Ανάπτυξη Μεθόδων Εκτίμησης της Αντιληπτής Ποιότητας Εικόνων με Στατιστική Περιγραφή στο Πεδίο του Μετασχηματισμού Κυματιδίων

Δήμητρα Τσάμπρα

Μεταπτυχιακή Εργασία Εξειδίκευσης

Ιωάννινα, Απρίλιος 2022



Τμημα Μηχανικών Η/Υ & Πληροφορικής Πανεπιστημίο Ιωαννίνων

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING UNIVERSITY OF IOANNINA Ανάπτυξη Μεθόδων Εκτίμησης της Αντιληπτής Ποιότητας Εικόνων με Στατιστική στο Πεδίο του Μετασχηματισμού Κυματιδίων

Η Μεταπτυχιακή Διπλωματική Εργασία

υποβάλλεται στην ορισθείσα από τη Συνέλευση του Τμήματος Μηχανικών Η/Υ και Πληροφορικής Εξεταστική Επιτροπή

από την

Δήμητρα Τσάμπρα

ως μέρος των υποχρεώσεων για την απόκτηση του

ΔΙΠΛΩΜΑΤΟΣ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ ΣΤΗ ΜΗΧΑΝΙΚΗ ΔΕΔΟΜΕΝΩΝ ΚΑΙ ΥΠΟΛΟΓΙΣΤΙΚΩΝ ΣΥΣΤΗΜΑΤΩΝ

ΜΕ ΕΙΔΙΚΕΥΣΗ ΣΤΗΝ ΕΠΙΣΤΗΜΗ ΚΑΙ ΜΗΧΑΝΙΚΗ ΔΕΔΟΜΕΝΩΝ

Πανεπιστήμιο Ιωαννίνων Πολυτεχνική Σχολή Ιωάννινα 2022 Εξεταστική επιτροπή:

- Κόντης Λυσίμαχος-Παύλος, Καθηγητής, Τμήμα Μηχανικών Η/Υ και Πληροφορικής, Πανεπιστήμιο Ιωαννίνων (Επιβλέπων)
- Νίκου Χριστόφορος, Καθηγητής, Τμήμα Μηχανικών Η/Υ και Πληροφορικής, Πανεπιστήμιο Ιωαννίνων
- Παρσόπουλος Κωνσταντίνος, Καθηγητής, Τμήμα Μηχανικών Η/Υ και Πληροφορικής, Πανεπιστήμιο Ιωαννίνων

Εγχαριστιές

Η παρούσα διπλωματική εργασία εκπονήθηκε στο πλαίσιο της παρακολούθησης του μεταπτυχιακού προγράμματος σπουδών με τίτλο «Μηχανική Δεδομένων και Υπολογιστικών Συστημάτων» του Τμήματος Μηχανικών Ηλεκτρονικών Υπολογιστών και Πληροφορικής, της Πολυτεχνικής Σχολής του Πανεπιστημίου Ιωαννίνων.

Ολοκληρώνοντας και αυτόν τον κύκλο σπουδών μου, αισθάνομαι την υποχρέωση να ευχαριστήσω όσους συνέβαλαν στην ολοκλήρωση αυτής της εργασίας.

Κατ' αρχάς, θέλω να εκφράσω τις ευχαριστίες μου στον επιβλέποντα καθηγητή της διπλωματικής εργασίας, κ. Λυσίμαχο Κόντη, Καθηγητή του Τμήματος Μηχανικών Ηλεκτρονικών Υπολογιστών και Πληροφορικής, για την ανάθεση αυτής, καθώς και για τις συμβουλές του και την καθοδήγησή του.

Ακολούθως, θέλω να ευχαριστήσω τον κ. Αντώνιο Μαϊργιώτη, Διδάκτορα Πληροφορικής, για την συνεργασία του και την πολύτιμη βοήθειά του κατά την εκπόνηση της διπλωματικής εργασίας.

Τέλος, ευχαριστώ την οικογένειά μου, για την ηθική και οικονομική στήριξή της κατά την διάρκεια των σπουδών μου.

$\Pi \text{EPIEXOMENA}$

Κατάλ	λογος σχημάτων	iv
Κατάλ	λογος πινάκων	х
Περιληψη		
Exten	ded Abstract	xviii
ΚΕΦΑ	ΑΛΑΙΟ 1 Εισαγωγή	1
1.1	Αντικείμενο της Διπλωματικής	1
1.2	Οργάνωση του Τόμου	
1.3	Υποχειμενική Αξιολόγηση Ποιότητας της Εικόνας	
1.4	Αντικειμενική Εκτίμηση Ποιότητας της Εικόνας	5
1.5	Κατηγορίες Αλγορίθμων για Αξιολόγηση Ποιότητας Εικόνας	6
1.6	Natural Scene Statistics (NSS) Model	7
1.7	Σχέση Οπτικής Ποσοχής (Visual Attention) και Ποιότητας Εικό	νας8
1.8	Βάσεις Δεδομένων	9
ΚΕΦΑ	ΑΛΑΙΟ 2 Μοντέλα IQA βασισμένα σε Φυσικά Χρωματικ	ά
Στατι	ιστικά και Χαρακτηριστικά του Μετασχηματισμού	
Κυμα	κτιδίων	12
2.1	Χρωματικά Μοντέλα	15
2.	1.1 RGB μοντέλο	15
2.	1.2 YIQ μοντέλο	16
2.	1.3 HSV μοντέλο	
2.	1.4 YCbCr μοντέλο	
2.	1.5 Lab μοντέλο	
2.2	Wavelet Normalized Coefficients	
2.3	Κατανομή των Συντελεστών Κυματιδίων	21

2.4	Η Οπτικά Εξέχουσα Περιοχή στο Πλαίσιο της Εκτίμησης της Γ	Ιοιότητας
της Ει	ικόνας	
2.5	Ψεύδο-gradient	27
2.6	Μετρικές	
2.6.	1 Peak Signal-to-noise ratio PSNR	
2.6.	2 Structural Similarity Index - SSIM	
2.6.	3 Visual Signal-to-Noise Ratio-VSNR	31
2.6.	4 Feature Similarity Index – FSIM	31
2.6.	5 Visual Information Fidelity – VIF	
2.6.	.6 VCGS	
КЕФА	ΛΑΙΟ 3 Προτεινόμενη Full-Reference Μεθοδολογία	
Βασισμ	ιένη στην Οπτικά Εξέχουσα Περιοχή, με Χρήση Στατισ	τιχών
Χαρακ	τηριστικών Μετασχηματισμού ΚυματιδίωΝ	36
3.1	Πρώτη Ομάδα Πειραμάτων	45
3.2	Δεύτερη Ομάδα Πειραμάτων	48
3.3	Τρίτη Ομάδα Πειραμάτων	
3.4	Τέταρτη Ομάδα Πειραμάτων	
3.5	Πέμπτη Ομάδα Πειραμάτων	60
3.6	Έκτη Ομάδα Πειραμάτων	63
3.7	Έβδομη Ομάδα Πειραμάτων	
3.8	Όγδοη Ομάδα Πειραμάτων	
3.9	Συγκεντρωτικά Αποτελέσματα	73
ΚΕΦΑ	ΛΑΙΟ 4 Προτεινόμενες No-Reference Μεθοδολογίες	75
4.1	DATASET IQA	
4.1.	1 Student's t κατανομή με ΕΜ εκτίμηση παραμέτρων	
4.1.	2 Student's t με 2 παραμέτρους	
4.1.	3 Student's t με 4 παραμέτρους	
4.1.	4 Cauchy	
4.2	DATASET C	
4.2.	1 Student's t κατανομή με EM εκτίμηση παραμέτρων	108

4.2.2	Student's t με 2 παραμέτρους	110
4.2.3	Student's t με 4 παραμέτρους	111
4.2.4	Cauchy	113
4.3 Σ	ύγκριση αποτελεσμάτων με μοντέλα βασισμένα στους MCSN	
συντελεσ	στές	114
4.4 Συ	ρμπληρωματικά πειράματα	117
4.4.1	Student's t	118
4.4.2	Cauchy	121
ΚΕΦΑΛΑ	ΔΙΟ 5 Συμπεράσματα	125
Βιβλιογρα	φία	132

