Ψηφιακή Επεξεργασία Εικόνας
Αναφορά bonus project

Εργαστηριακή Ομάδα *76*

Γιουμερτάκης Απόστολος, 2017030142
Κατσούπης Ευάγγελος, 2017030077

Uniform scalar quantizer

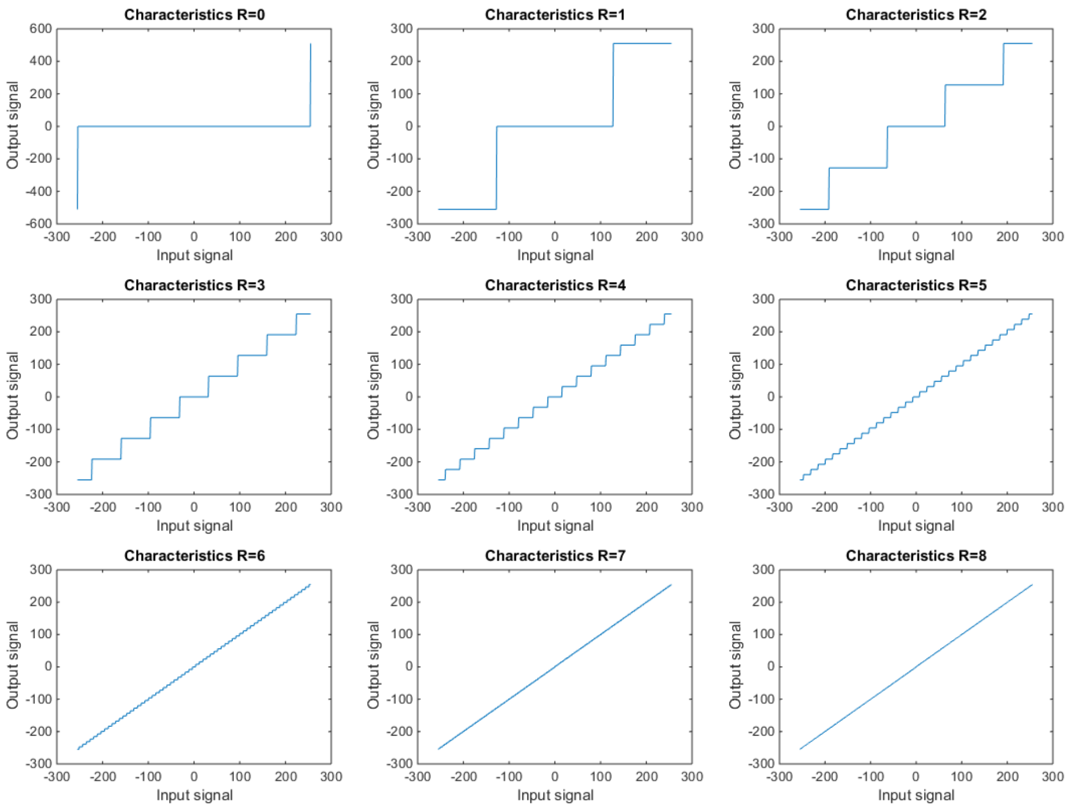
Στο 1ο μέρος της εργασίας, μας ζητήθηκε η κατασκευή της συνάρτησης “uni_scalar”, η οποία υλοποιεί έναν Ομοιόμορφο κβαντιστή.
Η έξοδος του κβαντιστή προκύπτει από την σχέση:

$$Q(x) = \Delta \times sign(x) \lfloor \frac{|x|}{\Delta} + \frac{1}{2} \rfloor$$
$$\Delta = \frac{A-(-A)}{L}, \quad L = 2^R$$

