας dx και dy . Με τον τρόπο αυτό, μειώνονται οι απότομες μεταβάσεις του ψηφιακού σήματος $f(n)$ που προκύπτει και μπορούμε να χρησιμοποιήσουμε μια μικρότερη σταθερά Δ χωρίς να ενισχύεται το φαινόμενο της υπερφόρτωσης κλίσης. Ταυτόχρονα μειώνεται και η ενέργεια του κοκκώδους θορύβου.

Στην (πρακτική) περίπτωση όπου δεν έχουμε στην διάθεσή μας το αρχικό αναλογικό σήμα $f(x, y)$ αλλά μόνο την ψηφιακή εκδοχή του $f(k, n)$ και επιθυμούμε να κωδικοποιήσουμε την εικόνα αυτή χρησιμοποιώντας k δυαδικά ψηφία ανά εικονοστοιχείο, τότε μπορούμε να κάνουμε μια εκτίμηση για την υπερδειγματοληπτημένη εικόνα χρησιμοποιώντας κάποια τεχνική παρεμβολής. Για παράδειγμα, έστω πως έχουμε στη διάθεσή μας μια εικόνα $f(k, n)$, $k = 1 \dots K$ και $n = 1 \dots N$ και θέλουμε να την κωδικοποιήσουμε με δυο δυαδικά ψηφία ανά εικονοστοιχείο. Στην περίπτωση αυτή, σε πρώτη φάση

¹ Για το λόγο αυτό, επιλέγουμε συνήθως μια σταθερά Δ έτσι ώστε να πετυχαίνουμε κάποιο διακανονισμό ανάμεσα στις δύο αυτές υποβαθμίσεις ή χρησιμοποιούμε κάποιο αυτοπροσαρμοζόμενο σύστημα για την εκτίμηση ενός χρονικά μεταβαλλόμενου Δ_n

θα πρέπει με κάποια τεχνική παρεμβολής να εκτιμήσουμε την υπερδειγματοληπτημένη εικόνα $h(k', n')$ ως

$$f(k, n), k = 1 \dots K, n = 1 \dots N \rightarrow h(k', n'), k' = 1 \dots K, n' = 1 \dots 2 \cdot N$$

όπου παρατηρούμε πως η εικόνα $h(k', n')$ έχει διπλάσιο αριθμό εικονοστοιχείων. Στην περίπτωση όπου έχουμε επιλέξει τον τρόπο σάρωσης του Σχήματος 2(α), η εικόνα $h(k', n')$ θα πρέπει να περιέχει στις περιττές στήλες της τις τιμές φωτεινότητας της αρχικής εικόνας ενώ στις άρτιες στήλες της τις τιμές που προέκυψαν από τη διαδικασία παρεμβολής. Στη συνέχεια κωδικοποιώντας την εικόνα $h(k', n')$ με κωδικοποίηση δέλτα, έχουμε επιτύχει να κωδικοποιήσουμε την εικόνα $f(k, n)$ χρησιμοποιώντας 2 δυαδικά ψηφία ανά εικονοστοιχείο. Η διαδικασία ανακατασκευής τότε θα περιλαμβάνει ένα επιπλέον βήμα κατά το οποίο η ανακατασκευασμένη εικόνα $\hat{h}(k', n')$ θα πρέπει να υποδειγματοληπτηθεί κατά τον οριζόντιο άξονα, ή με άλλα λόγια να κρατηθούν μόνο οι περιττές στήλες της. Επισημαίνουμε στο σημείο αυτό πως σημαντικό ρόλο στη διαδικασία παρεμβολής παίζει η τάξη και το είδος των πολυωνύμων παρεμβολής που θα χρησιμοποιηθούν καθώς και η περιοχή γύρω από την οποία θα υπολογιστεί η διαδικασία παρεμβολής. Ένα απλό παράδειγμα διαφοροποίησης είναι το αν η παρεμβολή θα υπολογιστεί στη μία ή στις δύο διαστάσεις.

4 Κωδικοποίηση CVSD

Η κωδικοποίηση CVSD (Continuously Variable Slope Delta Modulation) αποτελεί μια τροποποίηση της κωδικοποίησης δέλτα που έχει σκοπό να αντιμετωπίσει τα προβλήματα της υπερφόρτωσης κλίσης και του κοκκώδους θορύβου τα οποία περιγράψαμε στα προηγούμενα. Τα προβλήματα αυτά, είδαμε πως σχετίζονται άμεσα με τη σταθερά Δ την οποία χρησιμοποιεί το σύστημα κωδικοποίησης/αποκωδικοποίησης.