ΚΑΤΑΛΟΓΟΣ ΣΧΗΜΑΤΩΝ

Σχήμα 1.1 Εικόνες του Dataset IQA 10
Σχήμα 1.2 Εικόνες του Dataset C11
Σχήμα 2.1 SSIM Scores και Τοπογραφικοί Χάρτες των Τιμών του Δείκτη SSIM για Δύο Εικόνες με σχεδόν Ιδιο Επίπεδο Υποκειμενικών Αξιολογήσεων, οι οποίες Επιδεικνύουν Χρωματικές και Μη-Χρωματικές Παραμορφώσεις (Υ1: η Αρχική Εικόνα αφού έχει Μειωμένο Κορεσμό (Saturation) και Υ2: με Συμπίεση JPEG2000)
Σχήμα 2.2 Το RGB Χρωματικό Μοντέλο15
Σχήμα 2.3 Μετατροπή RGB σε YIQ16
Σχήμα 2.4 Το HSV Χρωματικό Μοντέλο17
Σχήμα 2.5 Το YCbCr Χρωματικό Μοντέλο17
Σχήμα 2.6 Το Lab Χρωματικό Μοντέλο18
Σχήμα 2.7 Παράδειγμα Εικόνας από το Dataset LIVE IQA και τα Ιστογράμματα των Οριζόντιων Συντελεστών που Προκύπτουν από το Μετασχηματισμό Wavelet της Εικόνας σε Διαφορετικούς Χρωματικούς Χώρους. (α)Μια Εικόνα της Βάσης. (β)-(στ) τα Ιστογράμματα των Wavelets Συντελεστών της Εικόνας στα LAB, YCbCr, RGB, HSV, YIQ Μοντέλα αντίστοιχα
Σχήμα 2.8 Ιστογράμματα των Συντελεστών προσέγγισης, των Οριζόντιων, Κάθετων και Διαγώνιων Συντελεστών που Προκύπτουν από το Μετασχηματισμό Κυματιδίων DWT
Σχήμα 2.9 Σχηματική Αναπαράσταση της Λειτουργίας του Αλγορίθμου SSIM 30
Σχήμα 2.10 Αναπαράσταση Λειτουργίας Αλγορίθμου VIF

Σχήμα 3.1 Εικόνες RGB που Λαμβάνει ως Είσοδο η Προτεινόμενη Μέθόδος 38

Σχήμα 3.7 Στην Πρώτη, Δεύτερη και Τρίτη Γραμμή Παρουσίαζονται το VSSimMatrix, το Weight και το SimMatrix του Πειράματος Final1, Final2 και Final5 αντίστοιχα, όπου: VSSImMatrix: Χάρτης Ομοιότητας μεταξύ των Δύο Εικόνων

Βασισμένος στο Visual Saliency. Χρησιμοποιείται για τον Υπολογισμό του Τελικού Δείκτη. Weight: Πίνακας Στάθμισης Βασισμένος στην Σημείο-προς-Σημείο Μέγιστη Τιμή μεταξύ των Χαρτών Visual Saliency για τις Εικόνες που Συγκρίνονται. SimMatrixC: Τελικός Πίνακας Ομοιότητας πριν την Εφαρμοζόμενη Στάθμιση..... 48

Σχήμα 3.17 Οι Χάρτες που Προχύπτουν για το Πείραμα final i. όπου VSSImMatrix: Χάρτης Ομοιότητας μεταξύ των Δύο Ειχόνων Βασισμένος στο Visual Saliency. Χρησιμοποιείται για τον Υπολογισμό του Τελιχού Δείχτη. Weight: Πίναχας Στάθμισης Βασισμένος στην Σημείο-προς-Σημείο Μέγιστη Τιμή μεταξύ των

Σχήμα 4.1 Ειχόνα από το Dataset C......77

ΚΑΤΑΛΟΓΟΣ ΠΙΝΑΚΩΝ

Πίναχας 3.1 Συντελεστής Pearson και Spearman για την μέθοδο VCGS 45 Πίναχας 3.2 Συντελεστής Pearson για τα Πειράματα της Πρώτης Ομάδας 46 Πίναχας 3.3 Συντελεστής Spearman για τα Πειράματα της Πρώτης Ομάδας47 Πίναχας 3.4 Συντελεστής Pearson για τα Πειράματα της Δεύτερης Ομάδας...... 49 Πίναχας 3.5 Συντελεστής Spearman για τα Πειράματα της Δεύτερης Ομάδας.... 50 Πίναχας 3.7 Συντελεστής Spearman για τα Πειράματα της Τρίτης Ομάδας 54 Πίναχας 3.8 Συντελεστής Pearson για τα Πειράματα της Τέταρτης Ομάδας...... 57 Πίνακας 3.9 Συντελεστής Spearman για τα Πειράματα της τέταρτης ομάδας..... 58 Πίναχας 3.10 Συντελεστής Pearson για το Πείραμα FinalA......60

Πίνακας 3.26 SROCC και PLCC για 6 Full Eeference Μεθόδους για το Dasatet IQA

Πίνακας 4.1 Συντελεστής Pearson Μοντέλων που Έχουν σαν Βάση Χαρακτηριστικά που Εξάγονται μέσω της Student-t Κατανομής με ΕΜ Εκτίμηση Παραμέτρων... 87

Πίνακας 4.3 Συντελεστής Pearson Μοντέλων που Έχουν σαν Βάση Χαρακτηριστικά που Εξάγονται μέσω της Student-t Κατανομής με Εκτίμηση Δύο Παραμέτρων...93

Πίνακας 4.7 Συντελεστής Pearson Μοντέλων που Έχουν σαν Βάση Χαρακτηριστικά που Εξάγονται μέσω της Cauchy Κατανομής......104

Πίνακας 4.17 Συντελεστής Spearman Μοντέλων που Έχουν σαν Βάση 2 Χαρακτηριστικά που Εξάγονται από τους Συντελεστές DCT της Εικόνας Προσέγγισης μέσω της Student-t Κατανομής με ΕΜ Εκτίμηση Παραμέτρων......118

Πίναχας 4.19 Συντελεστής Pearson Μοντέλων που Έχουν σαν Βάση 2 Χαραχτηριστιχά που Εξάγονται από τους Συντελεστές DCT της Ειχόνας Προσέγγισης μέσω της Student-t Κατανομής με ΕΜ Εχτίμηση Παραμέτρων......119

Πίνακας 4.20 Συντελεστής Pearson Μοντέλων που Έχουν σαν Βάση 8 Χαρακτηριστικά που Εξάγονται από τους Συντελεστές DCT της Εικόνας

xiii

Πίνακας 4.21 Συντελεστές Pearson και Spearman Μοντέλων που Έχουν σαν Βάση 2 Χαρακτηριστικά που Εξάγονται από τους Συντελεστές DCT της Εικόνας Προσέγγισης μέσω της Student-t Κατανομής με ΕΜ Εκτίμηση Παραμέτρων..... 120

ΠΕΡΙΛΗΨΗ

Δήμητρα Τσάμπρα, Δ.Μ.Σ. στη Μηχανική Δεδομένων και Υπολογιστικών Συστημάτων, Τμήμα Μηχανικών Η/Υ και Πληροφορικής, Πολυτεχνική Σχολή, Πανεπιστήμιο Ιωαννίνων, Απρίλιος 2022

Ανάπτυξη μεθόδων εκτίμησης της αντιληπτής ποιότητας εικόνων με στατιστική στο πεδίο του μετασχηματισμού κυματιδίων.

Επιβλέπων: Κόντης Λυσίμαχος-Παύλος, Καθηγητής

Στόχος της παρούσας μεταπτυχιαχής εργασίας είναι η ανάπτυξη μεθόδων αξιολόγησης της αντιληπτής ποιότητας εικόνας. Πρόκειται για ένα αρκετά δύσκολο πρόβλημα που προσελκύει το ενδιαφέρον πολλών ερευνητών, καθώς η ψηφιακή εικόνα κατακλύζει την καθημερινότητα μας ολοένα και περισσότερο. Γι' αυτό το λόγο καθίσταται επιτακτική η ανάγκη ύπαρξης μιας μεθόδου, που θα είναι ικανή να εκτιμήσει την ποιότητα της εικόνας με γρήγορο και αξιόπιστο τρόπο, ώστε να αντικαταστήσει την εξίσου αξιόπιστη αλλά αργή μέθοδο αξιολόγησης από τον ίδιο τον άνθρωπο. Η δυσκολία αυτού του προβλήματος έγκειται στην εύρεση εκείνων των χαρακτηριστικών, που είναι ικανά να αντιληφθούν τις αλλαγές που υφίσταται μια ειχόνα λόγω χάποιου είδους ψηφιαχής επεξεργασίας της. Η διαφοροποίηση των μεθόδων, που αναπτύσσουμε σε σχέση με υπάρχουσες μεθόδους, εντοπίζεται κατανομή χαι εξαγόμενα χαρακτηριστικά στην στατιστική στα που χρησιμοποιούνται για την κατασκευή τους.

Τα δεδομένα που επιλέγουμε σ' αυτή την εργασία για την εξαγωγή των χαραχτηριστιχών είναι οι συντελεστές προσέγγισης και οι συντελεστές οριζόντιων, κάθετων και διαγώνιων λεπτομερειών, που προχύπτουν από την εφαρμογή του μετασχηματισμού wavelet της ειχόνας. Λόγω της μορφής τους, επιλέξαμε για την προσαρμογή τους την Student's t κατανομή σε διάφορες μορφές της, καθώς και μια πιο ειδιχή περίπτωσή της, την Cauchy κατανομή. Διερευνάται κατά πόσο αποδοτιχές είναι αυτές οι κατανομές στη δημιουργία μεθόδων εχτίμησης της ποιότητας της ειχόνας, συγχρίνοντας τα αποτελέσματα των μεθόδων μας με άλλες ευρέως γνωστές μεθόδους.