όπου

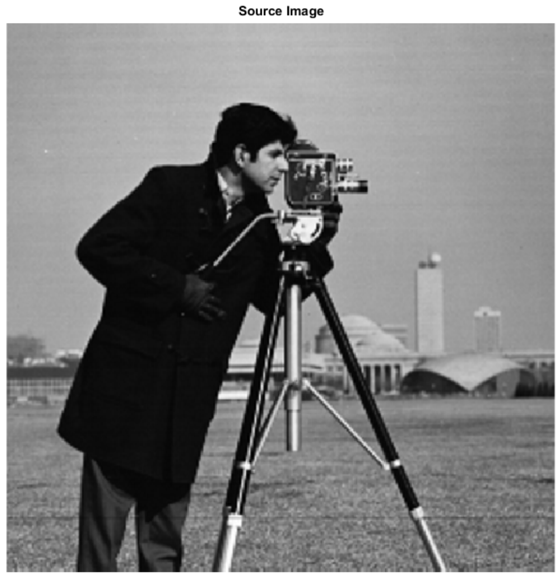
- x: το σήμα εισόδου
- Δ: το βήμα της κβάντισης
- $sign(x)$: το πρόσημο του σήματος εισόδου
- L: τα επίπεδα κβάντισης του σήματος
- R: τα απαιτούμενα bits ανά επίπεδο
- A: το πλάτος του σήματος εισόδου

Χρησιμοποιώντας την παραπάνω συνάρτηση, μπορούμε να κατασκευάσουμε τις χαρακτηριστικές γραφικές για σήμα εισόδου με πεδίο τιμών $[-255, 255]$ και για $R = 0...8$, οι οποίες φαίνονται στην πρακάτω εικόνα:

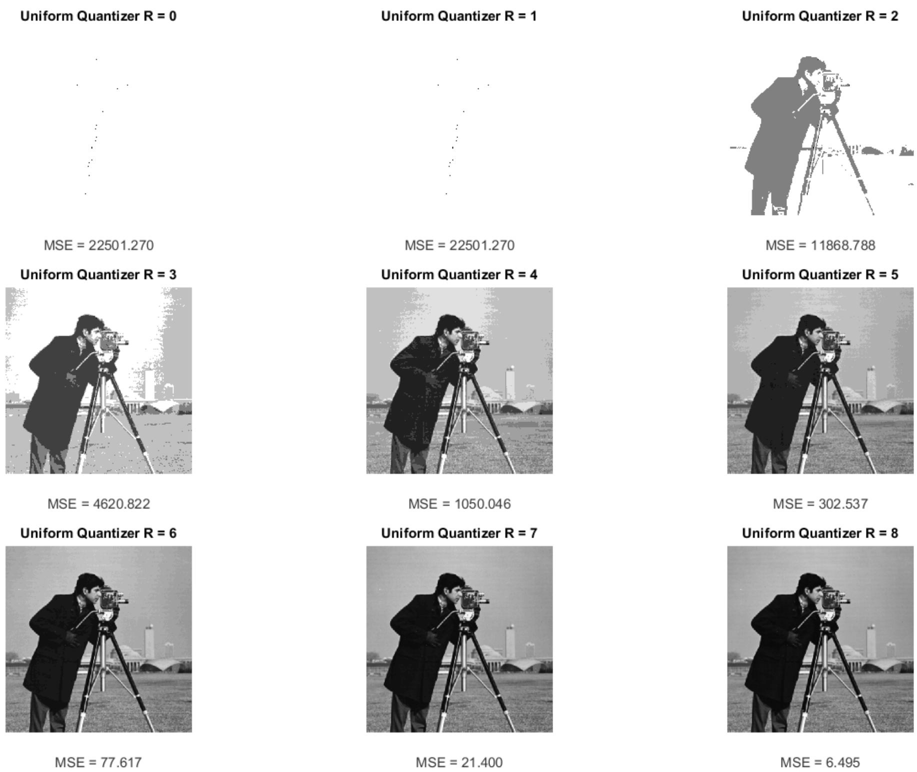


Παρατηρούμε πως όσο αυξάνεται η τιμή του R, τόσο περισσότερα τα επίπεδα κβάντισης που δημιουργούνται, οπότε αυξάνεται και η “ανάλυση” του σήματος εξόδου.

