

Ψηφιακή Επεξεργασία Εικόνας – Project

Σοφία Καφρίτσα Γεωργαντά, 2016030136

Part A

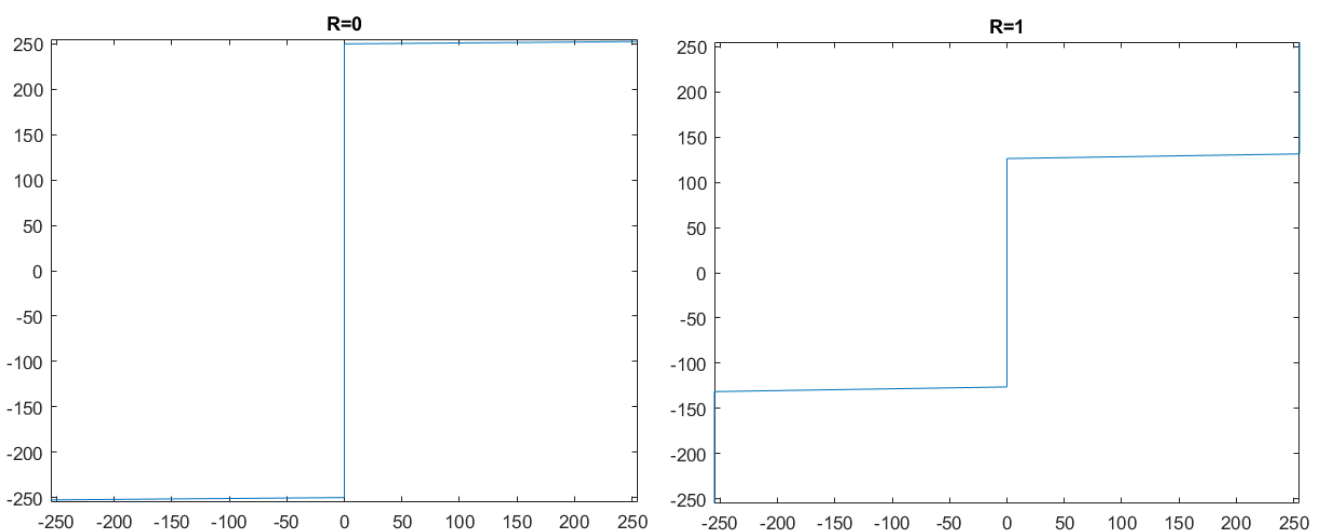
Στο 1^ο μέρος του project, υλοποίησα έναν uniform scalar quantizer, οποίος δουλεύει σύμφωνα με τη συνάρτηση:

$$Q(x) = \Delta \times \text{sgn}(x) \left\lfloor \frac{|x|}{\Delta} + \frac{1}{2} \right\rfloor$$

Αρχικά, δούλεψα με έναν πίνακα με τιμές από το -255 ως το 255. Πρώτα απ' όλα, ορίζω τα επίπεδα και το βήμα κβάντισης, κι έπειτα, καλώ τη συνάρτηση `uni_scalar` με όρισμα τον πίνακα και το βήμα κβάντισης Δ , για R από το 0 μέχρι το 8. Παράλληλα, υπολογίζω και την παραμόρφωση (distortion), για κάθε διαφορετικό R , όπως φαίνεται παρακάτω:

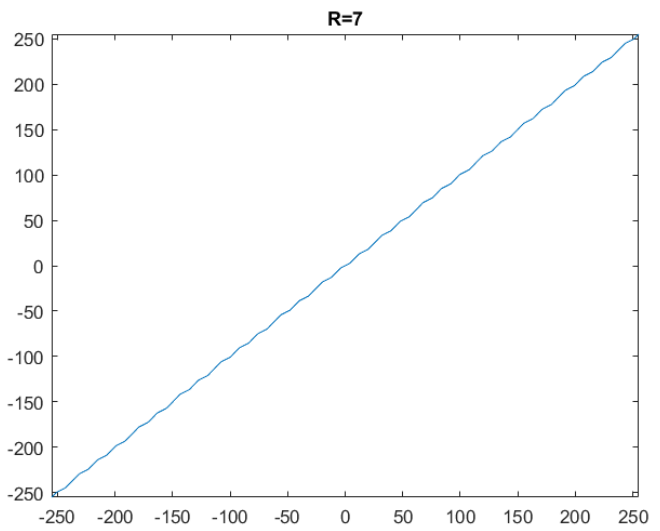
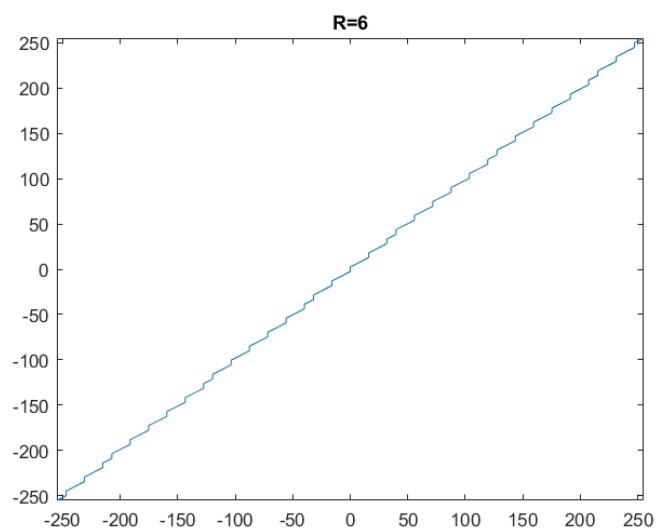
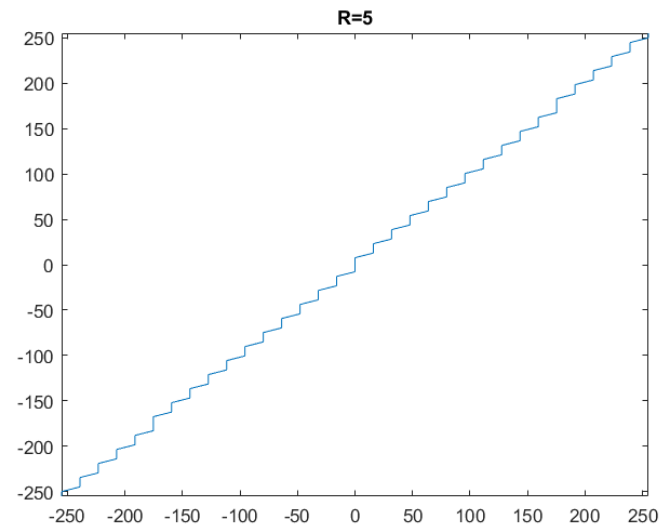
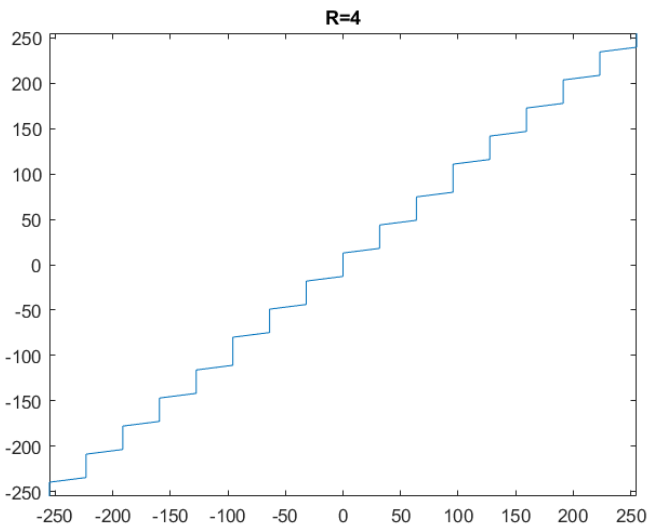
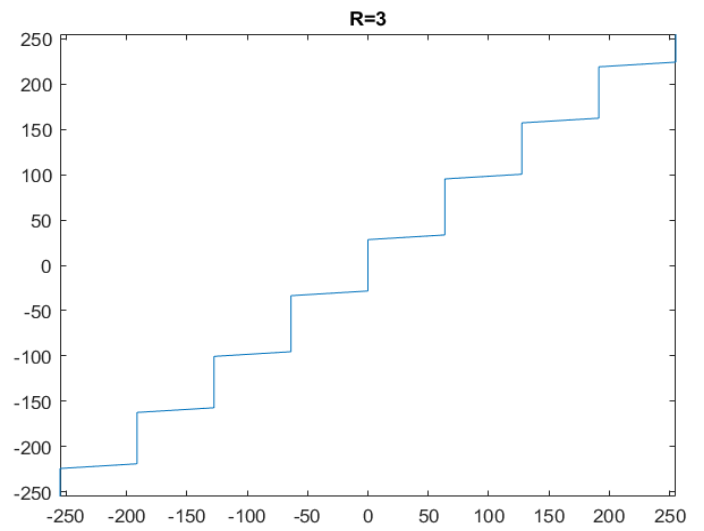
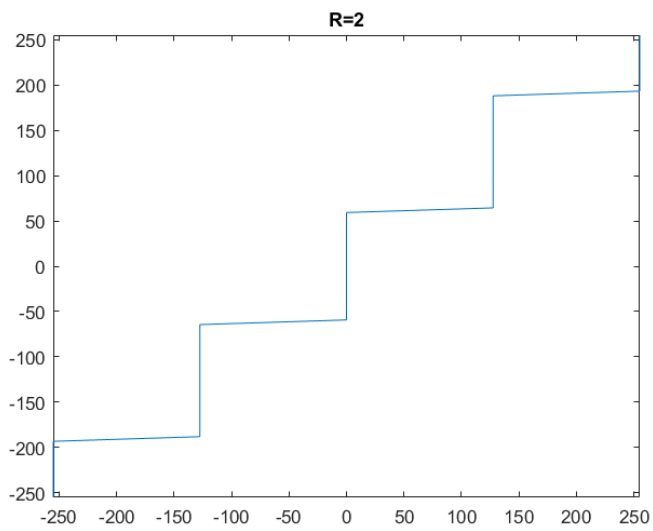
```
x = linspace(-A,A);  
  
%----- for vector as input signal  
for R=0:8  
    L = 2^R;  
    D = (2*A)/L;  
    quantized = uni_scalar(x,D);  
    distortion(R+1) = immse(x,quantized)  
    figure(),plot(quantized,x);  
    title(sprintf('R=%d',R));  
    axis([-A A -A A]);  
end
```

Οι χαρακτηριστικές συναρτήσεις του κβαντιστή από $R=0$ έως $R=8$, φαίνονται παρακάτω:



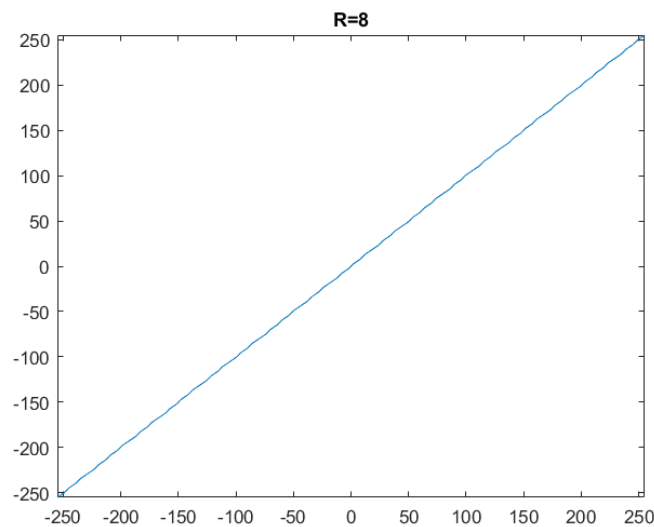
Ψηφιακή Επεξεργασία Εικόνας – Project

Σοφία Καφρίτσα Γεωργαντά, 2016030136

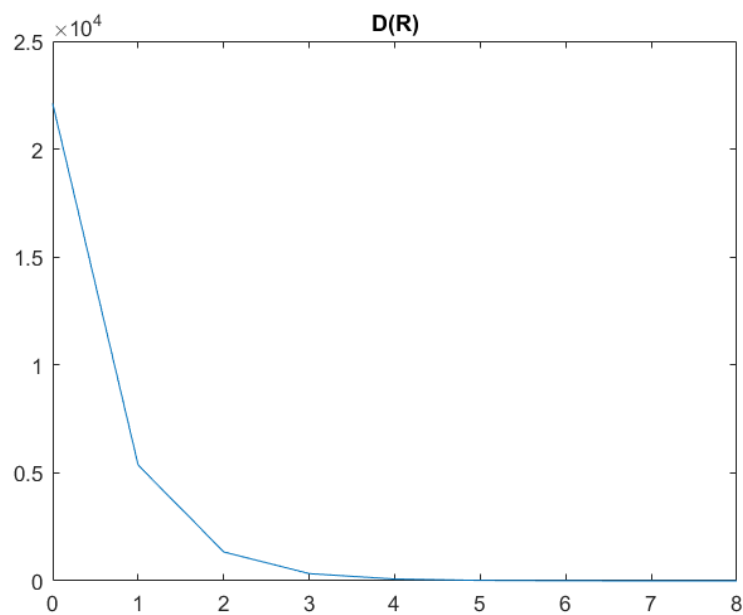


Ψηφιακή Επεξεργασία Εικόνας – Project

Σοφία Καφρίτσα Γεωργαντά, 2016030136



Η καμπύλη $D(R)$ φαίνεται παρακάτω:

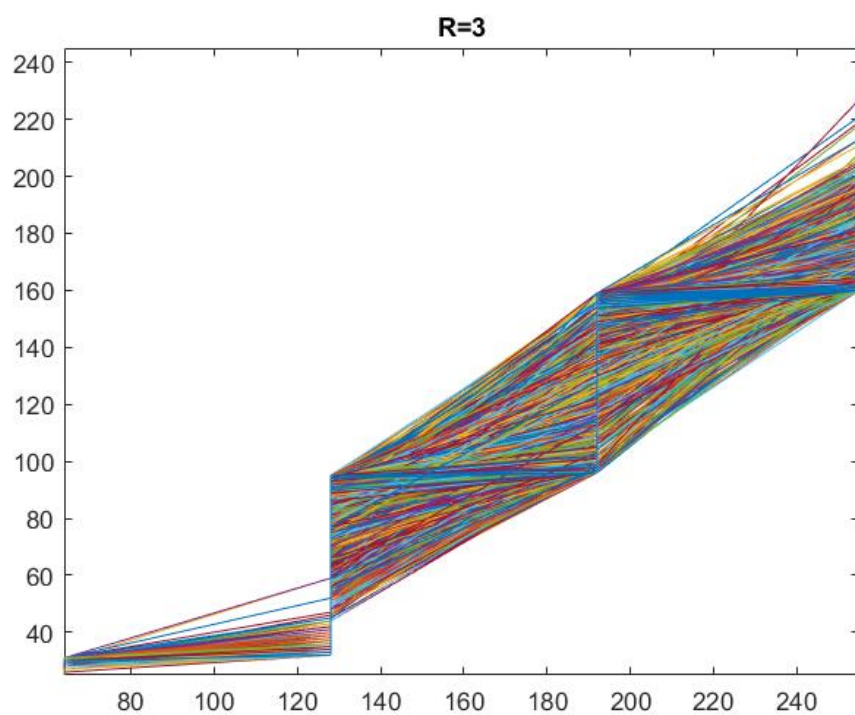
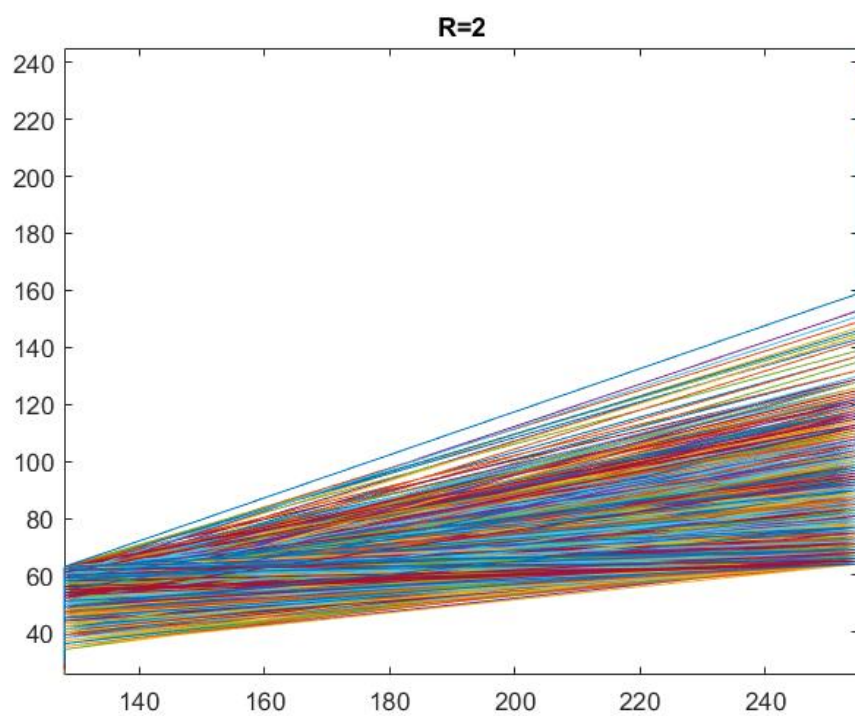


Από τις τιμές της distortion D , παρατηρούμε ότι έχει μέγιστη τιμή 2.21×10^4 για $R=0$ και ελάχιστη τιμή 0.327 για $R=8$. Αυτό συμβαίνει, γιατί όσο περισσότερη πληροφορία έχουμε, τόσο καλύτερη ανάλυση παίρνουμε, καθώς έχουμε στη διάθεσή μας περισσότερα bits.

Όσον αφορά την κβάντιση της εικόνας, η διαδικασία είναι παρόμοια, αλλά αλλάζει το R και πλέον παίρνει τιμές από 2 έως 9, ώστε να είναι αποδεκτό το αποτέλεσμα. Σημειώνουμε ότι η εικόνα είναι 8bits/pixel, άρα κβαντίζεται με $L=2^8=256$ levels. Τα αποτελέσματα από την κβάντιση της εικόνας φαίνονται παρακάτω:

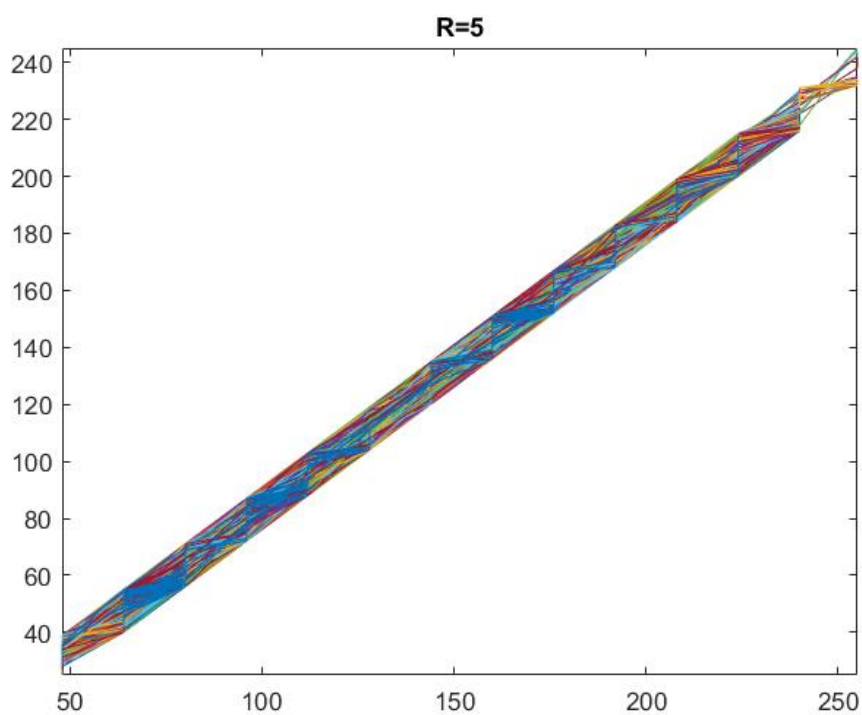
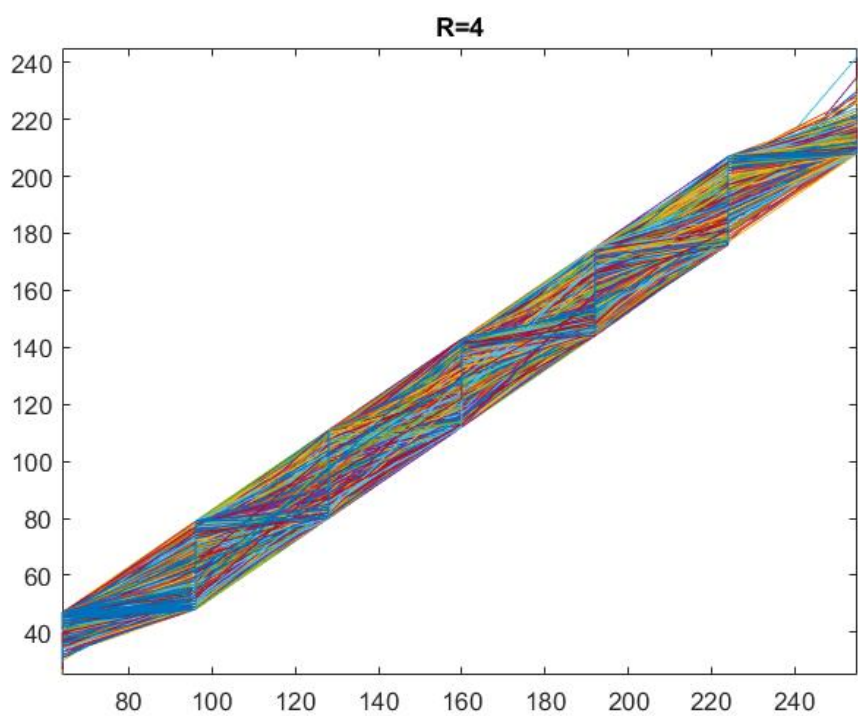
Ψηφιακή Επεξεργασία Εικόνας – Project

Σοφία Καφρίτσα Γεωργαντά, 2016030136



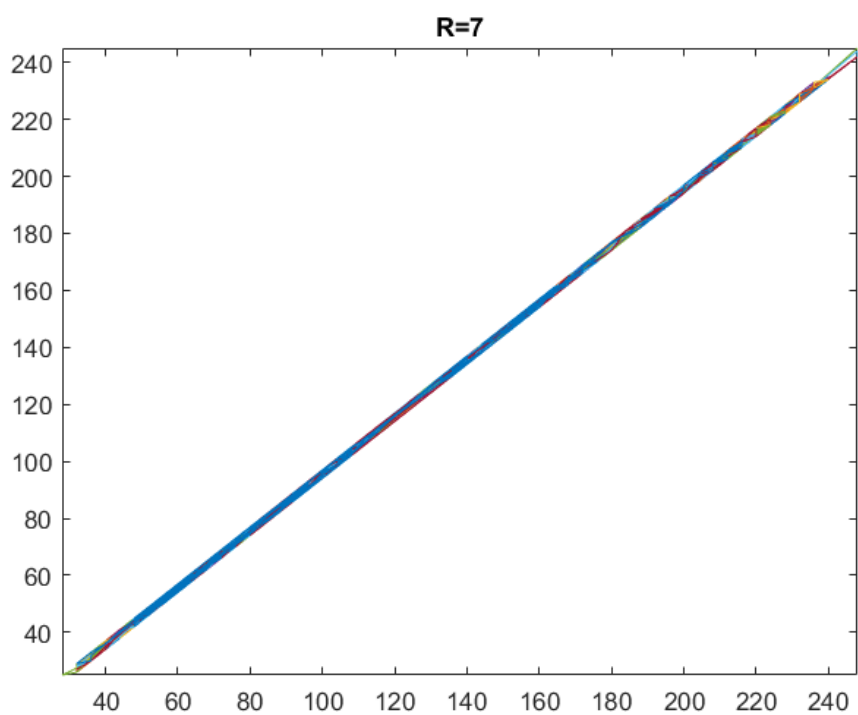
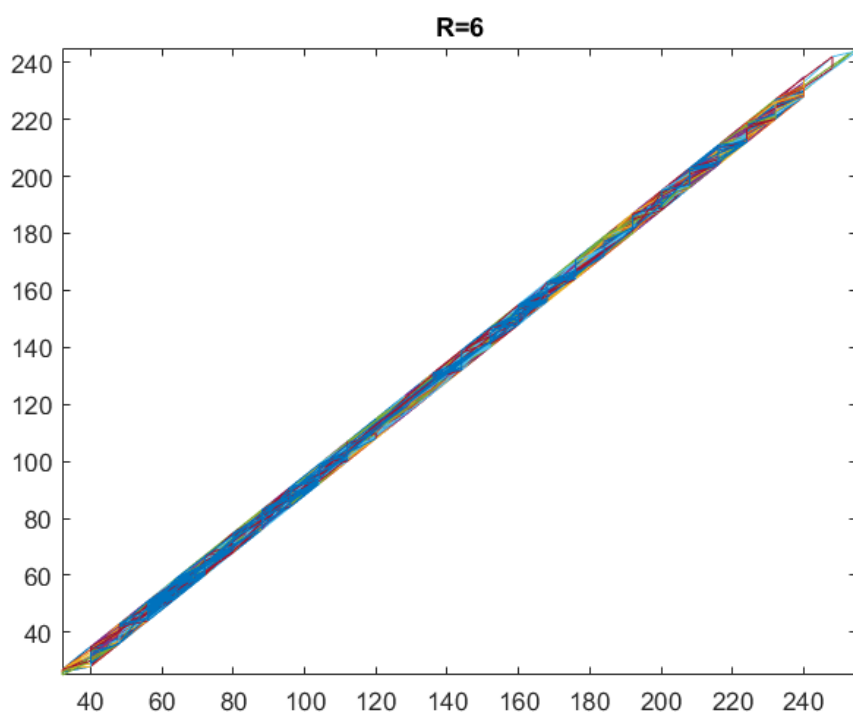
Ψηφιακή Επεξεργασία Εικόνας – Project