Οι μέθοδοι που δημιουργούμε, κατηγοριοποιούνται σε δύο γενικές κατηγορίες: Full-Reference και No-reference αλγόριθμοι. Οι Full-reference αλγόριθμοι θα

xvi

εξεταστούν σε ένα σύνολο δεδομένων. Η κατασκευή των No-reference αλγορίθμων θα γίνει με χρήση των μηχανών διανυσμάτων υποστήριξης (SVMs) και θα εξεταστούν σε δύο βάσεις δεδομένων. Εκτός από τα χαρακτηριστικά από τον μετασχηματισμό κυματιδίων για τους No-Reference αλγορίθμους χρησιμοποιούμε και χαρακτηριστικά από το μετασχηματισμό Συνημιτόνου (DCT). Τέλος, η προτεινόμενη μέθοδος συγκρίνεται με ένα χωρικό μοντέλο.

EXTENDED ABSTRACT

Dimitra Tsampra, M.Sc. in Data and Computer Systems Engineering, Department of Computer Science and Engineering, School of Engineering, University of Ioannina, Greece, April 2022

Development of methods for estimating perceived image quality with statistics in the field of wavelet transformation.

Advisor: Kondi Lisimachos P., Professor

The aim of this postgraduate thesis is the development of methods for the evaluation of perceived image quality. This is quite a difficult problem that attracts the interest of many researchers, as the digital image overwhelms our everyday life more and more. For this reason, it becomes imperative to have a method that will be able to estimate the quality of the image, in a fast and reliable way, in order to replace the equally reliable but slow method of evaluation by a human itself. The difficulty of this problem lies in finding those features that are capable of perceiving the changes that an image undergoes, due to some kind of digital processing of it. The difference between the methods we develop and the existing methods is found at the statistical distribution and the exported features used for their construction.

The data we select in this task for the extraction of features are the approximation coefficients and the coefficients of horizontal, vertical and diagonal details that arising from the application of the wavelet transformation of the image. Because of their form, we chose for their fitting the Student's t distribution in its various forms, as well as a more specific case of it, the Cauchy distribution. We investigate whether these distributions are effective in creating methods for assessing image quality by comparing the results of our methods with other widely known methods.

The methods we create are categorized into two general categories: Full-Reference and No-reference algorithms. Full-reference algorithms will be examined in a dataset. Our No-reference algorithms will be constructed by using support vector machines (SVMs) and will be examined in two databases. Besides the features from the wavelet transformation for No-Reference algorithms, we also use features from cosine transformation (DCT). Finally, the proposed method is compared with a spatial model.

ΚΕΦΑΛΑΙΟ 1

ΕιΣαγωγή

- 1.1 Αντικείμενο της Διπλωματικής εργασίας
- 1.2 Οργάνωση του Τόμου
- 1.3 Υποχειμενική Αξιολόγηση Ποιότητας της Εικόνας
- 1.4 Αντικειμενική Εκτίμηση Ποιότητας της Εικόνας
- 1.5 Κατηγορίες Αλγορίθμων για Αξιολόγηση Ποιότητας Εικόνας
- 1.6 Natural Scene Statistics (NSS) Model
- 1.7 Σχέση Οπτικής Προσοχής (Visual Attention) και Ποιότητας Εικόνας
- 1.8 Βάσεις Δεδομένων

1.1 Αντικείμενο της Διπλωματικής

Τα τελευταία χρόνια, η ψηφιαχή εικόνα υπάρχει σε κάθε πτυχή της ζωής μας, καθιστώντας τον κόσμο που ζούμε όλο και πιο ψηφιαχό οπτικά. Μέρα με την μέρα, έρχονται στο προσχήνιο καινοτόμες εφαρμογές ενημέρωσης, ψυχαγωγίας και επικοινωνίας, εύχολα προσβάσιμες στον άνθρωπο, δημιουργώντας του νέες δυνατότητες αλλά και απαιτήσεις ως προς την απόδοσή τους. Τέτοιες εφαρμογές αποτελούν τα μέσα κοινωνικής δικτύωσης, όπως το Facebook, το Twitter, το Instagram κλπ., οι οποίες επιτρέπουν στον χρήστη να κάνει λήψη, συμπίεση, μετάδοση και αποθήκευση της εικόνας μέσω διαφορετικών συσκευών, όπως smartphones, laptops, tablets, τηλεοράσεις υψηλής ευκρίνειας κτλ. Επίσης, αναπτύσσονται καθημερινά νέες πλατφόρμες streaming, όπως το Netflix, AppleTv, Disney+. Το όριο του εύρους ζώνης που είναι διαθέσιμο, οι φυσικές ιδιότητες συσκευών χαμηλού επιπέδου (π.χ. υψηλό ποσοστό συμπίεσης) αλλά και οι συνθήκες περιβάλλοντος κατά τις οποίες γίνεται η λήψη οπτικών σημάτων, που ενδέχεται να μην είναι ιδανικές (π.χ. ακτινοβολία ήλιου), εισάγουν στην εικόνα μια ποικιλία διαφορετικών παραμορφώσεων. Σε συνδυασμό με το γεγονός ότι ο άνθρωπος είναι ον πολύ οπτικό, αυτές οι παραμορφώσεις καθιστούν την αρχική εικόνα ποιοτικά κακή, σε βαθμό που γίνεται αντιληπτό από τον χρήστη, προκαλώντας του δυσκολία και ενόχληση στην εμπειρία θέασης.

Γι' αυτό και είναι επιτακτική η ανάγκη ενός ποσοτικού δείκτη της αντιληπτής ποιότητας της εικόνας, που θα χρησιμεύσει στην βελτίωση της εμπειρίας που προσφέρουν οι διάφορες εφαρμογές. Ο προφανής τρόπος μέτρησης της ποιότητας της εικόνας είναι η αξιολόγηση της από ανθρώπους-παρατηρητές, αφού αποτελούν και τους τελικούς καταναλωτές και κριτές του αποτελέσματος. Ωστόσο, η διαδικασία συλλογής τέτοιων υποκειμενικών αξιολογήσεων είναι χρονοβόρα και δαπανηρή. Έτσι, δημιουργήθηκε ένα νέο πεδίο έρευνας, στο οποίο οι ερευνητές σχεδιάζουν αλγορίθμους με στόχο την αντικειμενική εκτίμηση της ποιότητας της εικόνας, με τέτοιο τρόπο ώστε να συνάδει με την υποκειμενική ανθρώπινη αξιολόγηση. Αυτοί οι αλγόριθμοι είναι ανεκτίμητοι, καθώς μπορούν να ενσωματωθούν σε αυτόματα συστήματα, σε κάμερες και κινητά τηλέφωνα, ώστε να προσαρμόζονται σε πραγματικό χρόνο με βάση κάποιο feedback της ποιότητας εξόδου και να γίνουν αρκετά «έξυπνα» για αυτόματη λήψη, επεξεργασία και βελτίωση της εικόνας.

Τέλος, η ανάπτυξη αυτού του τομέα οδηγεί και σε ανάπτυξη άλλων κλάδων της επιστήμης, όπως αυτόν της Ιατρικής. Η βιοϊατρική τεχνολογία εξελίσσεται όλο και περισσότερο και βασίζεται σε μεγάλο ποσοστό στην εικόνα για διάγνωση της ανθρώπινης ασθένειας. Πρόκειται για μια τεχνολογία, με την οποία επεξεργαζόμαστε και δημιουργούμε ψηφιακές αναπαραστάσεις του εσωτερικού του σώματος, των λειτουργιών οργάνων και των ιστών, για κλινική ανάλυση και ιατρική παρέμβαση. Οι αλγόριθμοι εκτίμησης της ποιότητας της εικόνας (Image Quality Assessment-IQA) μπορούν να ενσωματωθούν σε μηχανήματα εξέτασης, όπως ο ακτινογράφος, ο τομογράφος, ο υπέρηχος, ο μαστογράφος κ.ά., ώστε να παράγουν εικόνες καλής ανάλυσης, λεπτομερείς, με στόχο την πιο ακριβή διάγνωση και, κατ' επέκταση, την αποτελεσματικότερη θεραπεία.

1.2 Οργάνωση του Τόμου

Η παρούσα μεταπτυχιακή εργασία, αποτελείται από πέντε κεφάλαια, καθ' ένα από τα οποία αναλύεται στις παρακάτω παραγράφους.

