

ΠΟΛΥΤΕΧΝΕΙΟ ΚΡΗΤΗΣ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ
ΥΠΟΛΟΓΙΣΤΩΝ



ΠΑΡΟΥΣΙΑΣΗ ΔΙΠΛΩΜΑΤΙΚΗΣ ΕΡΓΑΣΙΑΣ

ΑΝΑΣΤΑΣΙΟΥ ΠΑΝΑΓΙΩΤΗ ΠΑΠΑΝΑΣΤΑΣΙΟΥ

με θέμα

**Η Μελέτη ενός Νέου Κβαντιστή Εμπνευσμένου από τις Νευροεπιστήμες
σε μια Ολοκληρωμένη Αρχιτεκτονική Συμπίεσης Εικόνας**

The study of a novel neuro-inspired quantizer in an integrated image compression
architecture.

Εξεταστική Επιτροπή :

Καθηγητής Ζερβάκης Μιχαήλ (Επιβλέπων)

Καθηγητής Πετράκης Ευριπίδης (ΔΕΠ)

Δούτση Ευφροσύνη (Συνεργαζόμενη Ερευνήτρια)

22/12/2022

Περίληψη

Σε αυτήν τη διπλωματική εργασία μελετάται η απόδοση ενός κβαντιστή που προτάθηκε το 2021, ονόματι DUAL-SIM, εντός μιας ολοκληρωμένης αρχιτεκτονικής συμπίεσης εικόνας. Αυτός ο κβαντιστής εμπνέεται από τη δομή και τις λειτουργίες του ανθρώπινου οπτικού συστήματος και των βιολογικών νευρώνων οι οποίοι έχουν την ικανότητα να μετατρέπουν και να κωδικοποιούν ένα οπτικό ερέθισμα σε μια ακολουθία διακριτών γεγονότων ονόματι ακολουθία αιχμών βάσει των οποίων γίνεται η κβάντιση μιας εικόνας. Η ακολουθία αυτή των αιχμών επιτρέπει να γίνεται καταγραφή ενός δυναμικού σήματος, δηλαδή σήματος χρονικά μεταβαλλόμενου, και να καταγράφει μόνο την πληροφορία που αλλάζει, δίνοντας λιγότερη σημασία σε ό,τι είναι στατικό. Ο κβαντιστής λοιπόν, αξιοποιεί το χαρακτηριστικό του οπτικού συστήματος να ξεχωρίζει τη στατική από τη δυναμική πληροφορία και να καταγράφει μόνο τα νέα γεγονότα που προκύπτουν.

Στην εργασία αυτή, ο εμπνευσμένος από τις νευροεπιστήμες κβαντιστής συνδυάζεται με σύγχρονα φίλτρα, όπως τον διακριτό μετασχηματισμό συνημιτόνου (DCT) και τον διακριτό μετασχηματισμό κυματιδίων (DWT), προκειμένου να μελετηθεί τόσο το αντιστάθμισμα μεταξύ της απόδοσης και της παραμόρφωσης που δημιουργεί η προτεινόμενη αρχιτεκτονική συμπίεσης εικόνας όσο και η συμπεριφορά του κβαντιστή αν αυτός εισαχθεί σε ένα συμβατικό περιβάλλον αρχιτεκτονικής συμπίεσης εικόνων, όπως αυτού των αλγορίθμων JPEG και JPEG2000.

Abstract

In this thesis we study the performance of a of a quantizer proposed in 2021, called DUAL-SIM, within an integrated image compression architecture. This quantizer is motivated by the structure and functions of the human visual system and biological neurons which have the ability to convert and encode a visual stimulus into a sequence of discrete events called a spike train by which an image is quantized. This sequence of spikes allows a dynamic signal to be recorded, that is, a time-varying signal, and to record only the information that changes, giving less importance to what is static. The quantizer therefore exploits the characteristic of the visual system to distinguish static from dynamic information and to record only the events that emerge.

In this work, the neuro-inspired quantizer is combined with state-of-the-art filters, such as the Discrete Cosine Transform (DCT) and the Discrete Wavelet Transform (DWT), in order to study both the trade-off between performance and distortion created by the proposed image compression architecture and the behavior of the quantizer if it is introduced into a conventional image compression architecture environment, such as that of the JPEG and JPEG2000 algorithms.

Ευχαριστίες

Ξεκινώντας, θα ήθελα να ευχαριστήσω από καρδιάς όλους τους ανθρώπους που με βοήθησαν στην εκπόνηση της συγκεκριμένης διπλωματικής εργασίας. Πρώτα απ'όλα, οφείλω να εκφράσω ένα μεγάλο ευχαριστώ στον κύριο Ζερβάκη Μιχαήλ για την ανάθεση αυτής της διπλωματικής εργασίας. Ακολούθως, θα ήθελα να ευχαριστήσω εξίσου θερμά την κυρία Δούτση Ευφροσύνη για την εμπιστοσύνη, τη βοήθεια και την απεριόριστη συνεργασία και στήριξη που έλαβα από αυτήν όλο αυτό το διάστημα, προκειμένου να επιτευχθεί το τελικό επιθυμητό αποτέλεσμα. Η καθοδήγηση, οι συμβουλές και τα καίρια σχόλιά της ήταν καθοριστικά για την αποπεράτωση αυτής εδώ της εργασίας. Τέλος, θα ήθελα να ευχαριστήσω την οικογένειά μου και τους φίλους μου, που με στήριζαν και μου συμπαραστάθηκαν σε όλη τη διάρκεια διεξαγωγής της διπλωματικής μου εργασίας, ενθαρύνοντάς με σε κάθε δυσκολία

Πίνακας περιεχομένων

<i>Περίληψη</i>	3
<i>Abstract</i>	3
<i>Ευχαριστίες</i>	5
<i>Κατάλογος Εικόνων</i>	8
<i>Κεφάλαιο 1ο Εισαγωγή Στις Τεχνικές Συμπίεσης</i>	13
1.1. Γιατί είναι σημαντική η συμπίεση ;	13
1.1.1. Μη Απωλεστική Συμπίεση	14
1.1.2. Απωλεστική Συμπίεση	15
1.2 Θεμελιώδης Αρχή Της Κωδικοποίησης	16
1.3 Πρότυπα Συμπίεσης Εικόνας.....	18
1.3.1 JPEG	18
1.3.1.1 Μετατροπή του Χρωματικού Χώρου	21
1.3.1.2 Υποδειγματοληψία Χρωματικότητας	22
1.3.1.3 DCT	24
1.3.1.4 Κβάντιση	28
1.3.1.4.1 Ομοιόμορφη και Μη Ομοιόμορφη Κβάντιση.....	29
1.3.1.5 Συμπίεση Εντροπίας	33
1.3.2 JPEG2000	33
1.3.2.1 Κατάτμηση	34
1.3.2.2 DWT	34
1.3.2.3 Κβάντιση και Συμπίεση Εντροπίας.....	37
1.4 Ανάγκη Εύρεσης Νέων Μεθόδων Συμπίεσης Εικόνας.....	37
<i>Κεφάλαιο 2ο Εισαγωγή Της Νευροεπιστήμης Στην Επεξεργασία Εικόνων</i>	39
2.1 Εισαγωγή στο Οπτικό Σύστημα	39
2.2 Νευρώνες: Δομικά Κύτταρα του Νευρικού Συστήματος	40
2.3 Μηχανισμοί Παραγωγής Αιχμών.....	42
2.4 Το Μοντέλο LIF (Μηχανισμός Παραγωγής Αιχμών).....	43
2.5 Χρήση και ερμηνεία του μοντέλου LIF για τη δημιουργία κβαντιστή με απώτερο στόχο τη συμπίεση της εικόνας.....	47
<i>Κεφάλαιο 3ο Συμπεριφορά και Τρόπος Λειτουργίας Του Προτεινόμενου DUAL-SIM Κβαντιστή</i>	49
3.1 Τρόπος Λειτουργίας Προτεινόμενου Κβαντιστή.....	49
3.2 Συμπεριφορά Προτεινόμενου Κβαντιστή	50
3.3 Εισαγωγή Του DUAL-SIM Κβαντιστή στη Συμβατική Αρχιτεκτονική Συμπίεσης Εικόνων.....	51
3.3.1 Χωρίς Κάποιον Μετασχηματισμό	53
3.3.2. Με Μετασχηματισμό Συνημιτόνου (DCT).....	62
3.3.3 Με Μετασχηματισμό Κυματιδίων (DWT).....	71
3.4 Συμπεράσματα-Συγκρίσεις & Οπτικά Αποτελέσματα	80

3.5 Μελλοντική Δουλειά	83
Βιβλιογραφία	85

Κατάλογος Εικόνων

Εικόνα 1.1	Θεμελιώδης Αρχή Κωδικοποίησης - Ροή Σήματος	16
Εικόνα 1.2	Θεμελιώδης Αρχή Κωδικοποίησης - Ροή Σήματος	17
Εικόνα 1.3	Το Φως και τα Χρώματα μιας εικόνας	18
Εικόνα 1.4	Φωτεινότητα (αριστερά) και Χρωματικότητα (μέση) μιας εικόνας (δεξιά)	19
Εικόνα 1.5	Εικονοστοιχεία εικόνας και τα τρία δομικά τους στοιχεία	20
Εικόνα 1.6	Υπολογισμός Φωτεινότητας, Μπλε και Κόκκινης Χρωματικότητας μέσω των RGB	20
Εικόνα 1.7	Υπολογισμός Μέσης Τιμής της Μπλε και Κόκκινης Χρωματικότητας	21
Εικόνα 1.8	Συρρίκνωση (μόνο) της Μπλε και Κόκκινης Χρωματικότητας	21
Εικόνα 1.9	Εύκολα σημεία αντίληψης (πράσινο πλαίσιο - χαμηλές συχνότητες) και δύσκολα σημεία αντίληψης (κόκκινο πλαίσιο - υψηλές συχνότητες) από το μάτι	22
Εικόνα 1.10	Μετατόπιση του εύρους τιμών των πίξελ μιας εικόνας από το [0,255] στο [-128,127]	23
Εικόνα 1.11	Θεμελιώδεις Συχνότητες (βάση αποσύνθεσης)	24
Εικόνα 1.12	Συντελεστές DCT της εικόνας (αριστερός πίνακας)	24
Εικόνα 1.13	Διαίρεση του πίνακα συντελεστών DCT με τον Πίνακα Κβάντισης	26
Εικόνα 1.14	Κβαντοποιημένοι συντελεστές DCT	27
Εικόνα 1.15	Απλό Παράδειγμα Κβάντισης - Είσοδος (μπλε), Κβαντισμένη Έξοδος (κόκκινο) ..	27
Εικόνα 1.16	Σήμα εξόδου Κβαντιστή Midtread (κόκκινο) και Κβαντιστή Midriser (μπλε)	28
Εικόνα 1.17	Επίπεδα Κβάντισης (αριστερά) - Μέγεθος Βήματος Κβάντισης (δεξιά)	29
Εικόνα 1.18	Μεσοδιαστήματα και Επίπεδα Κβάντισης ενός Κβαντιστή	29
Εικόνα 1.19	Σήμα εξόδου Μη Ομοιόμορφου Κβαντιστή	30
Εικόνα 1.20	Κωδικοποίηση ζιγκ-ζαγκ	31
Εικόνα 1.21	Στάδια αποσύνθεσης εικόνας με εφαρμογή DWT	33
Εικόνα 1.22	Υποζώνες συχνοτήτων εικόνας με εφαρμογή DWT	34
Εικόνα 1.23	Οπτικό αποτέλεσμα εφαρμογής DWT	34
Εικόνα 2.1	Η διαδρομή του φωτός (οπτικό ερέθισμα) στο οπτικό σύστημα	37
Εικόνα 2.2	Διασυνδέσεις πολλών νευρώνων	38
Εικόνα 2.3	Βασικά Μέρη Ενός Νευρώνα	39
Εικόνα 2.4	Ηλεκτρικό Κύκλωμα του μοντέλου LIF	41
Εικόνα 2.5	Συμπεριφορά του μοντέλου LIF για τρεις διαφορετικές εισόδους I σε χρονικό διάστημα T με ένα συγκεκριμένο κατώφλι θ	44
Εικόνα 3.1	Ποσοτικοποίηση Εικόνας σε Αριθμό Αιχμών κατά την Νευρο-Εμπνευσμένη Κβάντιση	49
Εικόνα 3.2	Εισαγωγή του DUAL-SIM Κβαντιστή στη Συμβατική Αρχιτεκτονική	50

Εικόνα 3.3	Η (αρχική) εικόνα που χρησιμοποιήθηκε σε όλα τα πειράματα	51
Εικόνα 3.4	Παράμετρος C συναρτήσει του MSE για διάφορα θ και $T = 100\text{ ms}$ & $R = 0.35\ \Omega$	51
Εικόνα 3.5	Παράμετρος C συναρτήσει του MSE για διάφορα θ και $T = 100\text{ ms}$ & $R = 1\ \Omega$. .	52
Εικόνα 3.6	Παράμετρος C συναρτήσει του MSE για διάφορα θ και $T = 200\text{ ms}$ & $R = 0.35\ \Omega$	52
Εικόνα 3.7	Παράμετρος C συναρτήσει του MSE για διάφορα θ και $T = 200\text{ ms}$ & $R = 1\ \Omega$. .	53
Εικόνα 3.8	Παράμετρος R συναρτήσει του MSE για διάφορα C και $T = 100\text{ ms}$ & $\theta = 20\text{ V}$.	54
Εικόνα 3.9	Παράμετρος R συναρτήσει του MSE για διάφορα C και $T = 100\text{ ms}$ & $\theta = 50\text{ V}$.	54
Εικόνα 3.10	Παράμετρος R συναρτήσει του MSE για διάφορα C και $T = 200\text{ ms}$ & $\theta = 50\text{ V}$.	55
Εικόνα 3.11	Παράμετρος R συναρτήσει του MSE για διάφορα C και $T = 200\text{ ms}$ & $\theta = 100\text{ V}$	55
Εικόνα 3.12	Παράμετρος R συναρτήσει του MSE για διάφορα θ και $T = 200\text{ ms}$ & $C = 20\text{ F}$. .	56
Εικόνα 3.13	Παράμετρος R συναρτήσει του MSE για διάφορα θ και $T = 200\text{ ms}$ & $C = 50\text{ F}$. .	57
Εικόνα 3.14	Παράμετρος T συναρτήσει του MSE για διάφορα R και $\theta = 50\text{ V}$ & $C = 20\text{ F}$. . .	58
Εικόνα 3.15	Παράμετρος T συναρτήσει του MSE για διάφορα R και $\theta = 50\text{ V}$ & $C = 50\text{ F}$. . .	59
Εικόνα 3.16	Εντροπία κβαντιστή δίχως χρήση μετασχηματισμού συναρτήσει του θ για $T = 150\text{ ms}$ και διαφόρους συνδυασμούς R & C	60
Εικόνα 3.17	PSNR κβαντιστή δίχως χρήση μετασχηματισμού συναρτήσει του θ για $T = 150\text{ ms}$ και διαφόρους συνδυασμούς R & C	61
Εικόνα 3.18	Παράμετρος C συναρτήσει του MSE για διάφορα θ και $T = 100\text{ ms}$ & $R = 0.35\ \Omega$ με χρήση DCT	61
Εικόνα 3.19	Παράμετρος C συναρτήσει του MSE για διάφορα θ και $T = 100\text{ ms}$ & $R = 1\ \Omega$ με χρήση DCT	61
Εικόνα 3.20	Παράμετρος C συναρτήσει του MSE για διάφορα θ και $T = 200\text{ ms}$ & $R = 0.35\ \Omega$ με χρήση DCT	62
Εικόνα 3.21	Παράμετρος C συναρτήσει του MSE για διάφορα θ και $T = 200\text{ ms}$ & $R = 1\ \Omega$ με χρήση DCT	62
Εικόνα 3.22	Παράμετρος R συναρτήσει του MSE για διάφορα C και $T = 100\text{ ms}$ & $\theta = 1\text{ V}$ με χρήση DCT	63
Εικόνα 3.23	Παράμετρος R συναρτήσει του MSE για διάφορα C και $T = 100\text{ ms}$ & $\theta = 50\text{ V}$ με χρήση DCT	64
Εικόνα 3.24	Παράμετρος R συναρτήσει του MSE για διάφορα C και $T = 200\text{ ms}$ & $\theta = 50\text{ V}$ με χρήση DCT	64
Εικόνα 3.25	Παράμετρος R συναρτήσει του MSE για διάφορα C και $T = 200\text{ ms}$ & $\theta = 100\text{ V}$ με χρήση DCT	65
Εικόνα 3.26	Παράμετρος R συναρτήσει του MSE για διάφορα θ και $T = 200\text{ ms}$ & $C = 20\text{ F}$ με χρήση DCT	65
Εικόνα 3.27	Παράμετρος R συναρτήσει του MSE για διάφορα θ και $T = 200\text{ ms}$ & $C = 50\text{ F}$ με χρήση DCT	66
Εικόνα 3.28	Παράμετρος T συναρτήσει του MSE για διάφορα R και $\theta = 50\text{ V}$ & $C = 20\text{ F}$ με χρήση DCT	67
Εικόνα 3.29	Παράμετρος T συναρτήσει του MSE για διάφορα R και $\theta = 50\text{ V}$ & $C = 50\text{ F}$ με	

	χρήση DCT	67
Εικόνα 3.30	Εντροπία κβαντιστή με χρήση DCT μετασχηματισμού συναρτήσει του θ για $T = 150\text{ ms}$ και διαφόρους συνδυασμούς R & C	68
Εικόνα 3.31	PSNR κβαντιστή με χρήση DCT μετασχηματισμού συναρτήσει του θ για $T = 150\text{ ms}$ και διαφόρους συνδυασμούς R & C	68
Εικόνα 3.32	Παράμετρος C συναρτήσει του MSE για διάφορα θ και $T = 100\text{ ms}$ & $R = 0.35\ \Omega$ με χρήση DWT	70
Εικόνα 3.33	Παράμετρος C συναρτήσει του MSE για διάφορα θ και $T = 100\text{ ms}$ & $R = 1\ \Omega$ με χρήση DWT	70
Εικόνα 3.34	Παράμετρος C συναρτήσει του MSE για διάφορα θ και $T = 200\text{ ms}$ & $R = 0.35\ \Omega$ με χρήση DWT	71
Εικόνα 3.35	Παράμετρος C συναρτήσει του MSE για διάφορα θ και $T = 200\text{ ms}$ & $R = 1\ \Omega$ με χρήση DWT	71
Εικόνα 3.36	Παράμετρος R συναρτήσει του MSE για διάφορα C και $T = 100\text{ ms}$ & $\theta = 1\text{ V}$ με χρήση DWT	72
Εικόνα 3.37	Παράμετρος R συναρτήσει του MSE για διάφορα C και $T = 100\text{ ms}$ & $\theta = 50\text{ V}$ με χρήση DWT	73
Εικόνα 3.38	Παράμετρος R συναρτήσει του MSE για διάφορα C και $T = 200\text{ ms}$ & $\theta = 50\text{ V}$ με χρήση DWT	73
Εικόνα 3.39	Παράμετρος R συναρτήσει του MSE για διάφορα C και $T = 200\text{ ms}$ & $\theta = 100\text{ V}$ με χρήση DWT	74
Εικόνα 3.40	Παράμετρος R συναρτήσει του MSE για διάφορα θ και $T = 200\text{ ms}$ & $C = 20\text{ F}$ με χρήση DWT	74
Εικόνα 3.41	Παράμετρος R συναρτήσει του MSE για διάφορα θ και $T = 200\text{ ms}$ & $C = 50\text{ F}$ με χρήση DWT	75
Εικόνα 3.42	Παράμετρος T συναρτήσει του MSE για διάφορα R και $\theta = 50\text{ V}$ & $C = 20\text{ F}$ με χρήση DWT	76
Εικόνα 3.43	Παράμετρος T συναρτήσει του MSE για διάφορα R και $\theta = 50\text{ V}$ & $C = 50\text{ F}$ με χρήση DWT	76
Εικόνα 3.44	Εντροπία κβαντιστή με χρήση DWT μετασχηματισμού συναρτήσει του θ για $T = 150\text{ ms}$ και διαφόρους συνδυασμούς R & C	77
Εικόνα 3.45	PSNR κβαντιστή με χρήση DWT μετασχηματισμού συναρτήσει του θ για $T = 150\text{ ms}$ και διαφόρους συνδυασμούς R & C	78
Εικόνα 3.46	Ο αντίκτυπος των μεταβλητών R και C καθώς και των μετασχηματισμών DCT και DWT στην απόδοση του DUAL-SIM κβαντιστή	79
Εικόνα 3.47	Οπτικά αποτελέσματα του DUAL-SIM κβαντιστή χωρίς τη χρήση κάποιου μετασχηματισμού	80
Εικόνα 3.48	Οπτικά αποτελέσματα του DUAL-SIM κβαντιστή με τη χρήση του DWT μετασχηματισμού	80
Εικόνα 3.49	Οπτικά αποτελέσματα του DUAL-SIM κβαντιστή με τη χρήση του DCT μετασχηματισμού	81

Κεφάλαιο 1ο

Εισαγωγή Στις Τεχνικές Συμπίεσης

1.1. Γιατί είναι σημαντική η συμπίεση ;

Τις τελευταίες δεκαετίες η ραγδαία ανάπτυξη της τεχνολογίας έχει διευκολύνει σημαντικά την ζωή των ανθρώπων προσφέροντας πολλαπλές ανέσεις με σκοπό τη βελτίωση της καθημερινότητάς τους. Δείγματα αυτής της ραγδαίας εξέλιξης είναι μεταξύ άλλων η αντικατάσταση της αναλογικής από την ψηφιακή τηλεόραση, η χρήση των κινητών τηλεφώνων όχι πλέον μόνο ως μέσο επικοινωνίας και αποστολής μηνυμάτων (SMS) αλλά και ως μέσο λήψης βίντεο και φωτογραφιών, η αντικατάσταση των βιντεοκασετών (VHS) από τους ψηφιακούς δίσκους βίντεο (DVD) και αργότερα από τους ψηφιακούς δίσκους βίντεο πολύ μεγάλης χωρητικότητας (Blu-Ray Disc) καθώς και η δημιουργία βίντεο και ψηφιακών φωτογραφιών υψηλής ανάλυσης και ευκρίνειας αντίστοιχα. Επίσης, η ενίσχυση της χωρητικότητας των ήδη υπαρχόντων ηλεκτρονικών συσκευών (σκληρός δίσκος, USB, κάρτες SIM κ.ά.) για την αποθήκευση ολοένα και περισσότερων δεδομένων, όπως και η δημιουργία συσκευών ψυχαγωγίας (οικιακές και φορητές κονσόλες βιντεοπαιχνιδιών, τάμπλετ, ηλεκτρονικά βιβλία κ.ά.) που εισχωρούν όλο και περισσότερο στην καθημερινότητα των ανθρώπων. Αλλά φυσικά, το κύριο χαρακτηριστικό γνώρισμα αυτής της τεχνολογικής εξέλιξης είναι το διαδίκτυο η χρήση του οποίου έχει γίνει αναπόσπαστο κομμάτι από τη ζωή των ανθρώπων κυρίως λόγω της ηχηρής άνθισης των μέσων κοινωνικής δικτύωσης (YouTube, Instagram, Twitter κ.ά.) και των πλατφορμών μετάδοσης ζωντανής ροής (Zoom, Twitch, Discord κ.ά.). Ως εκ τούτου, παρατηρείται ότι ο αριθμός των χρηστών που καταγράφουν οπτικά δεδομένα, όπως βίντεο και εικόνες, με σκοπό τόσο την αποθήκευσή τους για μελλοντική χρήση όσο και την διάχυσή τους στα μέσα κοινωνικής δικτύωσης, αυξάνεται εκθετικά. Πρόσφατη έρευνα έχει δείξει πως εν έτει 2018 παράγονταν πάνω από 2.5 πεντάκις εκατομμύρια byte δεδομένων κάθε ημέρα και μέχρι το τέλος του 2022, θα υπάρχουν περίπου 97 zettabyte δεδομένων στον κόσμο ($1 \text{ zettabyte} = 1.0 \times 10^{21} \text{ bytes}$). Συνεπώς, βάσει των προαναφερθέντων, γίνεται άμεσα αντιληπτό πως η συμπίεση των δεδομένων, κυρίως των εικόνων και βίντεο, γίνεται όλο και πιο επιτακτική, αναγκαία και χρήσιμη.

Με τον όρο συμπίεση εννοείται η μετατροπή δεδομένων σε μια μορφή που απαιτεί λιγότερα bit έτσι ώστε τα δεδομένα να μπορούν να μεταδοθούν γρηγορότερα ή να αποθηκευτούν καταναλώνοντας λιγότερο χώρο. Η επιδίωξη της συμπίεσης δεδομένων είναι να αφαιρέσει τον πλεονασμό που πιθανώς υπάρχει στην πληροφορία προς συμπίεση διατηρώντας ταυτόχρονα την καλύτερη δυνατή ποιότητά της. Η αντίστροφη και συμπληρωματική διαδικασία της συμπίεσης ονομάζεται αποσυμπίεση και στοχεύει στην ανάκτηση των δεδομένων που έχουν υποστεί συμπίεση. Το επιθυμητό για ένα σύστημα συμπίεσης εικόνας είναι αφενός η επίτευξη της όσο δυνατόν μεγαλύτερης μείωσης του τρόπου αναπαράστασης των δεδομένων, με σκοπό τη δέσμευση λιγότερου αποθηκευτικού χώρου, και αφετέρου στην ταυτόχρονη διατήρηση της οπτικής ποιότητας για τη βέλτιστη απεικόνιση. Αυτά τα δύο στοιχεία είναι αντικρουόμενα μεταξύ τους καθώς όσο αυξάνεται η συμπίεση για τη χρήση λιγότερου αποθηκευτικού χώρου,

τόσο μειώνεται η ποιότητα της εικόνας, αλλά και αντιστρόφως. Οπότε, ανάλογα με την εφαρμογή για την οποία προορίζεται το σύστημα, δίνεται παραπάνω έμφαση είτε στη συμπίεση (π.χ. αποστολή εικόνας μέσω διαδικτύου) είτε στην ποιότητα (π.χ. αποθήκευση εικόνας υψηλής ευκρίνειας).

Ο λόγος του μεγέθους των αρχικών δεδομένων προς το μέγεθος των συμπίεσμένων δεδομένων ονομάζεται λόγος συμπίεσης. Υπάρχουν δύο βασικές μέθοδοι συμπίεσης δεδομένων: η απωλεστική και η μη απωλεστική. Όταν η αποσυμπίεση παράγει ένα ακριβές αντίγραφο των αρχικών δεδομένων τότε η συμπίεση λέγεται μη απωλεστική (lossless), ενώ αν παράγει προσεγγιστικά τα αρχικά δεδομένα λέγεται απωλεστική (lossy). Με την απωλεστική συμπίεση επιτυγχάνονται μεγαλύτεροι λόγοι συμπίεσης από ό,τι με τη μη απωλεστική, ενώ η πιστότητα των δεδομένων μετά την αποσυμπίεση γίνεται υψηλότερη όσο ο λόγο συμπίεσης μειώνεται. [1]

Τα τελευταία χρόνια υπάρχει μια στροφή ως προς τη μελέτη και χρήση νευρωνικών δικτύων στον χώρο της συμπίεσης και κυρίως του ανθρώπινου οπτικού συστήματος. Ένα από τα βασικότερα χαρακτηριστικά του οπτικού συστήματος, που δικαιολογεί αυτήν τη στροφή σε αυτό, είναι η ευαισθησία του σε ό,τι αλλαγές προκύπτουν σε ένα ήδη υπάρχον οπτικό ερέθισμα. Το οπτικό σύστημα επομένως, έχει τη δυνατότητα να διαχωρίζει τη σημαντική πληροφορία, δηλαδή την όποια πιθανή αλλαγή, από ένα στατικό φόντο το οποίο κατά βάση παραμένει ακίνητο. Ο νέος κβαντιστής λοιπόν που θα μελετηθεί παρακάτω, εμπνεόμενος από το οπτικό σύστημα, έχει την ικανότητα να εκτιμά την κίνηση με έναν έξυπνο και αποτελεσματικό τρόπο, μια ικανότητα στην οποία υπολείπονται τα τρέχοντα συστήματα συμπίεσης.

Στη συνέχεια του πρώτου κεφαλαίου αναφέρονται τα είδη και τα σύγχρονα μοντέλα της συμπίεσης. Στο δεύτερο κεφάλαιο παρουσιάζονται κάποιες βασικές πληροφορίες για το ανθρώπινο οπτικό σύστημα. Στο τρίτο και τελευταίο κεφάλαιο γίνεται η πρόταση του νέου κβαντιστή, στον οποίο βασίζεται η εργασία και τα πειράματα που αφορούν αυτόν, καθώς και μια πιθανή επέκταση αυτής της διπλωματικής εργασίας.

1.1.1. Μη Απωλεστική Συμπίεση

Για μία διεργασία μη απωλεστικής συμπίεσης τα ανακατασκευασμένα δεδομένα και τα αρχικά δεδομένα πρέπει να είναι πανομοιότυπα, διαδικασία αναφερόμενη και ως αντιστρεπτή διεργασία. Ένας απλοϊκός παραλληλισμός της διαδικασίας συμπίεσης και αποσυμπίεσης της μη απωλεστικής διεργασίας είναι αυτός του ταχυδρομικού γράμματος. Η συμπίεση είναι ανάλογη με το δίπλωμα ενός γράμματος πριν τοποθετηθεί σε φάκελο έτσι ώστε να μπορεί να μεταφερθεί πιο εύκολα και εργονομικά. Ενώ, όπως ακριβώς το γράμμα που είναι κλεισμένο μέσα στον φάκελο είναι δύσκολο να διαβαστεί και πρώτα πρέπει να βγει από αυτόν, έτσι και τα συμπίεσμένα δεδομένα πρέπει πρώτα να αποσυμπίεστούν για να επανέλθουν στην αρχική τους μορφή. Η μη απωλεστική συμπίεση βασίζεται στο αντιστάθμισμα μεταξύ τριών αλληλοσυγκρουόμενων παραγόντων, (α) την καθυστέρηση κωδικοποίησης, (β) την πολυπλοκότητα κωδικοποίησης και (γ) την αποδοτικότητα κωδικοποίησης. Η αποδοτικότητα της κωδικοποίησης μετριέται βάσει της εντροπίας της πηγής. Η εντροπία μιας τυχαίας

μεταβλητής είναι το μέσο επίπεδο «αβεβαιότητας» που είναι εγγενές στα πιθανά αποτελέσματα αυτής της μεταβλητής. [2]

Συγκεκριμένα, για μια διακριτή τυχαία μεταβλητή X , που παίρνει τιμές από ένα σύνολο A και κατανέμεται σύμφωνα με $p: A \rightarrow [0,1]$ η εντροπία ορίζεται ως εξής :

$$H(X) := - \sum_{x \in A} p(x) \log p(x) = \mathbb{E}[-\log p(X)] \quad (1.1)$$

Συνεπώς, η εντροπία είναι ένας δείκτης που ορίζει πόσο εύκολα μπορεί να συμπίεστεί μια πηγή για μια δεδομένη τυχαιότητα, επομένως προκύπτει πως ένα σήμα εισόδου με χαμηλή εντροπία, δηλαδή με λιγοστό θόρυβο, μπορεί να συμπίεστεί αρκετά εύκολα. Η πολυπλοκότητα της κωδικοποίησης είναι ανάλογη με την υπολογιστική επίδοση που απαιτείται για την υλοποίηση των λειτουργιών κωδικοποίησης και αποκωδικοποίησης. Η υπολογιστική επίδοση μετριέται συνήθως με βάση τις απαιτήσεις μνήμης καθώς και το πλήθος αριθμητικών λειτουργιών που απαιτούνται για τη συμπίεση του σήματος εισόδου. Μια σύνθετη και πολύπλοκη διεργασία συμπίεσης είθισται να οδηγεί σε αυξημένες καθυστερήσεις κωδικοποίησης στον κωδικοποιητή και στον αποκωδικοποιητή. Αυτές οι καθυστερήσεις μπορούν να μειωθούν αυξάνοντας την επεξεργαστική ισχύ της υπολογιστικής συσκευής, τροποποίηση που μπορεί να μην είναι πρακτική σε περιβάλλον με περιορισμούς ενέργειας και ισχύος (π.χ. ασύρματες κάμερες χαμηλής κατανάλωσης). Η μη απωλεστική συμπίεση ενδείκνυται για κείμενα αλλά και για διάδοση ιατρικών δεδομένων (π.χ. εικόνες διάγνωσης) στα οποία δεν πρέπει να υπάρχει απώλεια πληροφορίας καθώς μπορεί να προκύψουν λανθασμένα αποτελέσματα (π.χ. εσφαλμένη διάγνωση). [2]

1.1.2. Απωλεστική Συμπίεση

Ο βασικός στόχος της απωλεστικής συμπίεσης είναι η ελαχιστοποίηση του αριθμού των απαιτούμενων bits για την απεικόνιση του σήματος εισόδου με την καλύτερη δυνατή ποιότητα. Στην ουσία πρόκειται για μια διεργασία που αναπαριστά το σήμα εισόδου με λιγότερα bits, γεγονός που οδηγεί σε απώλεια πληροφορίας. Μια τέτοιου είδους διαδικασία συμπίεσης που καταλήγει σε μια μη τέλεια ανακατασκευή του αρχικού σήματος αναφέρεται και ως μη αντιστρεπτή. Η χρήση της απωλεστικής συμπίεσης γίνεται από συσκευές ή εφαρμογές που δεν απαιτούν να είναι ακριβώς ίδια τα ανακατασκευασμένα δεδομένα με τα αρχικά. Εντούτοις, η απωλεστική συμπίεση επιδιώκει στην εύρεση στοιχείων από το αρχικό σήμα που όταν αφαιρεθούν το τελικό οπτικό αποτέλεσμα δεν θα επηρεαστεί σημαντικά. Λόγω του επιπρόσθετου βαθμού ελευθερίας λάθους κατά την ανακατασκευή, η διεργασία της απωλεστικής συμπίεσης πετυχαίνει υψηλότερους λόγους συμπίεσης από αυτούς μιας μη απωλεστικής συμπίεσης. Η απωλεστική συμπίεση βασίζεται στο αντιστάθμισμα μεταξύ τεσσάρων αλληλοσυγκρουόμενων παραγόντων. Των τριών που αναφέρθηκαν στη μη απωλεστική συμπίεση καθώς και στην ποιότητα του ανακατασκευασμένου σήματος (παραμόρφωση). Η απωλεστική συμπίεση χρησιμοποιείται κυρίως από εφαρμογές

επεξεργασίας δεδομένων εικόνας ή και βίντεο, όπως το κλειστό κύκλωμα τηλεόρασης (CCTV : closed-circuit television) και τα DVD. [3]

Ο όρος ποιότητα ανακατασκευασμένου σήματος χρησιμοποιείται συχνά για να χαρακτηρίσει το σήμα εξόδου στον αποκωδικοποιητή. Υπάρχουν διάφορες μέθοδοι μέτρησης ποιότητας που είναι υπεύθυνες για τη μέτρηση της ποιότητας του ανακατασκευασμένου σήματος (σήμα εξόδου) σε σχέση με το σήμα εισόδου και οι πιο ευρέως χρησιμοποιούμενες είναι το MSE και το PSNR.

Συγκεκριμένα, το MSE (Μέσο Τετραγωνικό Σφάλμα - Mean Squared Error) ορίζεται ως εξής:

$$MSE(I, \tilde{I}) = \frac{1}{n} \sum_{i=1}^n (I_i - \tilde{I}_i)^2 \quad (1.2)$$

όπου I το σήμα εισόδου, \tilde{I} το σήμα εξόδου και n το μέγεθος του σήματος εισόδου.