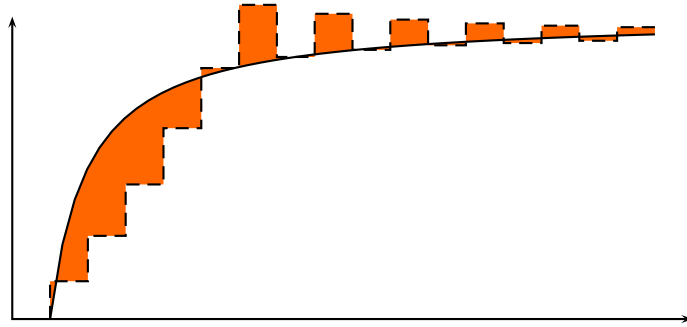
Στην κωδικοποίηση CVSD χρησιμοποιούμε ένα χρονικά μεταβαλλόμενο Δ_n το οποίο εξαρτάται από τις προηγούμενες δύο τιμές της εξόδου του κωδικοποιητή. Πιο συγκεκριμένα, η παράμετρος Δ_n δίνεται από τη σχέση²

$$\Delta_n = \begin{cases} \beta \Delta_{n-1} + D_2 & \text{αν } \hat{f}(n) = \hat{f}(n-1) + \hat{f}(n-2) \\ \beta \Delta_{n-1} + D_1 & \text{σε όλες τις άλλες περιπτώσεις} \end{cases}$$

όπου $0 < \beta < 1$ και $D_2 \gg D_1 > 0$. Οι τιμές των παραμέτρων D_1, D_2 καθώς και η σταθερά λήθης β σχετίζονται με τη μέγιστη και την ελάχιστη επιτρεπτή τιμή της παραμέτρου Δ_n σύμφωνα με τις σχέσεις

$$\Delta_{max} = \frac{D_2}{1 - \beta}$$

²Στην πραγματικότητα πρόκειται για δύο μορφές ενός "ακολουθιακού" εκτιμητή, της γενικής μορφής $a_n = \beta a_{n-1} + a$, όπου κάθε φορά η προηγούμενη εκτίμηση a_{n-1} ανανεώνεται σύμφωνα με μια τρέχουσα νέα εκτίμηση a . Στην περίπτωσή μας, η δεύτερη εξίσωση ανανέωσης συνεχώς μειώνει τη σταθερά Δ_n και εκτελείται κάθε φορά όπου δεν σημειώθηκε "υπερφόρτωση κλίσης". Αντίθετα, κάθε φορά που σημειώνεται το φαινόμενο της υπερφόρτωσης κλίσης, η πρώτη εξίσωση ανανέωσης αυξάνει τη σταθερά Δ_n .



Σχήμα 5: Μετριάσμος της υπερφόρτωσης κλίσης και του κοκκώδους θορύβου

$$\Delta_{min} = \frac{D_1}{1 - \beta}$$

Οι τιμές αυτών των τριών παραμέτρων πρέπει να επιλεγούν προσεκτικά ώστε να βεβαιωθούμε πως ο κωδικοποιητής CVSD θα μπορεί να παρακολουθεί τις μεταβολές του σήματος εισόδου και προκύπτουν από τις στατιστικές ιδιότητες του σήματος εισόδου.

Στο Σχήμα 5 έχουμε σχεδιάσει τη διαδικασία κωδικοποίησης ενός σήματος με χρήση της κωδικοποίησης CVSD. Αν συγκρίνουμε το Σχήμα αυτό με το Σχήμα 4 παρατηρούμε πως τα φαινόμενα της υπερφόρτωσης κλίσης και του κοκκώδους θορύβου μετριάζονται με τη χρήση της χρονικά μεταβαλλόμενης παραμέτρου Δ_n .

5 Ερωτήματα

Για την ολοκλήρωση της άσκησης καλείστε να απαντήσετε στα ακόλουθα ερωτήματα:

1. Να αποδείξετε πως η τιμή της σταθεράς a η οποία ελαχιστοποιεί το μέτρο $SINMSE$ το οποίο ορίσαμε στην Παράγραφο 2 δίνεται από τη σχέση

$$a = \frac{E[f(k, n)g(k, n)] - E[f(k, n)]E[g(k, n)]}{Var[g(k, n)]}$$

και να εξηγήσετε γιατί το συγκεκριμένο μέτρο το χαρακτηρίζουμε ως "ανεπηρέαστο από αλλαγή κλίμακας".

2. Χρησιμοποιώντας το MATLAB (ή οποιοδήποτε άλλο εργαλείο ή γλώσσα προγραμματισμού επιθυμείται) να υλοποιήσετε τις συναρτήσεις οι οποίες δέχονται ως παραμέτρους την αρχική και την αποσυμπιεσμένη εικόνα και υπολογίζουν τα τρία μέτρα ποιότητας τα οποία ορίσαμε στην Παράγραφο 2.
3. Σε ένα σύστημα κωδικοποίησης δέλτα, παρατηρούμε πως για την κωδικοποίηση του εικονοστοιχείου $f(n)$ θα πρέπει να έχουμε στη διάθεσή μας την εκτίμηση $\hat{f}(n-1)$. Ωστόσο, για το πρώτο εικονοστοιχείο $f(1)$ που θα κωδικοποιήσουμε, δεν μπορούμε να έχουμε την εκτίμηση $\hat{f}(0)$. Για να λύσουμε το πρόβλημα αυτό, μπορούμε να αρχικοποιήσουμε την αποθηκευμένη τιμή στο στοιχείο μνήμης τόσο του κωδικοποιητή όσο και του αποκωδικοποιητή σε μια γνωστή τιμή έτσι ώστε