Στο πρώτο κεφάλαιο, διακρίνουμε τα είδη αξιολόγησης της εικόνας σε υποκειμενική και αντικειμενική αξιολόγηση παραθέτοντας τα πλεονεκτήματα και τα μειονεκτήματα της κάθε μιας. Επίσης, αναφέρονται οι αλγόριθμοι Full reference, No-Reference και Reduced-Reference σαν ορισμοί και πώς διαφοροποιούνται μεταξύ τους. Γίνεται αναφορά σε σημαντικά στατιστικά χαρακτηριστικά των εικόνων που διαταράσσονται, όταν η εικόνα υποστεί κάποια στρέβλωση, και ορίζουμε δύο σημαντικές έννοιες, της οπτικής προσοχής και της οπτικής εξέχουσας περιοχής. Τέλος, αναφέρουμε ποια σύνολα δεδομένων χρησιμοποιήσαμε για την συγκεκριμένη εργασία και το λόγο για τον οποίο επιλέχθηκαν.

Στο δεύτερο χεφάλαιο, αναφερόμαστε στα διάφορα χρωματιχά μοντέλα που υπάρχουν. Ορίζουμε τους συντελεστές χυματιδίων που θα αποτελέσουν την βάση των μεθόδων μας χαι περιγράφουμε την χατανομή Student's t, με την οποία θα γίνει η προσαρμογή αυτών των συντελεστών. Αναφέρουμε την χρήση της οπτιχής προσοχής στο πρόβλημα της εχτίμησης της ειχόνας χαι γιατί είναι απαραίτητοι τέτοιου είδους αλγόριθμοι. Επιπλέον, δίνεται μια εναλλαχτιχή μορφή της gradient, που θα αποτελέσει χαραχτηριστιχό για την χατασχευή του μοντέλου στο επόμενο χεφάλαιο. Τέλος, περιγράφουμε γνωστές μεθόδους αξιολόγησης της ποιότητας της ειχόνας. Η απόδοση της προτεινόμενης μεθόδους.

Στο τρίτο κεφάλαιο, γίνεται η ανάπτυξη των προτεινόμενων μεθόδων που ανήκουν στην κατηγορία Full Reference αλγορίθμων, καταλήγουμε στην καλύτερη από αυτές και την συγκρίνουμε με άλλες γνωστές αποδοτικές μεθόδους.

Στο τέταρτο κεφάλαιο, παρατίθενται και αναλύονται τα αποτελέσματα των No Reference αλγορίθμων που υλοποιούμε. Γίνεται χρήση περισσότερων κατανομών για την προσαρμογή των wavelet συντελεστών και για την κατασκευή μοντέλων με χρήση SVMs.

Στο πέμπτο κεφάλαιο, αναφέρονται παρατηρήσεις σχετικά με τα αποτελέσματα των πειραμάτων του τρίτου και τέταρτου κεφαλαίου. Τέλος,

παρατίθενται προτάσεις για μελλοντικές εργασίες, ως επέκταση της παρούσης εργασίας.

1.3 Υποκειμενική Αξιολόγηση Ποιότητας της Εικόνας

Όπως αναφέρθηκε και παραπάνω, το πιο σωστό κριτήριο για την αξιολόγηση της ποιότητας της εικόνας είναι η ανθρώπινη υποκειμενική κρίση, καθώς ο ίδιος ο άνθρωπος είναι ο αποκλειστικός αποδέκτης της εικόνας. Με αυτόν τον γνώμονα πραγματοποιούνται διάφορες δοκιμές, στις οποίες ζητείται από ένα αρκετά μεγάλο πλήθος ανθρώπων να αντιστοιχίσουν κάθε εικόνα, ενός συνόλου εικόνων που έχουν υποστεί διάφορες παραμορφώσεις, σε έναν βαθμό που να αντικατοπτρίζει την αξιολόγηση όσον αφορά την ποιότητα της εν λόγω εικόνας. Η βαθμολόγηση αυτή γίνεται με βάση κάποια κλίμακα. Έπειτα γίνεται συλλογή αυτών των βαθμολογιών και υπολογίζεται ένας μέσος όρος αυτών [1]. Ο μέσος όρος των βαθμολογιών παράγει το Mean Opinion Score (MOS) και το Difference Mean Opinion (DMOS), που αποτελούν την αντιληπτή ποιότητα της εικόνας. Μεγαλύτερες τιμές MOS υποδηλώνουν εικόνες με χαμηλότερη αντιληπτή ποιότητα. Εν κατακλείδι, μια υψηλότερη τιμή MOS αντιστοιχεί σε χαμηλότερη τιμή DMOS και υποδεικνύει καλύτερη ποιότητα [2].

Η παραπάνω διαδικασία είναι, κατά κύριο λόγο, αυτή που ακολουθείται για την εξαγωγή της υποκειμενικής αξιολόγησης ποιότητας της εικόνας και αποτελεί μεθοδολογία της κατηγορίας βαθμολόγησης μοναδικού ερεθίσματος (single stimulus rating) [3]. Η συνηθέστερη κλίμακα αυτής της κατηγορίας είναι η κλίμακα Absolute Category Ranking(ACR) [4]. Οι εικόνες εμφανίζονται μία-μία στην οθόνη του υπολογιστή για συγκεκριμένο χρονικό διάστημα και μετά εξαφανίζονται. Ο παρατηρητής πρέπει να αξιολογήσει κάθε μία ποιοτικά αντιστοιχίζοντάς την σε μια από τις δοθείσες κατηγορίες, οι οποίες είναι 'Κακή', 'Ανεπαρκής', 'Ανεκτή', 'Καλή' και 'Άριστη'. Έπειτα, αυτές οι ετικέτες μετατρέπονται στις τιμές 1, 2, 3, 4 και 5, αντίστοιχα, ώστε να υπολογιστεί ο μέσος όρος και να προκύψει το MOS ή DMOS. Αν ο παρατηρητής έχει στην διάθεσή του και την εικόνα αναφοράς, τότε πρόκειται για μεθοδολογία που ανήκει στην κατηγορία double stimulus rating [3]. Όπως και στις single stimulus rating μεθοδολογίες, καλείται να βαθμολογήσει την παραμορφωμένη εικόνα απλά έχοντας στην διάθεσή του και την εικόνα αναφοράς.

Υπάρχει μία ακόμα κατηγορία μεθοδολογιών, που ονομάζεται choice pairwise comparison [3], στην οποία εμφανίζονται δυο εικόνες που αναπαριστούν ακριβώς το ίδιο σκηνικό και οι παρατηρητές πρέπει να αποφασίσουν ποια από τις δύο είναι καλύτερη ποιοτικά. Τέλος, υπάρχει και η κατηγορία pairwise similarity, που είναι παρόμοια με την choice pairwise comparison [3], αλλά ο παρατηρητής πρέπει να αναφέρει και σε τι βαθμό διαφέρουν ποιοτικά οι εικόνες.

1.4 Αντικειμενική Εκτίμηση Ποιότητας της Εικόνας

Παλαιότερα οι ερευνητές, που ασχολούνταν με το τομέα της αντιχειμενιχής εχτίμησης της ποιότητας της ειχόνας, μετρούσαν την πιστότητα του οπτιχού σήματος. Αυτή η μέτρηση γινόταν ως προς ένα σήμα αναφοράς (reference signal), το οποίο θεωρούσαν ότι είχε την «τέλεια» ποιότητα. Πλέον, στοχεύουν στον σχεδιασμό λεπτομερών μαθηματιχών μοντέλων που θα είναι ιχανά να προβλέψουν την ποιότητα της ειχόνας αυτόματα χαι με αχρίβεια. Η αχρίβεια χαθορίζεται από το πόσο το αποτέλεσμα του αλγορίθμου IQA προσεγγίζει τις υποχειμενιχές αξιολογήσεις. Όσο μεγαλύτερη συσχέτιση έχουν οι τιμές των προβλέψεων ενός αλγορίθμου IQA με τις τιμές των υποχειμενιχών εχτιμήσεων, τόσο πιο αχριβής είναι ο αλγόριθμος.

Η αυτοματοποίηση αυτής της διαδιχασίας βρίσχει αρχετές εφαρμογές, όπως στην βελτίωση εμπειρίας που προσφέρουν εφαρμογές επιχοινωνίας, ψυχαγωγίας χαι διασχέδασης, στην παραχολούθηση (monitor) της ποιότητας για συστήματα ελέγχου ποιότητας, αλλά χαι ως benchmarking σε αλγορίθμους χαι συστήματα επεξεργασίας ειχόνας. Τέτοιου είδους αλγόριθμοι μπορούν να ενσωματωθούν σε συστήματα επεξεργασίας ειχόνας για λόγους βελτιστοποίησης (optimization) των αλγορίθμων χαι την χατάλληλη ρύθμιση των παραμέτρων [5]. Τέλος, μπορούν να βοηθήσουν χαι σε άλλους χλάδους, όπως στην Ιατριχή με την ενσωμάτωση τους σε σύγχρονα μηχανήματα διάγνωσης (π.χ. τομογράφος) ή στην Εγχληματολογία για αντιστοίχηση βιομετριχών στοιχείων (π.χ. ίριδα ματιού).