Η παραμόρφωση του σήματος εξόδου ελαχιστοποιείται όταν το MSE τείνει στο μηδέν. Αφετέρου, το PSNR (Λόγος Μέγιστου Σήματος προς Θόρυβο - Peak Signal to Noise Ratio) ορίζεται ως εξής :

$$PSNR(I, \tilde{I}) = 10 \log_{10} \frac{MAX_I^2}{MSE(I, \tilde{I})} \quad (1.3)$$

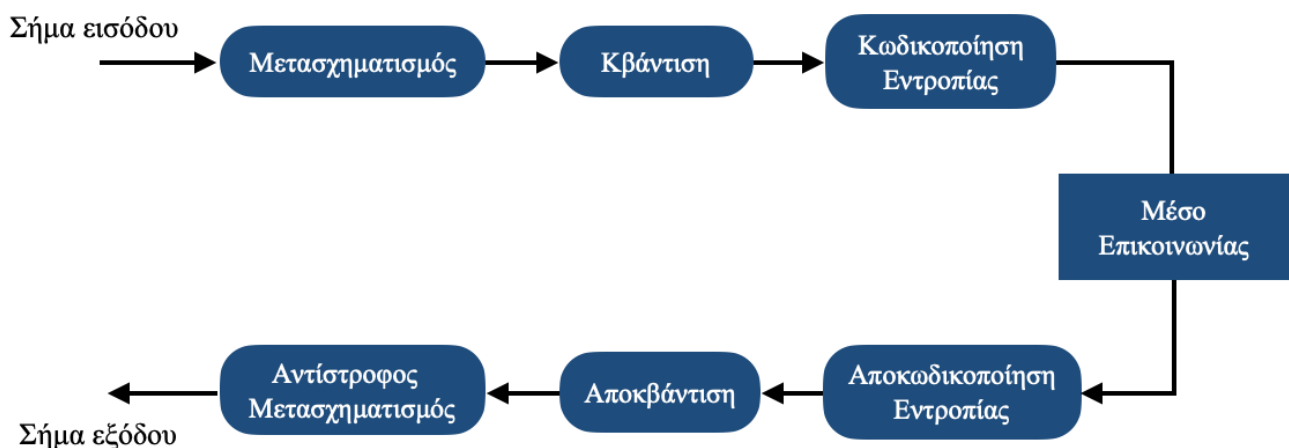
όπου I το σήμα εισόδου, \tilde{I} το σήμα εξόδου και MAX_I η μέγιστη πιθανή τιμή του εικονοστοιχείου του σήματος. Όταν τα εικονοστοιχεία αναπαρίστανται χρησιμοποιώντας 8 bit ανά δείγμα, το MAX_I είναι το 255. Γενικότερα, όταν τα δείγματα αναπαρίστανται χρησιμοποιώντας γραμμική παλμοκωδική διαμόρφωση (PCM - Pulse Code Modulation) το MAX_I είναι $2^b - 1$. Το PSNR εκφράζεται σε dB (ντεσιμπέλ) και βασίζεται στο απόλυτο σφάλμα μεταξύ των σημάτων εισόδου και εξόδου. Το PSNR τείνει στο άπειρο, πράγμα που σημαίνει ότι μια υψηλή τιμή PSNR (δηλαδή μικρές αριθμητικές διαφορές μεταξύ των δύο σημάτων) σημαίνει καλύτερη ποιότητα στο σήμα εξόδου. [4]

1.2 Θεμελιώδης Αρχή Της Κωδικοποίησης

Όλοι οι αλγόριθμοι απωλεστικής συμπίεσης έχουν υιοθετήσει την αρχή της κωδικοποίησης η οποία περιγράφει τη διαδικασία κωδικοποίησης αλλά και αποκωδικοποίησης ενός σήματος εισόδου, το οποίο μπορεί να είναι οποιασδήποτε μορφής δεδομένων, όπως ήχος ή εικόνα. Αρχικά, το σήμα εισόδου μετασχηματίζεται σε μια πιο «φιλική» προς συμπίεση μορφή μέσω κάποιου μετασχηματισμού. «Φιλική» θεωρείται η μορφή στην οποία η πληροφορία έχει

αναδιαμορφωθεί με τέτοιον τρόπο ώστε να διακρίνονται ποια κομμάτια της είναι πιο σημαντικά (εμπεριέχουν πολλή πληροφορία) από άλλα, ούτως ώστε κατά τη συμπίεση τα λιγότερα σημαντικά να έχουν μεγαλύτερη προτεραιότητα να απορριφθούν από ό,τι τα άλλα. Ο διακριτός μετασχηματισμός συνημιτόνου και ο διακριτός μετασχηματισμός κυματιδίων (DCT : Discrete Cosine Transform & DWT : Discrete Wavelet Transform) είναι κάποιοι από τους πιο συνήθεις μετασχηματισμούς. Εν συνεχεία, το αποτέλεσμα αυτού του μετασχηματισμού λαμβάνεται από έναν κβαντιστή ο οποίος του αφαιρεί πληροφορία που αυτός «θεωρεί» περιττή. Εξ αυτού γίνεται αντιληπτό πως η ύπαρξη παραμόρφωσης στο τελικό αποτέλεσμα προέρχεται από την κβάντιση και ακριβώς για αυτόν τον λόγο στις μη απωλεστικές συμπίεσεις δεν συμπεριλαμβάνεται το βήμα αυτής. Τέλος, το αποτέλεσμα του κβαντιστή κωδικοποιείται μέσω μιας μη απωλεστικής συνάρτησης ονόματι κωδικοποίηση εντροπίας. Η κβαντισμένη ποσότητα κωδικοποιείται σε κωδικές λέξεις το μέγεθος των οποίων εξαρτάται από τη συχνότητα εμφάνισής τους. Όσο πιο συχνά εμφανιζόμενη, τόσο πιο μικρή σε μέγεθος. Μετά την κωδικοποίηση το σήμα είτε αποθηκεύεται είτε αποστέλλεται σε κάποιον δέκτη διά ενός μέσου επικοινωνίας. Ο δέκτης λαμβάνει την κωδικοποιημένη μορφή την οποία και χρειάζεται για να ανακατασκευάσει το σήμα, δηλαδή να ξεκινήσει τη διαδικασία της αποκωδικοποίησης. Η διαδικασία αυτή αποτελείται από τα αντίστοιχα αντίστροφα βήματα της κωδικοποίησης, δηλαδή την αποκωδικοποίηση εντροπίας, την αποκβάντιση και τον αντίστροφο μετασχηματισμό.

Όλα αυτά τα βήματα, που απεικονίζονται φαίνονται και στην Εικόνα 1.1, περιγράφουν την αρχιτεκτονική της αρχής της κωδικοποίησης απώτερος σκοπός της οποίας είναι η εύρεση της «χρυσής τομής» μεταξύ της περιττής πληροφορίας που πρόκειται να απορριφθεί κατά την κβάντιση και της ποιότητας του ανακατασκευασμένου σήματος εξόδου που πρέπει να τείνει στο αρχικό.



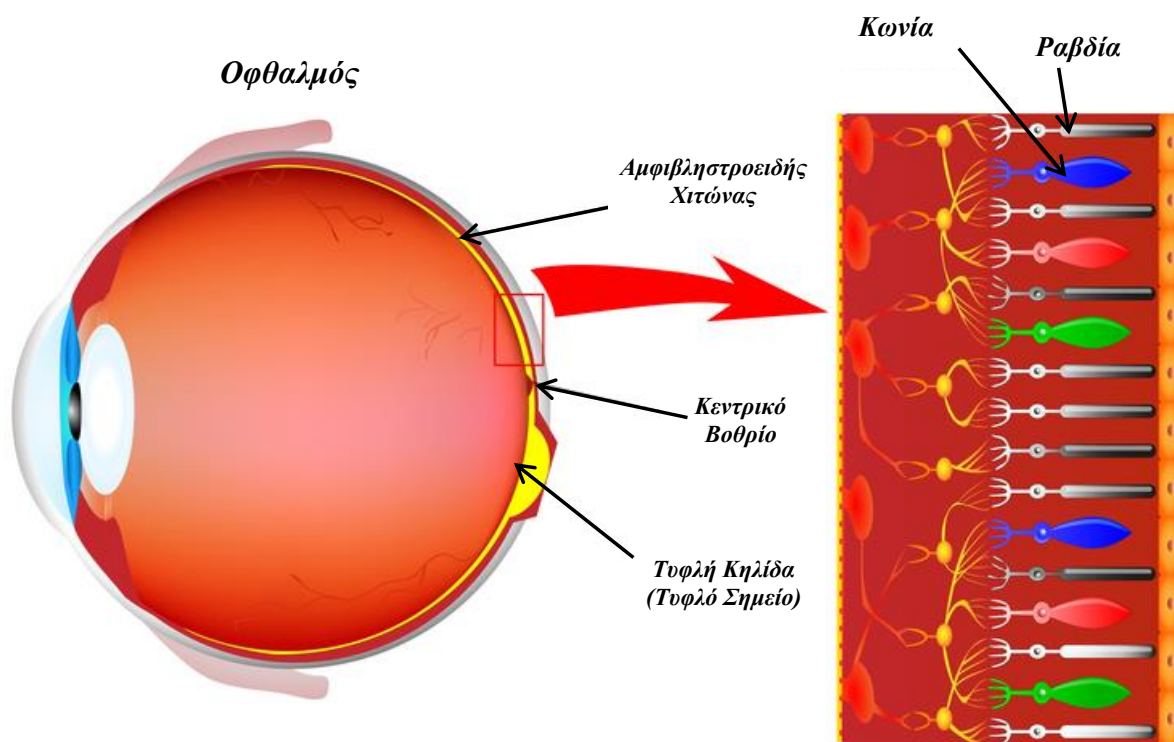
Εικόνα 1.1 : Θεμελιώδης Αρχή Κωδικοποίησης - Ροή Σήματος

1.3 Πρότυπα Συμπίεσης Εικόνας

Οι αλγόριθμοι συμπίεσης εικόνας χρησιμοποιούνται σε πολυάριθμες εφαρμογές της καθημερινότητάς μας όπως στο διαδίκτυο, στις ψηφιακές φωτογραφικές μηχανές, σε κάμερες παρακολούθησης κλπ. Η γενική μορφή ενός πρότυπου συμπίεσης εικόνας είναι αυτή της θεμελιώδους αρχής της κωδικοποίησης που περιγράφηκε στην *Εικόνα 1.1* . [5] Τα δύο πιο αποτελεσματικά πρότυπα συμπίεσης εικόνας είναι οι αλγόριθμοι των JPEG και JPEG2000.

1.3.1 JPEG

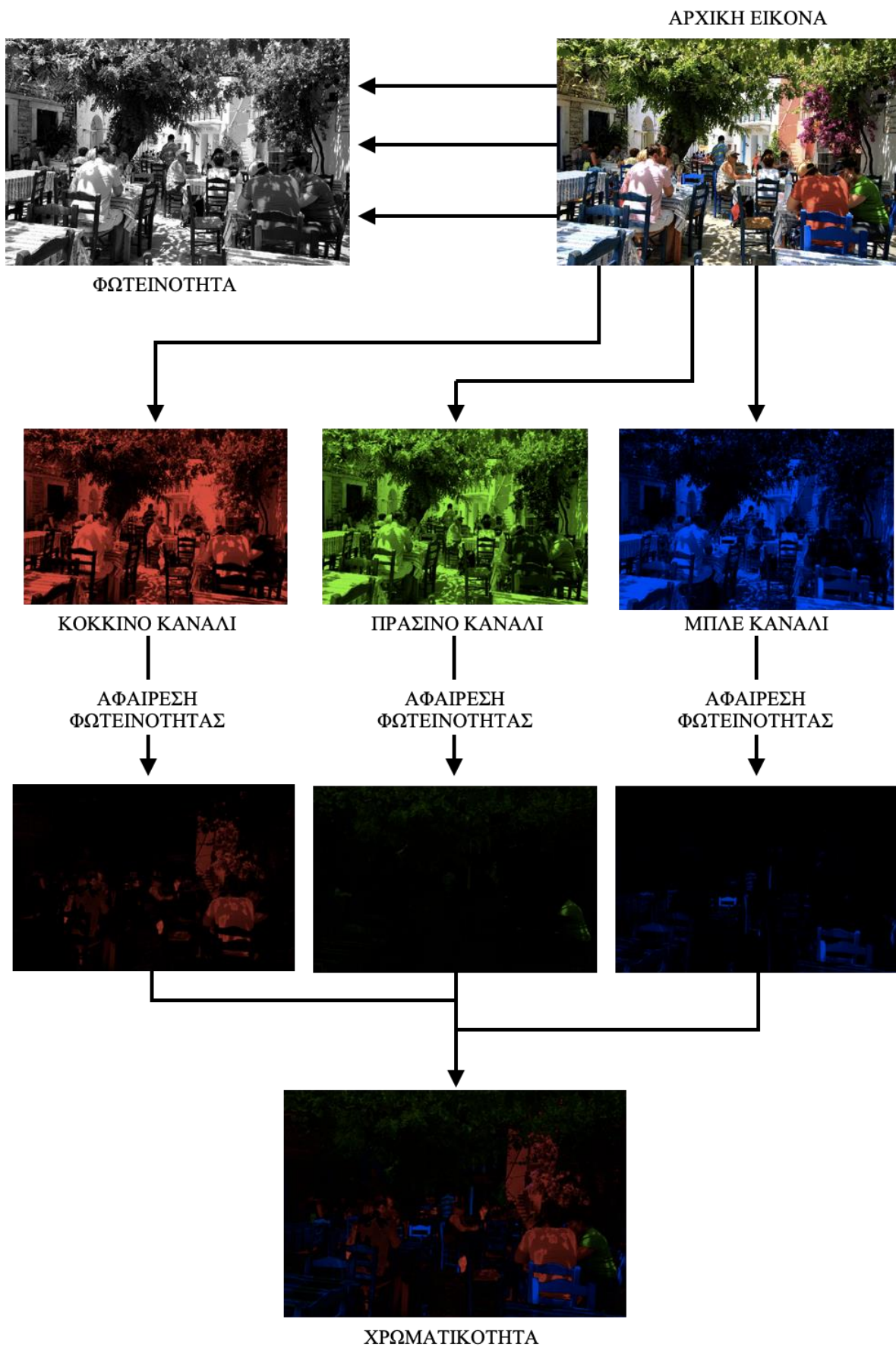
Εν γένει, ο αλγόριθμος JPEG αναλύει όλη την εικόνα που πρόκειται να συμπίεσει και αφαιρεί στοιχεία της, κάνοντας χρήση του μετασχηματισμού DCT, με σκοπό να πιάνει λιγότερο αποθηκευτικό χώρο, ενώ παράλληλα η οπτική ποιότητά της να μη αλλοιώνεται δραματικά. [6] Η αφαίρεση αυτών των στοιχείων επιτυγχάνεται καθώς ο αλγόριθμος εκμεταλλεύεται τις «αδυναμίες» που διέπουν το ανθρώπινο μάτι. Συγκεκριμένα στο ανθρώπινο μάτι υπάρχουν δύο διαφορετικά είδη φωτοϋποδοχέων κυττάρων, τα ραβδία και τα κωνία (βλ. *Εικόνα 1.2*).



Εικόνα 1.2 : Κύτταρα Φωτοϋποδοχείς - Ραβδία & Κωνία

Τα ραβδία είναι υπεύθυνα για τη σύλληψη του φωτός και συμβάλλουν σημαντικά στην όραση σε συνθήκες χαμηλού φωτισμού, αλλά παρ' όλα αυτά δεν είναι ευκόλως επηρεαζόμενα από το χρώμα. Από την άλλη, τα κωνία με τους χρωματοϋποδοχείς για το κόκκινο, πράσινο και μπλε

είναι ευαίσθητα στα χρώματα καθώς και υπεύθυνα για την αντίληψη αυτών. Να σημειωθεί πως σε κάθε ανθρώπινο μάτι υπάρχουν περίπου 100 εκατομμύρια κύτταρα ραβδίων σε αντίθεση με τα 6 εκατομμύρια κύτταρα κωνίων. Ως εκ τούτου, ο οφθαλμός είναι πολύ πιο ευεπίφορος στα λαμπερά και σκοτεινά σημεία μιας εικόνας, δηλαδή στη φωτεινότητά της (luminance), και όχι τόσο στα χρώματα που εμπεριέχει η εικόνα, δηλαδή στη χρωματικότητά της (chrominance). Στην ουσία η χρωματικότητα είναι τα αμιγή χρώματα, χρώματα δηλαδή δίχως φωτεινότητα, από τα οποία αποτελείται η εικόνα (βλ. *Εικόνα 1.3 στην επόμενη σελίδα*).



Εικόνα 1.3 : Το Φως και τα Χρώματα μιας εικόνας

Η σχέση μεταξύ φωτεινότητας και χρωματικότητας μπορεί να γίνει πιο ευδιάκριτη στο παρακάτω παράδειγμα της *Εικόνας 1.4*.

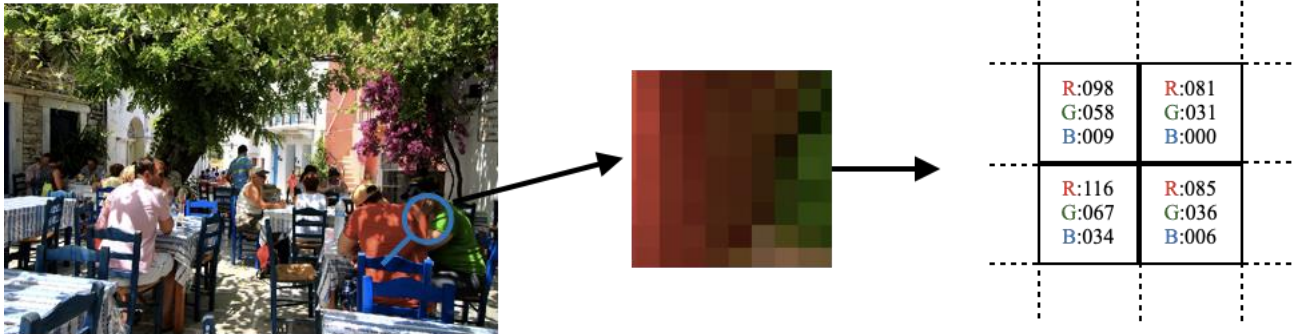


Εικόνα 1.4 : Φωτεινότητα (αριστερά) και Χρωματικότητα (μέση) μιας εικόνας (δεξιά)

Η δεξιά εικόνα (η πρωτότυπη) φαίνεται να είναι το ίδιο λεπτομερές με την ασπρόμαυρη εκδοχή της (αριστερή εικόνα) όπου απεικονίζεται μόνο η φωτεινότητα. Αν απεικονιστούν όμως μόνο τα χρώματα της εικόνας, δηλαδή η χρωματικότητα (μεσαία εικόνα), το οπτικό αποτέλεσμα είναι σημαντικά λιγότερο λεπτομερές, γεγονός που επιβεβαιώνει πως το ανθρώπινο μάτι είναι πιο ευαίσθητο στη φωτεινότητα παρά στα χρώματα. Έχοντας υπόψιν αυτήν την πληροφορία, ο αλγόριθμος JPEG ακολουθεί πέντε βασικά βήματα για τη συμπίεση μιας εικόνας, (α) μετατροπή του χρωματικού χώρου (colour space conversion), (β) υποδειγματοληψία της χρωματικότητας (chrominance downsampling), (γ) DCT, (δ) κβάντιση (quantization) και (ε) περαιτέρω συμπίεση εντροπίας.

1.3.1.1 Μετατροπή του Χρωματικού Χώρου

Όλες οι ψηφιακές εικόνες αποτελούνται από εικονοστοιχεία (pixels) και κάθε εικονοστοιχείο έχει από ένα κόκκινο, πράσινο και μπλε δομικό στοιχείο (RGB : Red, Green, Blue) που έκαστο έχει μια τιμή η οποία κυμαίνεται από το 0 έως το 255 (όπως φαίνεται στο παράδειγμα της *Εικόνας 1.5*), με τις χαμηλές τιμές να αντιπροσωπεύουν τις σκουρόχρωμες αποχρώσεις και με τις υψηλές τις ανοιχτόχρωμες. Ο συνδυασμός και των τριών αυτών δομικών στοιχείων (R-G-B) δημιουργεί το χρώμα για κάθε ένα από τα εικονοστοιχεία της εικόνας.



Εικόνα 1.5 : Εικονοστοιχεία εικόνας και τα τρία δομικά τους στοιχεία

Η διαδικασία της μετατροπής του χρωματικού χώρου λαμβάνει και τις τρεις τιμές των R-G-B του κάθε εικονοστοιχείου μιας εικόνας και υπολογίζει τρεις νέες τιμές (βλ. παράδειγμα της Εικόνας 1.6). Τη φωτεινότητα (Y), την μπλε χρωματικότητα (Cb) και την κόκκινη χρωματικότητα (Cr). Αυτή η διαδικασία είναι αντιστρέψιμη και καμία πληροφορία δεν χάνεται κατά τη μετατροπή.

$$\begin{array}{c}
 \begin{array}{|c|} \hline \text{R : 081} \\ \hline \end{array} + \begin{array}{|c|} \hline \text{G : 035} \\ \hline \end{array} + \begin{array}{|c|} \hline \text{B : 000} \\ \hline \end{array} = \begin{array}{|c|} \hline \text{Yellow} \\ \hline \end{array} \\
 Y = 0.299R + 0.587G + 0.114B \\
 Cb = -0.1687R + -0.3313G + 0.15B + 128 \\
 Cr = 0.5R - 0.4187G - 0.0813B + 128 \\
 \begin{array}{|c|} \hline \text{Y : 147} \\ \hline \end{array} + \begin{array}{|c|} \hline \text{Cb : 086} \\ \hline \end{array} + \begin{array}{|c|} \hline \text{Cr : 141} \\ \hline \end{array} = \begin{array}{|c|} \hline \text{Yellow} \\ \hline \end{array}
 \end{array}$$

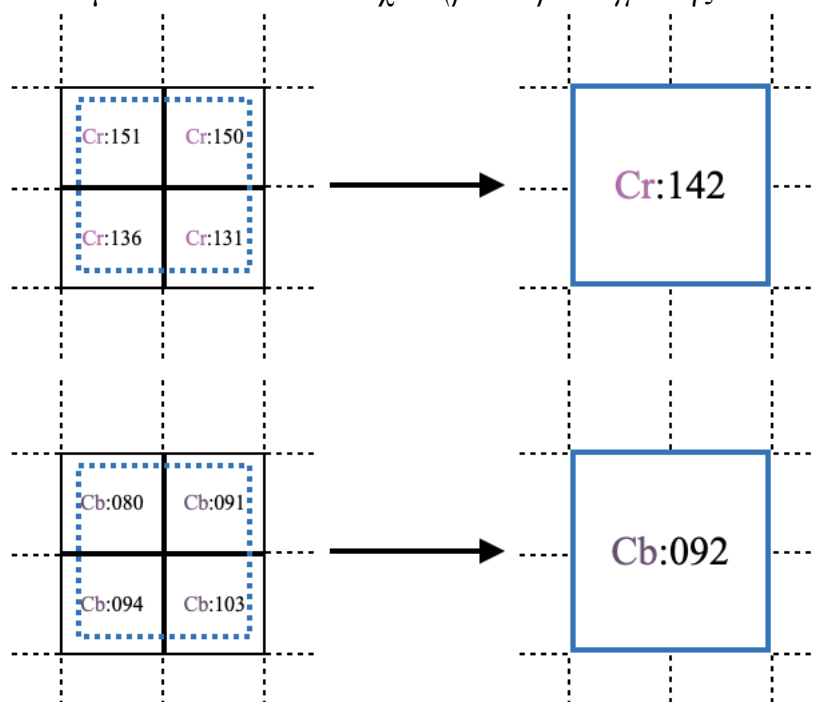
↑ ↓

Εικόνα 1.6 : Υπολογισμός Φωτεινότητας, Μπλε και Κόκκινης Χρωματικότητας μέσω των RGB

1.3.1.2 Υποδειγματοληψία Χρωματικότητας

Το βήμα της υποδειγματοληψίας χρωματικότητας, σε αντίθεση με το πρώτο βήμα, αφαιρεί αρκετή πληροφορία. Όπως πρωτοαναφέρθηκε στην υποενότητα 1.3.1 το ανθρώπινο μάτι είναι καλύτερο στην ανίχνευση της φωτεινότητας και όχι τόσο σε αυτήν των χρωμάτων, ως εκ τούτου σε αυτό το βήμα τα δύο δομικά στοιχεία της κόκκινης χρωματικότητας και της μπλε χρωματικότητας χωρίζονται σε τετράγωνα των 2×2 εικονοστοιχείων. Για κάθε τετράγωνο εικονοστοιχείων υπολογίζεται μια μέση τιμή και έπειτα η κόκκινη και μπλε χρωματικότητα

συρρικνώνονται ώστε το κάθε τετράγωνό τους, των 2×2 εικονοστοιχείων, να αντιπροσωπεύει πλέον μόνο ένα εικονοστοιχείο (βλ. παράδειγμα της Εικόνας 1.7).



Εικόνα 1.7 : Υπολογισμός Μέσης Τιμής της Μπλε και Κόκκινης Χρωματικότητας

Αυτό έχει ως αποτέλεσμα η πληροφορία την οποία το ανθρώπινο μάτι δεν αντιλαμβάνεται αρκετά λεπτομερώς, δηλαδή την κόκκινη και μπλε χρωματικότητα, να συρρικνώνεται στο $1/4$ της αρχικής της διάστασης, ενώ ταυτόχρονα η διάσταση της φωτεινότητας μένει ως έχει (Εικόνα 1.8). Επομένως, ο αλγόριθμος JPEG μέσα σε δύο βήματα μείωσε κατά το ήμισυ από πλευράς διαστάσεων την εικόνα προς συμπίεση.



$$\text{Πριν : } 1 + 1 + 1 = 3$$

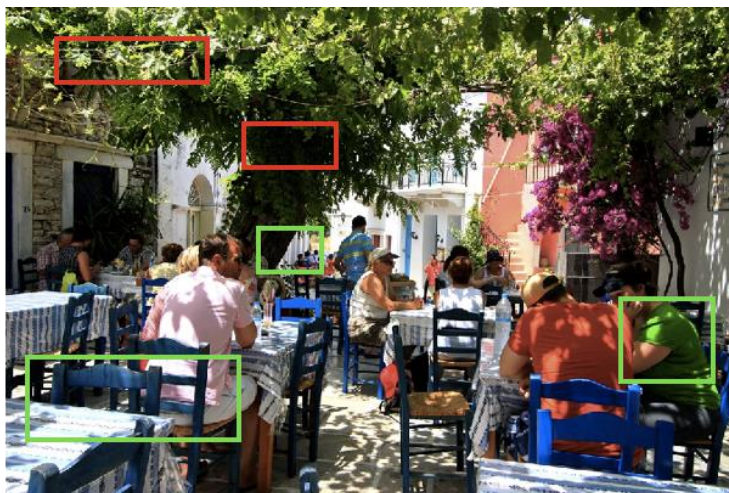
$$\text{Μετά : } \frac{1}{4} + \frac{1}{4} + 1 = 1,5$$

Εικόνα 1.8 : Συρρίκνωση (μόνο) της Μπλε και Κόκκινης Χρωματικότητας

Να σημειωθεί ότι στο επόμενο βήμα ο μετασχηματισμός DCT πρέπει να λάβει την εικόνα στις αρχικές της διαστάσεις, οπότε όταν η εικόνα επανασυναρμολογείται από τα Y, Cb και Cr δομικά της στοιχεία, τα Cb και Cr υφίστανται επανακλιμάκωση (rescaling process) ούτως ώστε να ταιριάζουν στο μέγεθος του Y δομικού στοιχείου. Συνεπώς όταν οι τιμές για το RGB μοντέλο επαναυπολογίζονται βάσει του τροποποιημένου πια YCbCr μοντέλου πρόκειται να είναι διαφορετικές σε σχέση με αυτές που υπήρχαν στο πρώτο βήμα, εξού και η απώλεια πληροφορίας.

1.3.1.3 DCT

Ο διακριτός μετασχηματισμός συνημιτόνου, το τρίτο βήμα του αλγορίθμου JPEG, μετασχηματίζει την πληροφορία της εικόνας από τον χώρο του χρόνου (time domain) στον χώρο των συχνοτήτων (frequency domain) και εκμεταλλεύεται το γεγονός ότι το ανθρώπινο μάτι είναι αρκετά αδύναμο στο να αντιλαμβάνεται τις υψηλές συχνότητες σε μια εικόνα. [7] Πρακτικά, για το παρακάτω παράδειγμα (Εικόνα 1.9), αυτό σημαίνει πως το μάτι αντιλαμβάνεται αποτελεσματικά τις άκρες του δέντρου, των τραπεζιών και των κτηρίων καθώς και τα περιγράμματα από τις σιλουέτες των ανθρώπων. Όμως, στην προσπάθειά του να διαχωρίσει χρωματικά δεδομένα υψηλών συχνοτήτων όπως οι διάφορες αποχρώσεις ήσκιου που δημιουργεί το δέντρο ή να ξεχωρίσει ένα φύλλο μέσα από μια συστάδα φύλλων πέφτει αρκετά έξω.



Εικόνα 1.9 : Εύκολα σημεία αντίληψης (πράσινο πλαίσιο - χαμηλές συχνότητες) και δύσκολα σημεία αντίληψης (κόκκινο πλαίσιο - υψηλές συχνότητες) από το μάτι

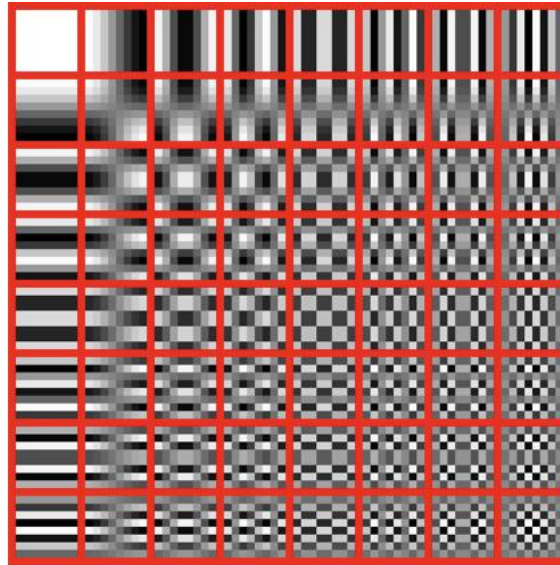
Στην ουσία ο διακριτός μετασχηματισμός συνημιτόνου αναλύει όλη την εικόνα προσπαθώντας να βρει περιοχές με υψηλή συχνότητα εναλλασσόμενης φωτεινότητας και χρωματικότητας και να τις αφαιρέσει. Όλη αυτή η διαδικασία της ανίχνευσης και αφαίρεσης λαμβάνει χώρα και στα τρία δομικά στοιχεία (Y-Cb-Cr) της εικόνας.

Αρχικά ο μετασχηματισμός χωρίζει την εικόνα σε τετράγωνα των 8×8 εικονοστοιχείων ούτως ώστε να είναι χωρισμένη σε τετράγωνα των 64 εικονοστοιχείων, οι τιμές των οποίων κυμαίνονται από το 0 έως το 255. Στη συνέχεια μετατοπίζει κάθε τιμή αφαιρώντας από κάθε μία το 128 με αποτέλεσμα το εύρος να κυμαίνεται πλέον από το -128 έως το 127, με τη μικρότερη τιμή να αντιπροσωπεύει το μαύρο χρώμα και τη μεγαλύτερη το λευκό (βλ. παράδειγμα της Εικόνας 1.10).

037	036	037	042	047	050	055	061	-091	-092	-091	-086	-081	-078	-073	-067
037	037	038	044	049	052	054	057	-091	-091	-090	-084	-079	-076	-074	-071
038	039	044	052	056	057	056	056	-090	-089	-084	-076	-072	-071	-072	-072
037	044	056	065	068	067	062	059	-091	-084	-072	-063	-060	-061	-066	-069
043	057	072	083	085	080	073	066	-085	-071	-056	-045	-043	-048	-055	-062
054	057	097	105	102	093	084	077	-074	-071	-031	-023	-026	-035	-044	-051
072	101	117	116	111	105	095	085	-056	-027	-011	-012	-017	-023	-033	-043
091	115	124	121	116	108	100	092	-037	-013	-004	-007	-012	-020	-028	-036

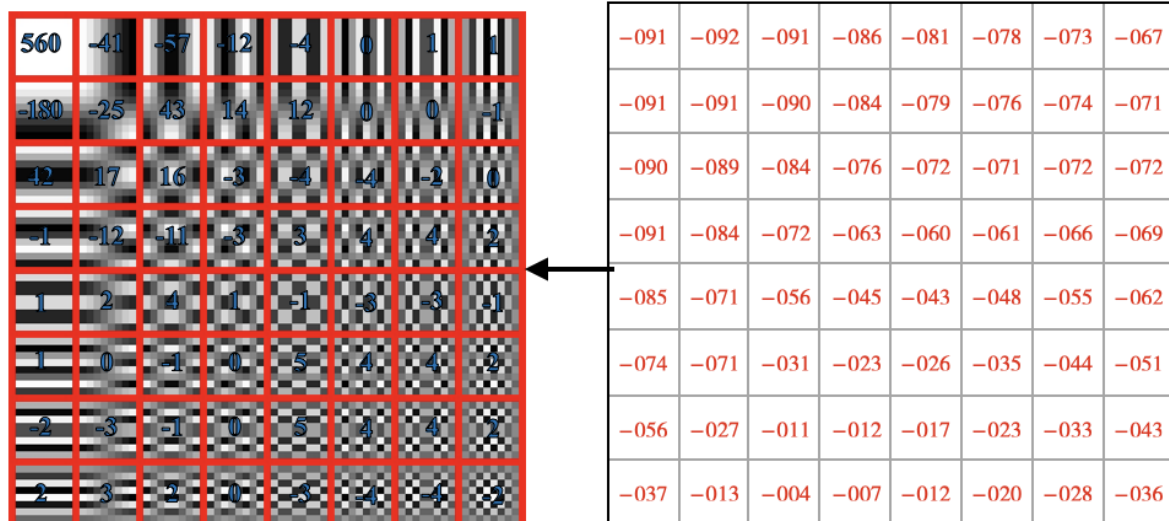
Εικόνα 1.10 : Μετατόπιση του εύρους τιμών των πίξελ μιας εικόνας από το [0,255] στο [-128,127]

Τέλος, προσπαθεί να ανακατασκευάσει κάθε τετράγωνο των 64 εικονοστοιχείων έχοντας ως βάση 64 συχνότητες αποσύνθεσης (Εικόνα 1.11).



Εικόνα 1.11 : Θεμελιώδεις Συχνότητες (βάση αποσύνθεσης)

Κατ'αυτόν τον τρόπο κάθε τιμή των εικονοστοιχείων των τετραγώνων της εικόνας μετασχηματίζεται σε έναν αριθμό (σταθερά - συντελεστής DCT) ο οποίος υποδεικνύει πόσες φορές χρησιμοποιήθηκε η συγκεκριμένη συχνότητα αποσύνθεσης (βλ. παράδειγμα της Εικόνας 1.12) ώστε να ξαναχτιστεί το κάθε τετράγωνο της εικόνας. Ο μετασχηματισμός DCT είναι μια αντιστρέψιμη διαδικασία, οπότε δεν υπάρχει απώλεια πληροφορίας.



Εικόνα 1.12 : Συντελεστές DCT της εικόνας (αριστερός πίνακας)

Οι συντελεστές DCT δηλώνουν τη σχετική βαρύτητα κάθε συχνότητας που συμβάλλει στη σύνθεση του διακριτού σήματος της εικόνας. Το μεγαλύτερο τμήμα της ενέργειας του σήματος (σημαντικότερη πληροφορία της εικόνας) εντοπίζεται στις χαμηλότερες συχνότητες (πάνω αριστερό μέρος του πίνακα της Εικόνας 1.12), ενώ οι συντελεστές στις τελευταίες θέσεις

(κάτω δεξιό μέρος του πίνακα της Εικόνας 1.12) αντιπροσωπεύουν τις υψηλότερες συχνότητες οι οποίες μπορούν να αγνοηθούν. Οι εξισώσεις που διέπουν τον DCT είναι οι εξής:

$$F(u, v) = \frac{1}{4} C(u) C(v) \sum_{x=0}^7 \sum_{y=0}^7 f(x, y) \cos \frac{(2x+1)u\pi}{16} \cos \frac{(2y+1)v\pi}{16} \quad (1.4)$$

όπου,

u η οριζόντια χωρική συχνότητα (spatial frequency)

v η κάθετη χωρική συχνότητα

$$C(u) = \begin{cases} \frac{1}{\sqrt{2}} & \text{αν } u = 0 \\ 1 & \text{αλλιώς} \end{cases} \quad \text{και} \quad C(v) = \begin{cases} \frac{1}{\sqrt{2}} & \text{αν } v = 0 \\ 1 & \text{αλλιώς} \end{cases}$$

οι παράγοντες κανονικοποίησης προκειμένου ο μετασχηματισμός να είναι ορθοκανονικός.

$f(x, y)$ η τιμή του εικονοστοιχείου στις συντεταγμένες (x, y)

$F(u, v)$ ο DCT συντελεστής στις συντεταγμένες (x, y)

$$f(x, y) = \frac{1}{4} \sum_{x=0}^7 \sum_{y=0}^7 C(u) C(v) \cos \frac{(2x+1)u\pi}{16} \cos \frac{(2y+1)v\pi}{16} \quad (1.5)$$

όπου,

x το εικονοστοιχείο της αντίστοιχης σειράς

y το εικονοστοιχείο της αντίστοιχης στήλης

$f(x, y)$ η ανακατασκευασμένη τιμή του εικονοστοιχείου στις συντεταγμένες (x, y)

1.3.1.4 Κβάντιση

Η κβάντιση, το τέταρτο βήμα του αλγορίθμου, λαμβάνει τα 8×8 τετράγωνα σταθερών (αποτέλεσμα μετασχηματισμού DCT) από τα οποία αποτελείται η εικόνα και διαιρεί κάθε τιμή των εικονοστοιχείων τους με την αντίστοιχη τιμή από έναν πίνακα ονόματι πίνακας κβάντισης και στη συνέχεια στρογγυλοποιεί κάθε αποτέλεσμα της διαίρεσης στον πλησιέστερο ακέραιο αριθμό (βλ. παράδειγμα της Εικόνας 1.13).