Στην συνέχεια, χρησιμοιώντας πάλι την συνάρτηση "uni_scalar" και για $R = 0...8$, κβαντίζουμε το σήμα είσοδου , το οποίο είναι η παρακάτω εικόνα:



Οπότε για κάθε τιμή του R προκύπτουν οι παρακάτω κβαντισμένες εικόνες:

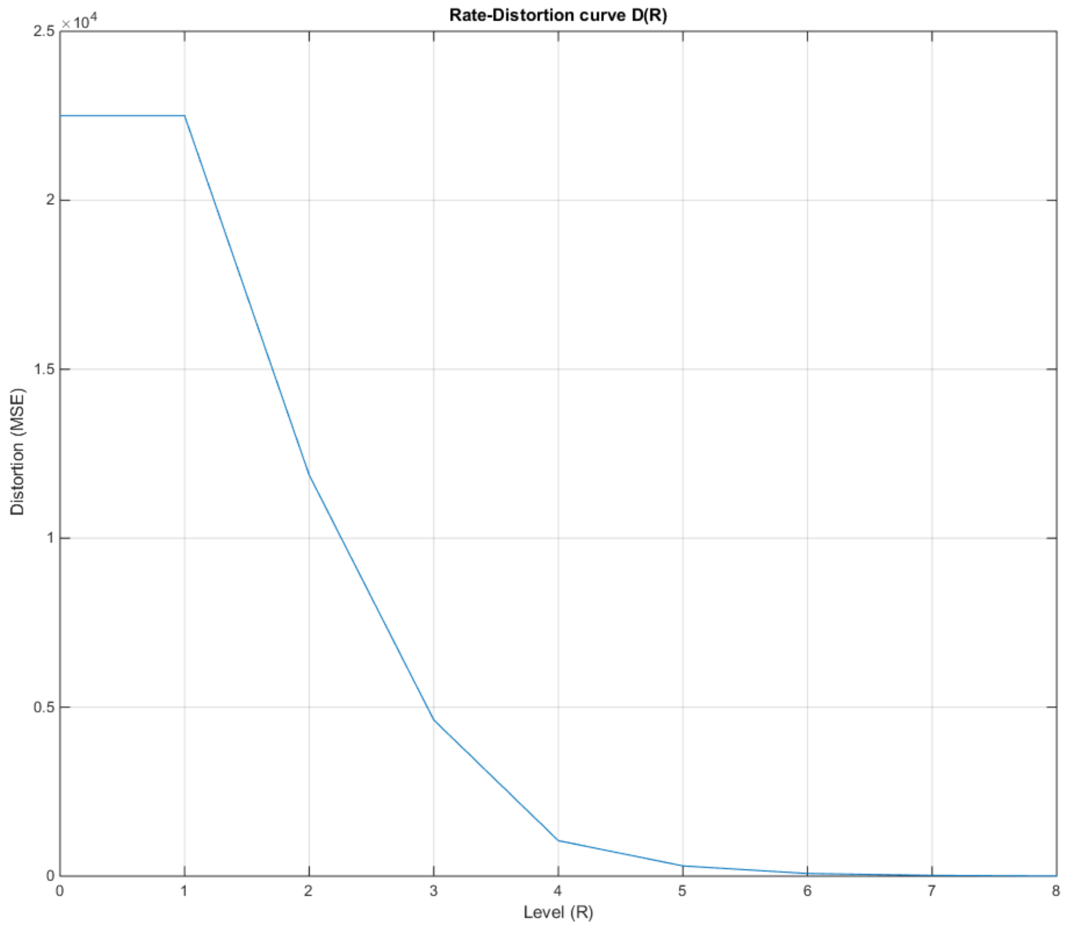


Παρατηρούμε ότι καθώς αυξάνεται το R, τόσο αυξάνονται και τα επίπεδα κβάντισης, τα οποία ταυτίζονται με τα επίπεδα φωτεινότητας κάθε pixel, οπότε αυξάνεται και η ανάλυση κάθε εικόνας.

Για την σύγκριση των παραπάνω εικόνων με την αρχική εικόνα, υπολογίστηκε το Μέσο Τετραγωνικό Σφάλμα (MSE), το οποιό δηλώνει το σφάλμα κβάντισης. Τα αποτελέσματα φαίνονται στον παρακάτω πίνακα:

R	0	1	2	3	4	5	6	7	8
MSE	22501.27	22501.27	11868.79	4620.82	1050.05	302.54	77.62	21.40	6.50

Αναπαριστώντας το MSE ως συνάρτηση του R, προκύπτει η Καμπύλη του Ρυθμού Παραμόρφωσης (rate-distortion curve), $D(R)$, η οποία απεικονίζεται στην παρακάτω γραφική:



Παρατηρείται πως η παραμόρφωση της εικόνας μειώνεται με την αύξηση της τιμής του R, όπως περιμένουμε.