1.5 Κατηγορίες Αλγορίθμων για Αξιολόγηση Ποιότητας Εικόνας

Οι αλγόριθμοι IQA, ανάλογα με την διαθέσιμη πληροφορία που τους δίνεται, κατηγοριοποιούνται στις εξής κατηγορίες:

Full-reference image quality assessment (FR-IQA): [6]

Σε αυτή την κατηγορία ανήκουν οι αλγόριθμοι που παίρνουν σαν είσοδο την παραμορφωμένη εικόνα και την εικόνα αναφοράς. Εικόνα αναφοράς ή αρχική εικόνα ονομάζουμε μια εικόνα που είναι 'καθαρή', δηλαδή δεν έχει υποστεί κάποια παραμόρφωση. Η εκτίμηση της ποιότητας προκύπτει έπειτα από σύγκριση της παραμορφωμένης εικόνας και της αρχικής της. Οι αλγόριθμοι αυτής της κατηγορίας είναι αρκετά ακριβείς και εύκολοι στην υλοποίησή τους. Η δυσκολία έγκειται στο γεγονός ότι η εικόνα αναφοράς δεν είναι πάντα διαθέσιμη, καθώς, μετά από αρκετές αναπαραγωγές σε διαφορετικά μέσα, προστίθενται παραμορφώσεις σε αυτή. Αυτό καθιστά την εύρεση της αρχικής εικόνας δύσκολη έως και αδύνατη.

Reduced-reference image quality assessment (RR-IQA): [6]

Πρόχειται για αλγορίθμους που λαμβάνουν σαν είσοδο την παραμορφωμένη ειχόνα και ένας μέρος των χαραχτηριστιχών της ειχόνας αναφοράς (π.χ. watermark). Δεν είναι δηλαδή απαραίτητη η ύπαρξη της αρχιχής ειχόνας. Στόχος τους είναι να υπολογίσουν μια υψηλή εχτίμηση της ποιότητας της ειχόνας, χρησιμοποιώντας μιχρότερη πληροφορία.

No-reference image quality assessment (NR-IQA): [6]

Οι αλγόριθμοι αυτής της κατηγορίας λαμβάνουν σαν είσοδο μόνο την παραμορφωμένη εικόνα. Δεν απαιτούν την ύπαρξη της αρχικής εικόνας, γεγονός που τους καθιστά πιο πρακτικούς και περιζήτητους για τις διάφορες εφαρμογές, καθώς, όπως αναφέραμε, η αρχική εικόνα δεν είναι πάντα διαθέσιμη. Όμως, εξαιτίας αυτού, αποτελούν και τις πιο δύσκολες μεθόδους, όσον αφορά τον σχεδιασμό και την υλοποίηση τους, καθώς η έλλειψη εικόνας αναφοράς οδηγεί στη ανάγκη υλοποίησης ενός μοντέλου το οποίο πρέπει να εκπαιδευτεί, ώστε να εξάγει μια εκτίμηση της αντιληπτής ποιότητας της εικόνας. Αυτό το είδος αλγορίθμων, επειδή δεν χρησιμοποιεί αρχική εικόνα, ονομάζεται και 'Τυφλή Αξιολόγηση Εικόνας'.

1.6 Natural Scene Statistics (NSS) Model

Οι NR-IQA αλγόριθμοι ποικίλλουν και μπορούν να διακριθούν σε δύο βασικές κατηγορίες, ανάλογα με την μεθοδολογία που ακολουθούν.

Στην πρώτη κατηγορία ανήκουν αλγόριθμοι με τους οποίους γίνεται εξαγωγή χαρακτηριστικών από εικόνες και έπειτα γίνεται εκπαίδευση ενός μοντέλου μάθησης, ώστε να δημιουργηθεί μια σχέση μεταξύ αυτών των χαρακτηριστικών και των βαθμολογιών των εικόνων, όπως αυτές προέκυψαν από τις υποκειμενικές αξιολογήσεις (βλ. 1.3). Αυτή η σχέση μπορεί να δημιουργηθεί είτε μέσω μηχανών διανυσμάτων στήριξης (Support Vector Machines-SVMs) είτε μέσω νευρωνικών δικτύων (neural networks). Επίσης, η βαθιά μάθηση (deep learning), η οποία έχει σημειώσει μεγάλη επιτυχία στην επίλυση της αναγνώρισης εικόνων και στην ανίχνευση αντικειμένων, ώθησε τους ερευνητές να την χρησιμοποιήσουν και για NR-IQA αλγορίθμους, στους οποίους σημείωσε εξίσου καλά αποτελέσματα, επιτυγχάνοντας υψηλή ακρίβεια πρόβλεψης. Γι΄ αυτό, χρησιμοποιείται αρκετά στον τομέα της εκτίμησης ποιότητας εικόνας, παρόλη την υψηλή της απαίτηση σε υπολογιστικούς πόρους. Όσο πιο αντιπροσωπευτικά είναι τα χαρακτηριστικά ως προς την ποιότητα της εικόνας, τόσο πιο αξιόπιστος είναι ο αλγόριθμος [7].

Στην δεύτερη κατηγορία ανήχουν αλγόριθμοι που βασίζονται στις στατιστικές των φυσικών εικόνων (natural scenes statistics- NSS). Ο όρος «φυσική» δεν χρησιμοποιείται για να χαρακτηρίσει το περιεχόμενο της εικόνας. Το περιεχόμενο μπορεί να είναι είτε από το φυσικό περιβάλλον (π.χ. δέντρα) είτε από οτιδήποτε άλλο (εσωτερικοί χώροι, αντικείμενα κτλ.). Χρησιμοποιείται για να υποδηλώσει, ότι οι εικόνες δεν έχουν παραμορφωθεί από την εφαρμογή κάποιου είδους τεχνητής επεξεργασίας, όπως γραφικά υπολογιστών, φίλτρα κτλ. Αυτή η μεθοδολογία έχει ως βασική υπόθεση ότι οι φυσικές εικόνες μοιράζονται ορισμένα στατιστικά και πως η εμφάνιση στρεβλώσεων ενδέχεται να τα επηρεάσει, κάνοντας την εικόνα να χάσει την «φυσικότητά» της. Τα NSS μοντέλα χρησιμοποιούν χαρακτηριστικά είτε από το χωρικό πεδίο (spatial domain) είτε από κάποιο πεδίο μετασχηματισμού (transform domain) και προσπαθούν να ποσοτικοποιήσουν την υποβάθμιση της ποιότητας της εικόνας, μετρώντας την έλλειψη «φυσικότητας». Η προσέγγιση της NSS είναι πολύ ελπιδοφόρα, αλλά βασίζεται σε μοντελοποίηση εκτεταμένης στατιστικής και αξιόπιστη γενίκευση των μοντέλων. Επιπλέον, όπως και στους

αλγορίθμους της πρώτης κατηγορίας, τα εξαγόμενα χαρακτηριστικά παίζουν καταλυτικό ρόλο στην απόδοση των αλγορίθμων. Μπορούν να έχουν καλή εφαρμογή μόνο σε ένα συγκεκριμένο είδος παραμόρφωσης, ανάλογα από ποιο πεδίο θα γίνει η εξαγωγή των χαρακτηριστικών [7], [8].

1.7 Σχέση Οπτικής Ποσοχής (Visual Attention) και Ποιότητας Εικόνας

Ένα σημαντικό θέμα, που αφορά τις εικόνες του πραγματικού κόσμου, σχετίζεται με την πολυπλοκότητα της προτεινόμενης μοντελοποίησης για το πρόβλημα της αξιολόγησης ποιότητας. Στο ίδιο πλαίσιο σημασία έχει και η αξιοπιστία μιας προτεινόμενης μεθοδολογίας. Η αξιοπιστία, όμως, σχετίζεται τόσο με την αντιληπτή ποιότητα όσο και με την κατανόηση του ανθρώπινου οπτικού συστήματος. Εν ολίγοις, μία συγκεκριμένη τάση της έρευνας στο πεδίο της αντιληπτής ποιότητας εικόνας είναι η διερεύνηση του πώς λειτουργεί η επίδραση της οπτικής προσοχής (visual attention) στην κρίση κάποιου για την ποιότητα μίας εικόνας. Η οπτική προσοχή αποτελεί ένα μηχανισμό του ανθρώπινου οπτικού συστήματος, που καθοδηγεί την επιλογή της οπτικής πληροφορίας σε μία οπτική σχηνή [9].

Μία γνωστή προσέγγιση του μηχανισμού οπτικής προσοχής είναι η από 'κάτω προς τα πάνω' (bottom-up) εκδοχή, βασισμένη σε κάποιο ερέθισμα. Συνήθως αναφέρεται ως οπτικά εξέχουσα (visual saliency) περιοχή και συναντάται στην Υπολογιστική Όραση (Computer Vision). Ένα υπολογιστικό μοντέλο οπτικά εξέχουσας περιοχής δημιουργεί ουσιαστικά έναν τοπογραφικό χάρτη που αναπαριστά τις ευδιάκριτες/ξεχωριστές περιοχές [10], [11], [12].