560	-41	-57	-12	-4	0	1	1
-180	-25	43	14	12	0	0	-1
-42	17	16	-3	-4	4	-2	0
-1	-12	-11	3	3	4	4	2
1	2	4	1	-1	3	-3	1
1	0	-1	0	5	4	4	2
-2	-3	-1	0	5	4	4	2
2	3	2	0	-3	-4	-4	2

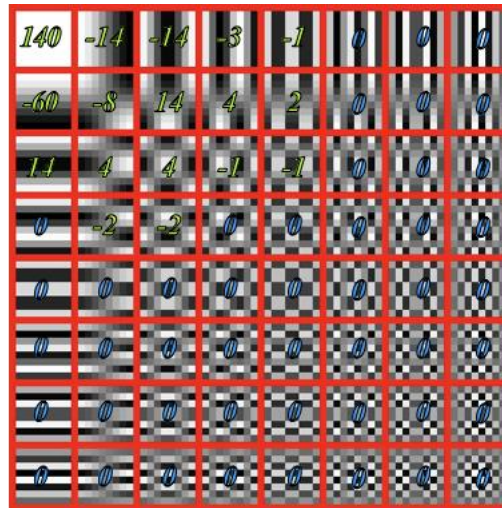
÷

04	03	04	04	04	06	11	15
03	03	03	04	05	08	14	19
03	04	04	05	08	12	16	20
04	05	06	07	12	14	18	20
06	06	09	11	14	17	21	23
09	12	12	18	23	22	25	21
11	13	15	17	21	23	25	21
13	12	12	13	16	19	21	21

Πίνακας Κβάντισης

Εικόνα 1.13 : Διαίρεση του πίνακα συντελεστών DCT με τον Πίνακα Κβάντισης

Ο πίνακας κβάντισης, κάθε κελί του οποίου περιέχει ένα βήμα κβάντισης, έχει μεγάλους αριθμούς στο κάτω δεξιό μέρος του, όπου εντοπίζονται τα δεδομένα υψηλών συχνοτήτων που το μάτι δύσκολα αντιλαμβάνεται και μικρότερους αριθμούς στο πάνω αριστερό μέρος του, όπου εντοπίζονται τα πιο ευδιάκριτα μοτίβα. Αυτή η κατανομή αριθμών στον πίνακα κβάντισης έχει ως αποτέλεσμα, μετά τη διαίρεση και την προαναφερθείσα στρογγυλοποίηση, οι περισσότερες τιμές των εικονοστοιχείων όλων των τετραγώνων της εικόνας να γίνουν μηδέν (βλ. Εικόνα 1.14). Το μηδέν στην ουσία πρόκειται για πληροφορία που μόλις αφαιρέθηκε, αλλά η αφαίρεσή της δεν έγινε αντιληπτή από το ανθρώπινο μάτι.

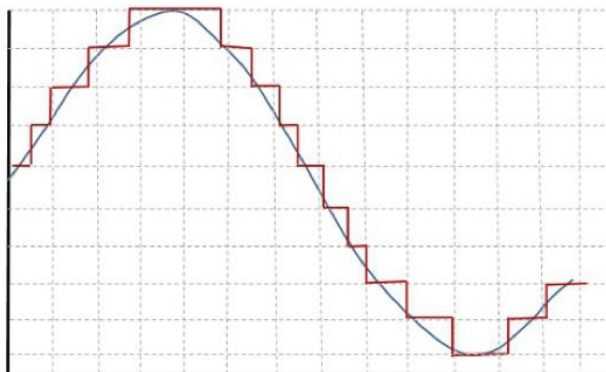


Εικόνα 1.14 : Κβαντοποιημένοι συντελεστές DCT

Κατ' ουσίαν, αυτό που επιτεύχθηκε μέσω των βημάτων του DCT μετασχηματισμού και της κβάντισης είναι η μετατροπή όλων των τιμών των εικονοστοιχείων όλων των τετραγώνων της εικόνας σε ολίγους μόνο αριθμούς και σε πολλά μηδενικά.

1.3.1.4.1 Ομοιόμορφη και Μη Ομοιόμορφη Κβάντιση

Εν γένει, η κβάντιση είναι η διαδικασία η οποία δέχεται ένα σήμα εισόδου και αντιστοιχίζει πολλές από τις τιμές του σε μία μόνο τιμή κατά την έξοδο. Μια διαδικασία μη αντιστρεπτή καθώς και η αιτία που υπάρχει απώλεια πληροφορίας στο σήμα εξόδου. [8]



Εικόνα 1.15 : Απλό Παράδειγμα Κβάντισης - Είσοδος (μπλε), Κβαντισμένη Έξοδος (κόκκινο)

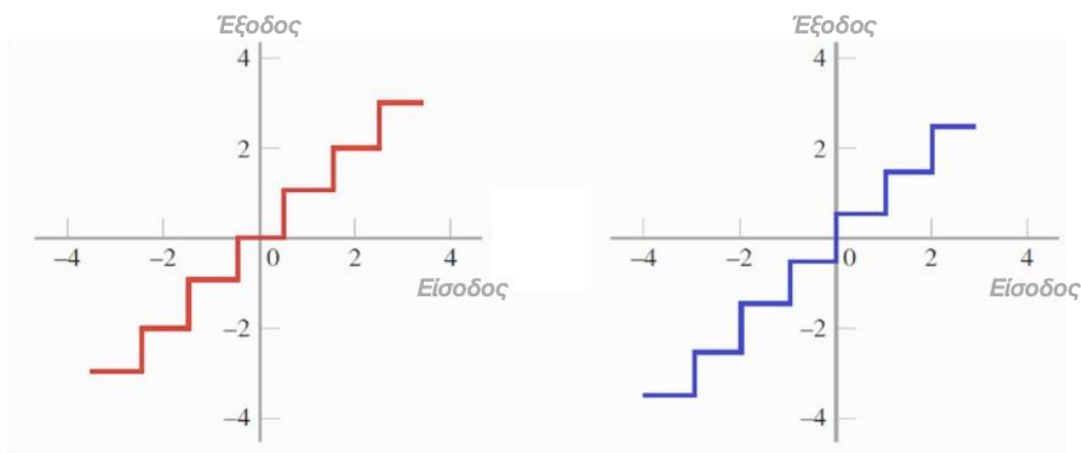
Όπως φαίνεται και από το παραπάνω παράδειγμα (Εικόνα 1.15), το σήμα εισόδου που είναι η μπλε γραμμή, μετά την κβάντιση αντιστοιχίστηκε στην κόκκινη γραμμή που είναι το σήμα εξόδου. Παρατηρείται δηλαδή μια διαφορά μεταξύ σήματος εξόδου και εισόδου. Η διαφορά αυτή μεταξύ μιας τιμής εισόδου και της κβαντισμένης τιμής της ονομάζεται σφάλμα κβάντισης.

Δύο είναι τα βασικά είδη κβάντισης, η ομοιόμορφη κβάντιση (uniform quantization), η οποία είναι η κβάντιση που χρησιμοποιείται από τους αλγορίθμους JPEG και JPEG2000, και η μη ομοιόμορφη κβάντιση (non-uniform quantization). Οι περισσότεροι ομοιόμορφοι κβαντιστές προσημασμένων δεδομένων εισόδου ταξινομούνται σε 2 τύπους, τον midriser και τον midtread (ελεύθερη αντίστοιχη απόδοση «μεσορίχι» και «μεσοπάτημα»). Η ορολογία βασίζεται παρομοιάζοντας το σήμα εξόδου του κβαντιστή σαν μια μικρή σκάλα αλλά και σε αυτό που συμβαίνει στην περιοχή γύρω από το μηδέν. Ο midriser κβαντιστής δεν εμπεριέχει την τιμή μηδέν ως μια από τις κβαντισμένες τιμές του (μοιάζει με ρίχι σκάλας), ενώ ο midtread κβαντιστής την εμπεριέχει (μοιάζει με πάτημα σκάλας). Ο midtread σχετίζεται με την πράξη της στρογγυλοποίησης και διέπεται από την εξής εξίσωση :

$$Q(x) = \Delta \cdot \left\lfloor \frac{x}{\Delta} + \frac{1}{2} \right\rfloor \quad (1.6)$$

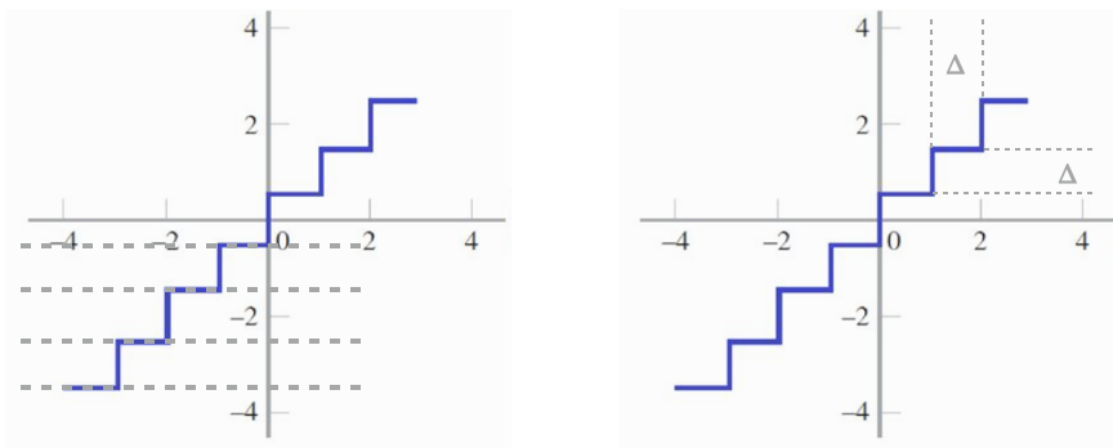
όπου x το σήμα εισόδου και Δ το βήμα κβάντισης, ενώ ο midriser σχετίζεται με την πράξη της περικοπής και διέπεται από την εξής εξίσωση :

$$Q(x) = \Delta \cdot \left(\left\lfloor \frac{x}{\Delta} \right\rfloor + \frac{1}{2} \right) \quad (1.7)$$



Εικόνα 1.16 : Σήμα εξόδου Κβαντιστή Midtread (κόκκινο) και Κβαντιστή Midriser (μπλε)

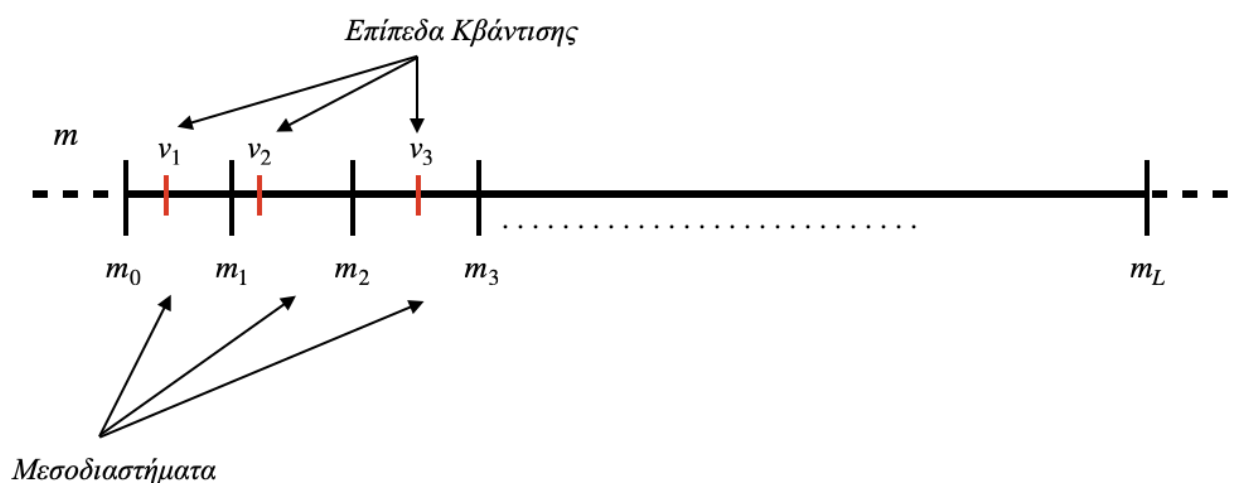
Η ποιότητα της εξόδου ενός κβαντιστή εξαρτάται από τον αριθμό των επιπέδων κβάντισης (όσα περισσότερα επίπεδα υπάρχουν, τόσο περισσότερο η έξοδος προσεγγίζει το σήμα εισόδου). Ο χώρος μεταξύ δύο γειτονικών επιπέδων κβάντισης είναι το μέγεθος του βήματος κβάντισης Δ . Στο παρακάτω παράδειγμα (Εικόνα 1.17) οι χοντρές παύλες αντιπροσωπεύουν τα επίπεδα κβάντισης (αριστερό διάγραμμα), ενώ οι λεπτές το μέγεθος του βήματος κβάντισης (δεξιό διάγραμμα).



Εικόνα 1.17 : Επίπεδα Κβάντισης (αριστερά) - Μέγεθος Βήματος Κβάντισης (δεξιά)

Οι διαφορές μεταξύ της ομοιόμορφης και μη ομοιόμορφης κβάντισης είναι στο βήμα κβάντισης. Το μέγεθος στο βήμα της ομοιόμορφης κβάντισης είναι ίδιο, ενώ στη μη ομοιόμορφη αλλάζει και προσαρμόζεται με τέτοιο τρόπο ώστε το κβαντισμένο σήμα να προσεγγίζει καλύτερα το σήμα εισόδου, μειώνοντας έτσι το σφάλμα κβάντισης. Ένας από τους πιο γνωστούς μη ομοιόμορφους κβαντιστές είναι ο κβαντιστής Lloyd-Max (Λόιντ-Μαξ). [9] [10]

Έστω ότι το σήμα εισόδου ενός κβαντιστή είναι το σήμα m . Ένας οποιοσδήποτε κβαντιστής χαρακτηρίζεται από τα μεσοδιαστήματα (intervals) και τα επίπεδα κβάντισης που θα επιλέξει (Εικόνα 1.18).



Εικόνα 1.18 : Μεσοδιαστήματα και Επίπεδα Κβάντισης ενός Κβαντιστή

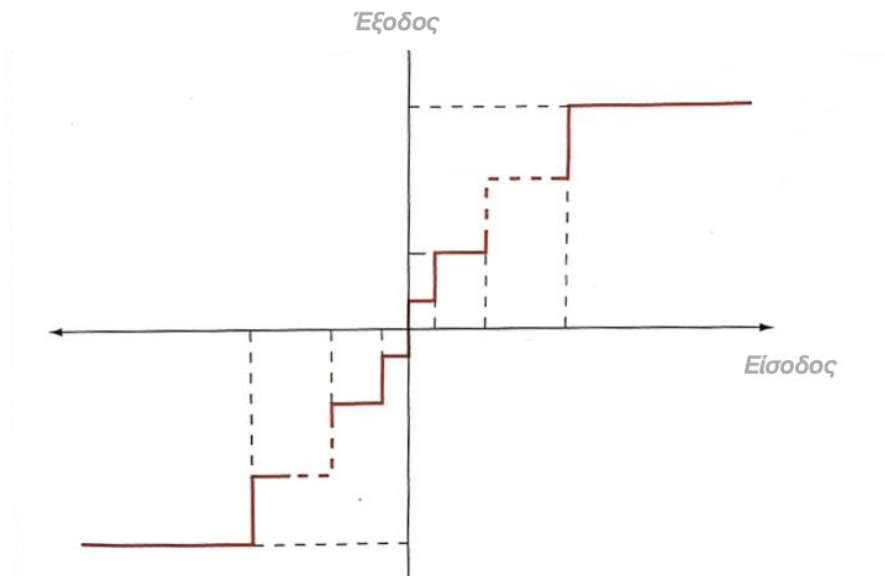
Ένας ομοιόμορφος κβαντιστής, έχοντας προκαθορισμένα επίπεδα κβάντισης, τοποθετεί τα επίπεδα κβάντισης στο μέσο των μεσοδιαστημάτων. Σε αντίθεση με τον ομοιόμορφο κβαντιστή, ο κβαντιστής Lloyd-Max επιλέγει τη βέλτιστη θέση για τα επίπεδα κβάντισης, και κατ'επέκταση τα βέλτιστα μεσοδιαστήματα, λαμβάνοντας υπόψιν του τη συνάρτηση πυκνότητας πιθανότητας $p(m)$ του σήματος εισόδου m . Ως κριτήριο βελτιστοποίησης ο κβαντιστής έχει την ελαχιστοποίηση του μέσου τετραγωνικού σφάλματος κβαντισμού, δηλαδή την παράμετρο $E[(m_i - t_i)^2]$, όπου t_i η κβαντισμένη τιμή του σήματος m εντός του διαστήματος $(m_i, m_{i+1}]$. Οι συνθήκες για το ελάχιστο σφάλμα είναι οι εξής:

$$\int_{-m_{i-1}}^{m_i} (m - t_i) p(m) \quad i = 1, 2, \dots, \frac{L}{2} \quad (1.8)$$

$$m_i = \begin{cases} 0 & i = 0 \\ \frac{t_i + t_{i+1}}{2} & i = 1, 2, \dots, \frac{L}{2} - 1 \\ \infty & i = \frac{L}{2} \end{cases} \quad (1.9)$$

$$m_{-1} = -m_i \quad t_{-1} = -t_i \quad (1.10)$$

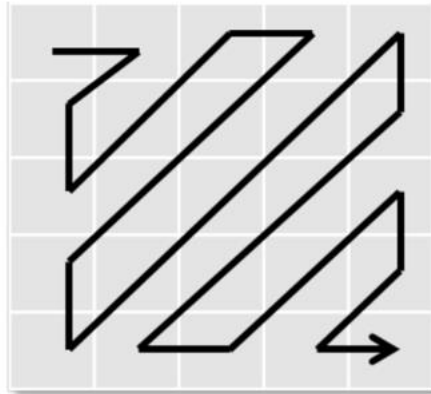
Για κάθε τιμή της παραμέτρου L , οι τιμές των μεγεθών m_i και t_i που ικανοποιούν τις τρεις παραπάνω εξισώσεις είναι βέλτιστες από την οπτική γωνία του μέσου τετραγωνικού σφάλματος. Μια τυπική συνάρτηση μη ομοιόμορφης κβάντισης είναι η εξής (Εικόνα 1.19).



Εικόνα 1.19 : Σήμα εξόδου Μη Ομοιόμορφου Κβαντιστή

1.3.1.5 Συμπίεση Εντροπίας

Μετά το βήμα κβάντισης πολλοί από τους συντελεστές DCT έχουν μηδενιστεί. Για τη συμπίεση αυτών των δεδομένων ο αλγόριθμος JPEG συνήθως χρησιμοποιεί κάποιον αλγόριθμο κωδικοποίησης εντροπίας (για ορισμό βλ. Εξ. 1.1), όπως ο αλγόριθμος Run Length Encoding (RLE) ή ο αλγόριθμος Huffman (Χάφμαν). Η σάρωση των συντελεστών DCT για την κωδικοποίησή τους γίνεται με μορφή ζιγκ-ζαγκ (βλ. Εικόνα 1.20).



Εικόνα 1.20 : Κωδικοποίηση ζιγκ-ζαγκ

1.3.2 JPEG2000

Ο αλγόριθμος JPEG2000 είναι μια εξελιγμένη μέθοδος συμπίεσης που έχει ως βάση την ίδια θεμελιώδη αρχή με τον αλγόριθμο JPEG με τη μόνη διαφορά να είναι η χρήση τού DWT μετασχηματισμού αντί του DCT και η χρήση της Αριθμητικής Κωδικοποίησης (Arithmetic Coding) για την κωδικοποίηση της εντροπίας. [11] Η αιτία δημιουργίας του αλγορίθμου JPEG2000 ήταν για την αντιμετώπιση των περιορισμών που χαρακτηρίζουν τον αλγόριθμο JPEG. Κάποιοι από αυτούς τους περιορισμούς είναι η χαμηλή απόδοση όταν η εφαρμογή του αλγορίθμου γίνεται σε δυαδικές (bi-level) εικόνες, τεχνητές εικόνες (computer generated) καθώς και σε εικόνες με χαμηλό ρυθμό μετάδοσης δεδομένων (π.χ. bit-rate < 0.25 bpp).

Ο αλγόριθμος JPEG2000, βασισμένος στον αλγόριθμο JPEG, ακολουθεί πέντε βασικά βήματα για τη συμπίεση μιας εικόνας, (α) μετατροπή του χρωματικού χώρου (colour space conversion), (β) κατάτμηση (tiling), (γ) DWT, (δ) κβάντιση (quantization) και (ε) περαιτέρω συμπίεση εντροπίας.

1.3.2.1 Κατάτμηση

Μετά το βήμα της μετατροπής του χρωματικού χώρου, που εκτελείται όπως και στον αλγόριθμο JPEG, η εικόνα χωρίζεται στα λεγόμενα πλακίδια (tiles), δηλαδή σε ορθογώνιες περιοχές κάθε μία από τις οποίες μετασχηματίζονται και κωδικοποιούνται ξεχωριστά. Τα πλακίδια μπορεί να είναι οποιουδήποτε μεγέθους και είναι επίσης δυνατόν ολόκληρη η εικόνα να θεωρηθεί ως ένα πλακίδιο. Μετά την επιλογή του μεγέθους όλα τα πλακίδια θα έχουν το ίδιο μέγεθος. Ο διαχωρισμός της εικόνας σε περιοχές-πλακίδια είναι επωφελής καθώς ο αποκωδικοποιητής θα χρειαστεί λιγότερη μνήμη για να αποκωδικοποιήσει την εικόνα και επιπλέον θα μπορεί να επιλέξει την αποκωδικοποίηση ορισμένων μόνο πλακιδίων για να επιτύχει μερική αποκωδικοποίηση της εικόνας. Το μειονέκτημα αυτής της προσέγγισης είναι ότι η ποιότητα της εικόνας μειώνεται λόγω του σηματοθορυβικού λόγου. Η χρήση πολλών πλακιδίων μπορεί να δημιουργήσει το φαινόμενο της τετραγωνοποίησης (blocking effect) παρόμοιο με αυτό που εμφανίζεται στον αλγόριθμο JPEG. [11]

1.3.2.2 DWT

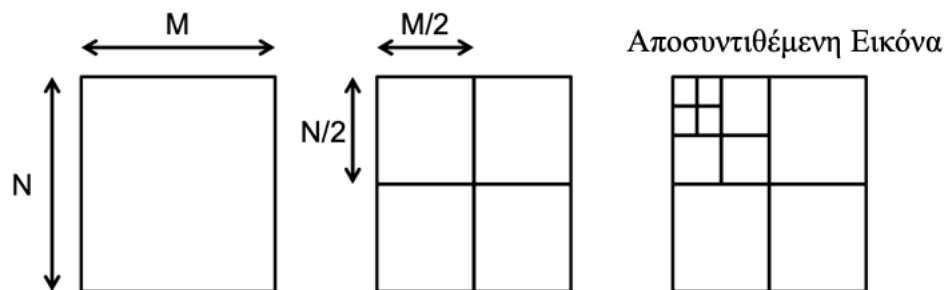
Ένα σήμα λέγεται στατικό (stationary) αν σε κάθε χρονική στιγμή το φάσμα συχνοτήτων του παραμένει το, ενώ μη στατικό (non-stationary) όταν το φάσμα συχνοτήτων στον χρόνο, δηλαδή οι διάφορες συχνότητες εμφανίζονται σε διαφορετικές χρονικές περιόδους, όπως για παράδειγμα σε μια ψηφιακή εικόνα. Το διακριτό σήμα μιας ψηφιακής εικόνας είναι μη στατικό γιατί εμφανίζει υψηλές συχνότητες στα όρια των περιοχών με απότομη αλλαγή της χρωματικής πληροφορίας και χαμηλές συχνότητες στις «συμπαγείς» περιοχές με ομαλές αλλαγές της χρωματικής πληροφορίας. Βάσει αυτών, η ανάγκη εφαρμογής κατάλληλου μετασχηματισμού που να δίνει την πληροφορία χρονικής εξέλιξης του συχνοτικού φάσματος ενός διακριτού σήματος και να επιτρέπει να προσαρμόζεται το επίπεδο συμπίεσης ανάλογα με την περιοχή της εικόνας είναι επιτακτική.

Ο μετασχηματισμός κυματιδίων διαχωρίζει την εικόνα σε υποζώνες συχνοτήτων (subbands) για να προσδιοριστεί το συχνοτικό φάσμα στη συγκεκριμένη περιοχή. Αυτή η ανάλυση προσφέρει σημαντικά επίπεδα ευελιξίας στη συμπίεση της εικόνας σε σχέση με τον αλγόριθμο JPEG. Ο μετασχηματισμός κυματιδίων προσφέρει δυνατότητα ρύθμισης :

- Στις υψηλές συχνότητες : Υψηλή χρονική ανάλυση και μικρή συχνοτική ανάλυση
- Στις χαμηλές συχνότητες : Χαμηλή χρονική ανάλυση και υψηλή συχνοτική ανάλυση

Η τεχνική αυτή είναι κατάλληλη για εικόνες όπου το σήμα έχει μεγάλες διαφορές σε μικρές χρονικές περιόδους δηλαδή σε μικρή έκταση όπως στα όρια διαφορετικών περιοχών της εικόνας και μικρές διαφορές σε μεγάλες χρονικές περιόδους δηλαδή σε μεγάλες εκτάσεις, όπως σε συνεχείς περιοχές της εικόνας με παρόμοια πληροφορία χρώματος.

Η εφαρμογή του DWT γίνεται με μία σειρά χαμηλοπερατών και υψηλοπερατών φίλτρων, εξού και η διαίρεση της εικόνας σε υποζώνες συχνοτήτων. Κατά την εφαρμογή του DWT σε μια εικόνα, η διαίρεση σε υποζώνες συχνοτήτων γίνεται πρώτα κατά σειρά και μετά κατά στήλη. Παραδείγματος χάριν, σε μια εικόνα $N \times M$ διαστάσεων κάθε σειρά φιλτράρεται και μετά το σήμα εξόδου υποδειγματοληπτείται για να δημιουργηθούν δύο $N \times \left(\frac{M}{2}\right)$ εικόνες. Στη συνέχεια φιλτράρεται και υποδειγματοληπτείται κάθε στήλη ώστε να δημιουργηθούν τέσσερις $\left(\frac{N}{2}\right) \times \left(\frac{M}{2}\right)$ υπο-εικόνες (βλ. *Εικόνα 1.21*).



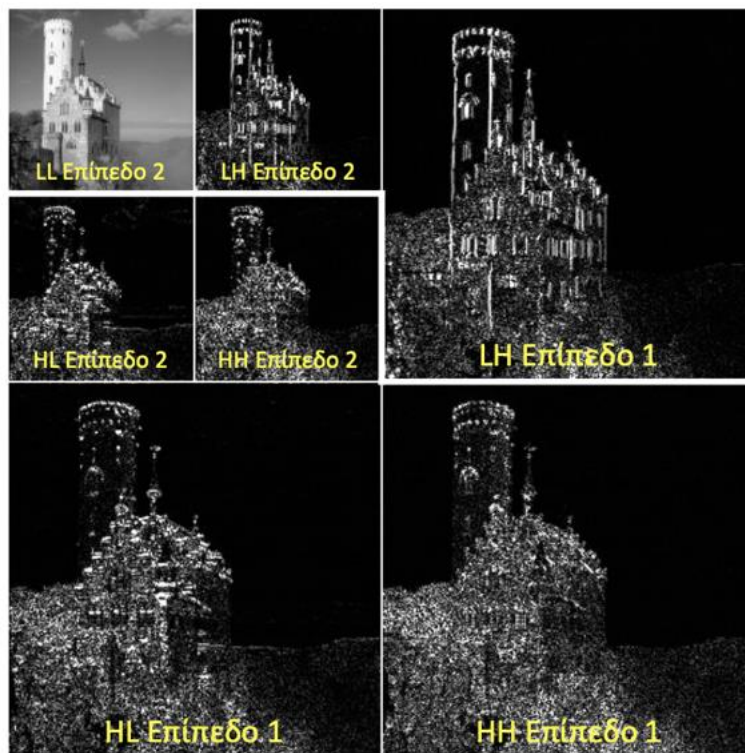
Εικόνα 1.21 : Στάδια αποσύνθεσης εικόνας με εφαρμογή DWT

Η εφαρμογή των φίλτρων στις δύο διαστάσεις διαιρεί την εικόνα σε περιοχές συχνοτήτων (βλ. *Εικόνες 1.22 και 1.23*). Πιο συγκεκριμένα :

- LL → Χαμηλή Οριζόντια και Χαμηλή Κάθετη (μόνο η LL υποζώνη μετασχηματίζεται περαιτέρω)
- HL → Υψηλή Οριζόντια και Χαμηλή Κάθετη
- LH → Χαμηλή Οριζόντια και Υψηλή Κάθετη
- HH → Υψηλή Οριζόντια και Υψηλή Κάθετη



Εικόνα 1.22 : Υποζώνες συχνότητας εικόνας με εφαρμογή DWT



Εικόνα 1.23 : Οπτικό αποτέλεσμα εφαρμογής DWT

Στην επεξεργασία εικόνας, μεταξύ των πολλών ειδών μετασχηματισμού κυματιδίων, έχουν χρησιμοποιηθεί κυρίως τα κυματίδια Haar, τα κυματίδια 5/3 και τα κυματίδια 9/7. Εν γένει, τα κυματίδια αποσυνθέτουν μια εικόνα σε διαφορετικές κλίμακες χρησιμοποιώντας μια αρχιτεκτονική πυραμιδικού αλγορίθμου. Η αποσύνθεση γίνεται κατά μήκος της κάθετης και της οριζόντιας κατεύθυνσης και διατηρεί σταθερό τον αριθμό των εικονοστοιχείων που απαιτούνται για την περιγραφή της εικόνας. Η συνάρτηση κυματιδίων δημιουργείται από διαστολές και μετατοπίσεις μιας συνάρτησης ψ η οποία ορίζεται ως εξής :

$$\psi_{a,b}(x) = |a|^{-\frac{1}{2}} \psi\left(\frac{x-b}{a}\right) \quad (1.11)$$

όπου,

$(a, b) \in \mathbb{R}$ και $a \neq 0$.

Τα κυματίδια υψηλής συχνότητας αντιστοιχούν σε $a < 1$, ενώ τα κυματίδια χαμηλής συχνότητας σε $a > 1$.

1.3.2.3 Κβάντιση και Συμπίεση Εντροπίας

Μετά την εφαρμογή τού DWT οι συντελεστές κυματιδίων υφίστανται βαθμωτή κβάντιση (Scalar Quantization) για να μειωθεί ο αριθμός των bit που χρειάζονται για την αναπαράστασή τους, εις βάρος όμως της τελικής οπτικής ποιότητας. Η παράμετρος που μπορεί να τροποποιηθεί για να οριστεί η τελική ποιότητα είναι το βήμα κβάντισης. Όσο μεγαλύτερο είναι το βήμα, τόσο μεγαλύτερη είναι η συμπίεση και η απώλεια ποιότητας. Όσο μικρότερο είναι το βήμα, τόσο καλύτερα θα προσεγγίζει η κβαντισμένη ποσότητα την αρχική. Το αποτέλεσμα της κβάντισης είναι ένα σύνολο ακέραιων αριθμών για τη συμπίεση των οποίων ο αλγόριθμος JPEG2000 συνήθως χρησιμοποιεί κάποιον αλγόριθμο κωδικοποίησης εντροπίας, όπως ο αλγόριθμος Αριθμητικής Κωδικοποίησης.

1.4 Ανάγκη Εύρεσης Νέων Μεθόδων Συμπίεσης Εικόνας

Όπως προαναφέρθηκε στην ενότητα 1.1, ζούμε σε μια εποχή όπου η παραγωγή δεδομένων, όπως εικόνες και βίντεο γίνεται με αρκετά ταχείς ρυθμούς. Παρατηρείται από τη μέχρι τώρα βιβλιογραφία ότι τα συστήματα συμπίεσης δεν μπορούν να συμβαδίσουν με τη ραγδαία αύξηση των δεδομένων και ανά 10 χρόνια περίπου εμφανίζεται καινούργιο πρότυπο συμπίεσης, με τα πιο διαδεδομένα στην τρέχουσα εποχή να είναι το JPEG και JPEG2000 για τις εικόνες και το AVC και HEVC για τα βίντεο. Με κάθε νέο πρότυπο ο αλγόριθμος συμπίεσης εικόνων/βίντεο γίνεται όλο και πιο πολύπλοκος αλλά και ενεργειακά απαιτητικός. Επιπρόσθετα, το κοινό μειονέκτημα όλων αυτών των προτύπων συμπίεσης είναι ότι υποβάλλουν σε επεξεργασία τα δεδομένα προς συμπίεση χωρίς καμία γνώση για το οπτικό τους περιεχόμενο δημιουργώντας έτσι τεχνικά σφάλματα στο τελικό οπτικό αποτέλεσμα πράγμα δυσάρεστο για το ανθρώπινο μάτι. Συνεπώς, η αργή εξέλιξη των συστημάτων συμπίεσης σε συνδυασμό με την καθημερινώς τεράστια παραγωγή οπτικών δεδομένων καθιστά αντιληπτό πως η δημιουργία αλγορίθμων συμπίεσης ικανών να επεξεργαστούν με έξυπνο και ενεργειακά αποδοτικό τρόπο όσο το δυνατόν περισσότερη πληροφορία είναι αρκετά αναγκαία.

Ένα σύστημα που μπορεί να ανταποκριθεί στις προαναφερθείσες ανάγκες και απαιτήσεις, σύμφωνα με τη βιολογία, είναι το ανθρώπινο οπτικό σύστημα. Στόχος αυτής της εργασίας είναι να μελετήσει μια νέα μέθοδο συμπίεσης εικόνων εμπνευσμένη από το οπτικό σύστημα. Μελετώντας το σύστημα αυτό ως μια μηχανή επεξεργασίας οπτικών ερεθισμάτων, φαίνεται να είναι ένα αρκετά έξυπνο και αποτελεσματικό μοντέλο προς μίμηση. Κυριότερος λόγος γι'

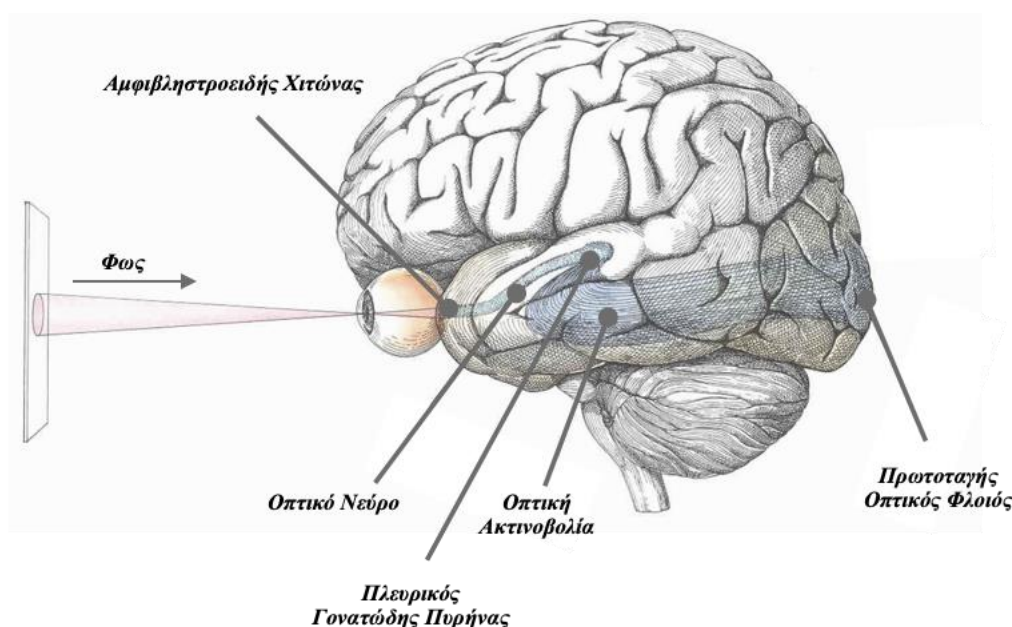
αυτόν τον ισχυρισμό είναι πως ο τρόπος με τον οποίο το οπτικό σύστημα μετασχηματίζει και κωδικοποιεί τα οπτικά ερεθίσματα είναι δυναμικός αλλά και ενεργειακά μη απαιτητικός σε σύγκριση με τα τρέχοντα πρότυπα συμπίεσης. Το οπτικό σύστημα είναι μέρος του νευρωνικού συστήματος, οπότε η προτεινόμενη μέθοδος συμπίεσης του μοντέλου βασίστηκε στη μελέτη των νευρώνων.

Κεφάλαιο 2ο

Εισαγωγή Της Νευροεπιστήμης Στην Επεξεργασία Εικόνων

2.1 Εισαγωγή στο Οπτικό Σύστημα

Το ανθρώπινο οπτικό σύστημα είναι μέρος του κεντρικού νευρικού συστήματος το οποίο αποτελείται από πολλές διαφορετικές περιοχές που συμμετέχουν στην κωδικοποίηση ενός οπτικού ερεθίσματος. Οι πιο σημαντικές από αυτές τις περιοχές είναι ο αμφιβληστροειδής χιτώνας, το οπτικό νεύρο, ο έξω γονατώδης πυρήνας ή πλευρικός γονατώδης πυρήνας (LGN) και ο πρωτοταγής οπτικός φλοιός (V1), μεταξύ άλλων. Ο αμφιβληστροειδής είναι ένα στρώμα ιστού, που καλύπτει την εσωτερική επιφάνεια του ματιού, και είναι υπεύθυνο για τη σύλληψη, μετατροπή και κωδικοποίηση των οπτικών ερεθισμάτων σε μια ακολουθία ηλεκτρικών παλμών (ακολουθία αιχμών - *spike train*). Αυτός ο «κώδικας» αιχμών μεταδίδεται μέσω του οπτικού νεύρου στα κύτταρα του πλευρικού γονατώδους πυρήνα τα οποία συσχετίζουν χωρικά αλλά και χρονικά το σήμα εξόδου και στα δύο μάτια προκειμένου να επιτευχθεί μια τρισδιάστατη αναπαράσταση του χώρου. Τα κύτταρα στον πρωτοταγή οπτικό φλοιό είναι πιο ευαίσθητα στην ανίχνευση προσανατολισμού και κίνησης αλλά και στη διάκριση του σχήματος ή του χρώματος (βλ. Εικόνα 2.1).