Σοφία Καφρίτσα Γεωργαντά, 2016030136



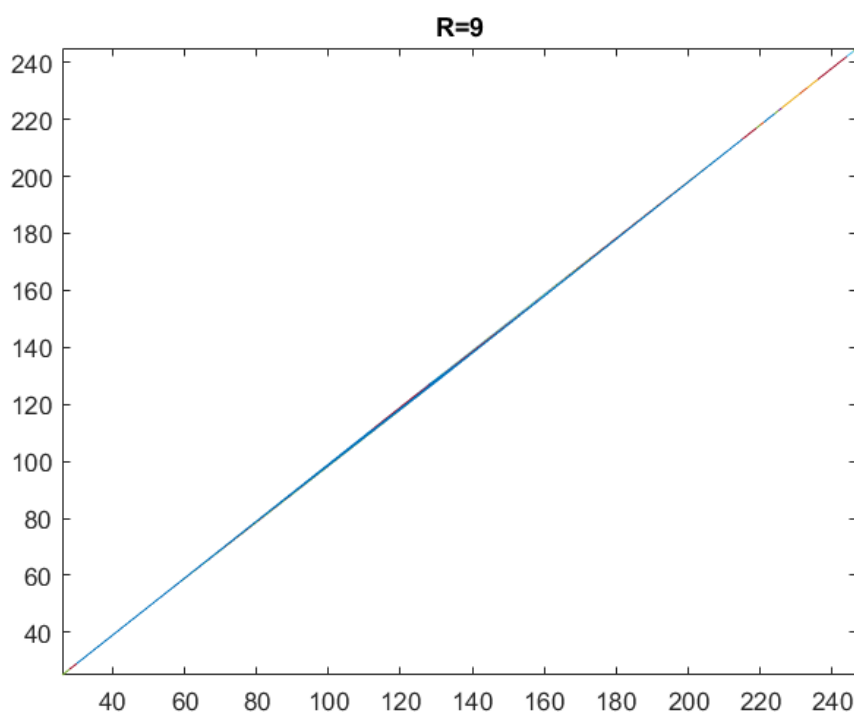
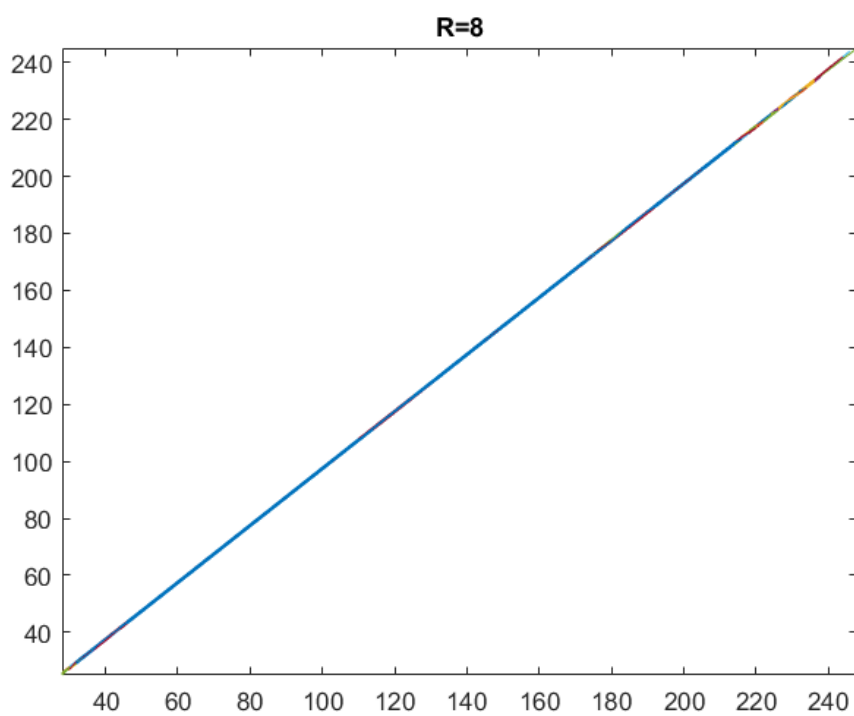
Ψηφιακή Επεξεργασία Εικόνας – Project