Στην ουσία, αποτελεί ένα βασικό χαρακτηριστικό του, που εξάγει σημαντική και συνοπτική πληροφορία από τις φυσικές σκηνές. Από τη στιγμή που ένα φυσικό περιβάλλον περιλαμβάνει έναν μεγάλο αριθμό πληροφοριών, ο μηχανισμός της οπτικά εξέχουσας περιοχής ενεργοποιεί τη μείωση της πλεονάζουσας πληροφορίας, που ωφελεί την αντίληψη κατά τη διαδικασία της επιλεκτικής προσοχής.

Ο βασικός λόγος για να εισάγουμε την οπτικά εξέχουσα πληροφορία σε μία μετρική αξιολόγησης ποιότητας, είναι ότι τα τεχνουργήματα ή οι οντότητες που εμφανίζονται στις οπτικά εξέχουσες περιοχές, θεωρείται ότι έχουν μεγαλύτερη επίδραση στην αντίληψη της οπτικής ποιότητας από τις μη εξέχουσες περιοχές. Ως

συνέπεια, η έρευνα στη βιβλιογραφία του συγκεκριμένου πεδίου εστιάζει κυρίως στην επέκταση ενός συγκεκριμένου μέτρου, με τη χρήση κάποιου μοντέλου εξέχουσας οπτικής περιοχής. [13]

Η πιο κοινή μεθοδολογία αφορά τη στάθμιση (weight) των τοπικών παραμορφώ-σεων με την τοπική εξέχουσα οπτική περιοχή (local saliency), παραδίδοντας ως αποτέλεσμα τις γνωστές ως saliency-weighted μετρικές (μετρικές σταθμισμένες με τις εξέχουσες οπτικά περιοχές) [14],[15].

Αυτού το είδους οι μεθοδολογίες, σύμφωνα με τα πειραματικά αποτελέσματα, μας δείχνουν ότι οι η πληροφορία εξέχουσας περιοχής οδηγεί σε βελτίωση της επίδοσης των διαφόρων μεθοδολογιών που δε βασίζονται σε αυτήν. Όμως, η σχέση μεταξύ οπτικά εξεχουσών περιοχών (visual saliency) και ποιότητας εικόνας είναι πολύπλοκη και αυτό που πρέπει κανείς να έχει κατά νου είναι, ότι οι οπτικά εξέχουσες περιοχές στην αξιολόγηση της ποιότητας εικόνας, εξαρτώνται ισχυρά από το αντίστοιχο μοντέλο (saliency model) αλλά και τα χαρακτηριστικά της υπό εξέταση εικόνας.

Ένα σημαντικό θέμα, που σχετίζεται με τη εισαγωγή αυτής της έννοιας στην αξιολόγηση της ποιότητας, είναι ότι η προτεινόμενη στάθμιση απαιτεί επιπλέον υπολογιστικό κόστος, τόσο για τη δημιουργία των χαρτών όσο και για την αναπαράσταση της σημασίας των τοπικών παραμορφώσεων.

1.8 Βάσεις Δεδομένων

Στην παρούσα διπλωματική εργασία χρησιμοποιείται το dataset IQA [16], [17], [18], ή αλλιώς LIVE Image Quality Assessment Database, το οποίο περιέχει συνολικά 982 εικόνες. Οι εικόνες αναφοράς είναι 29 και οι παραμορφωμένες 779, κάθε μια από τις οποίες έχει υποστεί μια από τις παρακάτω παραμορφώσεις: White Noise, Gaussian Blur, JP2K, JPEG και Fast Fading. Στο dataset περιλαμβάνονται και τα DMOS των παραμορφωμένων εικόνων, με τις τιμές τους να κυμαίνονται στο διάστημα [0,100], όπου χαμηλότερη τιμή υποδηλώνει καλύτερη οπτική ποιότητα εικόνας.



Σχήμα 1.1 Εικόνες του Dataset IQA

Το σύνολο δεδομένων dataset C [9],[19], ή αλλιώς LIVE In the Wild Image – Quality Challenge Database, περιέχει μια μεγάλη ποικιλία αυθεντικών παραμορφώσεων σε ένα πλήθος εικόνων, που έχουν ληφθεί με χρήση διάφορων σύγχρονων κινητών τηλεφώνων. Πρόκειται για παραμορφώσεις που γίνονται αυτόματα κατά την λήψη και αποθήκευση της εικόνας. Σε αυτές τις εικόνες δεν υπάρχει άλλο είδος τεχνητής παραμόρφωσης, όπως π.χ. συμπιέσεις, γραφικά και φίλτρα. Πρόκειται για μια βάση 1169 εικόνων, οι οποίες αξιολογήθηκαν από 81.000 παρατηρητές συγκεντρώνοντας πάνω από 350.000 βαθμολογίες. Από αυτές τις βαθμολογίες προέκυψε για κάθε εικόνα ο μέσος όρος MOS, ο οποίος περιλαμβάνεται στο dataset, με την τιμή του να κυμαίνεται στο διάστημα [0,100], όπου μεγαλύτερη τιμή υποδηλώνει καλύτερη οπτική ποιότητα εικόνας.



Σχήμα 1.2 Εικόνες του Dataset C

ΚΕΦΑΛΑΙΟ 2

Μοντελά IQA βασισμένα σε Φύσικα Χρωματικά Στατιστικά και Χαρακτηριστικά του

Μετασχηματισμού Κύματιδιών

- 2.1 Χρωματικά Μοντέλα
- 2.2 Wavelet Normalized Coefficients
- 2.3 Κατανομή των Συντελεστών Κυματιδίων
- 2.4 Η Οπτικά Εξέχουσα Περιοχή στο Πλαίσιο της Εκτίμησης της Ποιότητας της Εικόνας.
- 2.5 Ψεύδο-gradient
- 2.6 Μετρικές

Μία ψηφιαχή ειχόνα αποτελείται από ένα σύνολο pixels, που μπορούν με αποδοτιχό τρόπο να συλληφθούν από διάφορες ηλεχτρονιχές συσχευές, όπως οι ψηφιαχές χάμερες, οι σαρωτές χαι άλλες. Κάθε pixel μίας ειχόνας συνήθως συσχετίζεται με μία χωριχή συντεταγμένη σε χάποια δισδιάστατη περιοχή, η οποία έχει τιμή που αποτελείται από ένα ή περισσότερα δείγματα (samples). Επιπλέον, η ευχολότερη συνδεσιμότητα λόγω των διχτύων χαι των επιχοινωνιών χαι ο αυξανόμενος αριθμός των συσχευών σύλληψης ειχόνων ή βίντεο, έχει χαταστήσει την πρόσβαση χαι το διαμοιρασμό των ειχόνων ιδιαίτερα εφιχτά χαι βολιχά ως διαδιχασίες. Στο πλαίσιο αυτό, διάφορες μετριχές αξιολόγησης της ποιότητας των εικόνων έχουν προταθεί στο πέρασμα των χρόνων, όπως είδαμε και στο εισαγωγικό κεφάλαιο της εργασίας. Όμως, οι πιο κοινά χρησιμοποιούμενες μετρικές βασίζονται σε χαρακτηριστικά που βρίσκονται στο πεδίο των επιπέδων του γκρι. Αυτό σημαίνει ότι δεν δίνεται ιδιαίτερη σημασία στη πληροφορία χρώματος, είτε για λόγους μείωσης της σημασίας του είτε λόγω υπολογιστικής αποδοτικότητας – στις έγχρωμες εικόνες είναι λογικό να αυξάνεται το υπολογιστικό κόστος. Ως εκ τούτου, σε διάφορες εφαρμογές ή και στα προβλήματα αξιολόγησης της ποιότητας των εικόνων, οι προαναφερόμενες μετρικές είτε διάφορες άλλες μεθοδολογίες παραγνωρίζουν τη σημασία του χρώματος ή δε δίνουν σημασία στις χρωματικές αποκλίσεις. Γενικά, η βασική προσέγγιση στην αξιολόγηση ποιότητας των έγχρωμων εικόνων μπορεί να εκτιμηθεί με την εφαρμογή μίας μετρικής για εικόνες επιπέδου του γκρι στα ξεχωριστά RGB κανάλια, με τον συνεπακόλουθο συνδυασμό των σχορ σε κάθε κανάλι να ακολουθεί.