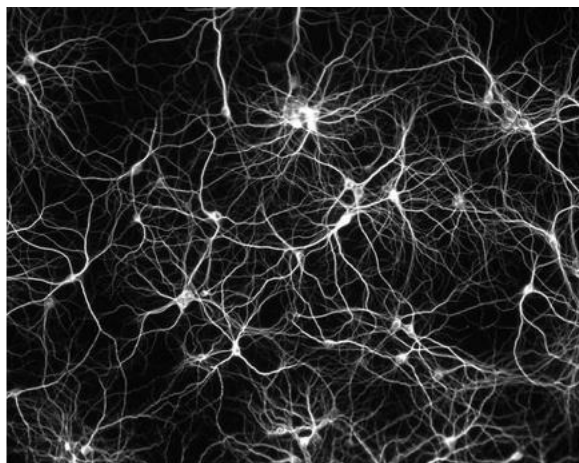


Εικόνα 2.1 : Η διαδρομή του φωτός (οπτικό ερέθισμα) στο οπτικό σύστημα

Κυριότερος σκοπός του οπτικού συστήματος είναι ο μετασχηματισμός του φωτός, που είναι ένα χωροχρονικά μεταβαλλόμενο ερέθισμα, για να εφαρμοστεί και να αναγνωριστεί ορθά στον εγκέφαλο. Αυτός ο μετασχηματισμός λαμβάνει χώρα μέσα στον αμφιβληστροειδή χιτώνα. Ο αμφιβληστροειδής είναι μια πολύπλοκη δομή υπεύθυνη τόσο για την απορρόφηση όσο και για τον μετασχηματισμό του φωτός σε ηλεκτρικούς παλμούς. Αποτελείται από πολυάριθμα κύτταρα τα οποία διαφέρουν μεταξύ τους ως προς το σχήμα αλλά και ως προς τον ρόλο που επιτελούν. [12] [13]

2.2 Νευρώνες: Δομικά Κύτταρα του Νευρικού Συστήματος

Οι στοιχειώδεις μονάδες επεξεργασίας πληροφορίας στο νευρικό σύστημα είναι νευρώνες (ή νευρικά κύτταρα). Όπως παρατηρείται στην *Εικόνα 2.2* οι νευρώνες απαντούν σε διάφορα σχήματα και μεγέθη ενώ διαθέτουν μακριές προεκτάσεις, που θυμίζουν καλωδιώσεις, μέσω των οποίων συνδέονται μεταξύ τους. [14]

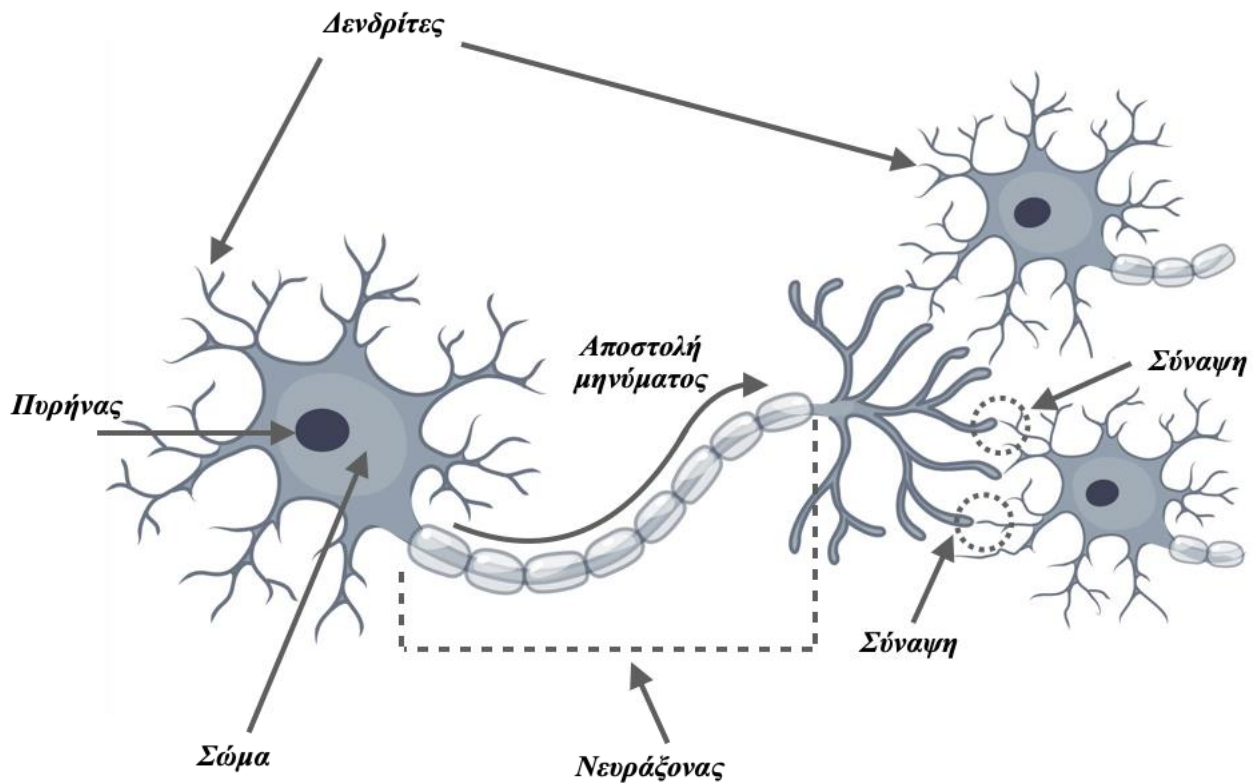


Εικόνα 2.2 : Διασυνδέσεις πολλών νευρώνων

Το κεντρικό νευρικό σύστημα δεν αποτελείται αμιγώς από νευρώνες (αλλά και από τον εγκέφαλο και τον νωτιαίο μυελό), αλλά στο πλαίσιο αυτής της διπλωματικής εργασίας θα γίνει επικέντρωση μόνο στους νευρώνες καθώς αυτού του είδους τα κύτταρα εμπλέκονται άμεσα στην επεξεργασία πληροφοριών και την παραγωγή των ηλεκτρικών παλμών (τις αιχμές). Τα σημαντικότερα από αυτά τα κύτταρα είναι οι φωτοϋποδοχείς (ειδική κατηγορία νευρώνων) καθώς αυτοί είναι που μετασχηματίζουν το οπτικό ερέθισμα σε μια αλληλουχία ηλεκτρικών παλμών. Οι φωτοϋποδοχείς χωρίζονται σε δύο κατηγορίες, τα κωνία, υπεύθυνα για την αντίχνευση των χρωμάτων, και τα ραβδία, υπεύθυνα για τη σύλληψη του φωτός. [13]

Ένας τυπικός νευρώνας μπορεί να χωριστεί σε τρία λειτουργικά διακριτά μέρη, τους δενδρίτες, το σώμα και τον νευράξονα (ή απλώς άξονα) που αναλαμβάνουν τον ρόλο της συλλογής, της επεξεργασίας και της μεταφοράς (σε γειτονικούς νευρώνες) αντίστοιχα. Συγκεκριμένα, οι δενδρίτες συλλέγουν σήματα από άλλους νευρώνες και τα μεταδίδουν στο σώμα το οποίο είναι και το μεγαλύτερο τμήμα του κυττάρου. Το σώμα με τη σειρά του επεξεργάζεται το σήμα-

πληροφορία και παράγει ένα σήμα εξόδου μόνο εάν η συνολική είσοδος που μόλις έλαβε υπερβαίνει ένα ορισμένο όριο (κατώφλι ή αλλιώς ουδός). Το σήμα εξόδου εν συνεχεία μεταφέρεται στον νευράξονα, ο οποίος το μεταδίδει σε άλλους γειτονικούς νευρώνες. [15]



Εικόνα 2.3 : Βασικά Μέρη Ενός Νευρώνα

Το σημείο στο οποίο ενώνονται δύο νευρώνες ονομάζεται σύναψη (βλ. δεξιό τμήμα Εικόνας 2.3). Ένας νευρώνας μεταφέρει ένα μήνυμα (νευρωνικό σήμα) μέσω του νευράξονά του και το αποστέλλει σε έναν άλλον νευρώνα διά του σημείου σύναψης με το οποίο αυτοί οι δύο ενώνονται. Ο νευρώνας που αποστέλλει το σήμα ονομάζεται προσυναπτικό κύτταρο, ενώ ο νευρώνας που λαμβάνει το σήμα ονομάζεται μετασυναπτικό κύτταρο. [16]

Τα νευρωνικά σήματα αποτελούνται από βραχείς ηλεκτρικούς παλμούς γεγονός που παρατηρήθηκε με την τοποθέτηση ενός λεπτού ηλεκτροδίου κοντά στον νευράξονα ή το σώμα ενός νευρώνα. Οι παλμοί αυτοί, οι λεγόμενες αιχμές, έχουν πλάτος περίπου $100mV$ και διάρκεια περίπου $1 - 2ms$ και είναι η στοιχειώδης μονάδα μετάδοσης σήματος. Η μορφή ενός παλμού καθώς αυτός διαδίδεται κατά μήκος ενός νευράξονα δεν αλλάζει. Μια σειρά από παλμούς που εκπέμπει ένας μεμονωμένος νευρώνας ονομάζεται ουρά αιχμών και είναι μια ακολουθία στερεότυπων γεγονότων που εμφανίζονται ανά τακτά χρονικά διαστήματα. Η μορφή μιας αιχμής δεν φέρει κάποια πληροφορία δεδομένου ότι όλες οι αιχμές που

εκπέμπονται από έναν νευρώνα είναι πανομοιότυπες μεταξύ τους. Αυτό που πραγματικά έχει σημασία είναι ο αριθμός των αιχμών και χρόνος στον οποίο αυτές εξεπέμφθησαν. [17] [18]

2.3 Μηχανισμοί Παραγωγής Αιχμών

Ο τρόπος λειτουργίας του ανθρώπινου οπτικού συστήματος, όντας ένας έξυπνος μηχανισμός, έχει γίνει βάση μελέτης και έρευνας για τη βελτίωση της απόδοσης των συστημάτων συμπίεσης. Το οπτικό σύστημα επιτρέπει την έξυπνη σύλληψη και μετασχηματισμό των οπτικών πληροφοριών σε μια ακολουθία γεγονότων, τις λεγόμενες αιχμές, τα οποία απαρτίζονται από έναν εξαιρετικά αποτελεσματικό, συμπαγή αλλά και ταυτόχρονα γεμάτο πληροφορία νευρικό κώδικα, ο οποίος μπορεί να μεταδοθεί άθικτος στον οπτικό φλοιό, μέρος του εγκεφάλου που επεξεργάζεται και αποκωδικοποιεί τις οπτικές πληροφορίες.

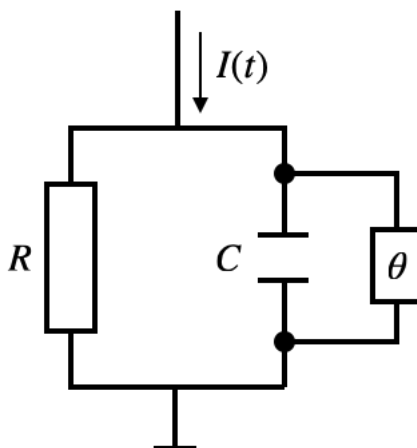
Στη βιβλιογραφία υπάρχουν πολλοί Μηχανισμοί Παραγωγής Αιχμών (*Spike Generation Mechanisms - SGM*) [19] [20] που προσεγγίζουν τον τρόπο με τον οποίο οι νευρώνες συμπεριφέρονται και μετασχηματίζουν ένα σταθερό οπτικό ερέθισμα εισόδου σε μια ακολουθία N διακριτών γεγονότων ονόματι ακολουθία αιχμών. Το μοντέλο των Hodgkin (*Χότζκιν*) και Huxley (*Χάξλεϋ*) είναι ένα από τα πιο ακριβή μοντέλα που αναπαράγουν τη συμπεριφορά των νευρώνων με αρκετά υψηλή ορθότητα και πιστότητα. Το μοντέλο αποτελείται από ένα ηλεκτρικό κύκλωμα με έναν πυκνωτή και διέπεται από τέσσερις πολύπλοκες διαφορικές εξισώσεις. Στη γενική του μορφή, το μοντέλο περιγράφεται από την εξίσωση (2.1),

$$I(t) = I_C(t) + \sum_k I_k(t) \quad (2.1)$$

η οποία περιγράφει πως το ρεύμα εισόδου I του κυκλώματος διασπάται στο ρεύμα I_C που φορτίζει τον πυκνωτή και σε επιμέρους I_k ρεύματα. Το άθροισμα αυτών των I_k ρευμάτων είναι ο τρόπος με τον οποίο το μοντέλο προσεγγίζει τη βιολογική συμπεριφορά των κυττάρων ενός νευρώνα, μία προσέγγιση που παρεκκλίνει από το πλαίσιο αυτής της πτυχιακής εργασίας καθώς στόχος της είναι να μελετήσει μια νέα μέθοδο συμπίεσης εικόνων εμπνευσμένη από τη φιλοσοφία του τρόπου λειτουργίας των νευρώνων και όχι να αναπαράγει πιστά τη λειτουργία των νευρώνων. Από προσπάθειες απλούστευσης του μοντέλου Hodgkin-Huxley [21] προέκυψαν διάφορα μοντέλα που περιγράφουν εξίσου καλά τη συμπεριφορά των νευρώνων [22], ένα από τα γνωστότερα και ευρέως χρησιμοποιούμενα είναι το μοντέλο Leaky Integrate-and-Fire (LIF) που προσεγγίζει με ακρίβεια τον τρόπο παραγωγής αιχμών των νευρώνων. [15] [23]

2.4 Το Μοντέλο LIF (Μηχανισμός Παραγωγής Αιχμών)

Το μοντέλο LIF (*Leaky Integrate and Fire model* – Μοντέλο Διαρρέουσας Ολοκλήρωσης και Πυροδότησης) είναι ένα γνωστό και απλό μοντέλο που προσεγγίζει την νευρωνική διαδικασία κωδικοποίησης χρησιμοποιώντας ένα κύκλωμα αποτελούμενο από έναν πυκνωτή και μία αντίσταση παράλληλα συνδεδεμένα (Εικόνα 2.4).



Εικόνα 2.4 : Ηλεκτρικό Κύκλωμα του μοντέλου LIF

Το μοντέλο LIF, στη γενική του μορφή, ενσωματώνει στον μηχανισμό του και την απειθή περίοδο (refractory period) των νευρώνων. Η απειθή περίοδος είναι το χρονικό διάστημα που χρειάζεται ένας νευρώνας για να επανέλθει στην αρχικά ήρεμη κατάστασή του αφού διεγέρθηκε και εξέπεμψε μια ουρά αιχμών ως κάποιο σήμα εξόδου. Συμπεριλαμβάνοντας αυτήν την περίοδο, το μοντέλο LIF προσεγγίζει λεπτομερώς, και από βιολογικής απόψεως, τη συμπεριφορά των νευρώνων το οποίο ξεφεύγει από τα πλαίσια αυτής της εργασίας (και δεν προσφέρει επιπλέον βοήθεια για τη συμπίεση της εικόνας κάνοντας χρήση του μοντέλου LIF) και για αυτό αυτή θεωρήθηκε αμελητέα.

Το κύκλωμα της Εικόνας 2.4 περιγράφεται από την εξής διαφορική εξίσωση πρώτου βαθμού:

$$I(t) = \frac{V(t)}{R} + C \frac{dV}{dt}, \quad (2.2)$$

όπου $I(t)$ το σήμα εισόδου, C ο πυκνωτής, R η σε παράλληλη σύνδεση με τον πυκνωτή αντίσταση και $V(t)$ η τάση που αναπτύσσεται στην αντίσταση. Ο πυκνωτής προσομοιώνει την χωρητικότητα της μεμβράνης ενός νευρώνα, ενώ η τάση το μεμβρανικό δυναμικό που αναπτύσσεται στον νευρώνα. Πολλαπλασιάζοντας την εξίσωση (2.2) με R , ορίζεται η χρονική σταθερά του μοντέλου, ο παράγοντας ολοκλήρωσης $\tau = RC$. Κάθε φορά που το μεμβρανικό

δυναμικό ενός νευρώνα ξεπερνάει ένα κατώφλι θ , όπου $\theta > 0$, ο νευρώνας εκπέμπει μία αιχμή. Η στιγμή αυτή της παραγωγής της αιχμής ονομάζεται στιγμή πυροδότησης και θα συμβολίζεται ως $t^k, k \geq 1$. Ένας νευρώνας, για ένα συγκεκριμένο κατώφλι θ , εκπέμπει αιχμή όταν πληρούται η εξής συνθήκη :

$$V(t) \geq \theta \quad (2.3)$$

Αμέσως μετά την εκπομπή της αιχμής το μεμβρανικό δυναμικό επαναφέρεται στο 0, δηλαδή σε τιμή μικρότερη του κατωφλίου, ώστε να ξεκινήσει εκ νέου η διέγερση του νευρώνα. Δεδομένου ότι οι αιχμές είναι πανομοιότυπα γεγονότα, δηλαδή σήματα σχεδόν ολόιδιου σχήματος, αυτό που τις διαφοροποιεί και χαρακτηρίζει είναι η χρονική στιγμή της πυροδότησής τους.

Έστω ότι το μοντέλο LIF δέχεται το εξής σήμα εισόδου, το οποίο θεωρείται σταθερό για μια συγκεκριμένη χρονική περίοδο T :

$$I(t) = I \mathbf{1}_{[0 \leq t \leq T]}(t), \quad (2.4)$$

όπου $\mathbf{1}$ είναι ένας δείκτης που ισούται με 1 αν $0 \leq t \leq T$, και 0 σε οποιαδήποτε άλλη περίπτωση. Το μοντέλο LIF θέτοντας ένα κατώφλι θ καθορίζει αν το σήμα εισόδου θα προκαλέσει παραγωγή αιχμών ή όχι. Συγκεκριμένα, αν η τιμή της εισόδου δημιουργήσει μεμβρανικό δυναμικό που ξεπερνάει την τιμή του κατωφλίου, ο νευρώνας διεγείρεται και εκπέμπει αιχμές, διαφορετικά ο νευρώνας παραμένει σιωπηλός.

Υποθέτοντας πως η εκπομπή της k -οστής αιχμής γίνεται τη χρονική στιγμή $t^{(k)}$, τότε λύνοντας τη διαφορική εξίσωση (2.2) προκύπτει :

$$V_k(t) = RI[1 - \exp(-\frac{t - t^{(k)}}{\tau})], \quad \forall t \geq t^{(k)} \quad (2.5)$$

Ο νευρώνας παράγει αιχμή όταν η τάση $V_k(t)$ ξεπεράσει την τιμή του κατωφλίου θ . Η χρονική στιγμή $t^{(k+1)}$ στην οποία γίνεται η επόμενη εκπομπή αιχμής ονομάζεται $(k + 1)$ -οστή στιγμή πυροδότησης και ικανοποιεί την εξής εξίσωση :

$$V_k(t^{(k+1)}) = \theta \quad (2.6)$$

Προκύπτει ότι :

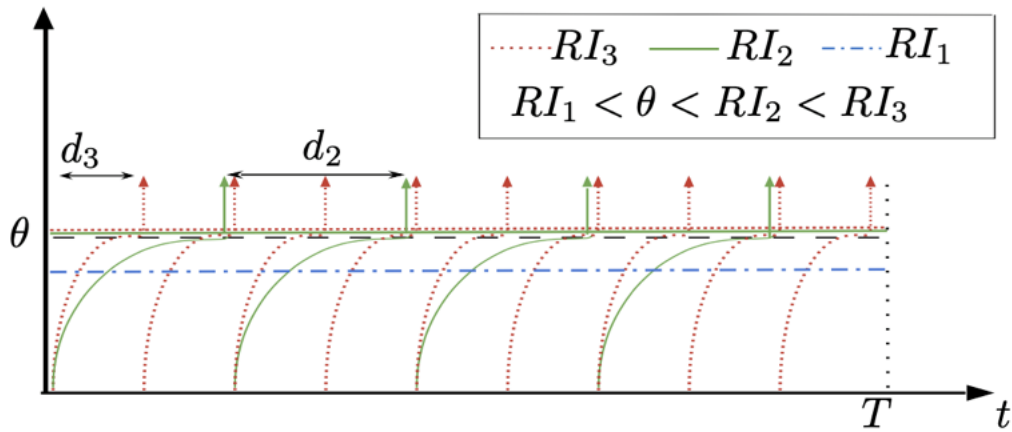
$$t^{(k+1)} = \begin{cases} +\infty & \text{αν } RI \leq \theta \\ t^{(k)} - \tau \ln[1 - \frac{\theta}{RI}] & \text{αν } RI > \theta \end{cases} \quad (2.7)$$

Ακριβώς μετά την $(k + 1)$ -οστή εκπομπή αιχμής την χρονική στιγμή $t^{(k+1)}$ το μεμβρανικό δυναμικό μηδενίζεται ($V_{k+1}(t^{(k+1)}) = 0$) και η διαδικασία ολοκλήρωσής του ξεκινά και πάλι από την αρχή τη χρονική στιγμή $t > t^{(k+1)}$ μέχρις ότου να γίνει εκπομπή της επόμενης αιχμής. Η ασυμπτωτική τιμή RI καθορίζει την παραγωγή των αιχμών. Αν ισχύει $RI \geq \theta$, τότε γίνεται εκπομπή μιας αιχμής, διαφορετικά δεν συμβαίνει καμία παραγωγή.

Βάσει παραπάνω, γίνεται εμφανές πως το μοντέλο LIF παράγει μια αλληλουχία αιχμών η οποία κωδικοποιεί το σήμα εισόδου I . Αυτή η αλληλουχία αιχμών απαιτεί την εκτίμηση της καθυστέρησης $d(I)$ μεταξύ δύο διαδοχικών στιγμών πυροδότησης t^k και t^{k+1} με την παραδοχή ότι η διαδικασία διέγερσης ξεκινάει τη χρονική στιγμή $t^0 = 0$. Από τον ορισμό των χρονικών στιγμών t^{k+1} (εξίσωση 2.7) προκύπτει πως η καθυστέρηση μεταξύ δύο διαδοχικών αιχμών συμπεριλαμβανομένης και της καθυστέρησης της πρώτης αιχμής από τη στιγμή t^0 είναι σταθερή καθώς το σήμα εισόδου I είναι σταθερό. Επομένως ο χρόνος άφιξης μιας αιχμής μετά την εκπομπή της αμέσως προηγούμενης της (δηλαδή το χρονοδιάστημα μεταξύ δύο διαδοχικών χρονικών στιγμών εκπομπής αιχμών, $t^{k+1} - t^k$) είναι ίδιος και υπολογίζεται ως εξής :

$$d(I) = \begin{cases} +\infty & \text{αν } RI < \theta \\ h(I) = -\tau \ln[1 - \frac{\theta}{RI}] & \text{αν } RI > \theta \end{cases} \quad (2.8)$$

Πέραν αυτών, η καθυστέρηση εξαρτάται σε μεγάλο βαθμό και από την ένταση I . Όσο ισχυρότερο είναι το σήμα εισόδου, τόσο πιο γρήγορη θα είναι η εκπομπή της πρώτης αιχμής, δηλαδή θα υπάρξει μικρή καθυστέρηση. Αντιθέτως, ένα αδύναμο σήμα εισόδου θα χρειαστεί περισσότερο χρόνο για να διεγείρει τον νευρώνα ώστε να γίνει η πυροδότηση της πρώτης αιχμής, ήτοι θα υπάρξει μεγάλη καθυστέρηση. Αυτό γίνεται πιο σαφές στην *Εικόνα 2.5*.



Εικόνα 2.5 : Συμπεριφορά του μοντέλου LIF για τρεις διαφορετικές εισόδους I σε χρονικό διάστημα T με ένα συγκεκριμένο κατώφλι θ

Όπως απεικονίζεται και παραπάνω, το μοντέλο LIF για ένα συγκεκριμένο κατώφλι θ δέχεται ως είσοδο τρία σταθερά διαφορετικής έντασης σήματα $I_1 < I_2 < I_3$. Βάση της εξίσωσης (2.8) οι εντάσεις I_3 και I_2 είναι ικανές να διεγείρουν τον νευρώνα και επομένως να προκαλέσουν εκπομπή αιχμών με χρονικές καθυστερήσεις $d_3 < d_2$ αντίστοιχα. Η τρίτη ένταση I_1 δεν είναι αρκετά δυνατή για να διεγείρει τον νευρώνα (η συνθήκη $RI_1 > \theta$ δεν πληρούται), με αποτέλεσμα ο χρόνος καθυστέρησης άφιξης αιχμών d_1 να γίνεται άπειρο και ο νευρώνας να παραμένει σιωπηλός.

Γίνεται αντιληπτό, πως το μοντέλο LIF εντός συγκεκριμένου χρονικού διαστήματος T , κωδικοποιεί ένα σήμα εισόδου I σε μια αλληλουχία αιχμών. Οι αιχμές αυτής της αλληλουχίας ισαπέχουν χρονικά μεταξύ τους κατά $h(I)$ όπως ορίζεται στην εξίσωση (2.8). Αν αυτή η χρονική καθυστέρηση μεταξύ των αιχμών ήταν γνωστή, τότε θα μπορούσε να γίνει ανακατασκευή της τιμής του σήματος εισόδου ως εξής :

$$\tilde{I} = \begin{cases} 0 & \text{αν } d > T \\ h^{-1}(d) & \text{αν } d \leq T \end{cases} \quad (2.9)$$

όπου \tilde{I} το ανακατασκευασμένο σήμα εισόδου και $h^{-1}(d)$ η αντίστροφη συνάρτηση της $h(I)$ που ορίζεται ως εξής :

$$h^{-1} = \frac{\theta}{R(1 - e^{-\frac{d}{\tau}})} \quad (2.10)$$

Βάσει εξίσωσης (2.9), όταν η χρονοκαθυστέρηση d είναι μεγαλύτερη από το διάστημα του χρονικού παραθύρου T δεν υπάρχει εκπομπή κάποιας αιχμής. Οπότε κατά την ανακατασκευή αποφασίζεται να επιστροφή μια αυθαίρετη τιμή, η τιμή μηδέν θεωρείται η πιο λογική απάντηση. Στην αντίθετη περίπτωση για την οποία ισχύει $d \leq T$, δεν υπάρχει σφάλμα κατά την ανακατασκευή. [24] [25]

2.5 Χρήση και ερμηνεία του μοντέλου LIF για τη δημιουργία κβαντιστή με απώτερο στόχο τη συμπίεση της εικόνας

Οι μηχανισμοί παραγωγής αιχμών, τις τελευταίες δεκαετίες, έχουν προσελκύσει το ενδιαφέρον πολλών ερευνητών, καθώς μπορεί να φανερώσουν τον τρόπο με τον οποίο μπορεί να σχεδιαστούν ενεργειακά αποδοτικοί κωδικοποιητές ικανοί να κωδικοποιούν τα αναλογικά σήματα σε αλληλουχίες αιχμών. Η ερμηνεία των αλληλουχιών αιχμών θεωρείται ένα εξαιρετικά σημαντικό ζήτημα για διάφορες εφαρμογές όπως η επεξεργασία σήματος και εικόνας με περιορισμένη ισχύ. Ως εκ τούτου, η ανακάλυψη του καλύτερου Μηχανισμού Ερμηνείας Αιχμών (*Spike Interpretation Mechanism* - SIM) είναι βαρύνουσα. Ένας τέτοιος μηχανισμός θα επιτρέψει τη χρήση κωδικών αιχμών ώστε σήματα εισόδου (π.χ. μια εικόνα) να κωδικοποιούνται και να αποκωδικοποιούνται με τις λιγότερες πιθανές απώλειες πληροφορίας.

Στη συγκεκριμένη περίπτωση του μοντέλου LIF, το οποίο επιλέχθηκε για τον λόγο ότι με έναν απλό τρόπο περιγράφει αρκετά πιστά τον μηχανισμό παραγωγής των αιχμών από τους νευρώνες, ο τρόπος λειτουργίας του μπορεί να ερμηνευτεί με δύο τρόπους. Ο ένας είναι από την άποψη του χρόνου και ο άλλος από την άποψη της πυκνότητας εμφάνισης των αιχμών που παράγει.

Η ερμηνεία από πλευράς χρόνου έχει γίνει ήδη κατά την περιγραφή του μοντέλου στην *Ενότητα 2.4*. Συγκεκριμένα η παραγωγή της ακολουθίας αιχμών γίνεται ερμηνεύοντας, μέσω της εξίσωσης (2.8), ότι το χρονικό μεσοδιάστημα μεταξύ δύο εκπομπών αιχμής (καθυστέρηση διαιχμής - interspike delay) είναι σταθερό όταν και η είσοδος είναι σταθερό ερέθισμα. Ως εκ τούτου αυτό που υπολογίζει το μοντέλο είναι ο ακριβής χρόνος εκπομπής της πρώτης αιχμής λόγω κάποιας συγκεκριμένης και σταθερής έντασης εισόδου και βάσει αυτού του χρόνου μπορεί να γίνει πλήρης ανάκτηση της έντασης που προκάλεσε αυτό το ερέθισμα, μέσω των εξισώσεων (2.9 και 2.10).

Το μοντέλο LIF αυτό καθ'αυτό δεν μπορεί να «συμβάλλει» στο πλαίσιο της συμπίεσης. Το μοντέλο παρόλο που κωδικοποιεί και αποκωδικοποιεί μια είσοδο με βάση τον χρόνο δεν προσφέρει τρόπο ώστε αυτή η κωδικοποιημένη πληροφορία να ομαδοποιηθεί (*κβαντιστεί*) και ως εκ τούτου να μειωθεί. Όσες διαφορετικές τιμές εντάσεων κι αν υπάρχουν ως είσοδοι, τόσες διαφορετικές τιμές χρονικών καθυστερήσεων θα παραχθούν. Γι'αυτόν τον λόγο το μοντέλο LIF συνδυάζεται με μεθόδους ερμηνείας της πυκνότητας των παραγόμενων αιχμών. Συγκεκριμένα, για ένα ορισμένο χρονικό παράθυρο T και με βάση της χρονικής

καθυστέρησης που υπολόγισε το μοντέλο LIF εντός αυτού του παραθύρου μπορεί να μετρηθεί ο αριθμός των αιχμών που παρήγαγε μια είσοδος. Συνεπώς, κατά τη διάρκεια ενός χρονικού διαστήματος T διαφορετικές εντάσεις μπορεί και να παραγάγουν τον ίδιο αριθμό από αιχμές παρόλο που αυτές θα έχουν διαφορετικές χρονοκαθυστερήσεις.

Εκμεταλλευόμενοι αυτές τις δύο ερμηνείες του μοντέλου LIF μπορούμε να δημιουργήσουμε έναν κβαντιστή εμπνευσμένο από τις νευροεπιστήμες ο οποίος θα μπορεί να βοηθήσει στον τρόπο με τον οποίο συμπίεζονται οι εικόνες. Ο κβαντιστής ονομάζεται DUAL-SIM Κβαντιστής (Κβαντιστής Διττής Ερμηνείας) [24] και η ανάλυσή του γίνεται στο *Κεφάλαιο 3*.

Κεφάλαιο 3ο

Συμπεριφορά και Τρόπος Λειτουργίας Του Προτεινόμενου DUAL-SIM Κβαντιστή

3.1 Τρόπος Λειτουργίας Προτεινόμενου Κβαντιστή

Σκοπός αυτής της εργασίας είναι να μελετήσει έναν νέο κβαντιστή εμπνευσμένο από το μοντέλο LIF, που έχει τις βάσεις του στη νευροεπιστήμη, ώστε ένα σήμα εισόδου να κωδικοποιείται σε μια αλληλουχία αιχμών και μετά η ανακατασκευή του να γίνεται μετρώντας τις αιχμές που προέκυψαν. Συγκεκριμένα, για ένα σταθερό σήμα εισόδου I θα υπάρχει ένας νευρώνας τον οποίο θα προσπαθεί να διεγείρει εντός ενός συγκεκριμένου χρονικού διαστήματος T . Αν η ένταση της εισόδου είναι αρκετά ισχυρή ο νευρώνας θα διεγερθεί και θα παράγει αιχμές οι οποίες θα ισαπέχουν χρονικά. Ο υπολογισμός του χρόνου καθυστέρησης άφιξης της πρώτης αιχμής γίνεται με την εξίσωση (2.8). Οι αιχμές που ακολουθούν την πρώτη εμφανίζονται με την ίδια καθυστέρηση d όταν το σήμα εισόδου είναι σταθερό. Γνωρίζοντας τον χρόνο καθυστέρησης d , υπολογίζεται ο αριθμός N των αιχμών που μόλις παρήχθησαν στο διάστημα T ως εξής :

$$N = \begin{cases} 0 & \text{αν } d > T \\ \lfloor \frac{T}{d} \rfloor & \text{αν } d < T \end{cases} \quad (3.1)$$

Μέσω αυτών των μηχανισμών, το σήμα εισόδου ποσοτικοποιείται (κβαντίζεται) σε αριθμό εκπεμπόμενων αιχμών. Για την αντίστροφη διαδικασία, δηλαδή αυτή της αποκβάντισης, το μοντέλο έχοντας τον αριθμό των αιχμών του σήματος έχει τη δυνατότητα να υπολογίσει κατά προσέγγιση τον χρόνο \tilde{d} ως εξής :

$$\tilde{d} = \frac{T}{N} \quad (3.2)$$

Κατά συνέπεια αυτού του υπολογισμού γίνεται η κατά προσέγγιση ανάκτηση του σήματος εισόδου \tilde{I} μέσω των εξισώσεων (2.10) και (2.9). Έτσι μέσω αυτών των βημάτων/υπολογισμών δημιουργείται ένας κβαντιστής εμπνευσμένος από τη λειτουργία των νευρώνων του οπτικού συστήματος του ανθρώπου.

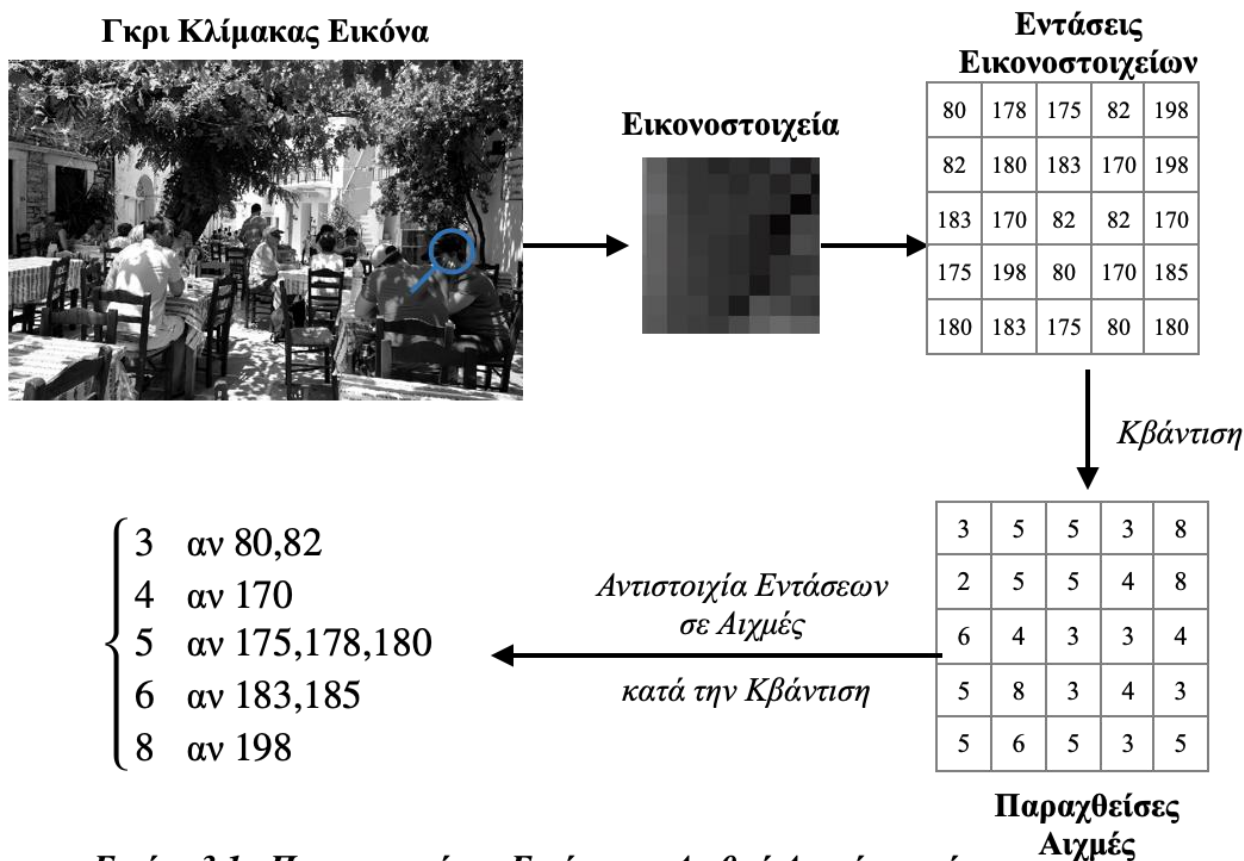
Η λειτουργία αυτού του προτεινόμενου κβαντιστή μπορεί να επεκταθεί ώστε να χρησιμοποιηθεί στη συμπίεση μιας εικόνας. Η μόνη διαφορά σε αυτήν την περίπτωση θα είναι ο αριθμός των νευρώνων, αντί να υπάρχει μόνο ένας νευρώνας θα υπάρχουν τόσοι όσα είναι

και τα εικονοστοιχεία της εικόνας που πρόκειται να συμπιεστεί. Αναλυτικότερα ο DUAL-SIM κβαντιστής δέχεται ως είσοδο μια εικόνα και αντιμετωπίζει την τιμή της έντασης (φωτεινότητα) κάθε εικονοστοιχείου της σαν ένταση ρεύματος I που προσπαθεί να διεγείρει έναν νευρώνα, ο οποίος όταν ερεθιστεί αρκετά εκπέμπει μία αιχμή. Ο κάθε νευρώνας χρειάζεται συγκεκριμένο χρόνο d , ο οποίος εξαρτάται από τη φωτεινότητα του εκάστοτε εικονοστοιχείου (εξίσωση 2.8), για να εκπέμψει την πρώτη αιχμή. Η εκπομπή της επόμενης αιχμής θα συμβεί μετά από τον ίδιο χρόνο d καθώς η εικόνα εισόδου είναι στατική. Γνωρίζοντας τον χρόνο καθυστέρησης d του κάθε εικονοστοιχείου, υπολογίζεται ο αριθμός N των αιχμών που παράγει κάθε εικονοστοιχείο (εξίσωση 3.1) σε ένα εξ αρχής ορισμένο χρονικό παράθυρο T . Μέσω αυτών των μηχανισμών το μοντέλο ποσοτικοποιεί (κβαντίζει) τις εντάσεις της εικόνας σε αριθμό εκπεμπόμενων αιχμών. Για την αντίστροφη διαδικασία, δηλαδή αυτή της αποκβάντισης, το μοντέλο έχοντας τον αριθμό των αιχμών κάθε εικονοστοιχείου έχει τη δυνατότητα να υπολογίσει κατά προσέγγιση τον αντίστοιχο χρόνο \tilde{d} (εξίσωση 3.2) και κατά συνέπεια (εξισώσεις 2.10 και 2.9) να ανακτήσει την ένταση (φωτεινότητα) \tilde{I} κάθε εικονοστοιχείου. [24] [25]

Ο DUAL-SIM όντας κβαντιστής μπορεί να εισαχθεί στην τρέχουσα αρχιτεκτονική των συστημάτων συμπίεσης που κατά μέσο όρο αποτελούνται από έναν μετασχηματισμό (DCT στο JPEG και DWT στο JPEG2000), έναν κβαντιστή και την κωδικοποίηση εντροπίας.