Haar transform

Για την υλοποίηση του δεύτερου μέρους αρχικά, δημιουργήκε η συνάρτηση haar_transform() η οποία δέχεται σαν ορίσματα τις τιμές του πίνακα αναφοράς αλλά και το επίπεδο της μορφοποίησης που θέλουμε κάθε φορά. Ανάλογα με το επίπεδο, η συνάρτηση επεξεργάζεται το αντίστοιχο sub-band υπολογίζοντας την μέση τιμή και διαφορά για κάθε ζευγαρι από pixels και αποθηκεύοντάς την στην αντίστοιχη θέση. Το παραπάνω συμβάινει τόσο για τις σειρές όσο και για τις στήλες του πίνακα εισαγωγής. Ο παραπάνω τρόπος μεταφράζει ουσιαστικά τον τρόπο με τον οποίο πραγματοποιείται ο μετασχηματισμός $Haar$ που ορίζεται ώς $T = HFH^T$.

Με βάση τα παραπάνω έχουμε:

Σχήμα 1: Transformation of, level 1 (left) and level 2 (right)

Να σημειωθεί ότι οι παραπάνω πίνακες δεν παρουσιάζουν όλη την πληροφορία που προέκυψε από την μορφοποίηση μιας και έχει αλλάξει ο τύπος δεδομένων (χωρίς να σημαίνει ότι χάθηκε η πληροφορία).

Σε επόμενη φάση πραγματοποιήθηκε η κβάντιση του κάθε subband με την χρήση της συνάρτησης uni_scalar(), που αναλύθηκε στο πρώτο μέρος. Για το επίπεδο 1 με $R = 2$ (4 επίπεδα κβάντισης) και για το επίπεδο 2 $R = 4$ (16 επίπεδα κβάντισης) έχουμε τα παρακάτω αποτελέσματα:

Σχήμα 2: Quantization with R = 2 (left) and R = 4 (right)

Μπορούμε να παρατηρήσουμε την διαφορά στο πλήθος των επιπέδων και για τις δύο περιπτώσεις με πιο 'ευδιάκριτη' την 2η μιας και εμφανίζονται παραπάνω λεπτομέρειες.

Entropy

Ο υπολογισμός της εντροπίας έγινε βάσει της παρακάτω εξίσωσης, όπως αυτή προκύπτει από το θεώρημα του Shannon:

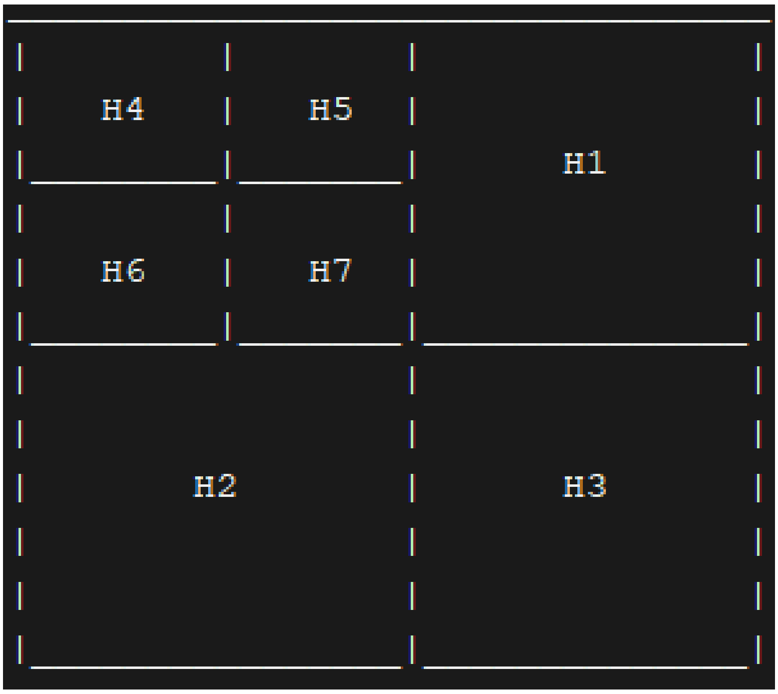
$$H = -\sum_{k=0}^{L-1} p(r_k) log_2(p(r_k))$$

$$p(r_k) = \frac{n_k}{MN}$$

όπου:

- r_k: διακριτή τυχαιά μεταβλητή στο διάστημα $[0, L-1]$ (επίπεδα κβάντισης)
- $p(r_k)$: η πιθανότητα εμφάνισης της μεταβλητής
- n_k: το πλήθος των εμφανίσεων του k-οστου επιπέδου κβάντισης σε ένα πίνακα
- $M \times N$: το μέγεθος του πίνακα

Οπότε, σύμφωνα με τα παραπάνω, κατασκευάστηκε η συνάρτηση "calc_entropy", η οποία υπολογίζει την πιθανότητα εμφάνισης κάθε επιπέδου κβάντισης καθώς και την εντροπία για ένα δεδομένο πίνακα ως είσοδο.
Δίνοντας ως είσοδο την εικόνα που πήραμε ως έξοδο από τον μετασχηματισμό Haar για $level = 2$, προκύπτουν τα εξής subbands με εντροπίες όπως φαίνεται παρακάτω:



Σχήμα 3: Subband entropies

entropy	H_1	H_2	H_3	H_4	H_5	H_6	H_7
bits/pixel	0.242	0.128	$\simeq 0$	2.495	0.139	0.272	0

Πρατηρούμε ότι η μεγαλύτερη εντροπία παρουσιάζεται για H_1, όπου υπάρχει και η περισσότερη πληροφορία της εικόνας. Οι υπόλοιπες περιοχές έχουν ελάχιστη πληροφορία και συγκεκριμένα οι περιοχές H_3, H_7 έχουν τιμές φωτεινότητας '0' σχεδόν σε όλα τα pixels.

Η συνολική εντροπία που παρουσιάζει η εικόνα είναι το άθροιασμα των επιμέρους εντροπιών για κάθε subband άρα:

$$H_{tot} = 3.277 \text{ bits/pixel}$$

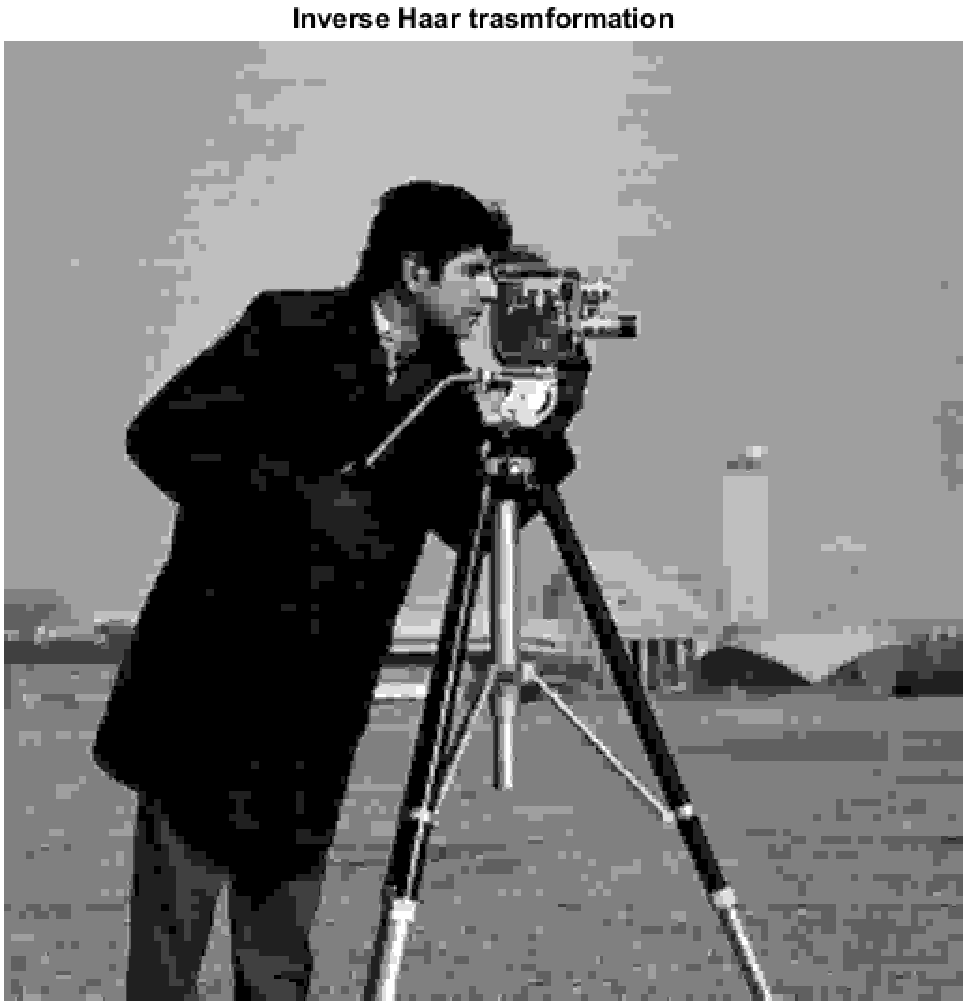
Οπότε ο Λόγος Συμπίεσης (Compression ratio) που προκύπτει είναι:

$$C = \frac{R}{H_{tot}} = 1.221$$

Inverse Transformation

Για την ανάκτηση της εικόνας δημιουργήθηκε η συνάρτηση inverse_haar_transform() η οποία όμοια παίρνει σαν ορίσματα έναν $(N*N)$ πίνακα που θέλουμε να ανακτήσουμε και το επίπεδο της μορφοποίησης. Αρχικά οριοθετείται το κομμάτι του πίνακα, από του οποίου θα αντλήσουμε τα δεδομένα, σύμφωνα με το επίπεδο που εισάγουμε. Πιο συγκεκριμένα, η πληροφορία που θέλουμε να εξάγουμε θα βρίσκεται πάντα στον πρώτο $(N/_{level} * N/_{level})$ υποπίνακα που συναντάμε. Για παράδειγμα, το πρώτο από το πρώτο ζευγάρι από pixel που θα δημιουργηθεί, θα αποτελείται από την μέση τιμή, που είναι αποθηκευμένη στο πρώτο pixel της εικόνας εισαγωγής προσθέτοντάς του, την διαφορά που έχει το $(N/_{level})/2+1$ στοιχείο. Αντίστοιχα το δεύτερο, θα αποτελείται από την μέση τιμή αφαιρώντας απο αυτήν, την παραπάνω διαφορά. Αξίζει να σημειωθεί πως για τον παραπάνω τρόπο δεν χρειάζονται περαιτέρω έλεγχοι για αρνητικές διαφορές μιας και το πρόσημο αυτών θα κανονίσει για την σωστή προσθαφαίρεση. Η παραπάνω διαδικασία γίνεται αρχικά, για τις στήλες και έπειτα για τις γραμμές.

Αν εφαρμόσουμε τον παραπάνω αλγόριθμο για την κβαντισμένη εικόνα (level 2) έχουμε το παρακάτω αποτέλεσμα:



Συγκρίνοντας της τελική εικόνα με την αρχική παρατηρούμε μέσο τετραγωνικό σφάλμα: $MSE \simeq 154$ και $PSNR \simeq -21.89$ (η επίδραση του κβάντισης είναι μεγάλη εξού και το αρνητικό πρόσημο). Δεδομένου ότι οι παραπάνω αλγόριθμοι λειτουργούν ορθά, το σφάλμα οφείλεται στην κβάντιση που δέχθηκε η εικόνα.

Αν εφαρμοστεί ο αντίστροφος μετασχηματισμός Haar στην αρχική εικόνα, παρατηρείται μηδαμινό σφάλμα μιας και γίνεται πλήρη ανάκτηση εικόνας ($MSE = PSNR = 0$).

Να σημειωθεί ότι για να προσεγκίσουμε μια πλήρη ανάκτηση της εικόνας που αποστάλθηκε, θα πρέπει να εφαρμώσουμε αποκωδικοποίηση της εντροπίας, ώστε, στην συνέχεια να εφαρμιστεί η αντίστραφη διαδηκασία της κβάντησης. Στην συνέχεια θα εφαρμόζαμε την αντίστροφη διαδηκασία του μετασχηματισμού Haar.