Σοφία Καφρίτσα Γεωργαντά, 2016030136



Ψηφιακή Επεξεργασία Εικόνας – Project

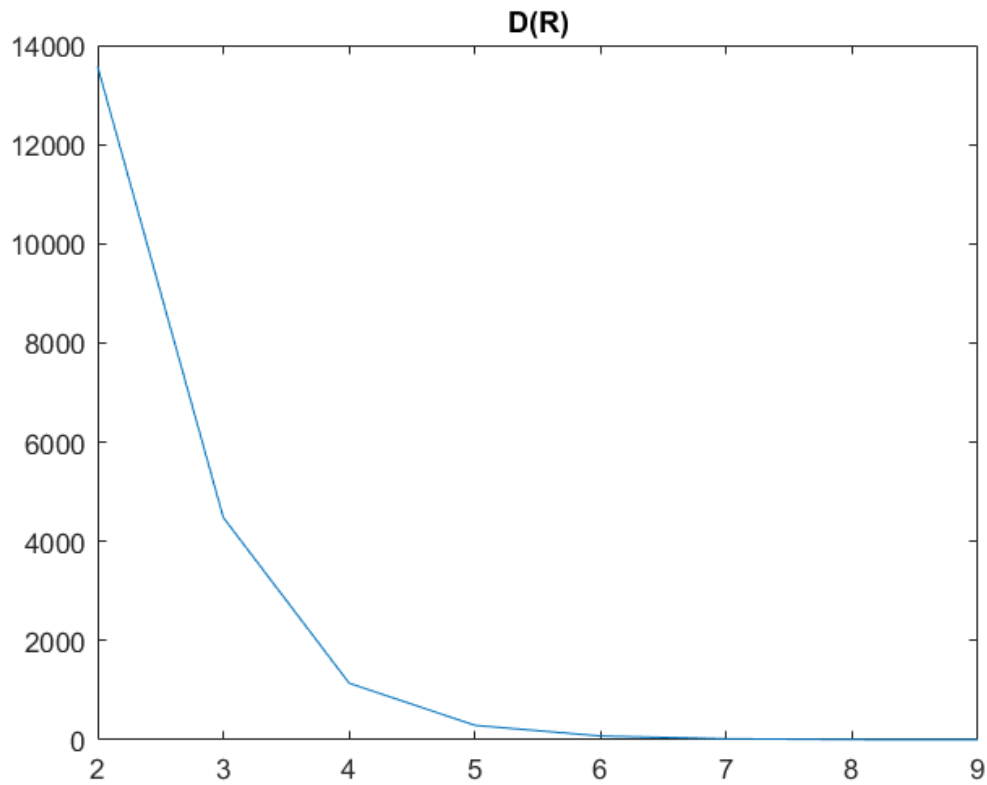
Σοφία Καφρίτσα Γεωργαντά, 2016030136



Ψηφιακή Επεξεργασία Εικόνας – Project

Σοφία Καφρίτσα Γεωργαντά, 2016030136

Όσον αφορά την αλλοίωση, παρατηρούμε και πάλι ότι όσο αυξάνεται ο αριθμός των bits, τόσο μειώνεται η αλλοίωση. Επίσης, σημειώνουμε μέγιστη κι ελάχιστη τιμή $1.35e+04$ και 2.53 αντίστοιχα. Τα αποτελέσματα αυτά συμπίπτουν και πάλι με τη θεωρία, καθώς όσο μεγαλύτερο βήμα κβάντισης έχουμε, τόσο λιγότερο χώρο χρειαζόμαστε για να αποθηκεύσουμε την πληροφορία και όσο περισσότερη πληροφορία έχουμε, τόσο καλύτερη ανάλυση παίρνουμε, καθώς έχουμε στη διάθεσή μας περισσότερα bits.



Ψηφιακή Επεξεργασία Εικόνας – Project

Σοφία Καφρίτσα Γεωργαντά, 2016030136

Part B

Στο 2^ο μέρος, ασχοληθήκαμε με την επεξεργασία βίντεο. Αρχικά, φόρτωσα την εικόνα με τη συνάρτηση `VideoReader` και μέσω από τις πληροφορίες που προσφέρονται από τη `MATLAB` για το video, παρατήρησα τα εξής δεδομένα:

Αριθμός frames: 141

Framerate του βίντεο: 30 frames/sec

Ανάλυση του κάθε frame: 320x240

Συνολική διάρκεια video: 4.7 sec

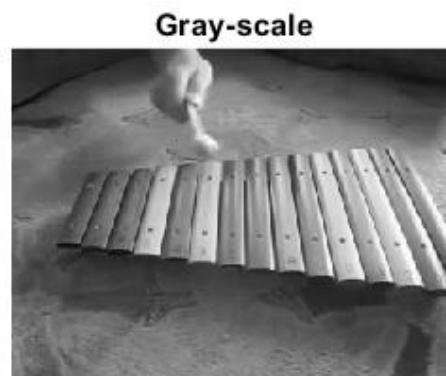
Στη συνέχεια, ακολουθεί ένα script που εκτυπώνει τα frames του video με το δεδομένο framerate. Συνολικά εκτυπώνονται και τα 141 frames.

```
%read frames
while hasFrame(vd)
    vidFrame = readFrame(vd);
    figure,imshow(vidFrame)
    pause(1/frameRate);
end
```

Έπειτα, εξάγω το 50^ο frame, και το μετατρέπω σε εικόνα gray-scale.

```
%extractin 50th frame
vd = VideoReader('xylophone3.mp4')
frame = read(vd,50);
framegray = rgb2gray(frame);
```

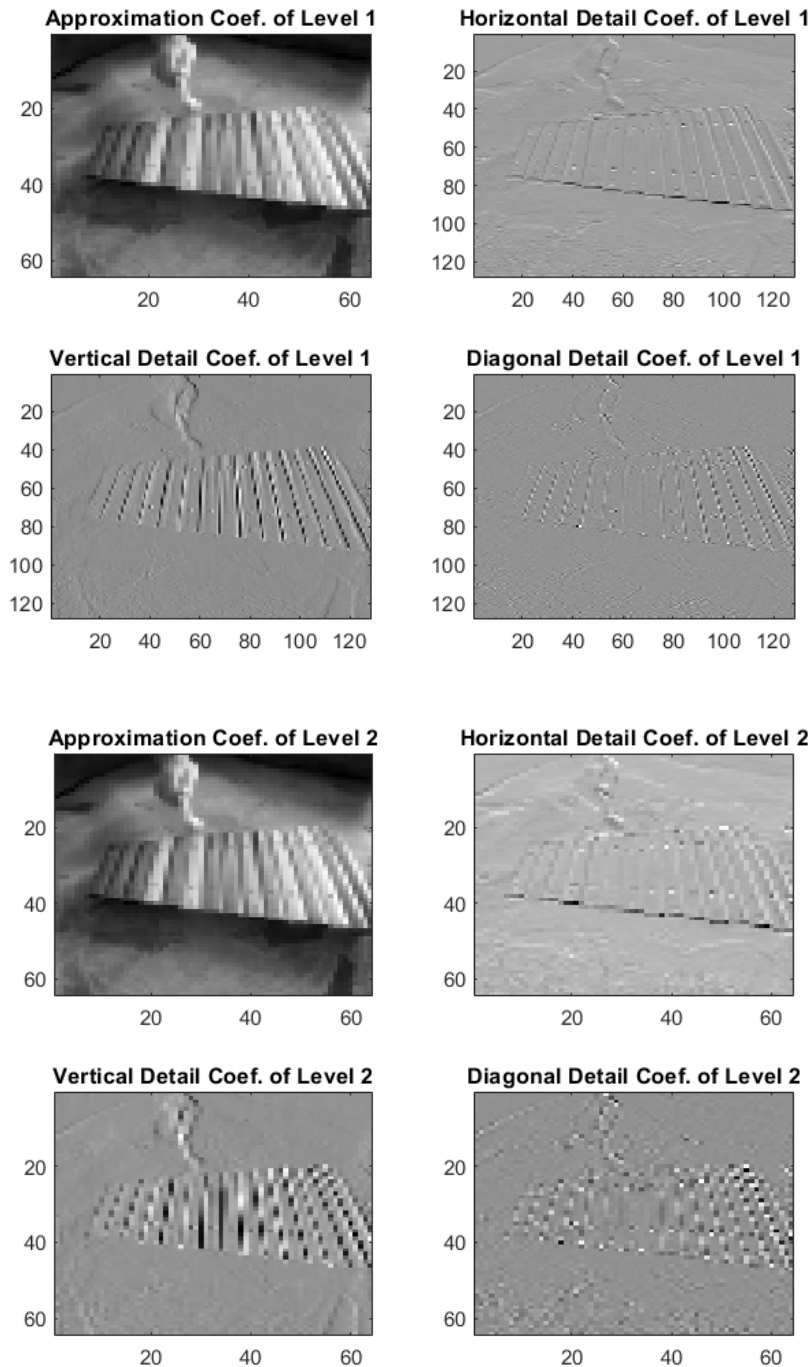
Τέλος, εκτυπώνω το 50^ο frame σε RGB και gray-scale version, όπως φαίνεται παρακάτω:



Part C

Στο 3^ο μέρος, άρχισα ως εξής:

Εξάγω το 50^ο frame από το βίντεο και ορίζω το μέγεθός του σε 256x256. Στη συνέχεια, με τη συνάρτηση *ihaar2*, εφαρμόζω μετασχηματισμό Haar για 2 decomposition levels. Τα αποτελέσματα φαίνονται παρακάτω:



Ψηφιακή Επεξεργασία Εικόνας – Project

Σοφία Καφρίτσα Γεωργαντά, 2016030136

Στη συνέχεια, με βήμα κβάντισης $R=4$, χρησιμοποιώντας τον κβαντιστή του part A, κβαντίζω κάθε subband ως εξής:

```
A=255;
R=4;
L = 2^R;
D = (2*A)/L;
sub_h1_R4 = uni_scalar(h{1},D);
sub_v1_R4 = uni_scalar(v{1},D);
sub_d1_R4 = uni_scalar(d{1},D);
sub_h2_R4 = uni_scalar(h{2},D);
sub_v2_R4 = uni_scalar(v{2},D);
sub_d2_R4 = uni_scalar(d{2},D);
```

Έπειτα, υπολογίζω την εντροπία κάθε subband και τη συνολική εντροπία ως εξής:

```
e_h1_R4 = entropy(sub_h1_R4);
e_v1_R4 = entropy(sub_v1_R4);
e_d1_R4 = entropy(sub_d1_R4);
e_h2_R4 = entropy(sub_h2_R4);
e_v2_R4 = entropy(sub_v2_R4);
e_d2_R4 = entropy(sub_d2_R4);

e_total_R4 == e_h1_R4+e_v1_R4+e_d1_R4+e_h2_R4+e_v2_R4+e_d2_R4
```

Τέλος, ανακατασκευάζω την εικόνα, χρησιμοποιώντας τα subbands και την συνάρτηση *ihaart2()*, και υπολογίζω και την αλλοίωση μέσω του MSE και του PSNR:

```
h_1 = {sub_h1_R4,sub_h2_R4};
v_1 = {sub_v1_R4,sub_v2_R4};
d_1 = {sub_d1_R4,sub_d2_R4};

Irec = ihaart2(a,h_1,v_1,d_1);
psnr_R4 == psnr(framegray,uint8(Irec))
mse_R4==immse(framegray,uint8(Irec))
```

Επαναλαμβάνω τη διαδικασία, αλλά αυτή τη φορά ορίζω βήμα κβάντισης $R=3$ για το 1^ο επίπεδο και $R=5$ για το 2^ο επίπεδο. Τα αριθμητικά αποτελέσματα από τις δύο διαδικασίες παρατίθενται παρακάτω:

1^η μέθοδος κβάντισης ($R=4$):

Συνολική Εντροπία: 1.95
MSE: 24.16
PSNR: 34.3

Ψηφιακή Επεξεργασία Εικόνας – Project

Σοφία Καφρίτσα Γεωργαντά, 2016030136

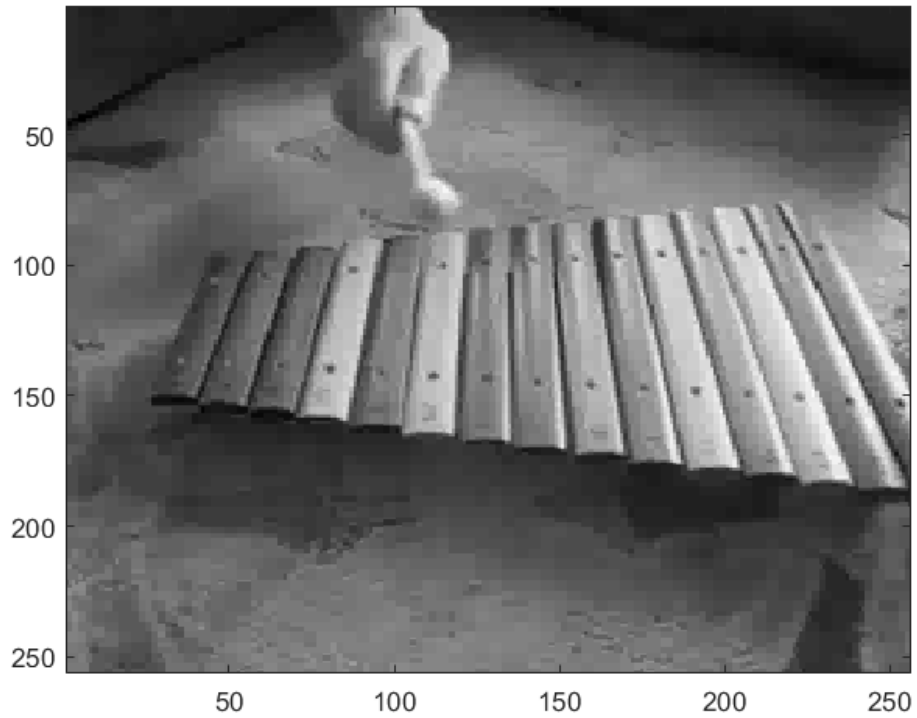
2^η μέθοδος κβάντισης (R=3, R=5):