3.2 Συμπεριφορά Προτεινόμενου Κβαντιστή

Έστω πως έχουμε μια εικόνα προς συμπίεση. Η τιμή της έντασης (φωτεινότητα) του κάθε εικονοστοιχείου της εικόνας θεωρείται πως είναι μια ξεχωριστή είσοδος (το ρεύμα) για τον κβαντιστή. Κάθε ρεύμα εισόδου προσπαθεί να διεγείρει από έναν νευρώνα μέσα σε ένα συγκεκριμένο χρονικό παράθυρο T που εξ αρχής έχει οριστεί. Ένας νευρώνας διεγείρεται όταν το ρεύμα εισόδου είναι «αρκετά ισχυρό» ώστε να καταφέρει να ξεπεράσει την τιμή ενός κατωφλίου θ που επίσης έχει οριστεί εξ αρχής. Όταν ο νευρώνας διεγερθεί παράγει έναν συγκεκριμένο αριθμό αιχμών. Ο κάθε νευρώνας χρειάζεται έναν συγκεκριμένο χρόνο d (εξίσωση 2.8) για να εκπέμψει την πρώτη αιχμή. Η εκπομπή της επόμενης αιχμής θα συμβεί μετά από τον ίδιο χρόνο d καθώς η εικόνα εισόδου είναι στατική. Συνεπώς, ο κβαντιστής αρκεί να υπολογίσει τον χρόνο άφιξης της πρώτης αιχμής, καθώς οι άλλες που θα έπονται, θα φθάσουν στον ίδιο χρόνο. Γνωρίζοντας τον χρόνο καθυστέρησης μεταξύ των ισαπεχουσών αιχμών ο κβαντιστής μπορεί να υπολογίσει τον αριθμό των αιχμών που παραγάγει το κάθε εικονοστοιχείο της εικόνας εντός του χρονικού παραθύρου T (εξίσωση 3.1). Το σημαντικότερο στην όλη μεθοδολογία είναι πως εικονοστοιχεία διαφορετικών εντάσεων παραγάγουν, για ένα συγκεκριμένο T και θ , τον ίδιο αριθμό αιχμών. Με αυτή τη διαδικασία ο κβαντιστής ποσοτικοποιεί την εικόνα σε αριθμό αιχμών βάσει των εντάσεων του κάθε εικονοστοιχείου της, όπως φαίνεται στο παρακάτω παράδειγμα στην Εικόνα 3.1 :



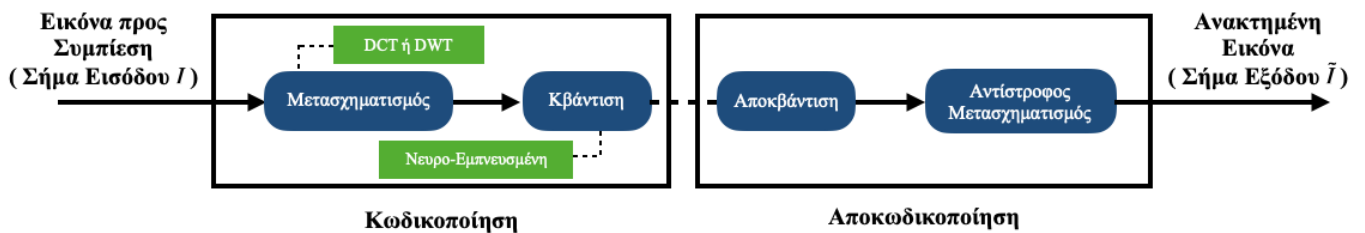
Εικόνα 3.1 : Ποσοτικοποίηση Εικόνας σε Αριθμό Αιχμών κατά την Νευρο-Εμπνευσμένη Κβάντιση

Με αντίστροφη λογική, γνωρίζοντας δηλαδή τον αριθμό των αιχμών και το χρονικό παράθυρο T , μπορεί να υπολογιστεί προσεγγιστικά η χρονική καθυστέρηση \tilde{d} μεταξύ των αιχμών (εξίσωση 3.2) και εν συνεχεία η κατά προσέγγιση ένταση \tilde{I} τού κάθε εικονοστοιχείου (εξισώσεις 2.10 και 2.9) της εικόνας. Αυτή αντίστροφη λειτουργία είναι η διαδικασία της αποκβάντισης, δηλαδή η διαδικασία ανάκτησης της εικόνας.

Να σημειωθεί πως ο νευρο-κβαντιστής έχει, προς το παρόν, δομηθεί έτσι ώστε να διαχειρίζεται μόνο εικόνες γκρι κλίμακας (Grayscale Images).

3.3 Εισαγωγή Του DUAL-SIM Κβαντιστή στη Συμβατική Αρχιτεκτονική Συμπίεσης Εικόνων

Παρόλο που ο DUAL-SIM είναι ένας κβαντιστής εμπνευσμένος από τις νευροεπιστήμες, μπορεί κάλλιστα να ενταχθεί στη συμβατική αρχιτεκτονικά συμπίεσης των εικόνων (βλ. *Εικόνα 3.2*).



Εικόνα 3.2 : Εισαγωγή του DUAL-SIM Κβαντιστή στη Συμβατική Αρχιτεκτονική

Ως πρώτη δοκιμή, έγινε χρήση μόνο του DUAL-SIM κβαντιστή χωρίς την ύπαρξη κάποιου μετασχηματισμού, καθαρά και μόνο για τον έλεγχο του τρόπου λειτουργίας του και τον τρόπο συμπεριφοράς των παραμέτρων του (R , C , T και θ). Στη συνέχεια, έγινε η εναλλάξ χρήση των μετασχηματισμών (κανονικών και αντίστροφων) DCT και DWT ώστε να ελεγχθεί η συμπεριφορά του κβαντιστή όταν εντάσσεται στο περιβάλλον της συμβατικής αρχιτεκτονικής συμπίεσης των εικόνων. [26] Η μελέτη της συμπεριφοράς του συστήματος έγινε κάνοντας χρήση γραφικών παραστάσεων που απεικονίζουν την παραμόρφωση (distortion) που έχει υποστεί το σήμα συναρτήσει του αριθμού δυφίων (bits). Η παραμόρφωση του σήματος υπολογίστηκε με τις μετρικές MSE και PSNR (για ορισμούς βλ. αντίστοιχα Εξ. 1.2 και 1.3), καθώς και υπολογίζοντας την εντροπία του ίδιου του συστήματος.

Όλα τα πειράματα και οι παρατηρήσεις έγιναν κβαντίζοντας την παρακάτω εικόνα και ως εκ τούτου συγκρίνοντας την ανακτημένη από την κβάντιση εικόνα με την αρχική της μορφή :

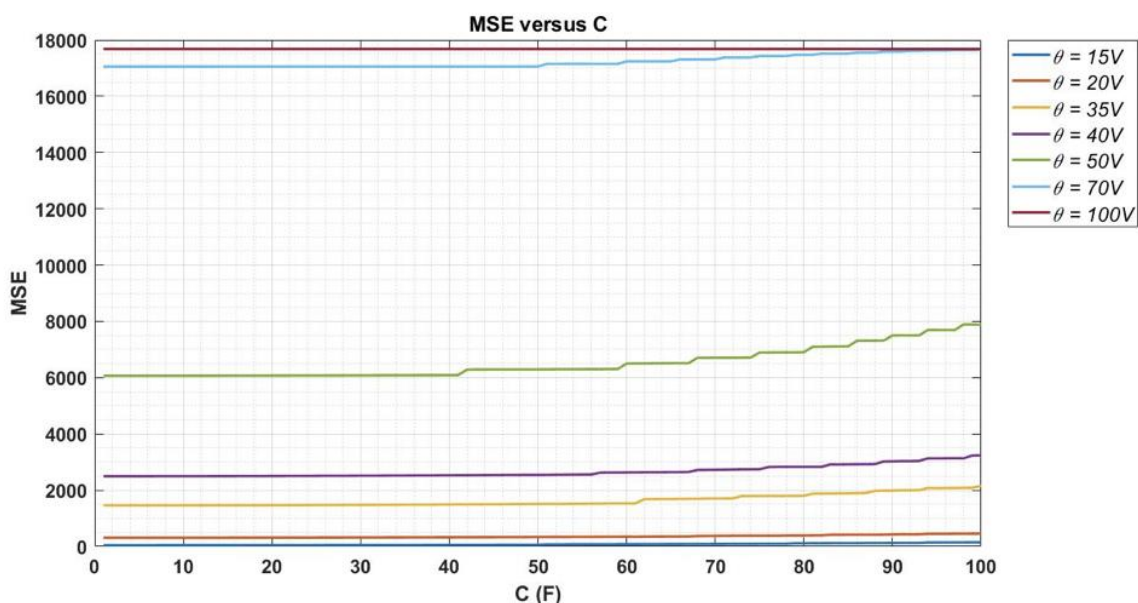


Εικόνα 3.3 : Η (αρχική) εικόνα που χρησιμοποιήθηκε σε όλα τα πειράματα

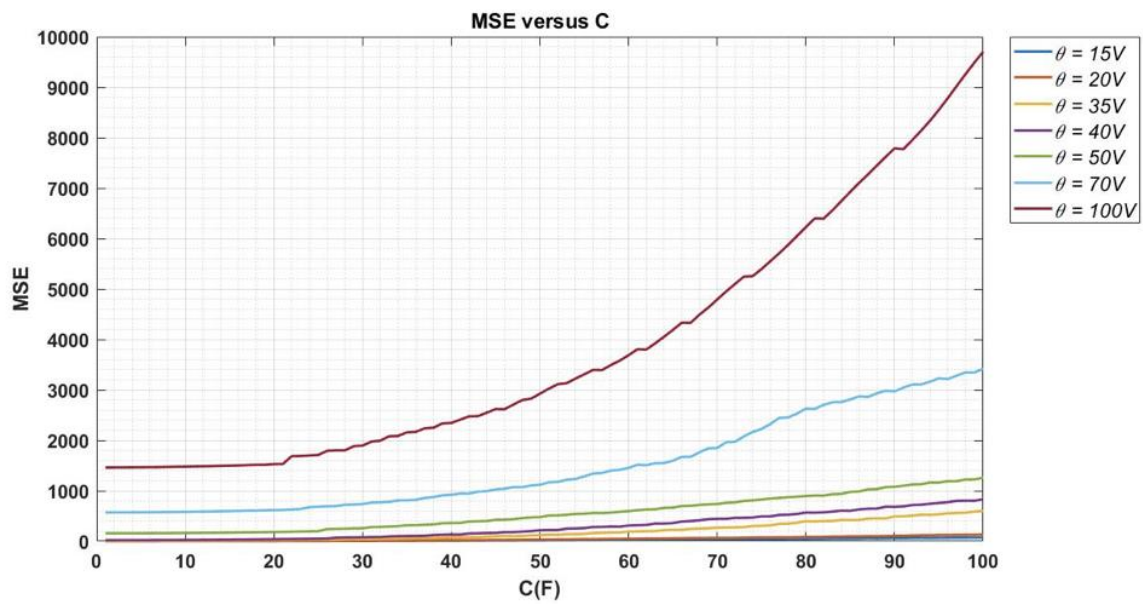
3.3.1 Χωρίς Κάποιον Μετασχηματισμό

Στα πρώτα πειράματα δεν έγινε κάποια χρήση μετασχηματισμού παρά μόνο ο ίδιος ο DUAL-SIM κβαντιστής. Αρχικά μελετήθηκε ξεχωριστά η συμπεριφορά των τριών βασικών παραμέτρων του κβαντιστή, η αντίσταση R , ο πυκνωτής C και το χρονικό παράθυρο παρατήρησης T . Κάθε φορά που μία από αυτές τις τρεις παραμέτρους ετίθετο υπό μελέτη οι άλλες δύο κρατιούνταν σταθερές. Συγκεκριμένα :

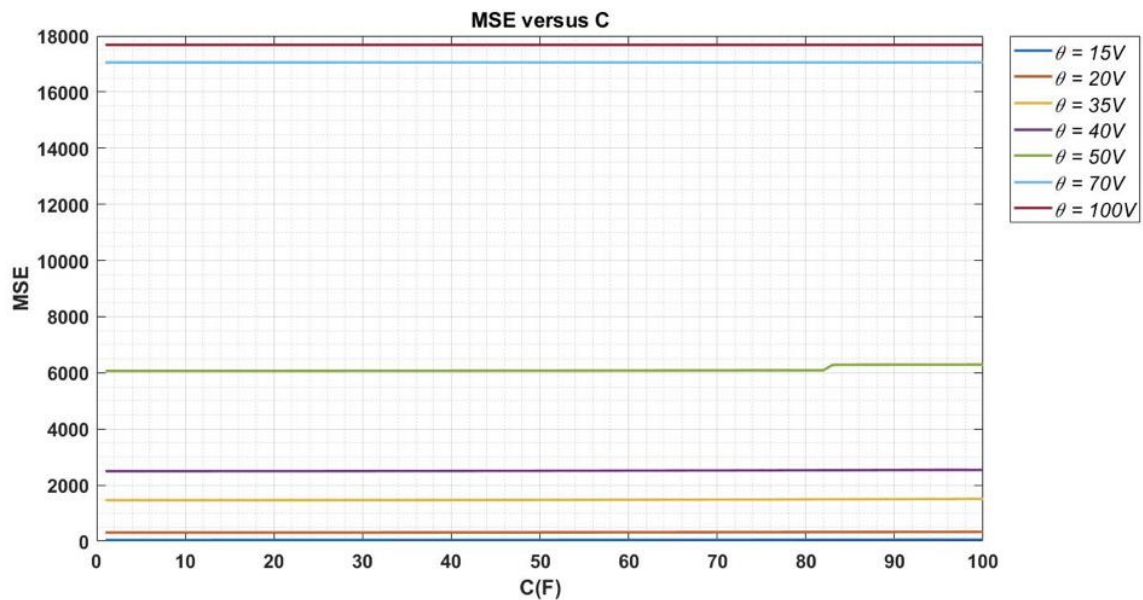
■ Μελέτη του πυκνωτή C



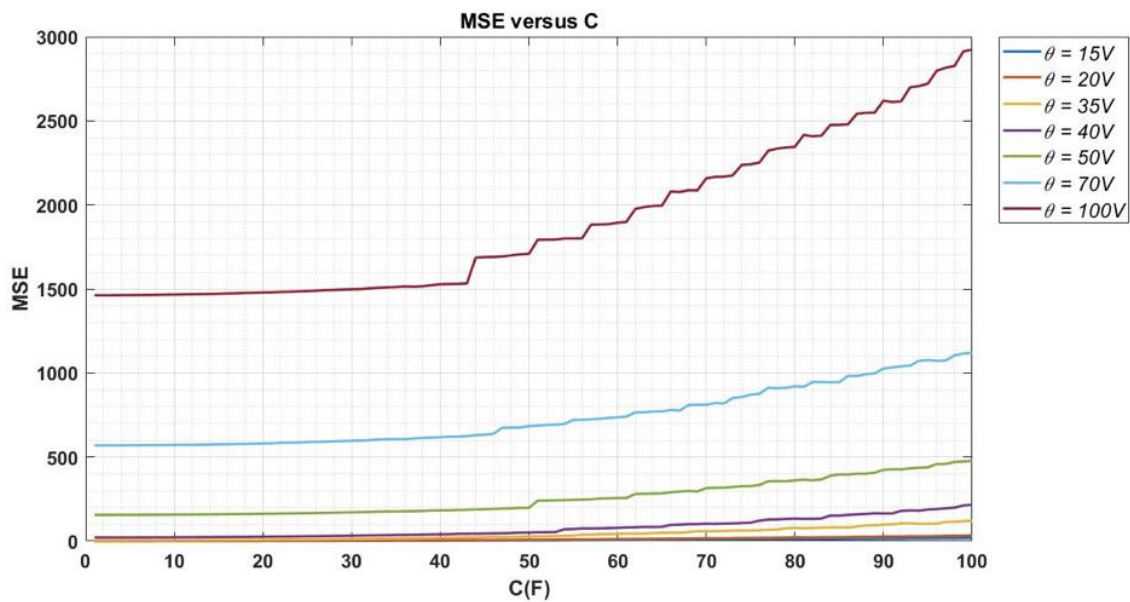
Εικόνα 3.4 : Παράμετρος C συναρτήσει του MSE για διάφορα θ και $T = 100 \text{ ms}$ & $R = 0.35 \Omega$



Εικόνα 3.5 : Παράμετρος C συναρτήσκει τον MSE για διάφορα θ και $T = 100\ ms$ & $R = 1\ \Omega$



Εικόνα 3.6 : Παράμετρος C συναρτήσκει τον MSE για διάφορα θ και $T = 200\ ms$ & $R = 0.35\ \Omega$



Εικόνα 3.7 : Παράμετρος C συναρτήσει του MSE για διάφορα θ και $T = 200\text{ ms}$ & $R = 1\ \Omega$

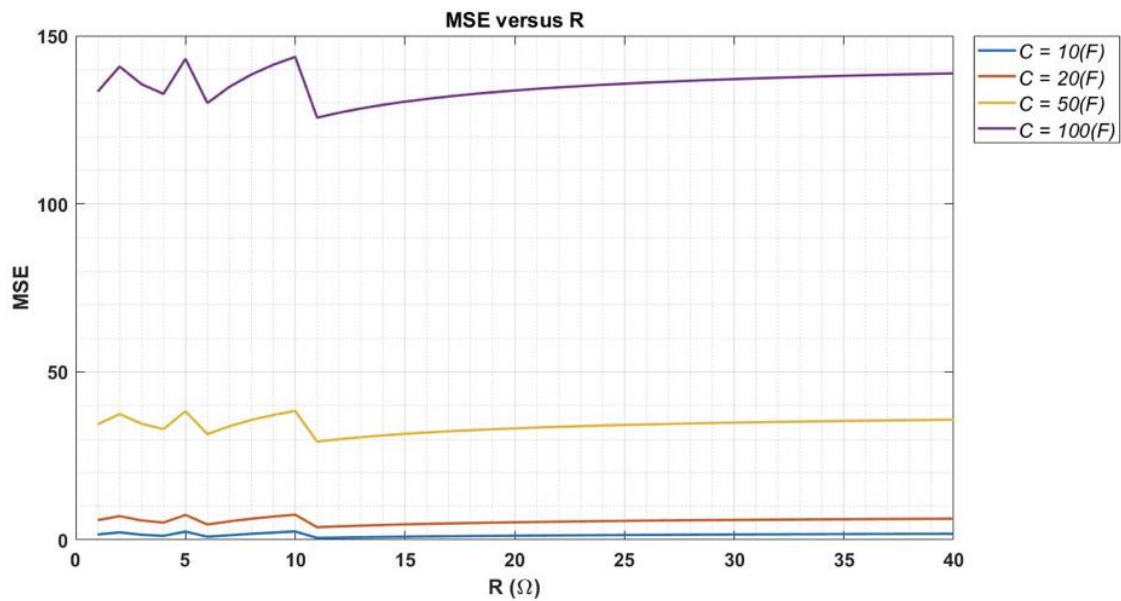
Στα παραπάνω τέσσερα πειράματα (Εικόνες 3.4 έως και 3.7) όλες οι παράμετροι του DUAL-SIM κβαντιστή κρατούνται σταθερές εκτός από τον πυκνωτή C , για τον οποίο γίνεται η μελέτη, και το ζεύγος «χρονικό παράθυρο T με αντίσταση R » βάσει του οποίου ζεύγους γίνονται συγκρίσεις μεταξύ των πειραμάτων. Συγκεκριμένα, για το παράθυρο T επιλέχθηκαν δύο τιμές (100ms και 200ms) για κάθε μία από τις οποίες επιλέχθηκαν δύο διαφορετικές τιμές για την αντίσταση R (0.35Ω και 1Ω). Για αυτούς τους συνδυασμούς των R και T χρησιμοποιήθηκαν επτά διαφορετικές τιμές κατωφλίου θ οι οποίες απεικονίζονται στα υπομνήματα των εικόνων των παραπάνω τεσσάρων πειραμάτων. Από τα εν λόγω πειράματα απορρέουν τα εξής συμπεράσματα :

Πρώτον, όσο αυξάνεται η τιμή του κατωφλίου θ , ανεξαρτήτως συνδυασμού των R και T , το MSE που δίνει ο κβαντιστής αυξάνεται.

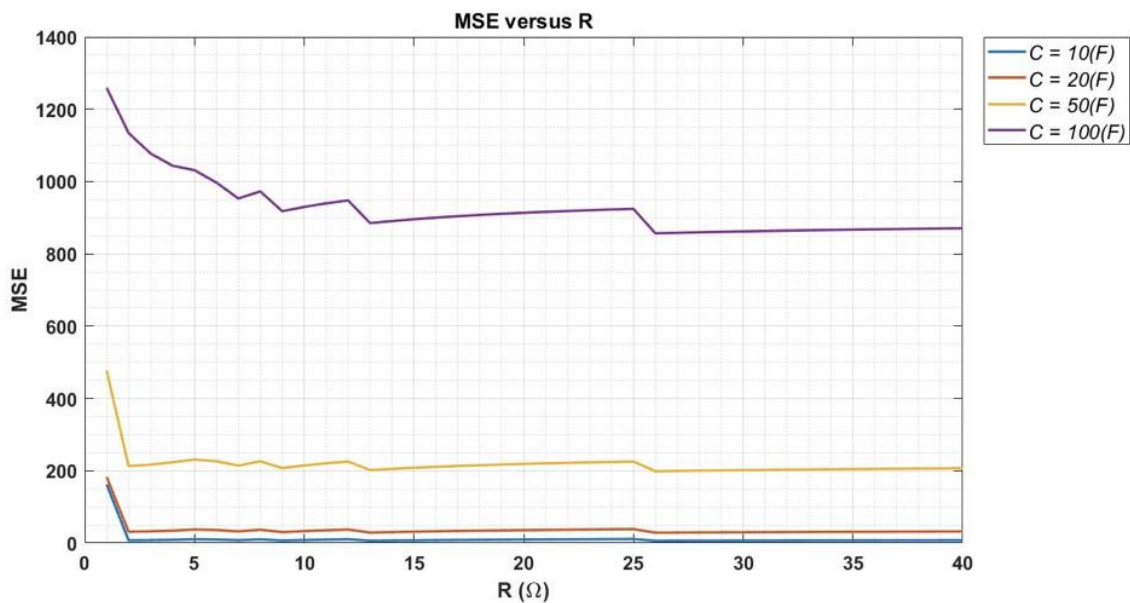
Δεύτερον, για σταθερό T , R και θ όσο αυξάνεται η τιμή του C τόσο αυξάνεται το MSE του κβαντιστή. Αυξάνοντας μόνο το R (από 0.35 σε 1) το MSE μειώνεται, αλλά συναρτήσει της αυξανόμενης τιμής του C το MSE συνεχίζει να αυξάνεται. Μάλιστα, για μεγάλη τιμή R , αυξάνοντας αρκετά την τιμή του C το MSE αυξάνεται απότομα (Εικόνες 3.5 και 3.7).

Τρίτον, για σταθερό R και θ αυξάνοντας μόνο το T (από 100 σε 200) παρατηρείται πως η συμπεριφορά του κβαντιστή δεν αλλάζει, παρά μόνο ότι έχοντας πλέον μεγαλύτερο χρονικό παράθυρο παρατήρησης το MSE αυξάνεται λιγότερο απότομα συναρτήσει της αυξανόμενης τιμής του πυκνωτή C .

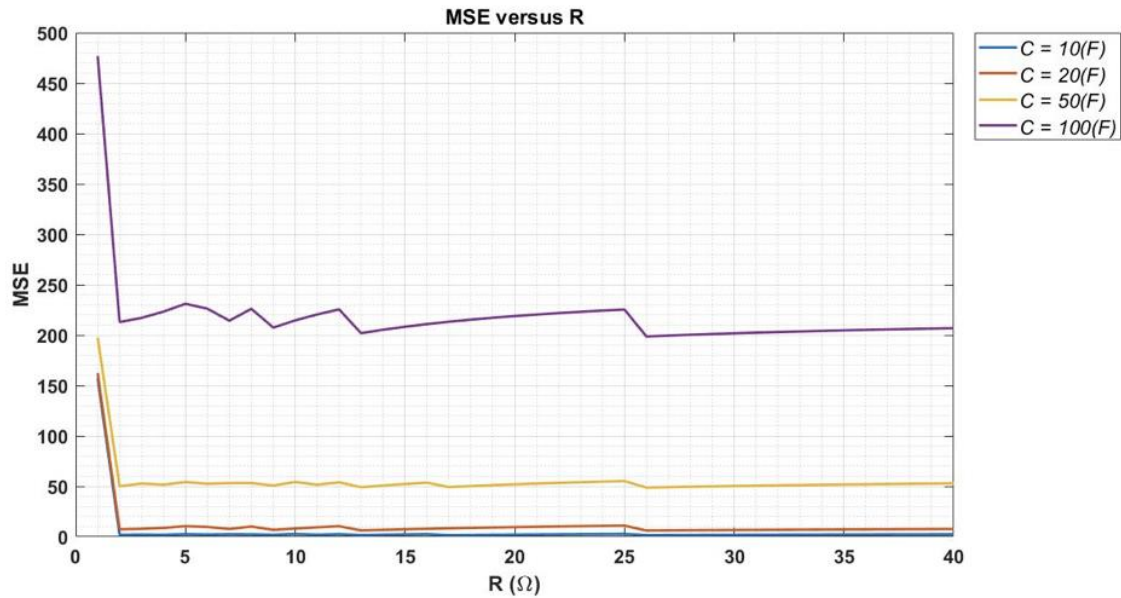
■ Μελέτη της αντίστασης R



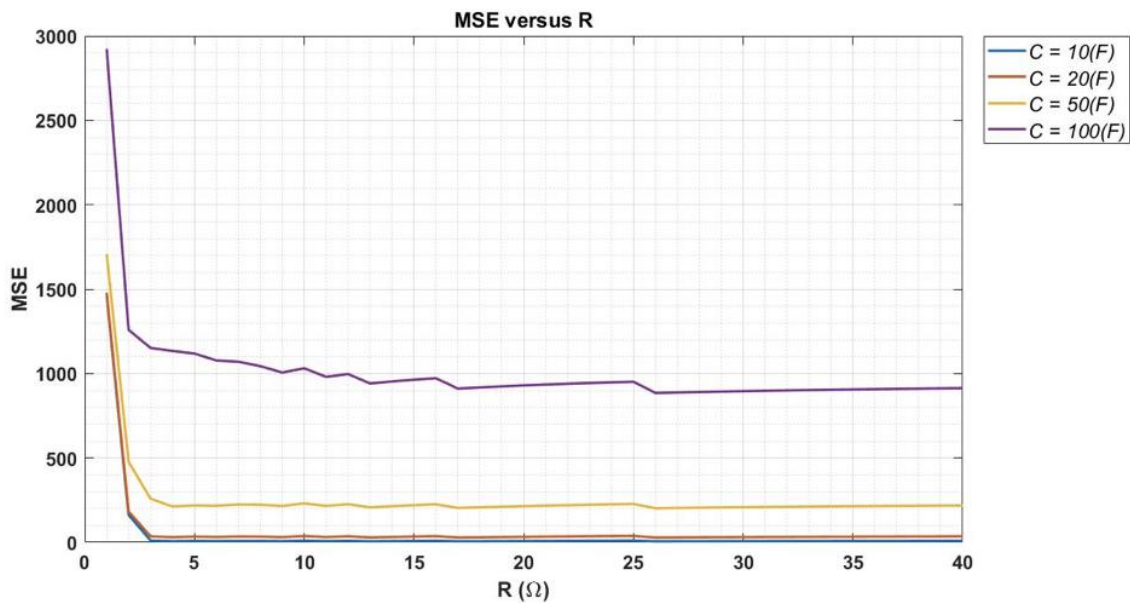
Εικόνα 3.8 : Παράμετρος R συναρτῆσει του MSE για διάφορα C και $T = 100\ ms$ & $\theta = 20\ V$



Εικόνα 3.9 : Παράμετρος R συναρτῆσει του MSE για διάφορα C και $T = 100\ ms$ & $\theta = 50\ V$



Εικόνα 3.10 : Παράμετρος R συναρτήσσει τον MSE για διάφορα C και $T = 200\ ms$ & $\theta = 50\ V$



Εικόνα 3.11 : Παράμετρος R συναρτήσσει τον MSE για διάφορα C και $T = 200\ ms$ & $\theta = 100\ V$

Στα παραπάνω τέσσερα πειράματα (Εικόνες 3.8 έως και 3.11) όλες οι παράμετροι του DUAL-SIM κβαντιστή κρατούνται σταθερές εκτός από την R , για την οποία γίνεται η μελέτη, και το ζεύγος «χρονικό παράθυρο T με κατώφλι θ » βάσει του οποίου ζεύγους γίνονται συγκρίσεις μεταξύ των πειραμάτων. Συγκεκριμένα, για το παράθυρο T επιλέχθηκαν δύο τιμές ($100ms$ και $200ms$) για κάθε μία από τις οποίες επιλέχθηκαν δύο διαφορετικές ζεύγη τιμών για το κατώφλι θ ($20V$ και $50V$ για τα $100ms$ και $50V$ και $100V$ για τα $200ms$). Για αυτούς τους συνδυασμούς των θ και T χρησιμοποιήθηκαν τέσσερις διαφορετικές τιμές για τον πυκνωτή C οι οποίες

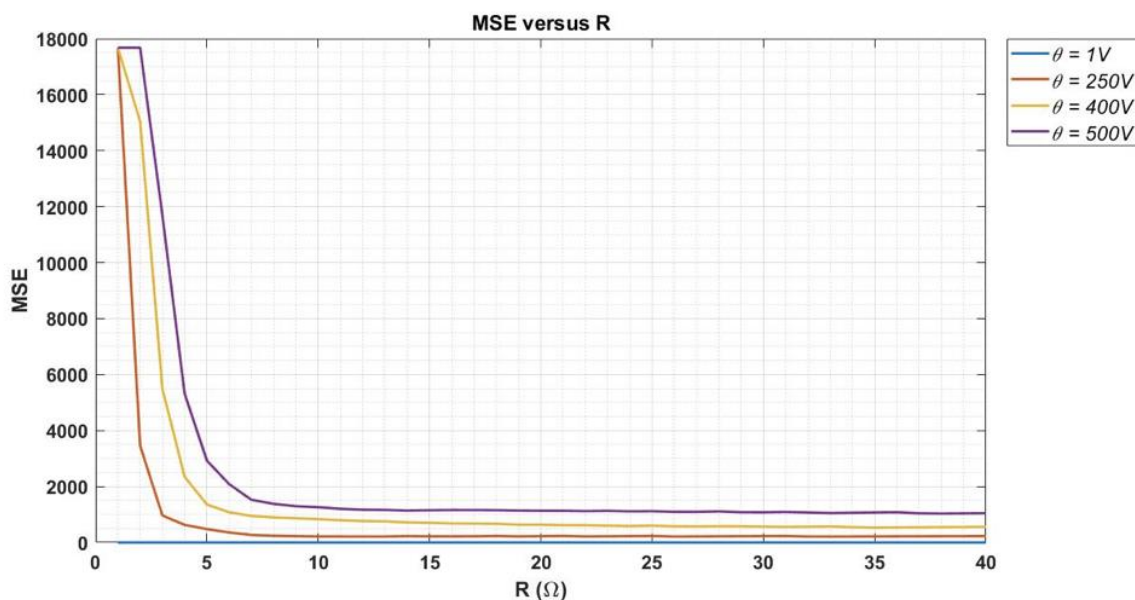
απεικονίζονται στα υπομνήματα των εικόνων των παραπάνω τεσσάρων πειραμάτων. Από τα εν λόγω πειράματα απορρέουν τα εξής συμπεράσματα :

Πρώτον, όσο αυξάνεται η τιμή του κατωφλίου θ (από 20 σε 100 ή από 50 σε 100), ανεξαρτήτως συνδυασμού των R , C και T , το MSE που δίνει ο κβαντιστής αυξάνεται. Ακριβώς το αντίστοιχο συμβαίνει και με τον πυκνωτή C , δηλαδή αυξάνοντας την τιμή του, ανεξαρτήτως των άλλων μεταβλητών, το MSE αυξάνεται.

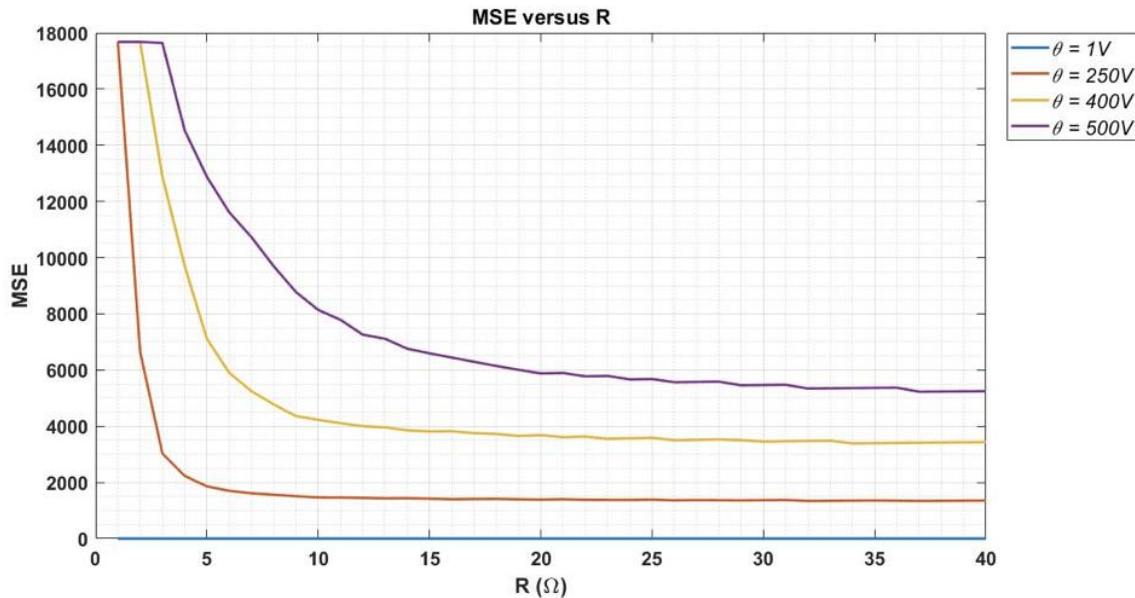
Δεύτερον, για σταθερό T , C και θ όσο αυξάνεται η τιμή της R τόσο μειώνεται το MSE του κβαντιστή. Παρατηρείται επίσης, πως όσο το R είναι αρκετά μικρότερο του C το MSE διατηρείται αρκετά υψηλό. Επιπροσθέτως, όσο αυτή η διαφορά μεταξύ R και C μικραίνει, καθώς η τιμή της R μεγαλώνει, υπάρχουν απότομες διακυμάνσεις στο MSE μέχρις ότου το R να φθάσει μια μεγάλη τιμή σε σχέση με το C (π.χ. στην Εικόνα 3.10 η μοβ γραμμή που αντιστοιχεί σε μεγάλο C σταθεροποιείται βραδύτερα από ό,τι η κίτρινη γραμμή που αντιστοιχεί σε μικρότερο C).

Τρίτον, για σταθερό C και θ αυξάνοντας μόνο το T (από 100 σε 200) παρατηρείται πως η συμπεριφορά του κβαντιστή δεν αλλάζει, παρά μόνο ότι έχοντας πλέον μεγαλύτερο χρονικό παράθυρο παρατήρησης το MSE αυξάνεται λιγότερο απότομα συναρτήσει της αυξανόμενης τιμής της αντίστασης R .