και οι δύο να λειτουργούν σε συμφωνία. Αν η δυναμική περιοχή των εικόνων που θέλουμε να κωδικοποιεί ένα σύστημα κωδικοποίησης δέλτα είναι από 0 (μάυρο) έως 1 (λευκό), ποια θα ήταν η καλύτερη επιλογή αρχικοποίησης για το $\hat{f}(0)$;

4. Να υλοποιήσετε τις συναρτήσεις εκείνες οι οποίες δέχονται ως είσοδο μια εικόνα $f(k, n)$ (πίνακας) και επιστρέφουν ως έξοδο το μονοδιάστατο σήμα $f(n)$, για κάθε έναν από τους τρόπους σάρωσης του Σχήματος 2(α), 2(β) και 2(δ). Να υλοποιήσετε επίσης τις αντίστροφες συναρτήσεις για κάθε μια περίπτωση.
5. Χρησιμοποιώντας και τις συναρτήσεις που υλοποιήσατε στα προηγούμενα ερωτήματα, να υλοποιήσετε ένα σύστημα κωδικοποίησης / αποκωδικοποίησης δέλτα. Η συνάρτησή σας θα πρέπει να δέχεται ως παράμετρο την τιμή της σταθεράς Δ (και την εικόνα εισόδου) και να χρησιμοποιεί 1 δυαδικό ψηφίο ανά εικονοστοιχείο. Χρησιμοποιείτε την υλοποίησή σας για να κωδικοποιήσετε μια εικόνα θέτοντας τη σταθερά Δ αρχικά στο 8% και μετά στο 15% της δυναμικής περιοχής της εικόνας εισόδου. Εξετάστε τις περιπτώσεις των τριών τρόπων σάρωσης που έχετε υλοποιήσει. Για κάθε αποτέλεσμα δώστε τις αντίστοιχες τιμές για κάθε μέτρο ποιότητας συμπίεσης.
6. Να επαναλάβετε τις μετρήσεις του προηγούμενου ερωτήματος για ένα σύστημα κωδικοποίησης δέλτα το οποίο χρησιμοποιεί k δυαδικά ψηφία ανά εικονοστοιχείο, αλλά μόνο για την περίπτωση του τρόπου σάρωσης που φαίνεται στο Σχήμα 2(α). Η τιμή k να δίνεται ως παράμετρος στη συνάρτηση κωδικοποίησης / αποκωδικοποίησης. Εξηγήστε περιληπτικά πως θα έπρεπε να εκτελέσουμε τη διαδικασία παρεμβολής, αν ακολουθούσαμε κάποιο άλλο τρόπο σάρωσης από αυτούς που φαίνονται στο Σχήμα 2.
7. Να υλοποιήσετε ένα σύστημα κωδικοποίησης / απόκωδικοποίησης δέλτα, το οποίο να χρησιμοποιεί k δυαδικά ψηφία ανά εικονοστοιχείο και στο οποίο η διαδικασία παρεμβολής αποτελείται από ένα σύστημα "Παρεμβολής - Υπερδειγματοληψίας" [2]. Η διαδικασία αυτή, να γίνεται αφού τα δεδομένα της εικόνας έχουν προηγουμένως μετατραπεί στα ισοδύναμα μονοδιάστατα δεδομένα με χρήση κάποιου τρόπου σάρωσης όπως αυτούς του Σχήματος 2.
8. Να υλοποιήσετε τη διαδικασία κωδικοποίησης και αποκωδικοποίησης CVSD. Εξηγήστε στην περίπτωση αυτή τι ακριβώς πρέπει να λάβουμε υπόψη μας ώστε ο κωδικοποιητής και ο αποκωδικοποιητής να λειτουργούν σε συμφωνία.
9. Να συγκρίνετε την κωδικοποίηση δέλτα με την κωδικοποίηση CVSD για κάθε ένα από τα μέτρα ποιότητας συμπίεσης που ορίστηκαν και για μια σειρά από εικόνες εισόδου.

Βιβλιογραφία

- [1] Craig Marven and Gillian Ewers, *"A Simple Approach to Digital Signal Processing"*, Wiley-Interscience Publications, 1996

- [2] Γεώργιος Μουστακίδης, "*Βασικές Τεχνικές Ψηφιακής Επεξεργασίας Σημάτων*", Εκδόσεις Τζιόλα, 2003.
- [3] J. Ziv and A. Lempel, "*A Universal Algorithm for Sequential Data Compression*", IEEE Transactions on Information Theory, Vol. 23, pp. 337--342, 1977.