Συνολική Εντροπία: 2.29

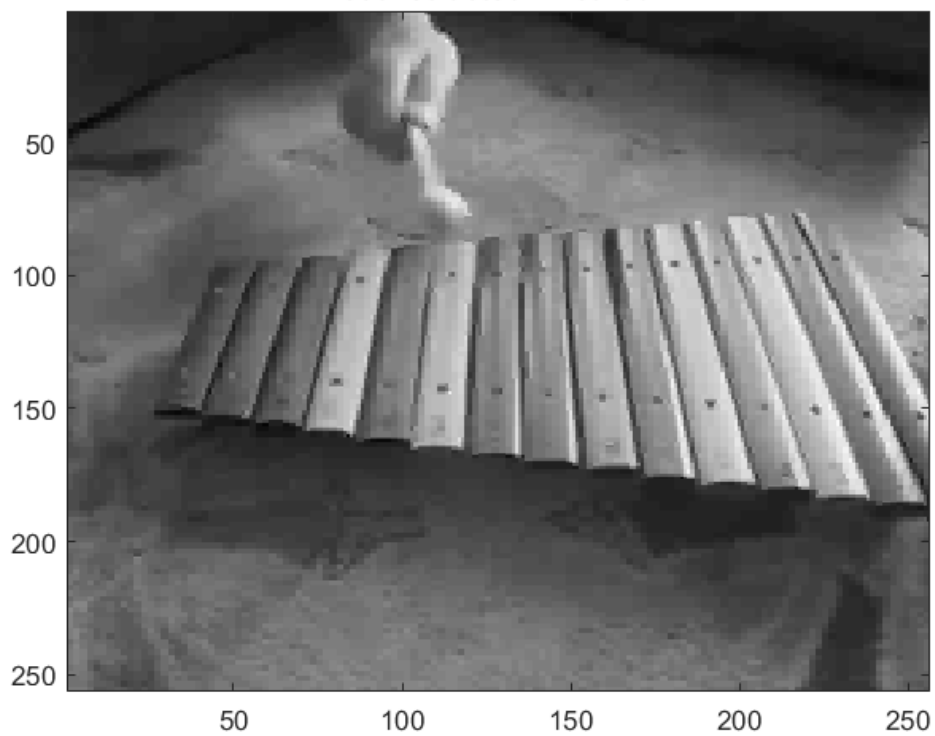
MSE: 31.61

PSNR: 33.13

Reconstructed - Method 1



Reconstructed - Method 2



Ψηφιακή Επεξεργασία Εικόνας – Project

Σοφία Καφρίτσα Γεωργαντά, 2016030136

Αξίζει να σημειώσουμε ότι, μικρότερο R σημαίνει μικρότερη ακρίβεια δειγματοληψίας, δηλαδή λιγότερα bits με τα οποία αναπαριστούμε την αρχική μας εικόνα. Η εντροπία εκφράζει την αβεβαιότητα και, συγκεκριμένα, στο δικό μας project, τον ελάχιστο αριθμό bits που απαιτούνται για να εκφραστεί μια τιμή pixel της εικόνας.

Επομένως, λιγότερες τιμές με μικρό R → λιγότερη αβεβαιότητα για την τιμή που μπορεί να πάρει ένα pixel → μικρότερη εντροπία.

Όπως παρατηρούμε και από τις τιμές της εντροπίας για τα διάφορα subbands, επιβεβαιώνεται το παραπάνω, καθώς όπως βλέπουμε και στην παρακάτω εικόνα, για αντίστοιχο subband οι τιμές της εντροπίας είναι μικρότερες όταν έχουμε μικρότερο R.

e_d1_R3	0
e_d1_R4	0.0268
e_d2_R4	0.2730
e_d2_R5	0.4724
e_h1_R3	0.0220
e_h1_R4	0.1280
e_h2_R4	0.5148
e_h2_R5	0.8387
e_v1_R3	0.1364
e_v1_R4	0.3404
e_v2_R4	0.6636
e_v2_R5	0.8245

Όσον αφορά το Compression Ratio, αν υπολογίσουμε την εντροπία Shannon στην εικόνα μας, βλέπουμε ότι είναι περίπου 7.56≈8. Επίσης, η συνολική εντροπία για τον 1^ο τρόπο κβάντισης ισούται με 1.95, ενώ με τον 2^ο τρόπο κβάντισης ισούται με 2.29. Επομένως, για το ρυθμό συμπίεσης έχουμε:

$$C_{R4} = \frac{\text{αριθμός αρχικών bits}}{\text{αριθμός bits μετα την συμπίεση}} = \frac{256*256*8}{256*256*1.95} = 4.1$$

$$C_{R3,5} = \frac{\text{αριθμός αρχικών bits}}{\text{αριθμός bits μετα την συμπίεση}} = \frac{256*256*8}{256*256*2.29} = 3.5$$

και για την περιττή πληροφορία έχουμε:

$$R_{R4} = 1 - \frac{1}{C_{R4}} = 76,6\%$$

$$R_{R3,5} = 1 - \frac{1}{C_{R3,5}} = 71.4\%$$

Συμπέρασμα:

Όπως παρατηρούμε απ' τα δεδομένα, η 1^η μέθοδος με ίδιο βήμα κβάντισης εμφανίζει μικρότερη παραμόρφωση συγκριτικά με τη 2^η μέθοδο που χρησιμοποιούμε διαφορετικό βήμα κβάντισης για κάθε επίπεδο, όπως μπορούμε να παρατηρήσουμε από το MSE και το PSNR. Όμως η 1^η μέθοδος εμφανίζει και περισσότερη περιττή πληροφορία. Δηλαδή, παρ' όλο που η 2^η μέθοδος εφαρμόζει και κβάντιση με R=5, το οποίο είναι μεγαλύτερο από το R=4 της 1^{ης} μεθόδου, είναι εν τέλει πιο αποδοτική

Ψηφιακή Επεξεργασία Εικόνας – Project

Σοφία Καφρίτσα Γεωργαντά, 2016030136

από την 1^η μέθοδο όσον αφορά την πληροφορία, καθώς συνυπολογίζοντας και την κβάντιση με $R=3$, καταλήγει σε λιγότερο περιττή πληροφορία.

Συνεπώς, μπορεί με την 1^η μέθοδο να έχουμε λιγότερη παραμόρφωση, όμως με τη 2^η μέθοδο έχουμε λιγότερη περιττή πληροφορία, και για το λόγο αυτό η τελευταία μπορεί να θεωρηθεί περισσότερο κατάλληλη και αποδοτική σε κάποιες περιπτώσεις. Μάλιστα, πολλές φορές το μάτι μπορεί να «ξεγελαστεί» και παρ' όλο που η εικόνα της 1^{ης} μεθόδου, όπως βλέπουμε παραπάνω, θεωρητικά, έχει λιγότερη παραμόρφωση, η 2^η εικόνα φαίνεται πιο «εξομαλυμένη» και καλαίσθητη.