Ακολουθώντας διαφορετική προσέγγιση (διαφορετικά ζεύγη τιμών κρατούνται σταθερά) έχουμε τα εξής δύο πειράματα :



Εικόνα 3.12 : Παράμετρος R συναρτήσει του MSE για διάφορα θ και $T = 200\text{ ms}$ & $C = 20\text{ F}$



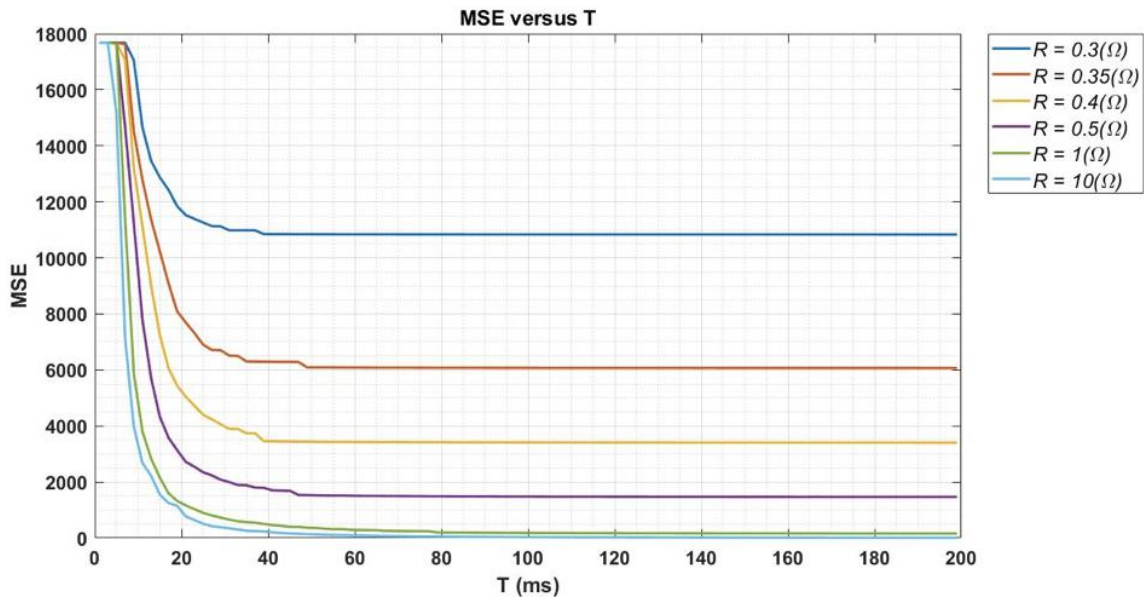
Εικόνα 3.13 : Παράμετρος R συναρτήσει του MSE για διάφορα θ και $T = 200\ ms$ & $C = 50\ F$

Στα δύο παραπάνω πειράματα (Εικόνες 3.12 και 3.13) το μόνο που αλλάζει σε σχέση με τα ακριβώς τέσσερα προηγούμενα παραδείγματα (Εικόνες 3.8 έως και 3.11) είναι πως για μία αυξανόμενη τιμή R επιλέχθηκε ένα μόνο παράθυρο T ($200ms$) για το οποίο επιλέχθηκαν δύο διαφορετικές τιμές C ($20F$ και $50F$). Για αυτούς τους συνδυασμούς των C και T χρησιμοποιήθηκαν τέσσερις διαφορετικές τιμές κατωφλίου θ οι οποίες απεικονίζονται στα υπομνήματα των εικόνων αυτών των πειραμάτων. Από τα εν λόγω πειράματα επιβεβαιώνονται τα εξής :

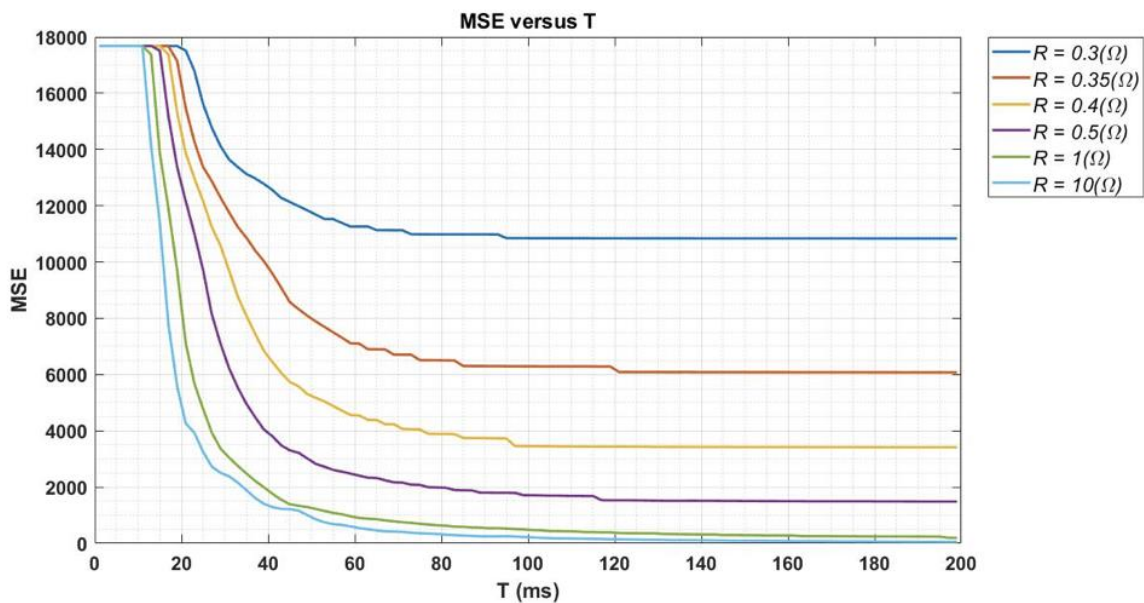
Πρώτον, όσο αυξάνεται η τιμή του κατωφλίου θ , ανεξαρτήτως των άλλων μεταβλητών, το MSE που δίνει ο κβαντιστής αυξάνεται.

Δεύτερον, όσο αυξάνεται η τιμή του πυκνωτή C (από $20F$ σε $50F$) το MSE του κβαντιστή αυξάνεται. Επιβεβαιώνεται επίσης πως όσο το R είναι αρκετά μικρότερο του C , το MSE είναι αρκετά υψηλότερο από ό,τι όταν αυτή η διαφορά μεταξύ R και C είναι μικρότερη (αυξάνοντας το R για μικρό C , το MSE μειώνεται απότομα -Εικόνα 3.12-, ενώ αυξάνοντας το R για μεγάλο C , το MSE μειώνεται ομαλότερα και βραδύτερα -Εικόνα 3.13-).

■ Μελέτη του χρονικού παραθύρου T



Εικόνα 3.14 : Παράμετρος T συναρτήσει του MSE για διάφορα R και $\theta = 50\text{ V}$ & $C = 20\text{ F}$



Εικόνα 3.15 : Παράμετρος T συναρτήσει του MSE για διάφορα R και $\theta = 50\text{ V}$ & $C = 50\text{ F}$

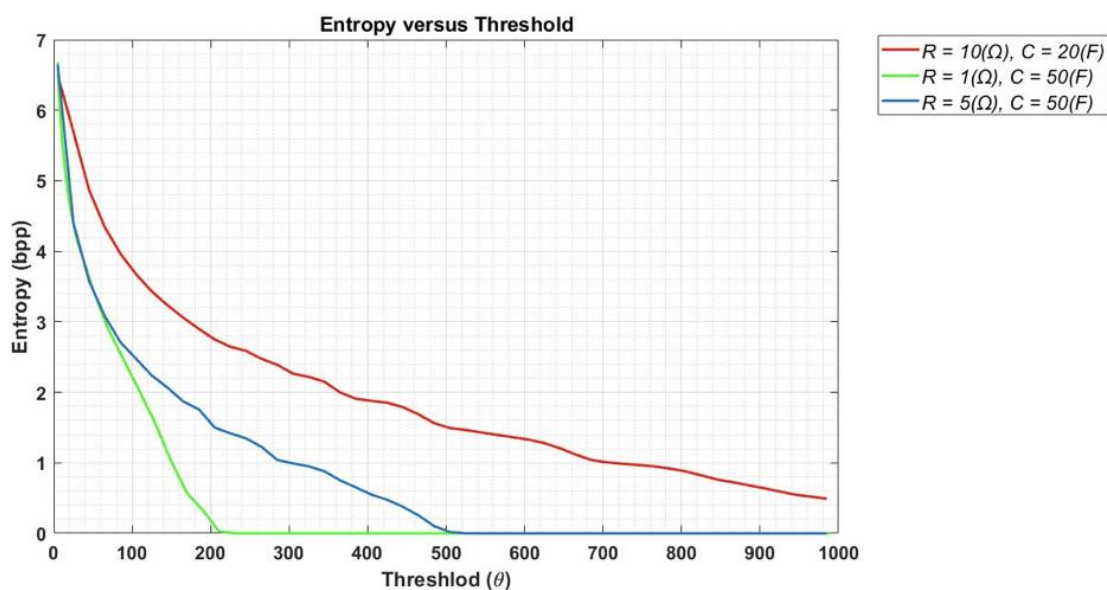
Στα δύο παραπάνω πειράματα (Εικόνες 3.14 και 3.15) όλες οι παράμετροι του DUAL-SIM κβαντιστή κρατούνται σταθερές εκτός από το T , για το οποίο γίνεται η μελέτη, και το ζεύγος «πυκνωτής C με κατώφλι θ » βάσει του οποίου ζεύγους γίνονται συγκρίσεις μεταξύ των πειραμάτων. Συγκεκριμένα, για τον πυκνωτή C επιλέχθηκαν δύο τιμές (20 F και 50 F) για κάθε μία από τις οποίες επιλέχθηκαν μία τιμή κατωφλίου θ (50 V). Για αυτούς τους συνδυασμούς

των θ και C χρησιμοποιήθηκαν έξι διαφορετικές τιμές για την αντίσταση R οι οποίες απεικονίζονται στα υπομνήματα των εικόνων των παραπάνω δύο πειραμάτων. Από τα εν λόγω πειράματα απορρέει το εξής συμπέρασμα :

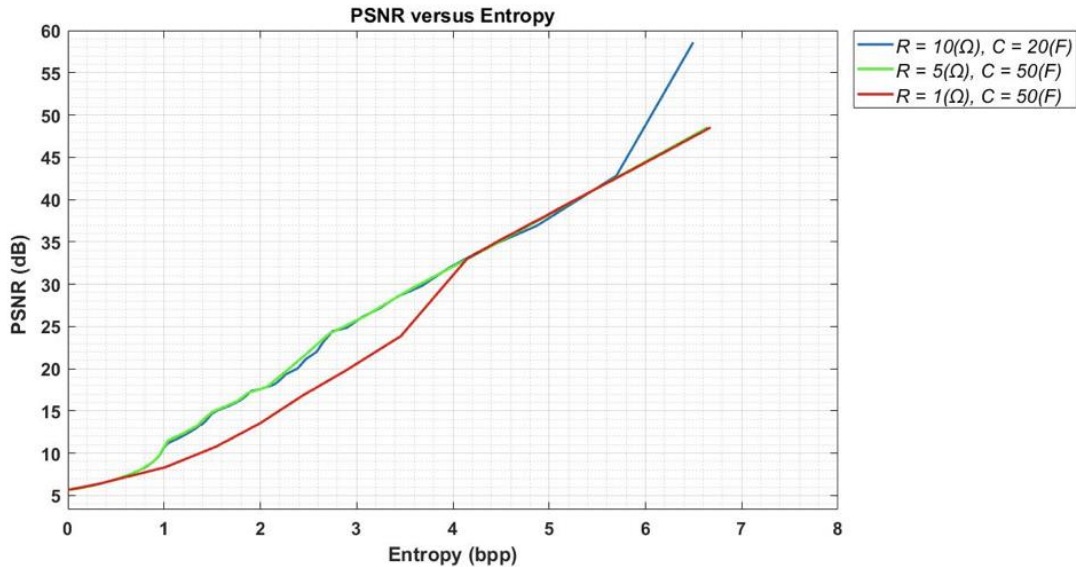
Όσο αυξάνεται η τιμή του χρονικού παραθύρου T , ανεξαρτήτως των άλλων μεταβλητών, το MSE που δίνει ο κβαντιστής μειώνεται, όμως από μια τιμή T και έκτοτε η αύξησή του δεν επιφέρει σχεδόν καμία μείωση στο MSE. Το T παύει να συνεισφέρει στη μείωση του MSE πιο γρήγορα όταν η αντίσταση R είναι μεγάλη και ο πυκνωτής C μικρός (Εικόνα 3.14) από ό,τι όταν και η αντίσταση R και ο πυκνωτής C έχουν μεγάλη τιμή (Εικόνα 3.15).

Επίσης επιβεβαιώνεται πως ο συνδυασμός «μικρή τιμή για τον πυκνωτή C και μεγάλη τιμή για την αντίσταση R » προκαλεί τον DUAL-SIM κβαντιστή να επιφέρει μικρό MSE.

Τέλος, μελετήθηκε η εντροπία που μπορεί να πετύχει ο κβαντιστής συναρτήσει του κατωφλίου θ για συγκεκριμένα R και C έχοντας το T στα $150ms$ (Εικόνα 3.16) καθώς και το PSNR που αποδίδει για αυτά τα R , C και T (Εικόνα 3.17). Συγκεκριμένα :



Εικόνα 3.16 : Εντροπία κβαντιστή δίχως χρήση μετασχηματισμού συναρτήσει του θ για $T = 150ms$ και διαφόρους συνδυασμούς R & C



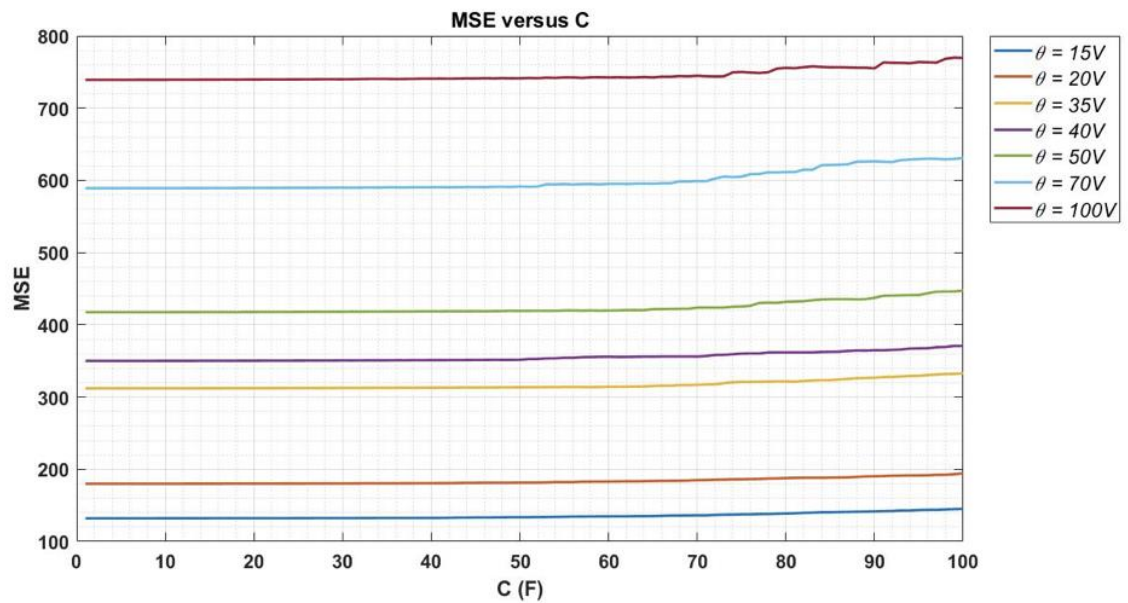
Εικόνα 3.17 : PSNR κβαντιστή δίχως χρήση μετασχηματισμού συναρτῆσει της Εντροπίας για $T = 150ms$ και διαφόρους συνδυασμούς R & C

Από τα δύο παραπάνω πειράματα προκύπτει πως ο κβαντιστής DUAL-SIM αποδίδει καλύτερα για μεγάλο R σε συνδυασμό με χαμηλό C και ότι για χαμηλό κατώφλι θ υπάρχει καλύτερη ανάκτηση εικόνας, δηλαδή υψηλό PSNR διατηρώντας ταυτόχρονα χαμηλά την Εντροπία.

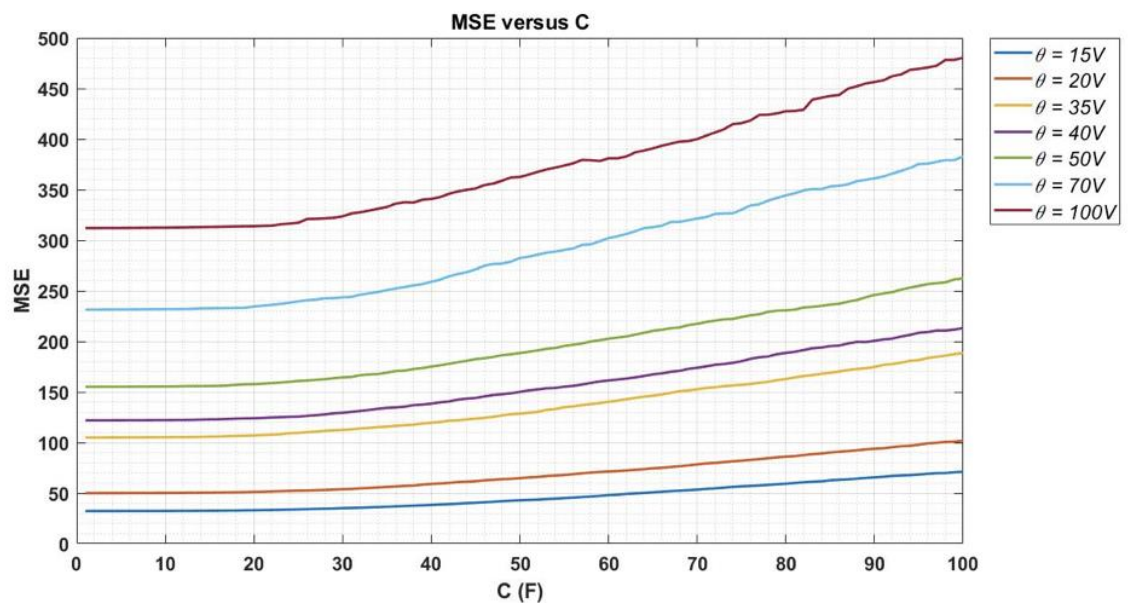
3.3.2. Με Μετασχηματισμό Συνημιτόνου (DCT)

Συνεχίζοντας τη μελέτη του DUAL-SIM κβαντιστή, στα πειράματα προστέθηκε ο DCT μετασχηματισμός, ο ίδιος μετασχηματισμός δηλαδή που χρησιμοποιεί ο JPEG αλγόριθμος. Μελετήθηκε ξεχωριστά η συμπεριφορά των τριών βασικών παραμέτρων του κβαντιστή, η αντίσταση R , ο πυκνωτής C και το χρονικό παράθυρο παρατήρησης T . Κάθε φορά που μία από αυτές τις τρεις παραμέτρους ετίθετο υπό μελέτη οι άλλες δύο κρατιούνταν σταθερές. Εν ολίγοις στα επόμενα πειράματα θα παρατηρηθεί πως η συμπεριφορά του κβαντιστή παραμένει ίδια και το μόνο που αλλάζει είναι η απόδοση (MSE , $PSNR$ και $Εντροπία$) που αυτός πετυχαίνει. Συγκεκριμένα :

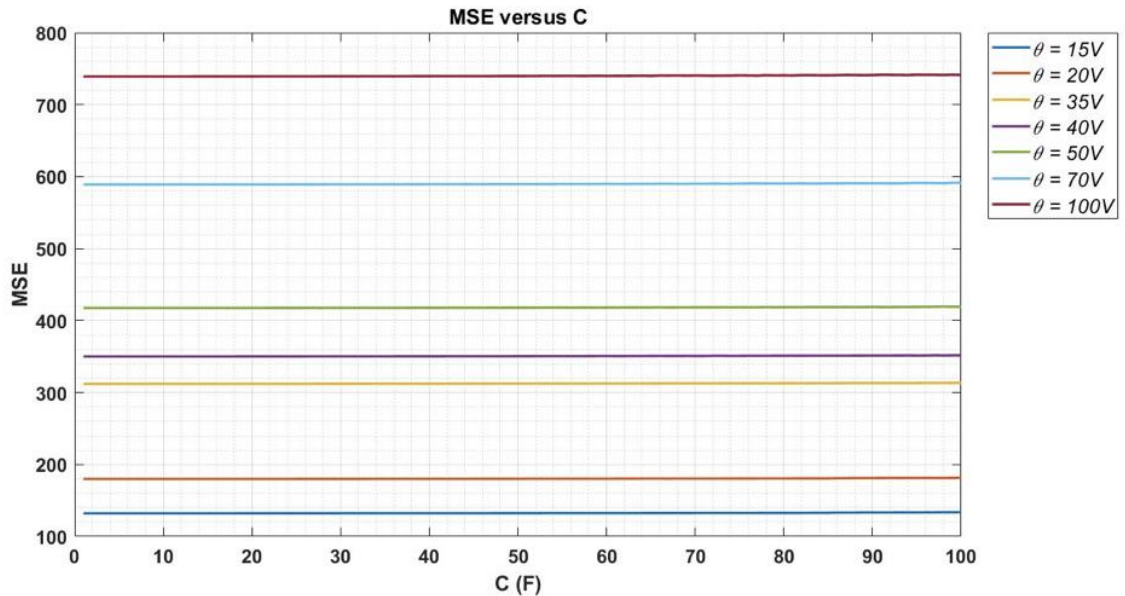
■ Μελέτη του πυκνωτή C



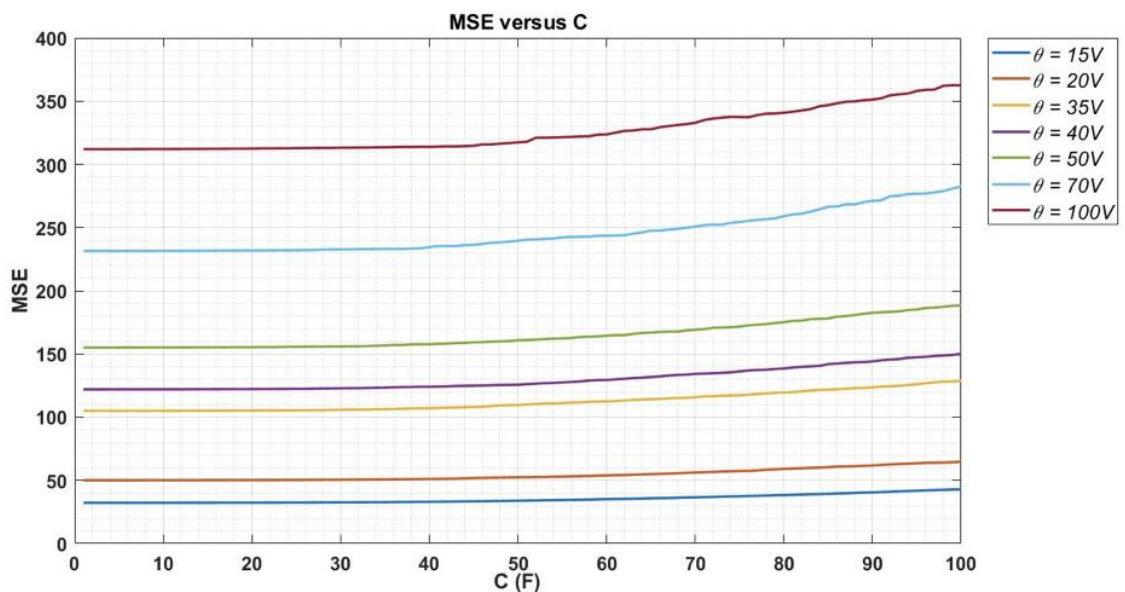
Εικόνα 3.18 : Παράμετρος C συναρτήσει του MSE για διάφορα θ και $T = 100 \text{ ms}$ & $R = 0.35 \Omega$ με χρήση DCT



Εικόνα 3.19 : Παράμετρος C συναρτήσει του MSE για διάφορα θ και $T = 100 \text{ ms}$ & $R = 1 \Omega$ με χρήση DCT



Εικόνα 3.20 : Παράμετρος C συναρτήσει του MSE για διάφορα θ και $T = 200 \text{ ms}$ & $R = 0.35 \Omega$ με χρήση DCT



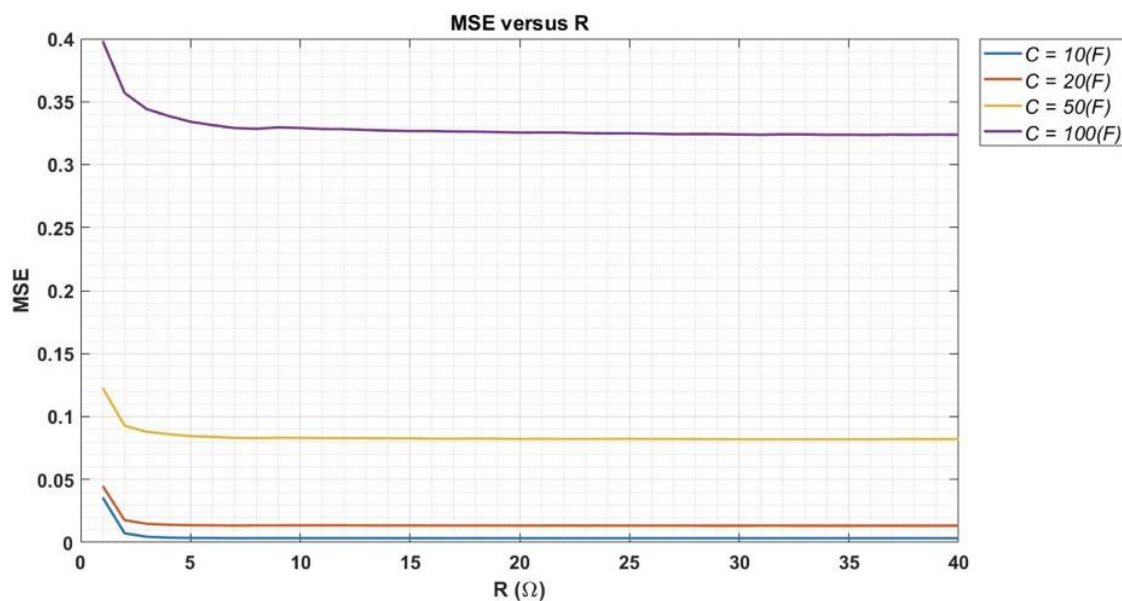
Εικόνα 3.21 : Παράμετρος C συναρτήσει του MSE για διάφορα θ και $T = 200 \text{ ms}$ & $R = 1 \Omega$ με χρήση DCT

Στα παραπάνω τέσσερα πειράματα (Εικόνες 3.18 έως και 3.21) όλες οι παράμετροι του DUAL-SIM κβαντιστή κρατούνται σταθερές εκτός από τον πυκνωτή C , για τον οποίο γίνεται η μελέτη, και το ζεύγος «χρονικό παράθυρο T με αντίσταση R » βάσει του οποίου γίνονται συγκρίσεις μεταξύ των πειραμάτων. Συγκεκριμένα, για το παράθυρο T επιλέχθηκαν δύο τιμές (100ms και 200ms) για κάθε μία από τις οποίες επιλέχθηκαν δύο διαφορετικές τιμές για την

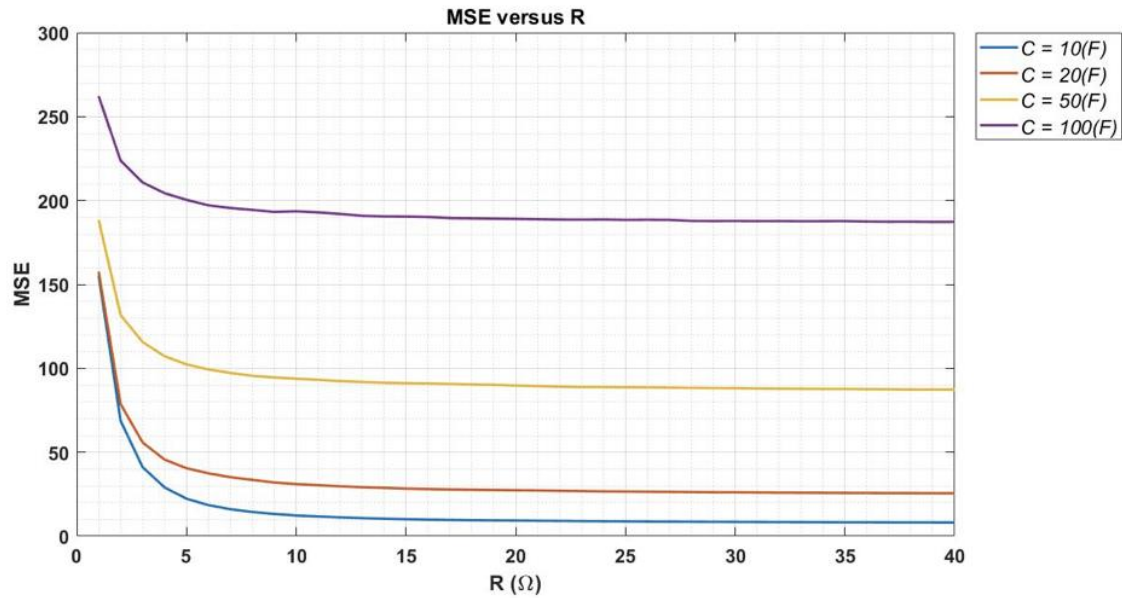
αντίσταση R (0.35Ω και 1Ω). Για αυτούς τους συνδυασμούς των R και T χρησιμοποιήθηκαν επτά διαφορετικές τιμές κατωφλίου θ οι οποίες απεικονίζονται στα υπομνήματα των εικόνων των παραπάνω τεσσάρων πειραμάτων. Από τα εν λόγω πειράματα επαληθεύονται τα συμπεράσματα που παρατηρήθηκαν από προηγούμενα πειράματα (τα πειράματα των Εικόνων 3.4 έως και 3.7). Συγκεκριμένα, όσο αυξάνεται η τιμή του κατωφλίου θ τόσο αυξάνεται και το MSE που δίνει ο κβαντιστής. Επίσης όσο αυξάνεται η τιμή της αντίστασης R τόσο μειώνεται το MSE ενώ, όσο αυξάνεται η τιμή του πυκνωτή C τόσο αυξάνεται το MSE. Τέλος, αυξάνοντας το χρονικό παράθυρο T , η συμπεριφορά του κβαντιστή μένει ίδια με τη μόνη διαφορά να είναι η λιγότερα απότομη αύξηση του MSE.

Επομένως προκύπτει το συμπέρασμα πως η προσθήκη του μετασχηματισμού DCT δεν αλλάζει ή παραποιεί τη συμπεριφορά του DUAL-SIM κβαντιστή. Εν αντιθέσει ο μετασχηματισμός ενισχύει τη διαδικασία της κβάντισης (και ακολούθως της αποκβάντισης) καθώς ο DUAL-SIM πετυχαίνει πλέον πολύ μικρότερα επίπεδα MSE.

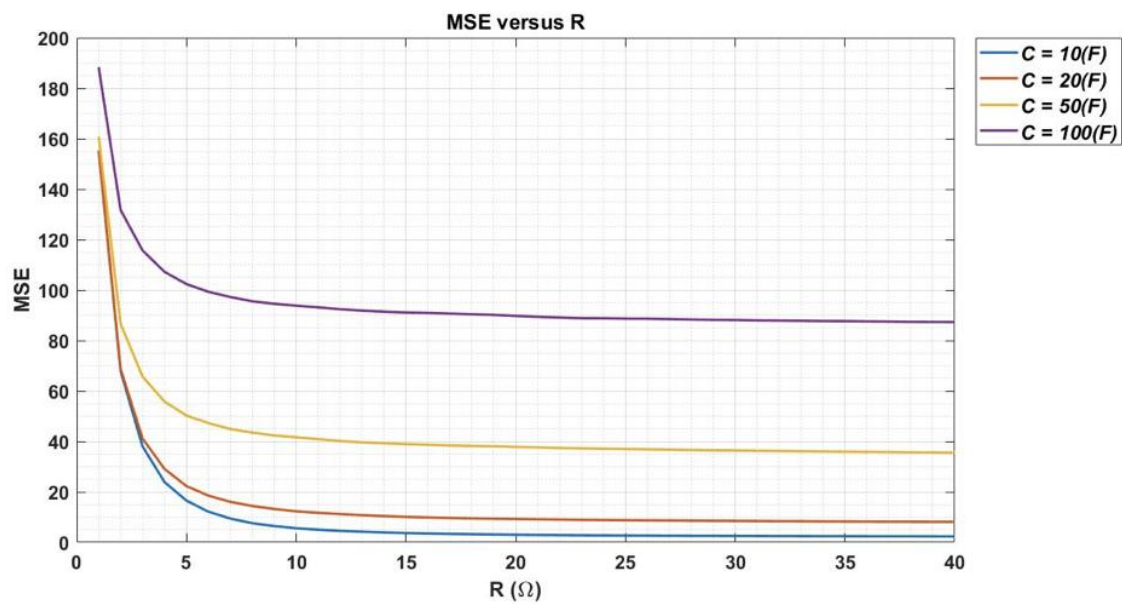
■ Μελέτη της αντίστασης R



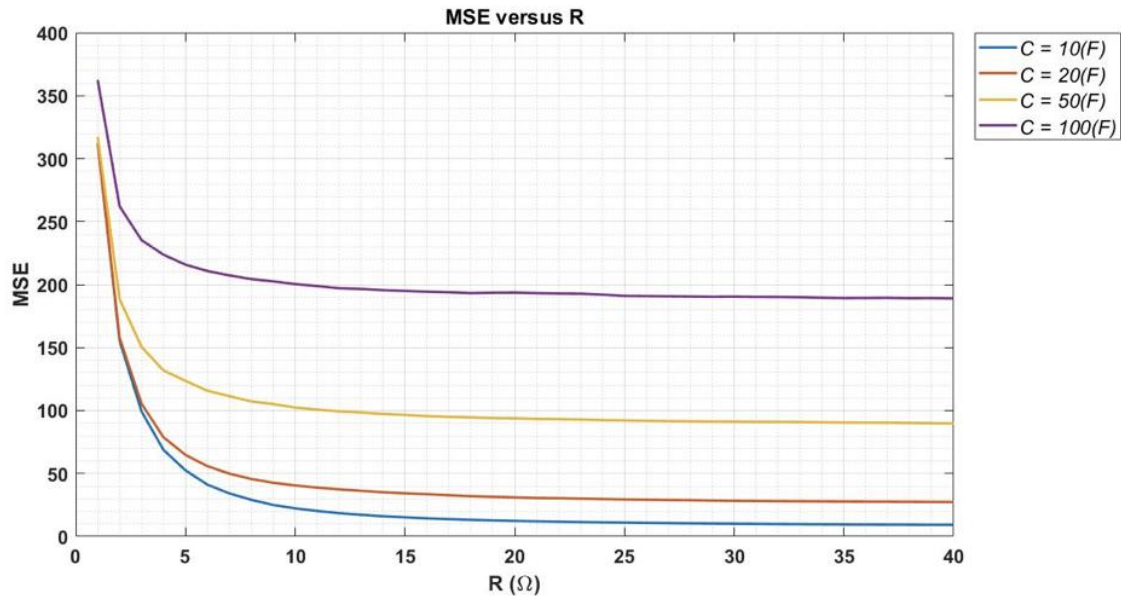
Εικόνα 3.22 : Παράμετρος R συναρτῆσει του MSE για διάφορα C και $T = 100\text{ ms}$ & $\theta = 1\text{ V}$ με χρήση DCT



Εικόνα 3.23 : Παράμετρος R συναρτῆσει του MSE για διάφορα C και $T = 100\ ms$ & $\theta = 50\ V$ με χρήση DCT

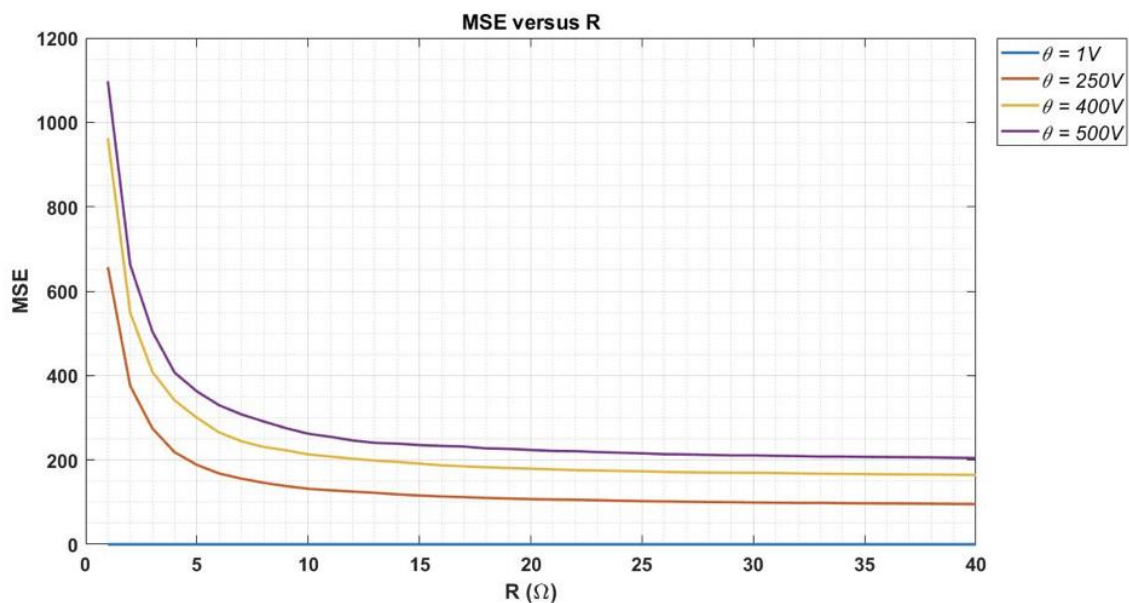


Εικόνα 3.24 : Παράμετρος R συναρτῆσει του MSE για διάφορα C και $T = 200\ ms$ & $\theta = 50\ V$ με χρήση DCT

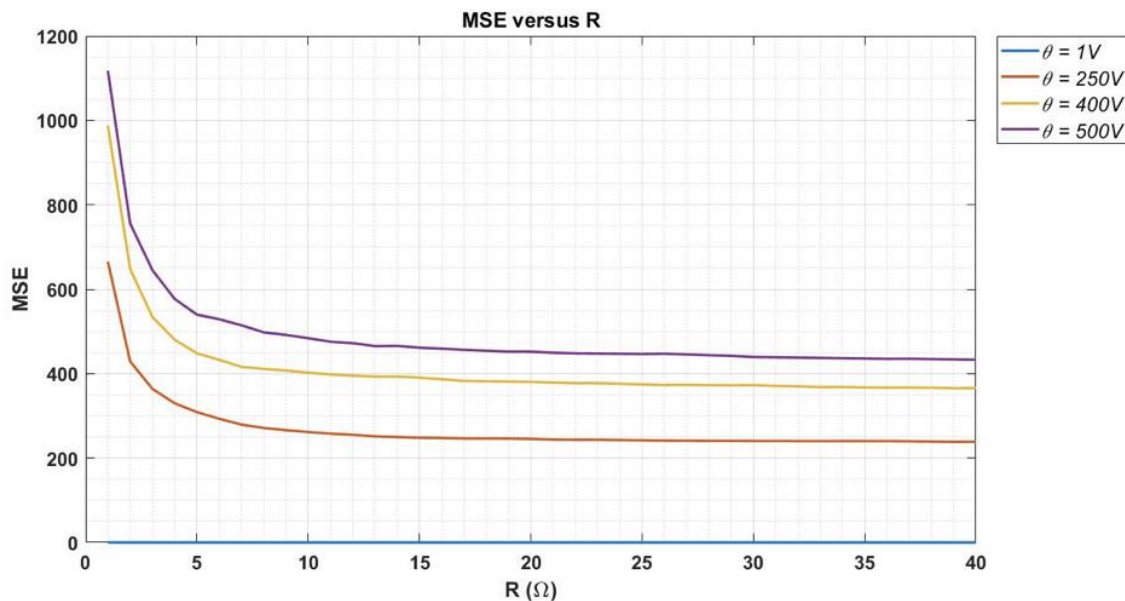


Εικόνα 3.25 : Παράμετρος R συναρτῆσει του MSE για διάφορα C και $T = 200\ ms$ & $\theta = 100\ V$ με χρήση DCT

Στα παραπάνω τέσσερα πειράματα (Εικόνες 3.22 έως και 3.25) όλες οι παράμετροι του DUAL-SIM κβαντιστή κρατούνται σταθερές εκτός από την R , για την οποία γίνεται η μελέτη, και το ζεύγος «χρονικό παράθυρο T με κατώφλι θ » βάσει του οποίου γίνονται συγκρίσεις μεταξύ των πειραμάτων. Συγκεκριμένα, για το παράθυρο T επιλέχθηκαν δύο τιμές ($100ms$ και $200ms$) για κάθε μία από τις οποίες επιλέχθηκαν δύο διαφορετικές ζεύγη τιμών για το κατώφλι θ ($1V$ και $50V$ για τα $100ms$ και $50V$ και $100V$ για τα $200ms$). Για αυτούς τους συνδυασμούς των θ και T χρησιμοποιήθηκαν τέσσερις διαφορετικές τιμές για τον πυκνωτή C οι οποίες απεικονίζονται στα υπομνήματα των εικόνων αυτών των τεσσάρων πειραμάτων.