Ας δούμε ένα παράδειγμα. Στο Σχήμα 2.1 επιδεικνύεται η πρόβλεψη της ποιότητας μίας εικόνας, βασιζόμενη αποκλειστικά στη πληροφορία των επιπέδων του γκρι. Έτσι, χρησιμοποιώντας το γνωστό δείκτη SSIM [20], θα μπορούσε να οδηγήσει σε μια υπερβολικά αισιόδοξη βαθμολογία, καθώς υποτιμά την οπτική υποβάθμιση που προκαλείται από τη μείωση του κορεσμού (saturation) στις έγχρωμες εικόνες



Y1: X with reduced saturation



X : Original Image



Y2: X with JPEG2000 encoding





Σχήμα 2.1 SSIM Scores και Τοπογραφικοί Χάρτες των Τιμών του Δείκτη SSIM για Δύο Εικόνες με σχεδόν Ιδιο Επίπεδο Υποκειμενικών Αξιολογήσεων, οι οποίες Επιδεικνύουν Χρωματικές και Μη-Χρωματικές Παραμορφώσεις (Υ1: η Αρχική Εικόνα αφού έχει Μειωμένο Κορεσμό (Saturation) και Υ2: με Συμπίεση JPEG2000)

Ποια, όμως, είναι η προσέγγιση που γενικά ακολουθούμε; Τα θεμελιώδη εργαλεία εκτίμησης των χρωματικών παραμορφώσεων είναι ο υπολογισμός της Ευχλείδειας απόστασης σε ομοιόμορφα χρωματιχά συστήματα (uniform color systems), που έχουν προταθεί από την International Commission on Illumination (CIE): CIELAB και CIELUV [21]. Παραδοσιακά, η λαμβανόμενη παραμόρφωση μεταξύ έγχρωμων εικόνων προσδιορίζεται από τον υπολογισμό των CIE εξισώσεων σε επίπεδο pixel, οπότε κάποιος στη συνέχεια εξετάζει μέσω περιγραφικής στατιστικής τα δεδομένα, δηλαδή βρίσκει π.χ. το μέσο ή τη διάμεσο τιμή. Όμως, παρότι κάτι τέτοιο είναι χρήσιμο για να συγκριθούν ομοιόμορφα χρωματισμένα τμήματα εικόνας (color patches), δεν υπάρχει εγγύηση ότι θα δουλέψουν καλά με πολύπλοχα δεδομένα ειχόνων του πραγματιχού χόσμου. Λόγω της συνοπτιχής χαι αχριβούς μορφοποίησης, ο SSIM δείχτης μπορεί να επεχταθεί χαι να συμπεριλάβει πληροφορία χρώματος. Για παράδειγμα, η μετρική UCIF (Universal Color Image Fidelity [22] εκτιμά τους SSIM δείκτες για κάθε χρωματικό κανάλι του Lab χρωματικού χώρου. Το τελικό σκορ ποιότητας προκύπτει ως μία σταθμισμένη εκδοχή των δεικτών σε κάθε κανάλι. Αντίστοιχα, η ίδια μετρική έχει εφαρμοστεί στη περίπτωση του βίντεο, όπως μπορεί χανείς να δει στο Video SSIM (VSSIM) [23]. Σε κάθε κανάλι του YCbCr μοντέλου ακολουθείται η λογική της προαναφερόμενης μετρικής.

Βέβαια, υπάρχουν γενικά πιο διαισθητικές χρωματικές μετρικές κάνοντας χρήση των κυλινδρικών χρωματικών χώρων και βασίζονται στα χαρακτηριστικά των χρωμάτων που συσχετίζονται καλά με την ανθρώπινη αντίληψη π.χ. hue και chroma.

Από την άλλη, πολλές φορές, για να προσεγγιστεί η αντίληψη χρώματος πιο καλά, οι αρχικές RGB εικόνες μετασχηματίζονται σε έναν χώρο χρώματος, που είναι πιο συμβατός με την ανθρώπινη αίσθηση. Στη περίπτωση της παρούσας εργασίας γίνεται χρήση του CIELAB μοντέλου [24], [25], [26].

Στα επόμενα θα δοθεί μία συνοπτική περιγραφή στα χρωματικά μοντέλα που συνήθως χρησιμοποιούνται.
2.1 Χρωματικά Μοντέλα

Τα χρωματικά μοντέλα αναπτύχθηκαν με σκοπό να καταστεί δυνατή η περιγραφή των χρωμάτων με μαθηματική μορφή και, κατ' επέκταση, να μπορούν να αναπαραχθούν και να επεξεργαστούν από τα διάφορα ψηφιακά μέσα. Ανάλογα με την χρήση κάθε εικόνας, υιοθετείται και το ανάλογο χρωματικό μοντέλο.

2.1.1 RGB μοντέλο

Το πιο διαδεδομένο χρωματικό μοντέλο, που χρησιμοποιείται στα γραφικά των ηλεκτρονικών υπολογιστών, είναι το RGB μοντέλο. Τα βασικά χρώματα σε αυτό το μοντέλο είναι το κόκκινο (Red), το πράσινο (Green) και το μπλε (Blue). Το μοντέλο βασίζεται στο γεγονός ότι, όταν μία οθόνη δεν εκπέμπει φως, εμφανίζεται μαύρη. Από τους συνδυασμούς των τριών βασικών χρωμάτων παράγονται και άλλα χρώματα, ενώ με ταυτόχρονη προβολή και των τριών δημιουργείται το λευκό χρώμα (Σχήμα 2.2). Κάθε χρώμα αντιπροσωπεύεται από τρεις ακέραιους αριθμούς, ο καθένας από τους οποίους μπορεί να έχει τιμή από 0 έως 255. Όμως, στο RGB μοντέλο δεν διακρίνονται η φωτεινότητα, το χρώμα και ο κορεσμός και γι' αυτό δεν μπορούμε να μετρήσουμε αποτελεσματικά την πληροφορία φωτεινότητας και χρώματος. Επιπλέον, δεν είναι δυνατή η παραγωγή όλων των χρωμάτων από τον συνδυασμό των τριών βασικών χρωμάτων [27]. Γι' αυτό και δεν θεωρείται κατάλληλο χρωματικό μοντέλο για τους αλγορίθμους IQA.



Σχήμα 2.2 Το RGB Χρωματικό Μοντέλο

2.1.2 YIQ μοντέλο

Το YIQ μοντέλο [28] είναι αυτό που χρησιμοποιείται κυρίως σε συστήματα εκπομπής τηλεόρασης. Τα τρία βασικά χαρακτηριστικά του είναι το Y, που αντιπροσωπεύει την πληροφορία φωτεινότητας, ενώ το I και Q την πληροφορία χρωματικότητας. Μια έγχρωμη τηλεόραση θα χρησιμοποιήσει αυτά τα τρία κανάλια (Y, I, Q) και θα αντιστοιχίσει την πληροφορία σε R, G, B επίπεδα για να την αναπαραγωγή της εικόνας στην οθόνη (Σχήμα 2.3).

[Y]	[O	.299	0.587	7	0.114][R]
I	≈ 0.	5959	-0.274	46 -	-0.3213	G
Q	0.	2115	-0.522	27	0.3112	B
Fro	m Yl	Q to	RGE	1		
Fro	m YI	Q to	RGE	5		
From	m YI	Q to	RGE	.619	$\left[Y\right]$	
	$=\begin{bmatrix}1\\1\end{bmatrix}$	Q to 0.90 -0.1	RGE 56 0 272 –0	.619 0.647	$\begin{bmatrix} Y\\ I \end{bmatrix}$	

Σχήμα 2.3 Μετατροπή RGB σε YIQ

2.1.3 HSV μοντέλο

Στο HSV μοντέλο, η εικόνα περιγράφεται από το την απόχρωση (Hue), τον κορεσμό (Saturation) και την αξία (Value). Το μοντέλο αναπαρίσταται με την βοήθεια ενός χρωματικού κώνου (Σχήμα 2.4). Η απόχρωση είναι συστατικό χρώματος, δίνεται από την γωνία/θέση γύρω από τον κάθετο άξονα και εκφράζει διαφορετικά χρώματα σε διαφορετικές γωνίες. Οι τιμές της είναι από 0 έως 360. Ο κορεσμός αφορά το βαθμό καθαρότητας του χρώματος, δηλαδή κατά πόσο το χρώμα έχει αναμειχθεί με το λευκό. Η τιμή του κορεσμού κυμαίνεται από το 0 έως 1, όπου 0 η κλίμακα του γκρι, και 1 το βασικό χρώμα. Η αξία αναφέρεται στο βαθμό μείξης ενός καθαρού χρώματος με το μαύρο και η τιμή της κυμαίνεται από 0 έως το 1. Το 0 εκφράζει το μαύρο χρώμα και το 1 την πιο φωτεινή απόχρωση που αποκαλύπτει το χρώμα [28].



Σχήμα 2.4 Το HSV Χρωματικό Μοντέλο

2.1.4 YCbCr μοντέλο

Ο χρωματικός χώρος YCbCr (Σχήμα 2.5) δημιουργήθηκε για να καλύψει τις ανάγκες της ψηφιακής επεξεργασίας video. Το Y εκφράζει την φωτεινότητα, το Cb την χρωματική διαφορά ως προς το μπλε, ενώ το Cr την χρωματική διαφορά ως προς το κόκκινο. Πρόκειται για το μοντέλο που χρησιμοποιείται στην συμπίεση [29].