Εικόνα 3.26 : Παράμετρος R συναρτῆσει του MSE για διάφορα θ και $T = 200\ ms$ & $C = 20\ F$ με χρήση DCT

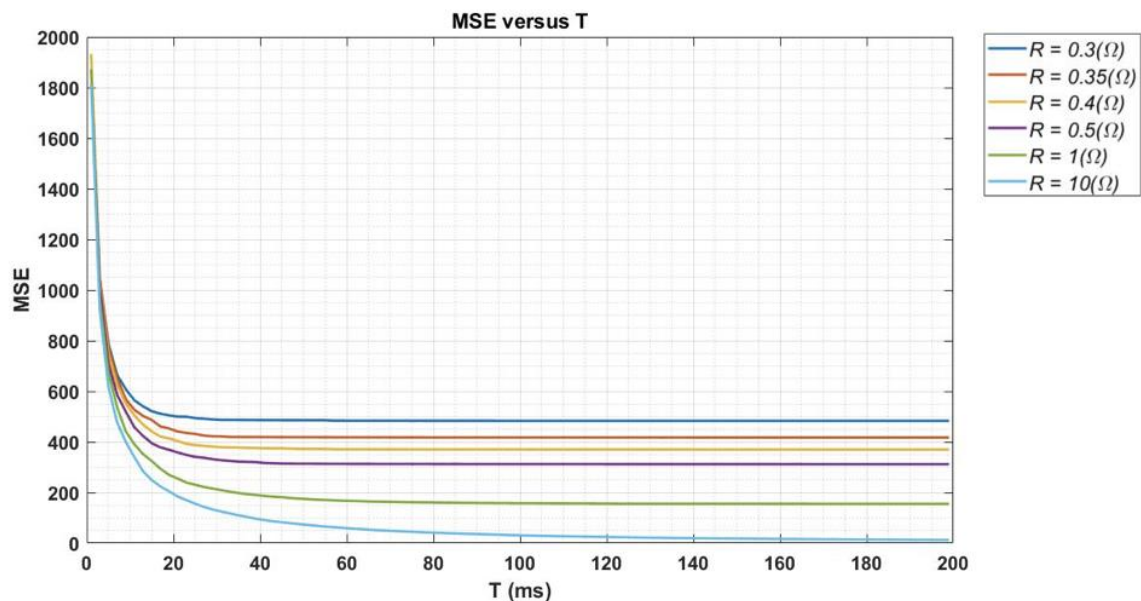


Εικόνα 3.27 : Παράμετρος R συναρτίζει του MSE για διάφορα θ και $T = 200 \text{ ms}$ & $C = 50 \text{ F}$ με χρήση DCT

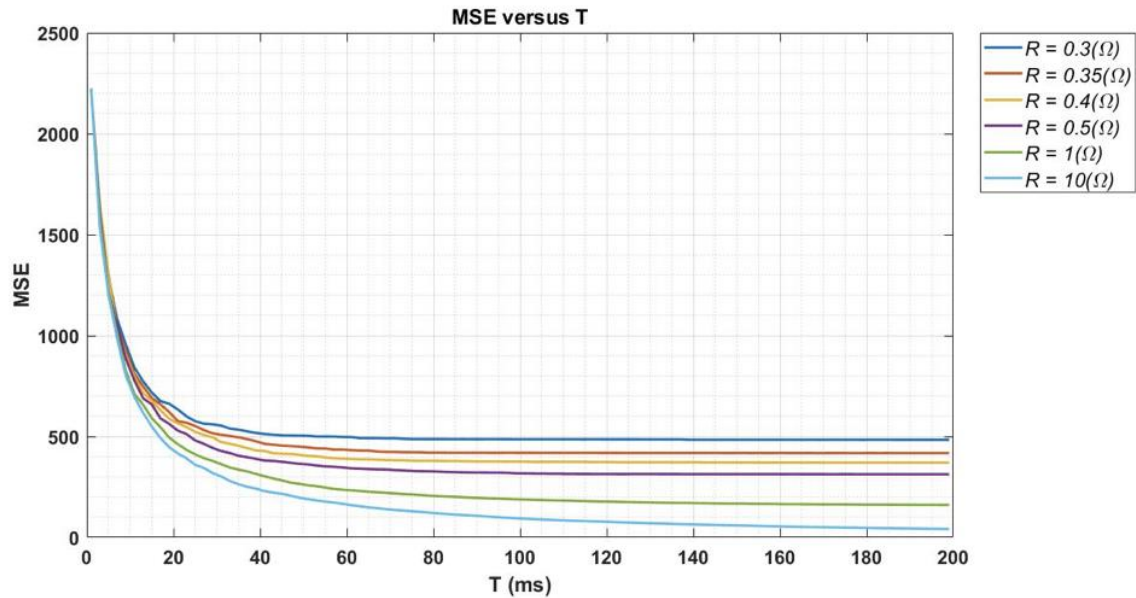
Στα επακόλουθα δύο πειράματα (Εικόνες 3.26 και 3.27) ακολουθήθηκε διαφορετική προσέγγιση, συγκεκριμένα για μία αυξανόμενη τιμή R επιλέχθηκε ένα μόνο παράθυρο T (200ms) για το οποίο επιλέχθηκαν δύο διαφορετικές τιμές C (20F και 50F). Για αυτούς τους συνδυασμούς των C και T χρησιμοποιήθηκαν τέσσερις διαφορετικές τιμές κατωφλίου θ οι οποίες απεικονίζονται στα υπομνήματα των εικόνων αυτών των πειραμάτων. Από αυτά τα έξι πειράματα επιβεβαιώνεται η ίδια συμπεριφορά του κβαντιστή με προηγούμενα πειράματα (τα πειράματα των Εικόνων 3.8 έως και 3.13). Αναλυτικότερα, αυξάνοντας είτε την τιμή του πυκνωτή C είτε την τιμή του κατωφλίου θ τόσο αυξάνεται και το MSE , ενώ αυξάνοντας την τιμή της αντίστασης R τόσο το MSE μειώνεται. Επιπλέον, αυξάνοντας το χρονικό παράθυρο T δεν παρατηρείται κάποια επιπλέον αλλαγή παρά μόνο ότι το MSE μειώνεται λιγότερο απότομα όσο το R αυξάνεται για ένα σταθερό C . Επομένως απορρέει το συμπέρασμα πως η χρήση του DCT μετασχηματισμού δεν άλλαξε τη συμπεριφορά του κβαντιστή, αλλά αντιθέτως την ωφέλησε καθώς τα επίπεδα MSE μειώθηκαν αξιοσημείωτα (π.χ. για $T = 200 \text{ ms}$, $C = 50 \text{ F}$, $\theta = 500 \text{ V}$ και αυξανόμενη R ο κβαντιστής χωρίς DCT πετυχαίνει MSE που κυμαίνεται από 18000 έως και 6000, ενώ με DCT πετυχαίνει MSE που κυμαίνεται από 1100 έως και 500).

■ Μελέτη του χρονικού παραθύρου T

Στα δύο παραπάνω πειράματα (Εικόνες 3.28 και 3.29) όλες οι παράμετροι του DUAL-SIM κβαντιστή κρατούνται σταθερές εκτός από το T , για το οποίο γίνεται η μελέτη, και το ζεύγος «πυκνωτής C με κατώφλι θ » βάσει του οποίου γίνονται συγκρίσεις μεταξύ των πειραμάτων. Συγκεκριμένα, για τον πυκνωτή C επιλέχθηκαν δύο τιμές ($20F$ και $50F$) για κάθε μία από τις οποίες επιλέχθηκαν μία τιμή κατωφλίου θ ($50V$). Για αυτούς τους συνδυασμούς των θ και C χρησιμοποιήθηκαν έξι διαφορετικές τιμές για την αντίσταση R οι οποίες απεικονίζονται στα υπομνήματα των εικόνων των παραπάνω δύο πειραμάτων. Από τα εν λόγω πειράματα επιβεβαιώνεται για άλλη μία φορά πως όσο αυξάνεται η τιμή του χρονικού παραθύρου T , ανεξαρτήτως των άλλων μεταβλητών, το MSE που δίνει ο κβαντιστής μειώνεται. Επίσης, επιβεβαιώνεται πως ο μετασχηματισμός DCT συνέβαλε στην απόδοση του κβαντιστή καθώς τα επίπεδα του MSE μειώθηκαν αξιόλογα (π.χ. για $R = 0.3 \Omega, C = 50 F, \theta = 50 V$ και αυξανόμενο T ο κβαντιστής χωρίς DCT πετυχαίνει MSE που κυμαίνεται από 18000 έως και 11000, ενώ με DCT πετυχαίνει MSE που κυμαίνεται από 2200 έως και 600).

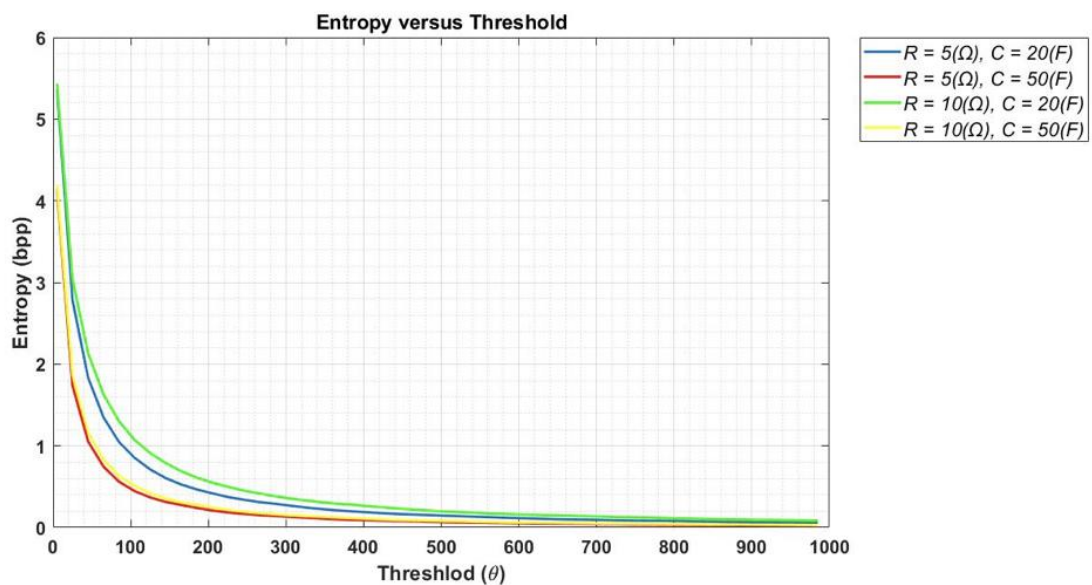


Εικόνα 3.28 : Παράμετρος T συναρτήσει του MSE για διάφορα R και $\theta = 50 V$ & $C = 20 F$ με χρήση DCT

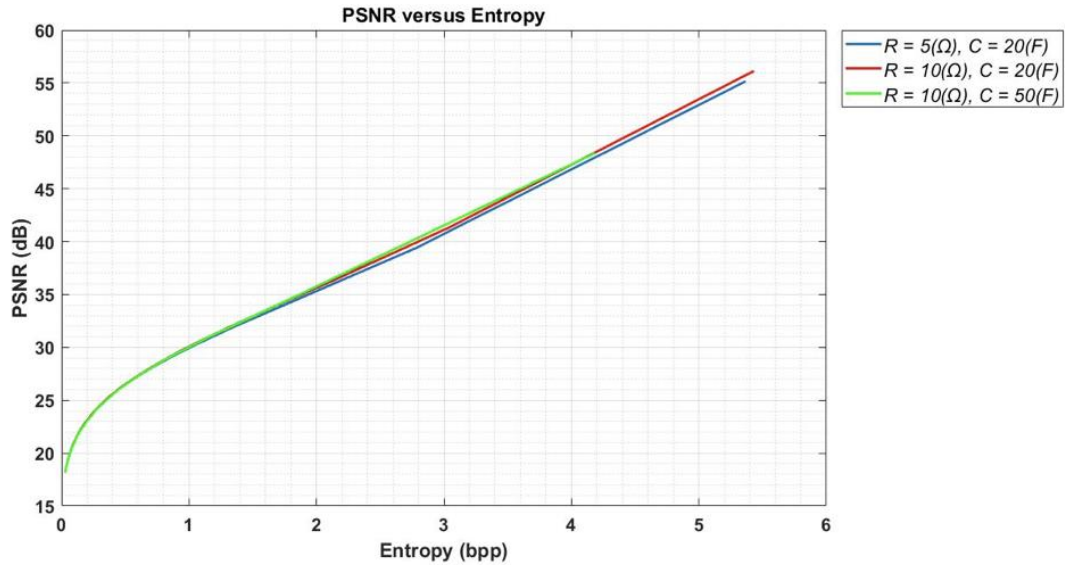


Εικόνα 3.29 : Παράμετρος T συναρτήσει του MSE για διάφορα R και $\theta = 50 \text{ V}$ & $C = 50 \text{ F}$ με χρήση DCT

Τέλος, μελετήθηκε η εντροπία που μπορεί να πετύχει ο κβαντιστής συναρτήσει του κατωφλίου θ για συγκεκριμένα R και C έχοντας το T στα $150ms$ (Εικόνα 3.30) καθώς και το PSNR που αποδίδει για αυτά τα R , C και T (Εικόνα 3.31) κάνοντας χρήση του μετασχηματισμού DCT . Συγκεκριμένα :



Εικόνα 3.30 : Εντροπία κβαντιστή με χρήση DCT μετασχηματισμού συναρτήσει του θ για $T = 150ms$ και διαφόρους συνδυασμούς R & C



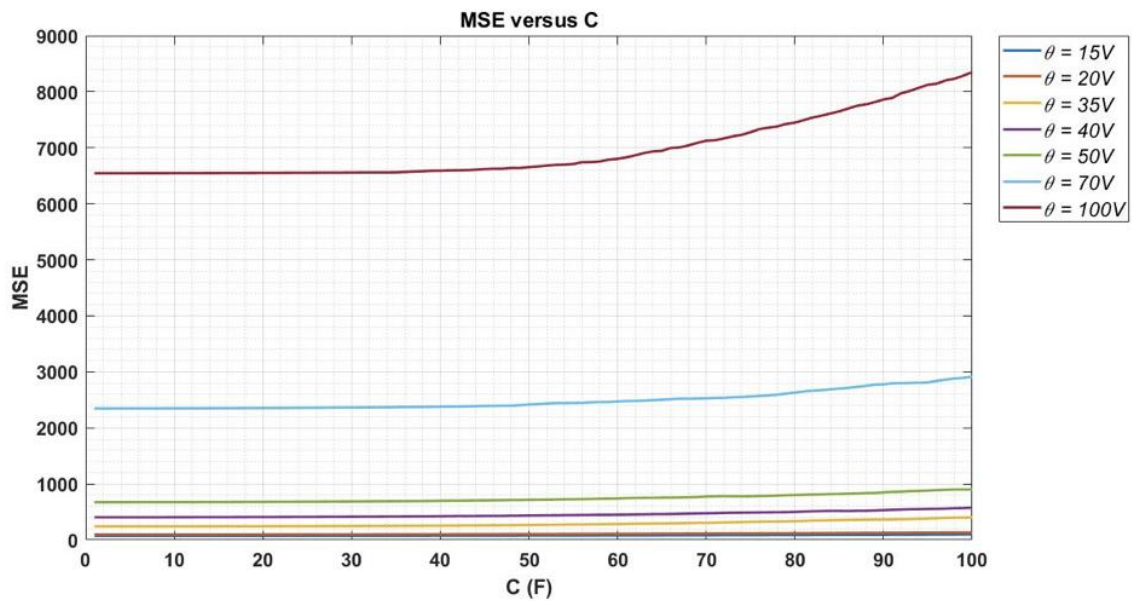
Εικόνα 3.31 : PSNR κβαντιστή με χρήση DCT μετασχηματισμού συναρτήσεως της Εντροπίας για $T = 150ms$ και διαφόρους συνδυασμούς R & C

Από τα δύο παραπάνω πειράματα προκύπτει πως ο κβαντιστής DUAL-SIM αποδίδει καλύτερα για μεγάλο R σε συνδυασμό με χαμηλό C και ότι για χαμηλό κατώφλι θ υπάρχει καλύτερη ανάκτηση εικόνας, δηλαδή υψηλό PSNR διατηρώντας ταυτόχρονα χαμηλά την Εντροπία. Σημειώνεται επίσης, πως η χρήση του DCT μετασχηματισμού συμβάλλει αρκετά στην απόδοση του κβαντιστή καθώς για μικρή εντροπία ($\sim 5.5 \text{ bpp}$) επιτυγχάνεται υψηλό PSNR ($\sim 57 \text{ dB}$) σε αντίθεση με τη μη χρήση του (όπου για $\sim 5.5 \text{ bpp}$ επιτυγχάνεται PSNR γύρω στα 43 dB). (Οι αριθμοί αυτοί αφορούν τον καλύτερο συνδυασμό R και C , εν προκειμένω $R = 5 \Omega$ και $C = 20 F$)

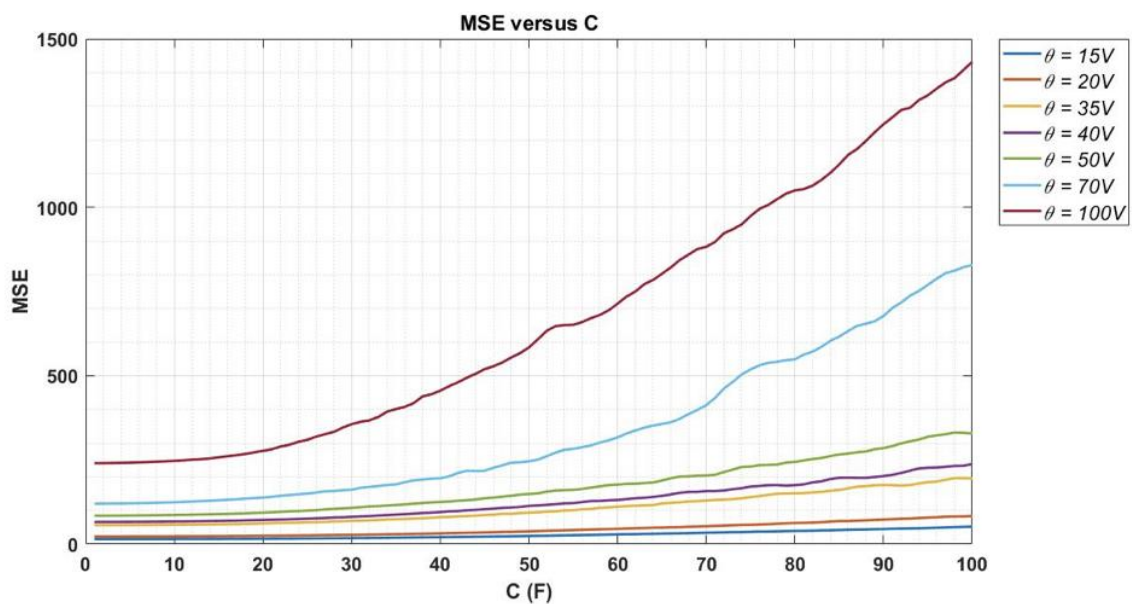
3.3.3 Με Μετασχηματισμό Κυματιδίων (DWT)

Στα τελευταία πειράματα της μελέτης του DUAL-SIM κβαντιστή προστέθηκε ο DWT μετασχηματισμός (στη θέση του DCT), ο ίδιος μετασχηματισμός δηλαδή που χρησιμοποιεί ο JPEG2000 αλγόριθμος και επαναμελετήθηκε ξεχωριστά η συμπεριφορά των τριών βασικών παραμέτρων του κβαντιστή (η αντίσταση R , ο πυκνωτής C και το χρονικό παράθυρο T). Εν ολίγοις στα επόμενα πειράματα θα παρατηρηθεί πως η συμπεριφορά του κβαντιστή παραμένει ίδια και το μόνο που αλλάζει είναι η απόδοση (MSE , $PSNR$ και $Εντροπία$) που αυτός πετυχαίνει. Συγκεκριμένα :

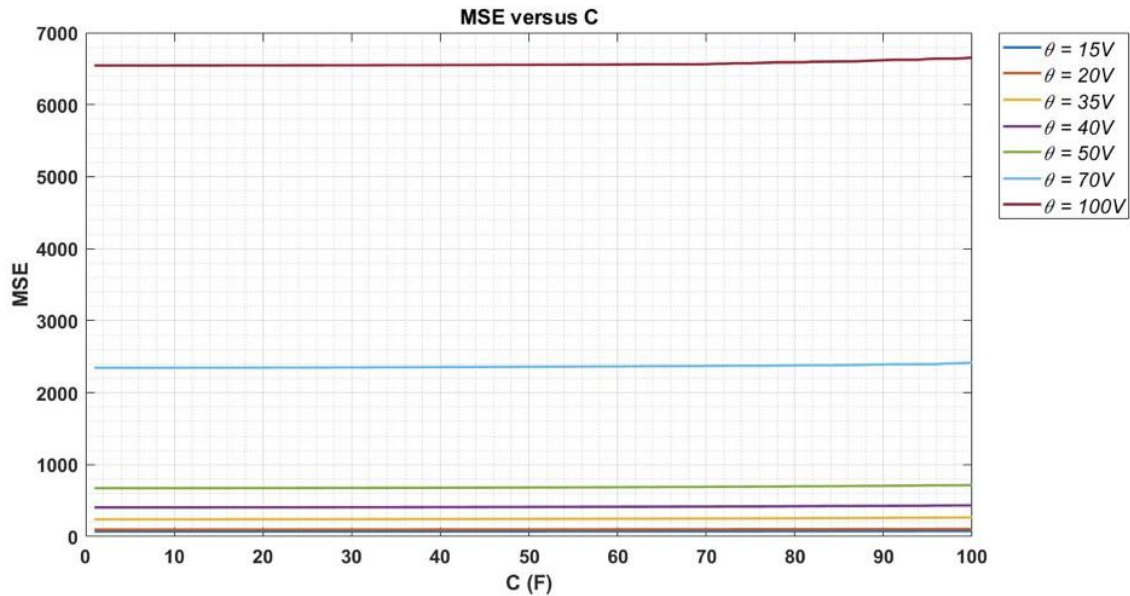
■ Μελέτη του πυκνωτή C



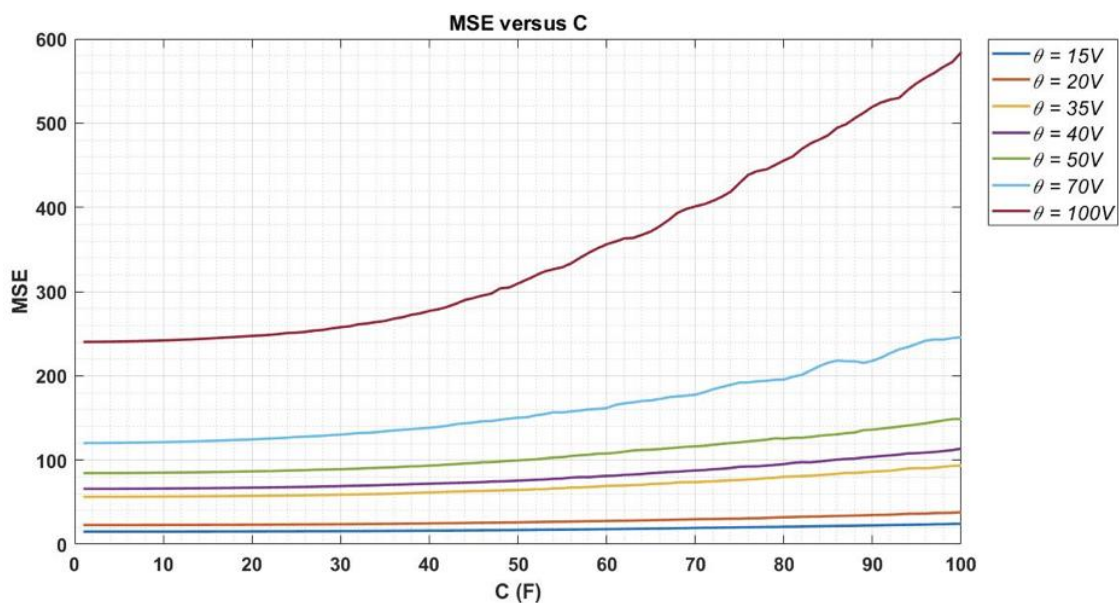
Εικόνα 3.32 : Παράμετρος C συναρτήσει του MSE για διάφορα θ και $T = 100 \text{ ms}$ & $R = 0.35 \Omega$ με χρήση DWT



Εικόνα 3.33 : Παράμετρος C συναρτήσει του MSE για διάφορα θ και $T = 100 \text{ ms}$ & $R = 1 \Omega$ με χρήση DWT



Εικόνα 3.34 : Παράμετρος C συναρτήσσει του MSE για διάφορα θ και $T = 200\text{ ms}$ & $R = 0.35\ \Omega$ με χρήση DWT



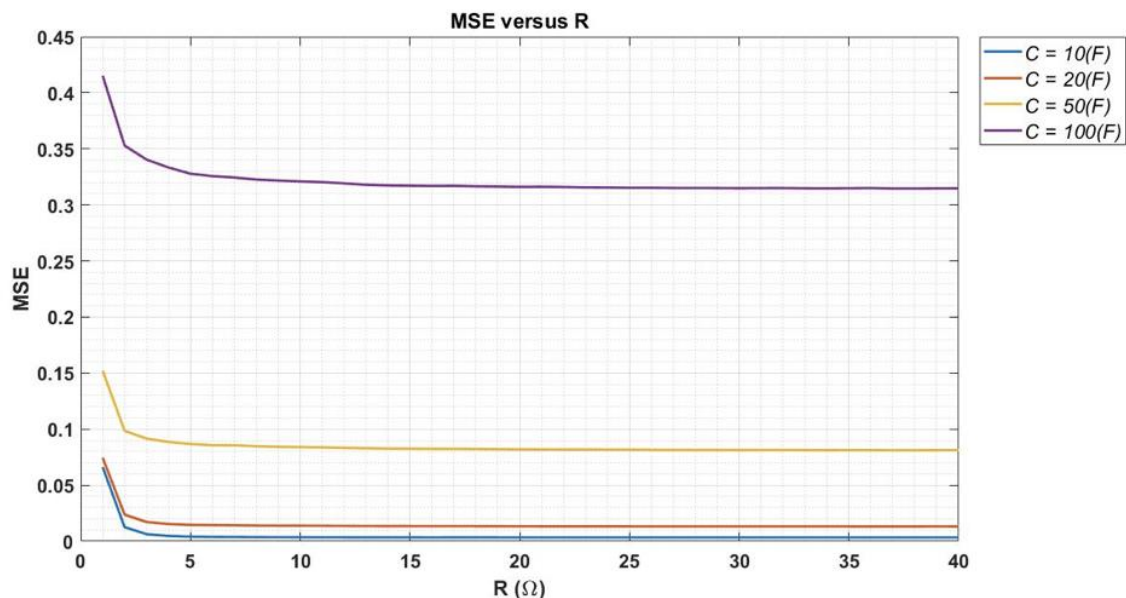
Εικόνα 3.35 : Παράμετρος C συναρτήσσει του MSE για διάφορα θ και $T = 200\text{ ms}$ & $R = 1\ \Omega$ με χρήση DWT

Στα παραπάνω τέσσερα πειράματα (Εικόνες 3.32 έως και 3.35) όλες οι παράμετροι του DUAL-SIM κβαντιστή κρατούνται σταθερές εκτός από τον πυκνωτή C , για τον οποίο γίνεται η μελέτη, και το ζεύγος «χρονικό παράθυρο T με αντίσταση R » βάσει του οποίου γίνονται συγκρίσεις μεταξύ των πειραμάτων. Συγκεκριμένα, για το παράθυρο T επιλέχθηκαν δύο τιμές

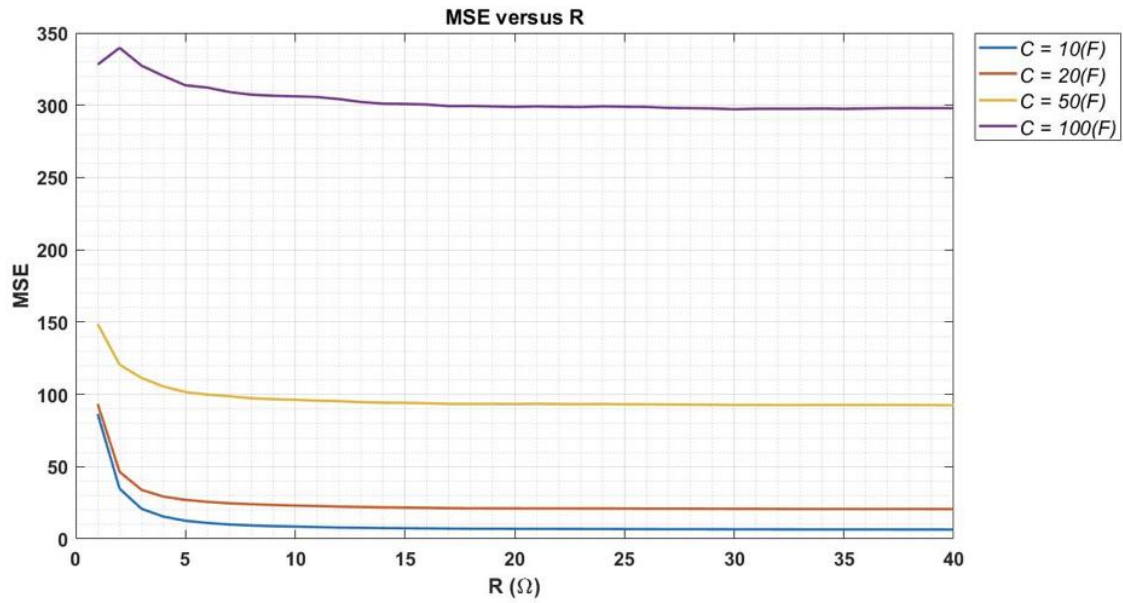
(100ms και 200ms) για κάθε μία από τις οποίες επιλέχθηκαν δύο διαφορετικές τιμές για την αντίσταση R (0.35Ω και 1Ω). Για αυτούς τους συνδυασμούς των R και T χρησιμοποιήθηκαν επτά διαφορετικές τιμές κατωφλίου θ οι οποίες απεικονίζονται στα υπομνήματα των εικόνων των παραπάνω τεσσάρων πειραμάτων. Από τα εν λόγω πειράματα επαληθεύονται τα συμπεράσματα που παρατηρήθηκαν από προηγούμενα πειράματα (τα πειράματα των Εικόνων 3.4 έως και 3.7). Συγκεκριμένα, όσο αυξάνεται η τιμή του κατωφλίου θ τόσο αυξάνεται και το MSE που δίνει ο κβαντιστής. Επίσης όσο αυξάνεται η τιμή της αντίστασης R τόσο μειώνεται το MSE ενώ, όσο αυξάνεται η τιμή του πυκνωτή C τόσο αυξάνεται το MSE. Τέλος, αυξάνοντας το χρονικό παράθυρο T , η συμπεριφορά του κβαντιστή μένει ίδια με τη μόνη διαφορά να είναι η λιγότερα απότομη αύξηση του MSE.

Επομένως προκύπτει το συμπέρασμα πως η προσθήκη του μετασχηματισμού DWT δεν αλλάζει ή παραποιεί τη συμπεριφορά του DUAL-SIM κβαντιστή. Εν αντιθέσει ο μετασχηματισμός ενισχύει τη διαδικασία της κβάντισης (και ακολούθως της αποκβάντισης) καθώς ο DUAL-SIM πετυχαίνει πλέον πολύ μικρότερα επίπεδα MSE, επίπεδα μικρότερα από όταν δεν χρησιμοποιείται μετασχηματισμός αλλά μεγαλύτερα από όταν χρησιμοποιείται ο μετασχηματισμός DCT (π.χ. για $R = 0.35\Omega$, $T = 100\text{ ms}$, $\theta = 100\text{ V}$ και αυξανόμενο C ο κβαντιστής χωρίς μετασχηματισμό πετυχαίνει MSE που κυμαίνεται από 18000 έως και 18200, με DCT πετυχαίνει MSE που κυμαίνεται από 740 έως και 790, ενώ με DWT πετυχαίνει MSE που κυμαίνεται από 6700 έως και 8300).

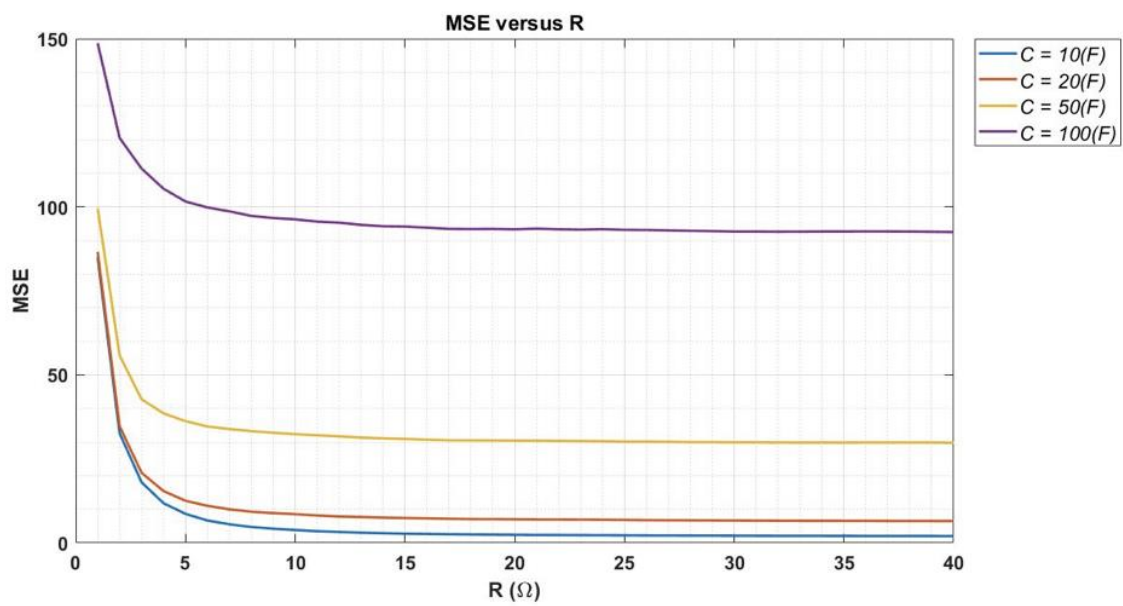
■ Μελέτη της αντίστασης R



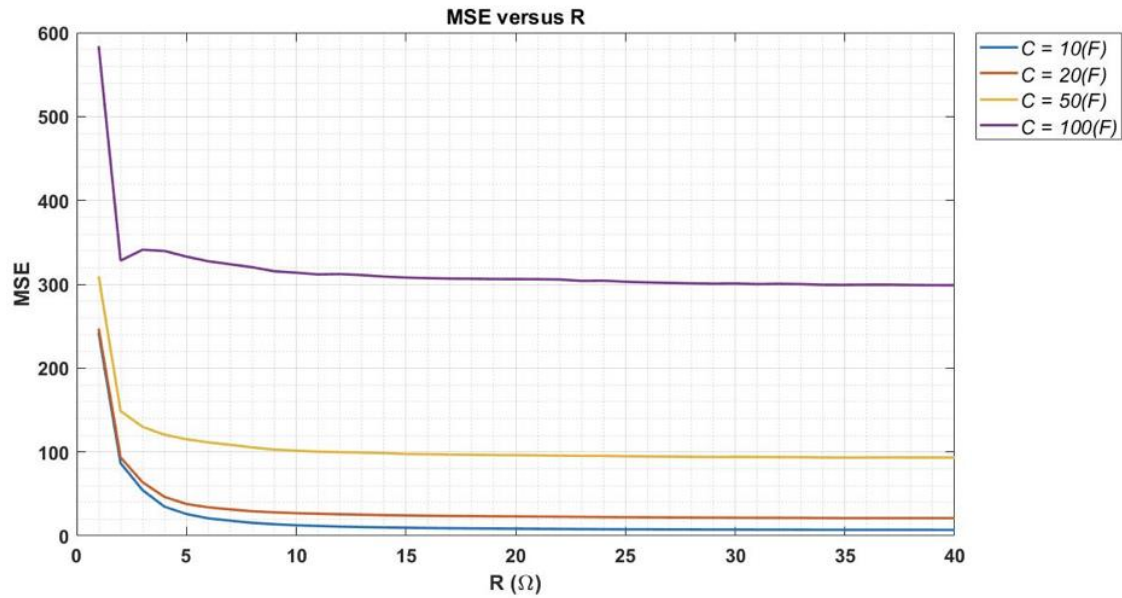
Εικόνα 3.36 : Παράμετρος R συναρτήσεως του MSE για διάφορα C και $T = 100\text{ ms}$ & $\theta = 1\text{ V}$ με χρήση DWT



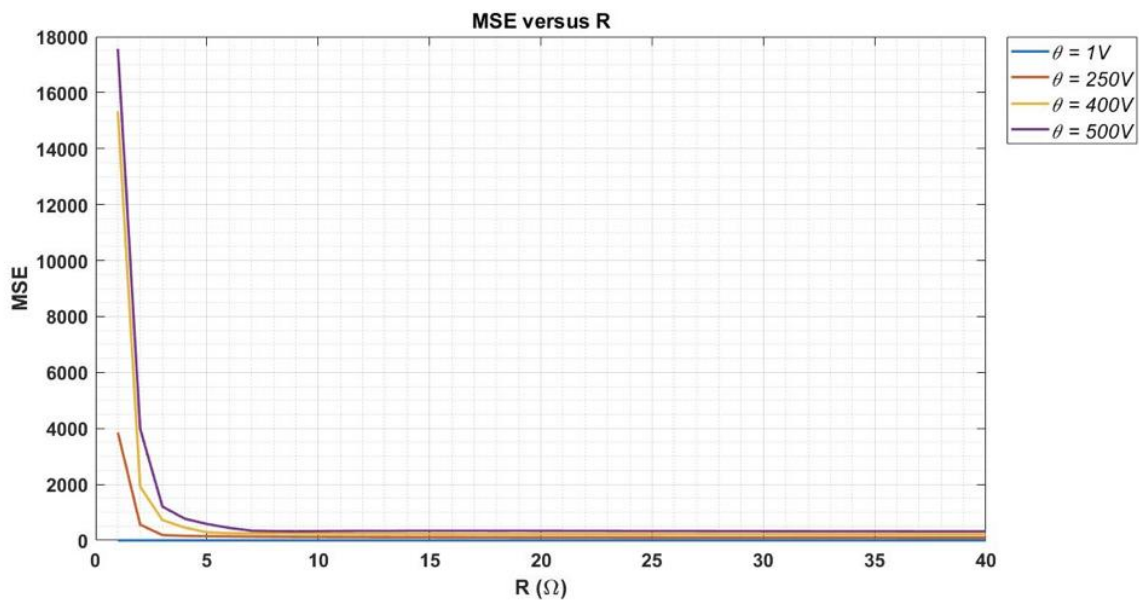
Εικόνα 3.37 : Παράμετρος R συναρτῆσει του MSE για διάφορα C και $T = 100\ ms$ & $\theta = 50\ V$ με χρήση DWT



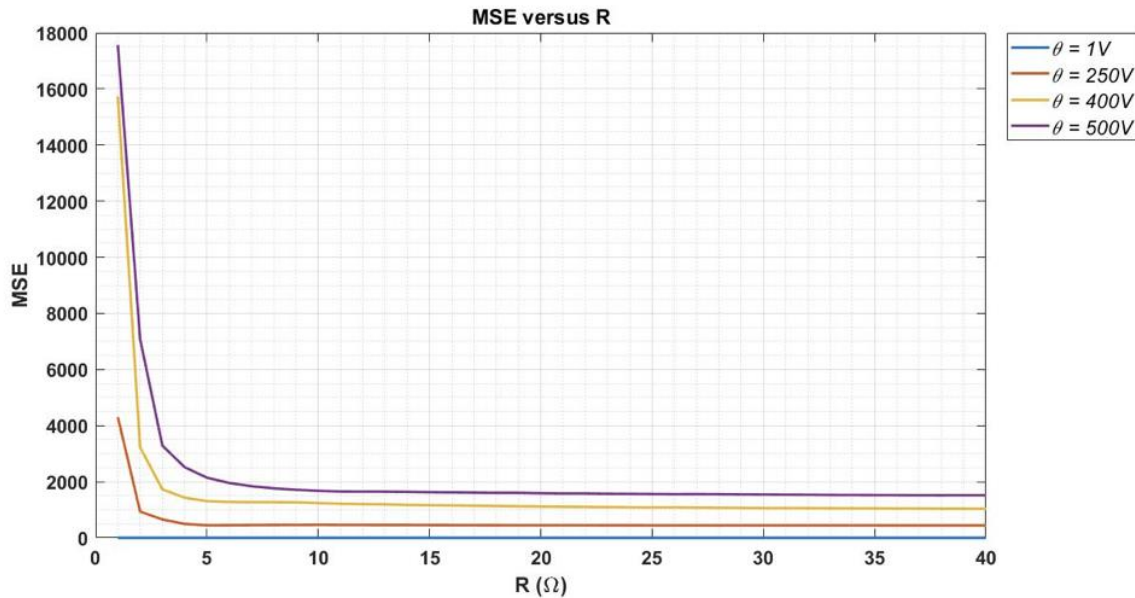
Εικόνα 3.38 : Παράμετρος R συναρτῆσει του MSE για διάφορα C και $T = 200\ ms$ & $\theta = 50\ V$ με χρήση DWT



Εικόνα 3.39 : Παράμετρος R συναρτήσει του MSE για διάφορα C και $T = 200\ ms$ & $\theta = 100\ V$ με χρήση DWT



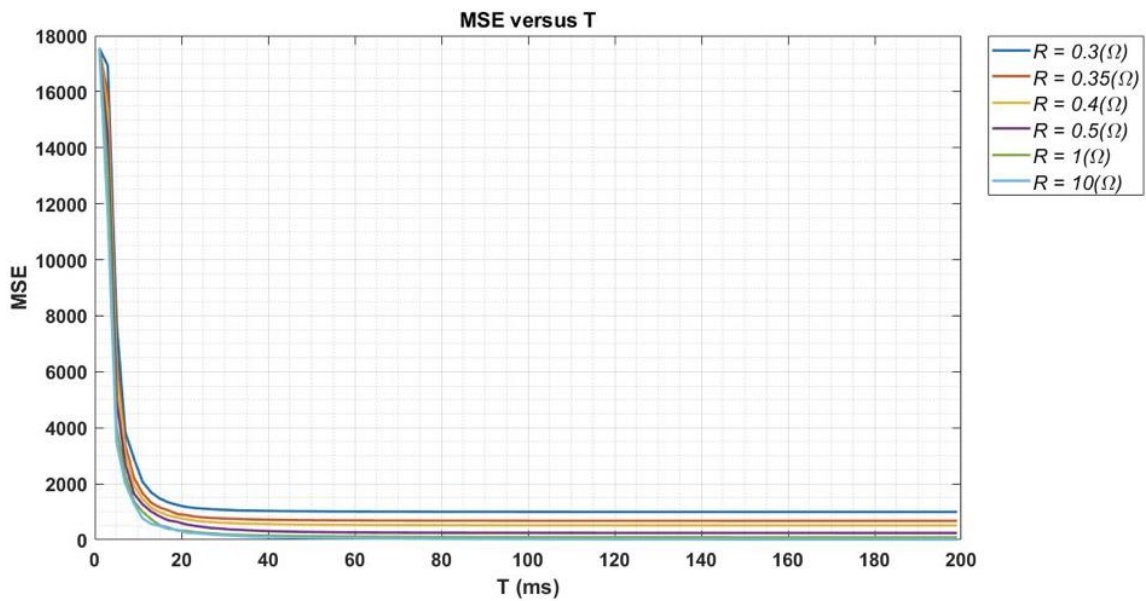
Εικόνα 3.40 : Παράμετρος R συναρτήσει του MSE για διάφορα θ και $T = 200\ ms$ & $C = 20\ F$ με χρήση DWT



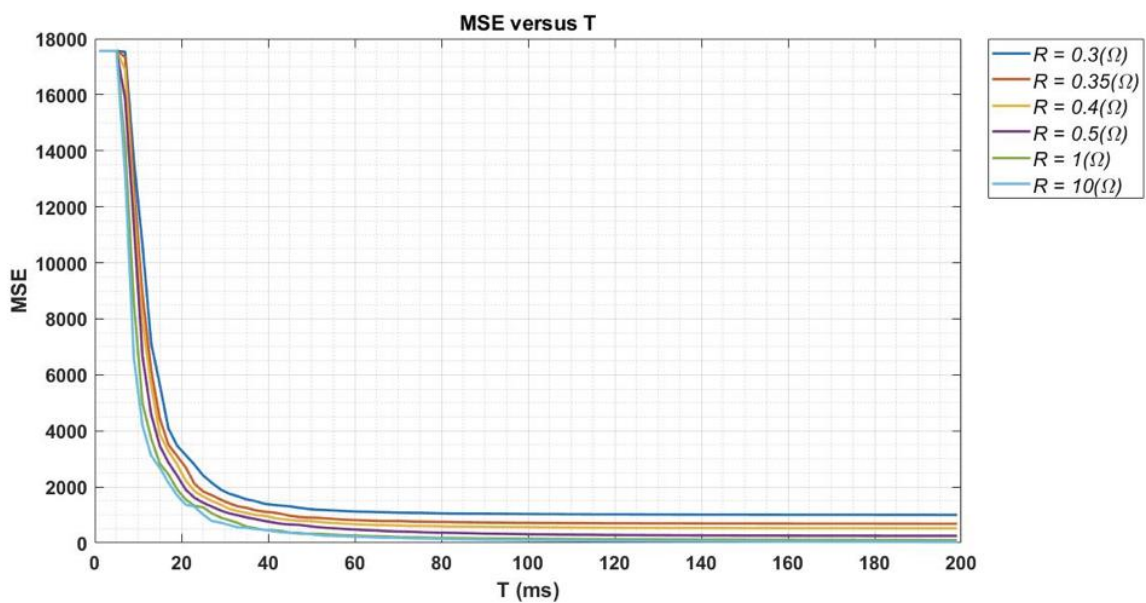
Εικόνα 3.41 : Παράμετρος R συναρτῆσει του MSE για διάφορα θ και $T = 200\text{ ms}$ & $C = 50\text{ F}$ με χρήση DWT

Στα παραπάνω τέσσερα πειράματα (Εικόνες 3.36 έως και 3.39) όλες οι παράμετροι του DUAL-SIM κβαντιστή κρατούνται σταθερές εκτός από την R , για την οποία γίνεται η μελέτη, και το ζεύγος «χρονικό παράθυρο T με κατώφλι θ » βάσει του οποίου γίνονται συγκρίσεις μεταξύ των πειραμάτων. Συγκεκριμένα, για το παράθυρο T επιλέχθηκαν δύο τιμές (100ms και 200ms) για κάθε μία από τις οποίες επιλέχθηκαν δύο διαφορετικές ζεύγη τιμών για το κατώφλι θ (1V και 50V για τα 100ms και 50V και 100V για τα 200ms). Για αυτούς τους συνδυασμούς των θ και T χρησιμοποιήθηκαν τέσσερις διαφορετικές τιμές για τον πυκνωτή C οι οποίες απεικονίζονται στα υπομνήματα των εικόνων αυτών των τεσσάρων πειραμάτων. Στα επακόλουθα δύο πειράματα (Εικόνες 3.40 έως και 3.41) ακολουθήθηκε διαφορετική προσέγγιση, συγκεκριμένα για μία αυξανόμενη τιμή R επιλέχθηκε ένα μόνο παράθυρο T (200ms) για το οποίο επιλέχθηκαν δύο διαφορετικές τιμές C (20F και 50F). Για αυτούς τους συνδυασμούς των C και T χρησιμοποιήθηκαν τέσσερις διαφορετικές τιμές κατωφλίου θ οι οποίες απεικονίζονται στα υπομνήματα των εικόνων αυτών των πειραμάτων. Από αυτά τα έξι πειράματα επιβεβαιώνεται η ίδια συμπεριφορά του κβαντιστή με προηγούμενα πειράματα (τα πειράματα των Εικόνων 3.8 έως και 3.13). Αναλυτικότερα, αυξάνοντας είτε την τιμή του πυκνωτή C είτε την τιμή του κατωφλίου θ τόσο αυξάνεται και το MSE , ενώ αυξάνοντας την τιμή της αντίστασης R τόσο το MSE μειώνεται. Επιπλέον, αυξάνοντας το χρονικό παράθυρο T δεν παρατηρείται κάποια επιπλέον αλλαγή παρά μόνο ότι το MSE μειώνεται λιγότερο απότομα όσο το R αυξάνεται για ένα σταθερό C . Επομένως απορρέει το συμπέρασμα πως η χρήση του DWT μετασχηματισμού δεν άλλαξε τη συμπεριφορά του κβαντιστή, αλλά αντιθέτως την ωφέλησε (λιγότερο όμως από ό,τι ο DCT) καθώς τα επίπεδα MSE μειώθηκαν (π.χ. για $T = 200\text{ ms}$, $C = 50\text{ F}$, $\theta = 500\text{ V}$ και αυξανόμενη R ο κβαντιστής χωρίς DWT πετυχαίνει MSE που κυμαίνεται από 18000 έως και 6000 , ενώ με DWT πετυχαίνει MSE που κυμαίνεται από 18000 έως και 1900).

■ Μελέτη του χρονικού παραθύρου T



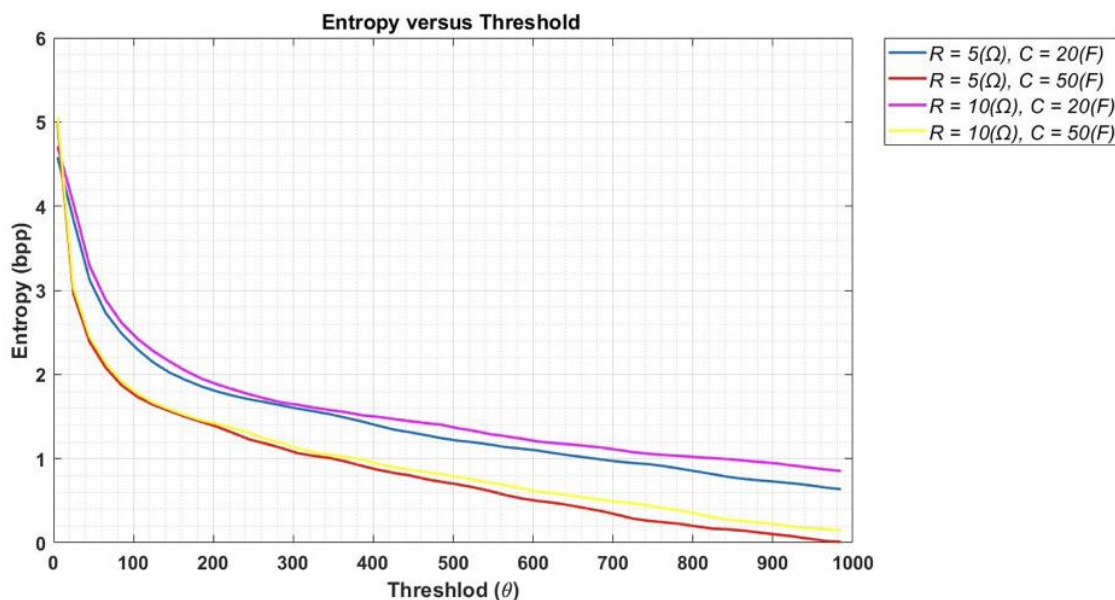
Εικόνα 3.42 : Παράμετρος T συναρτήσει του MSE για διάφορα R και $\theta = 50 \text{ V}$ & $C = 20 \text{ F}$ με χρήση DWT



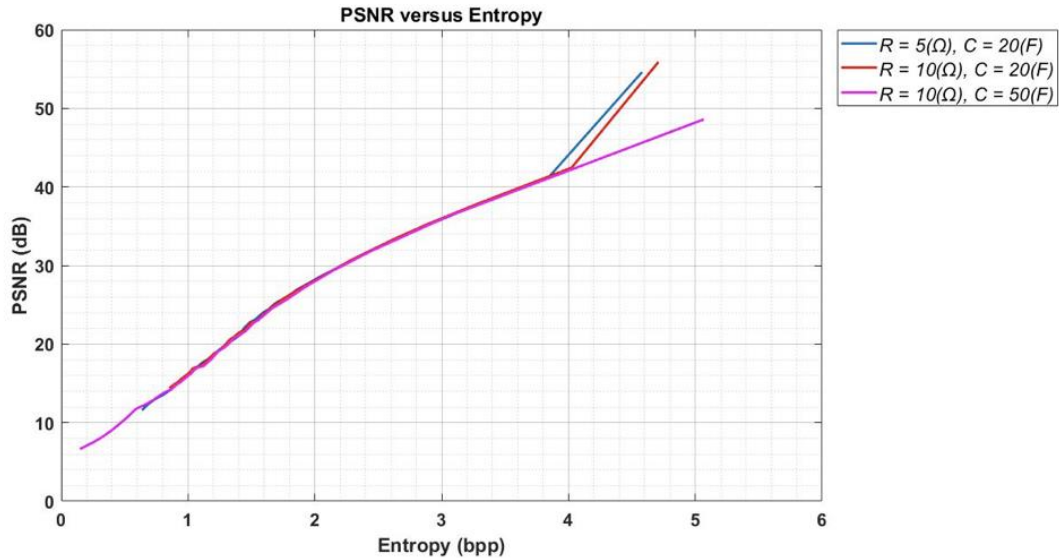
Εικόνα 3.43 : Παράμετρος T συναρτήσει του MSE για διάφορα R και $\theta = 50 \text{ V}$ & $C = 50 \text{ F}$ με χρήση DWT

Στα δύο παραπάνω πειράματα (Εικόνες 3.42 και 3.43) όλες οι παράμετροι του DUAL-SIM κβαντιστή κρατούνται σταθερές εκτός από το T , για το οποίο γίνεται η μελέτη, και το ζεύγος «πυκνωτής C με κατώφλι θ » βάσει του οποίου γίνονται συγκρίσεις μεταξύ των πειραμάτων. Συγκεκριμένα, για τον πυκνωτή C επιλέχθηκαν δύο τιμές ($20F$ και $50F$) για κάθε μία από τις οποίες επιλέχθηκαν μία τιμή κατωφλίου θ ($50V$). Για αυτούς τους συνδυασμούς των θ και C χρησιμοποιήθηκαν έξι διαφορετικές τιμές για την αντίσταση R οι οποίες απεικονίζονται στα υπομνήματα των εικόνων των παραπάνω δύο πειραμάτων. Από τα εν λόγω πειράματα επιβεβαιώνεται για άλλη μία φορά πως όσο αυξάνεται η τιμή του χρονικού παραθύρου T , ανεξαρτήτως των άλλων μεταβλητών, το MSE που δίνει ο κβαντιστής μειώνεται. Επίσης, επιβεβαιώνεται πως ο μετασχηματισμός DWT συνέβαλε στην απόδοση του κβαντιστή καθώς τα επίπεδα του MSE μειώθηκαν (π.χ. για $R = 0.3 \Omega$, $C = 50 F$, $\theta = 50 V$ και αυξανόμενο T ο κβαντιστής χωρίς DWT πετυχαίνει MSE που κυμαίνεται από 18000 έως και 11000, ενώ με DWT πετυχαίνει MSE που κυμαίνεται από 17900 έως και 2700).

Τέλος, μελετήθηκε η εντροπία που μπορεί να πετύχει ο κβαντιστής συναρτήσει του κατωφλίου θ για συγκεκριμένα R και C έχοντας το T στα $150ms$ (Εικόνα 3.44) καθώς και το PSNR που αποδίδει για αυτά τα R , C και T (Εικόνα 3.45) κάνοντας χρήση του μετασχηματισμού DWT. Συγκεκριμένα :



Εικόνα 3.44 : Εντροπία κβαντιστή με χρήση DWT μετασχηματισμού συναρτήσει του θ για $T = 150ms$ και διαφόρους συνδυασμούς R & C

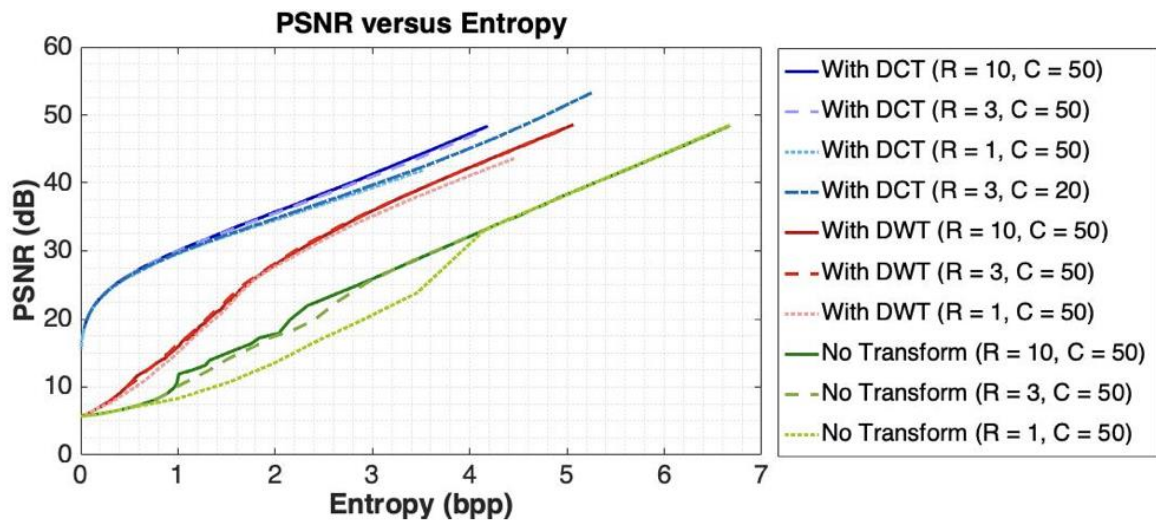


Εικόνα 3.45 : PSNR κβαντιστή με χρήση DWT μετασχηματισμού συναρτήσει της Εντροπίας για $T = 150ms$ και διαφόρους συνδυασμούς R & C

Από τα δύο παραπάνω πειράματα προκύπτει πως ο κβαντιστής DUAL-SIM αποδίδει καλύτερα για μεγάλο R σε συνδυασμό με χαμηλό C και ότι για χαμηλό κατώφλι θ υπάρχει καλύτερη ανάκτηση εικόνας, δηλαδή υψηλό PSNR διατηρώντας ταυτόχρονα χαμηλά την εντροπία. Σημειώνεται επίσης, πως η χρήση του DWT μετασχηματισμού συμβάλλει αρκετά στην απόδοση του κβαντιστή καθώς για μικρή εντροπία ($\sim 4.3 \text{ bpp}$) επιτυγχάνεται υψηλό PSNR ($\sim 50 \text{ dB}$) σε αντίθεση με τη μη χρήση του (όπου για $\sim 4.3 \text{ bpp}$ επιτυγχάνεται PSNR γύρω στα 35 dB). (Οι αριθμοί αυτοί αφορούν τον καλύτερο συνδυασμό R και C , εν προκειμένω $R = 5 \Omega$ και $C = 20 F$)

3.4 Συμπεράσματα-Συγκρίσεις & Οπτικά Αποτελέσματα

Συγκρίνοντας τα πειραματικά αποτελέσματα του DUAL-SIM κβαντιστή που παρουσιάστηκαν και αναλύθηκαν στην προηγούμενη υποενότητα 3.3 μπορούμε να καταλήξουμε σε δύο βασικά συμπεράσματα. Πρώτον, ότι το μοντέλο με ή χωρίς μετασχηματισμό αλλά κι ανεξαρτήτως του τύπου του μετασχηματισμού όταν αυτός χρησιμοποιείται λειτουργεί καλύτερα (απόδοση υψηλότερου PSNR) όταν γίνεται χρήση υψηλής τιμής στην αντίσταση R και χαμηλής τιμής στον πυκνωτή C και δεύτερον και πιο σημαντικό ότι το μοντέλο επιφέρει πιο επιθυμητά αποτελέσματα (υψηλό PSNR για χαμηλή εντροπία) όταν γίνεται η χρήση κάποιου μετασχηματισμού. Αυτό γίνεται πιο εμφανές συγκεντρώνοντας και συγκρίνοντας κάποια από τα αποτελέσματα σε ένα μόνο διάγραμμα όπως ακριβώς αυτό παρατίθεται στην παρακάτω εικόνα (Εικόνα 3.46).



Εικόνα 3.46 : Ο αντίκτυπος των μεταβλητών R και C καθώς και των μετασχηματισμών DCT και DWT στην απόδοση του $DUAL-SIM$ κβαντιστή

Παρατηρείται πως ο $DUAL-SIM$ κβαντιστής κερδίζει από τη χρήση DCT ή DWT μετασχηματισμού και πιο ειδικά η συμβολή του μετασχηματισμού συνημιτόνου είναι πολύ πιο αποτελεσματική από ό,τι αυτή του μετασχηματισμού κυματιδίων. Πιο συγκεκριμένα ο νευρο-εμπνευσμένος κβαντιστής επωφελείται αρκετά από τη «μετάβαση» της εικόνας στον χώρο των συχνοτήτων, δηλαδή από τη χρήση μετασχηματισμού, καθώς η χρήση του DCT ή DWT οργανώνει την πληροφορία της εικόνας σε ποσότητες οι οποίες μπορούν να χειριστούν και να υποστούν συμπίεση με πιο αποτελεσματικό τρόπο. Αναλυτικότερα, όπως ακριβώς αναλύθηκε και στις υποενότητες 1.3.1.3 και 1.3.2.2, η διαδικασία του μετασχηματισμού προσδιορίζει τα τμήματα πληροφοριών στο σήμα της εικόνας που μπορούν αποτελεσματικά να «πεταχτούν» κατά την κβάντιση χωρίς να διακυβεύεται σοβαρά η ποιότητα της εικόνας. Όσον αφορά την ποιότητα της εικόνας και την επίπτωση που έχει σε αυτή ο $DUAL-SIM$ κβαντιστής σε συνδυασμό με ή και χωρίς κάποιον μετασχηματισμό, αυτό γίνεται πιο ευδιάκριτο και κατανοητό από τις παρακάτω εικόνες (Εικόνες 3.47 έως και 3.49)

Χωρίς Μετασχηματισμό



$H = 1.009$ bpp
 $PSNR = 11.91$ dB



$H = 4.123$ bpp
 $PSNR = 32.91$ dB

Εικόνα 3.47 : Οπτικά αποτελέσματα του DUAL-SIM κβαντιστή χωρίς τη χρήση κάποιου μετασχηματισμού

Με DWT



$H = 1.016$ bpp
 $PSNR = 16.09$ dB



$H = 4.272$ bpp
 $PSNR = 43.92$ dB

Εικόνα 3.48 : Οπτικά αποτελέσματα του DUAL-SIM κβαντιστή με τη χρήση του DWT μετασχηματισμού

Με DCT



H = 1.056 bpp
PSNR = 30.38 dB



H = 4.187 bpp
PSNR = 48.39 dB

***Εικόνα 3.49 : Οπτικά αποτελέσματα του DUAL-SIM κβαντιστή
με τη χρήση του DCT μετασχηματισμού***

3.5 Μελλοντική Δουλειά

Στην παρούσα διπλωματική εργασία παρουσιάστηκε, αναλύθηκε και μελετήθηκε ένας νέος κβαντιστής ονόματι DUAL-SIM εμπνευσμένος από τον κλάδο της νευροεπιστήμης. Τα αποτελέσματα της εργασίας δείχνουν πως ο DUAL-SIM κβαντίζει αποτελεσματικά το σήμα εισόδου του, δηλαδή μια εικόνα, μιμούμενος τον τρόπο με τον οποίο το ανθρώπινο μάτι μαζί με τη βοήθεια του νευρωνικού συστήματος λαμβάνουν, κωδικοποιούν και μεταφέρουν μια εικόνα από το γύρο περιβάλλον στον εγκέφαλο. Παρουσιάστηκε επίσης, πως ο κβαντιστής DUAL-SIM μπορεί κάλλιστα να εισαχθεί στη συμβατική αρχιτεκτονική συμπίεσης των εικόνων, να συνδυαστεί και να αξιοποιήσει δηλαδή διάφορους μετασχηματισμούς, προσφέροντας αρκετά αξιόλογα αποτελέσματα.

Όσον αφορά περαιτέρω μελέτες πάνω στην απόδοση και λειτουργία του DUAL-SIM κβαντιστή αλλά και γενικά σε πιθανή μελλοντική εξέλιξή του, τρία είναι τα βασικά σημεία που πρέπει να διερευνηθούν. Πρώτον, είναι η σύγκρισή του με τα σύγχρονα μοντέλα συμπίεσης. Όντας πλέον εισαγμένος σε συμβατική αρχιτεκτονική συμπίεσης εικόνων ο κβαντιστής DUAL-SIM μπορεί να συγκριθεί με τα JPEG μοντέλα συμπίεσης, δηλαδή με τα

φίλτρα και τους κβαντιστές που διέπουν τις αντίστοιχες αρχιτεκτονικές αυτών των μοντέλων. Δεύτερον, είναι η διαφορετική αξιολόγηση των αποτελεσμάτων (*ανάκτηση εικόνας κατά την αποσυμπίεση*) που αυτός επιτυγχάνει. Στην παρούσα μελέτη, η απόδοση του κβαντιστή έγινε κάνοντας χρήση των ευρέως διαδεδομένων μετρικών, δηλαδή το MSE και το PSNR. Στόχος είναι αυτή η αξιολόγηση να γίνεται από νευρωνικά μοντέλα (*όπως το νευρωνικό δίκτυο*) που έχουν τη δυνατότητα να αναγνωρίζουν τα φυσικά χαρακτηριστικά των εικόνων όπως το ανθρώπινο οπτικό σύστημα. Και τρίτον και τελευταίο, είναι η προέκτασή του από τον δισδιάστατο χώρο στον τρισδιάστατο. Πιο συγκεκριμένα, είναι η χρήση του DUAL-SIM κβαντιστή για τη συμπίεση βίντεο, σε έναν χώρο δηλαδή που η σημασία της συμπίεσης είναι ακόμα πιο σημαντική λόγω των πληθώρα εικόνων από τις οποίες αποτελείται ένα βίντεο (*τα λεγόμενα καρέ*).

Βιβλιογραφία

- [1] Σ. Ν. ΔΗΜΗΤΡΙΑΔΗΣ , Ε. Γ. ΤΡΙΑΝΤΑΦΥΛΛΟΥ και Α. ΠΟΜΠΟΡΤΣΗΣ, ΤΕΧΝΟΛΟΓΙΑ ΠΟΛΥΜΕΣΩΝ ΘΕΩΡΙΑ ΚΑΙ ΠΡΑΞΗ, ΤΖΙΟΛΑΣ, 2004.
- [2] C. E. Shannon, , «A Mathematical Theory of Communication,» *Bell System Technical Journal*, τόμ. 27, αρ. 3, p. 379–423, 1948.
- [3] A. Bovic, Image & Video Processing, Academic Press, 2000.
- [4] R. C. Gonzalez και R. E. Woods, Ψηφιακή Επεξεργασία Εικόνας, Τζιόλα, 2020.
- [5] P. Havaladar, Multimedia Systems: Algorithms, Standards, and Industry Practices, Cengage Learning, 2009.
- [6] «T.81 – DIGITAL COMPRESSION AND CODING OF CONTINUOUS-TONE STILL IMAGES – REQUIREMENTS AND GUIDELINES,» *CCITT*, 1992, 2019.
- [7] N. Ahmed, T. Natarajan και K. R. Rao, «Discrete Cosine Transfom,» *IEEE Transactions on Computers*, τόμ. 23, αρ. 1, p. 90–93, 1974.
- [8] A. Gersho, "Quantization," *IEEE Communications Society Magazine*, vol. 15, no. 5, pp. 16-28, 1977.
- [9] J. Max, «Quantizing for minimum distortion,» *IEEE - IRE Transactions on Information Theory*, τόμ. 6, αρ. 1, pp. 7 - 12, 1960.
- [10] S. Lloyd, «Least squares quantization in PCM,» *IEEE Transactions on Information Theory*, τόμ. 28, αρ. 2, pp. 129 - 137, 1982.
- [11] D. Taubman και M. Marcellin, JPEG2000 Image Compression Fundamentals, Standards and Practice: Image Compression Fundamentals, Standards and Practice, Springer Science + Business Media, LLC, 2002.
- [12] D. H. Hubel, Eye, Brian, and Vision, Phoenix, USA: W H Freeman & Co., 1988.
- [13] A. R. Martin, P. A. Fuchs, D. A. Brown, M. E. Diamond και D. A. Weisblat, From Neuron to Brain, Massachusetts, USA: Sinauer Associates, Inc, 1984.
- [14] . R. F. Thompson , The Brain: A Neuroscience Primer, New York, USA: W H Freeman & Co, 1993.
- [15] W. Gerstner and W. M. Kistler, Spiking Neuron Models Single Neurons, Populations, Plasticity, Cambridge University Press, 2002.
- [16] S. Thorpe, D. Fize και C. Marlot , «Speed of processing in the human visual system,» *Nature*, 1996.
- [17] C. Koch , . O. Bernander και R. J. Douglas, «Do neurons have a voltage or a current threshold for action potential initiation?,» *J Comput Neurosci* ., τόμ. 2, αρ. 1, pp. 63-82, 1995.
- [18] S. J. Thorpe, «Spike arrival times: A highly efficient coding scheme for neural networks,» *North- Holland Elsevier*, 1990.
- [19] F. Theunissen και J. P. Miller, «Temporal encoding in nervous systems: a rigorous definition,» *J Comput Neurosci* ., τόμ. 2, αρ. 2, pp. 149-62, 1995.
- [20] W. Gerstner, «A framework for spiking neuron models: The spike response model,» *North-Holland*, τόμ. 4, 2001.
- [21] W. M. G. W. a. v. H. J. L. Kistler, «Reduction of the Hodgkin-Huxley Equations to a Single-VariableThreshold Model,» *Neural Computation*, τόμ. 9, p. 1015–1045, 1997.
- [22] C. F. Stevens και A. Zador, «Novel Integrate-and-fire-like Model of Repetitive Firing in Cortical Neurons,» 1998. [Ηλεκτρονικό]. Available:

https://www.researchgate.net/publication/2722512_Novel_Integrate-and-fire-like_Model_of_Repetitive_Firing_in_Cortical_Neurons.

- [23] G. Gestri, "Dynamics of a model for the variability of the interspike intervals in a retinal neuron," *Biol Cybern*, vol. 31, no. 2, pp. 97-8, 24 Nov 1978.
- [24] E. Doutsis, L. Fillatre, M. Antonini και P. Tsakalides, «Dynamic Image Quantization using Leaky Integrate-and-Fire Neurons,» 6 Nov 2019. [Ηλεκτρονικό]. Available: <https://hal.archives-ouvertes.fr/hal-02351180>.
- [25] E. Doutsis, G. Tzagkarakis και P. Tsakalides, «Neuro-Inspired Compression of RGB Images,» σε *27th European Signal Processing Conference (EUSIPCO)*, A Coruna, Spain, 2019.
- [26] I. E. Richardson, *The H.264 Advanced Video Compression Standard (Second Edition)*, A John Wiley and Sons, Ltd., Publication, 2010.