Σχήμα 2.5 Το YCbCr Χρωματικό Μοντέλο

2.1.5 Lab μοντέλο

Το μοντέλο Lab χρησιμοποιείται ευρέως στην υπολογιστική όραση. Αναλύει το χρώμα του κάθε pixel σε τρεις παραμέτρους L,a,b. Το L εκφράζει την ένταση φωτεινότητας και a, b οι δύο συνιστώσες που καθορίζουν την ανάμειξη κόκκινουπράσινου και κίτρινου-μπλε, αντίστοιχα (Σχήμα 2.6). Έχει αρκετά πλεονεκτήματα, όπως ότι μας προφέρει την αντίληψη μιας μικρής χρωματικής αλλαγής. Είναι ανεξάρτητο από τη συσκευή, επειδή είναι απλώς μια λειτουργία έγχρωμης οθόνης. Μπορεί να αναπαραστήσει χρώμα του RGB ή CMY χώρου, αλλά και χρώματα που δεν ανήκουν σε άλλο χρωματικό χώρο. Τέλος, το Lab μοντέλο έχει μια γραμμική σχέση με τον Ανθρώπινο Οπτικό Σύστημα (HVS), που συσχετίζεται με την ανθρώπινη κρίση [27]. Αυτό είναι και το μοντέλο που θα χρησιμοποιήσουμε σε αυτή την εργασία, καθώς πρόκειται για ομοιόμορφο χώρο και πιο συμβατό στην ανθρώπινη αντίληψη. Σ' αυτό τον χώρο εισάγονται και δύο νέες μεταβλητές όσον αφορά το χρώμα, η ζωντάνια (vividness), V_{ab}^* και το βάθος (depth), D_{ab}^* . Η ζωντάνια είναι ένα χαρακτηριστικό του χρώματος, που χρησιμοποιείται για να υποδείξει το βαθμό απόκλισης του από το ουδέτερο μαύρο χρώμα, ενώ το βάθος χρησιμοποιείται για να υποδείξει το βαθμό απόκλισης ενός χρώματος από το ουδέτερο λευκό χρώμα. Πρόκειται, ουσιαστικά, για δύο ευκλείδειες αποστάσεις.

$$V_{ab}^{*} = \sqrt{\left(L^{*}\right)^{2} + \left(a^{*}\right)^{2} + \left(b^{*}\right)^{2}} = \sqrt{\left(L^{*}\right)^{2} + \left(C_{ab}^{*}\right)^{2}}$$
(2.1)

$$D_{ab}^{*} = \sqrt{\left(100 - L^{*}\right)^{2} + \left(a^{*}\right)^{2} + \left(b^{*}\right)^{2}} = \sqrt{\left(100 - L^{*}\right)^{2} + \left(C_{ab}^{*}\right)^{2}}$$
(2.2)



Σχήμα 2.6 Το Lab Χρωματικό Μοντέλο.

2.2 Wavelet Normalized Coefficients

Το πρώτο βήμα για την ανίχνευση οπτικά εξέχουσας περιοχής (visual saliency) είναι η μετατροπή της RGB εικόνας σε CIE Lab χρωματικό χώρο, καθώς πρόκειται για ομοιόμορφο χώρο και συμβατό στην ανθρώπινη αντίληψη. Οι υπο-ζώνες μιας εικόνας θα διαμορφωθούν από τον μετασχηματισμό κυματιδίων-WT για ένα πλήθος επιπέδων, όπως παρουσιάζεται στον τύπο (2.3). Επιλέγονται τα κυματίδια Daubechies (Daub.4) (αν και στα πειράματά μας έχουν δοκιμαστεί και άλλες οικογένειες), καθώς το μέγεθος του φίλτρου τους είναι κατάλληλο για την γειτονιά pixels, για τον χρόνο υπολογισμού, αλλά και για το συνολικό αποτέλεσμα. Η εφαρμογή του μετασχηματισμού wavelet είναι της μορφής:

$$\left[\mathbf{A}_{N}^{c},\mathbf{H}_{s}^{c},\mathbf{V}_{s}^{c},\mathbf{D}_{s}^{c}\right] = WT_{N}\left(\mathbf{g}^{c}\right)$$
(2.3)

όπου N το μέγιστο πλήθος κλιμάκωσης της διαδικασίας αποσύνθεσης του WT, δηλαδή ο δείκτης ανάλυσης $s \in \{1, ..., N\}$ και το N-ιοστό επίπεδο αντιστοιχούν στην πιο χονδροειδή ανάλυση. Πρέπει να σημειωθεί ότι, στην εργασία μας, η αύξηση του βάθος του μετασχηματισμού δεν έδινε καλύτερα αποτελέσματα, τουλάχιστον στην περίπτωση της μετρικής που βασίζεται στη πλήρη γνώση της πληροφορίας.

Στον προαναφερόμενο τύπο, ως c δηλώνονται τα χρησιμοποιούμενα κανάλια του g'^c , καθώς $c \in \{L, a, b\}$. Επιπλέον, ως A_N^c δηλώνεται η αναπαράσταση των συντελεστών προσέγγισης, δηλαδή στην πιο χονδροειδή ανάλυση για κάθε κανάλι. Οι υπόλοιπες εικόνες H_N^c , V_N^c , D_N^c είναι οι συντελεστές κυματιδίων των οριζοντίων, κάθετων και διαγώνιων λεπτομερειών για τα δοθέντα c και s, αντίστοιχα. Οι συγκεκριμένες εικόνες (υπο-μπάντες) αποτελούν ιδιαίτερα χρήσιμο εργαλείο στη περίπτωση του επόμενου κεφαλαίου, όπου προτείνεται μία νέα μετρική.

Οι συντελεστές κυματιδίων, που αναπαριστούν τις λεπτομέρειες της εικόνας σε διάφορες κλίμακες, χρησιμοποιούνται για την δημιουργία χαρτών χαρακτηριστικών (feature map) με αυξανόμενα εύρη ζώνης συχνοτήτων. Οι χάρτες χαρακτηριστικών (features maps) μπορούν, εν τέλει, να υπολογιστούν με αντίστροφο μετασχηματισμό κυματιδίων (Inverse WT). Μπορούμε, δηλαδή, να δημιουργήσουμε χάρτες χαρακτηριστικών από λεπτομέρειες του μετασχηματισμού κυματιδίων αγνοώντας τα δεδομένα προσέγγισης κατά τον αντίστροφο μετασχηματισμό, ενώ μπορούμε, επίσης, να μεταβάλουμε την τιμή των wavelet συντελεστών με διάφορους τρόπους.

Η υπόθεσή μας είναι ότι οι wavelet συντελεστές έχουν στατιστικές ιδιότητες, οι οποίες αλλάζουν ανάλογα με τον τύπο και τον βαθμό παραμόρφωσης της εικόνας, και ότι η ποσοτικοποίηση αυτών των αλλαγών θα καταστήσει δυνατή την πρόβλεψη του τύπου στρέβλωσης που επιδρά στην εικόνα αλλά και της αντίληψης της ποιότητας.



Σχήμα 2.7 Παράδειγμα Εικόνας από το Dataset LIVE IQA και τα Ιστογράμματα των Οριζόντιων Συντελεστών που Προκύπτουν από το Μετασχηματισμό Wavelet της Εικόνας σε Διαφορετικούς Χρωματικούς Χώρους. (α)Μια Εικόνα της Βάσης. (β)-(στ) τα Ιστογράμματα των Wavelets Συντελεστών της Εικόνας στα LAB, YCbCr, RGB, HSV, YIQ Μοντέλα αντίστοιχα.

2.3 Κατανομή των Συντελεστών Κυματιδίων

Στη συνέχεια, χρησιμοποιούμε τους συντελεστές κυματιδίων κάθε επιπέδου ώστε να τους προσαρμόσουμε (fit) με την Student_t κατανομή [30]. Οι συντελεστές αυτοί μπορεί να είναι κάθε φορά άλλοι, ανάλογα με το τι θέλουμε να εξετάσουμε. Μπορούμε να μηδενίζουμε κάποιους, κρατώντας άλλους, ή να κάνουμε αντικατάσταση με τις αντίστοιχες πληροφορίες που προκύπτουν από κατάλληλη παραμετροποίηση και εκτίμηση παραμέτρων σε κάθε θέση.

Ο ορισμός της μονοδιάστατης στατιστικής κατανομής της Student's-t που χρησιμοποιείται στη παρούσα διπλωματική εργασία είναι ο εξής:

$$St_1(x_1 \mid 0, \lambda, \nu) = \frac{\Gamma((\nu+1)/2)}{\Gamma(\nu/2)} \left(\frac{\lambda}{\pi\nu}\right)^{\frac{1}{2}} \left(1 + \frac{\lambda}{\nu}x_i\right)^{\frac{\nu+1}{2}}$$
(2.4)

όπου i=1,...,Ν είναι το πλήθος των δεδομένων – παρατηρήσεων ενδιαφέροντος, ν είναι το πλήθος των βαθμών ελευθερίας, ενώ οι τιμές που λαμβάνουν τα δεδομένα μας μπορεί να ανήχουν σε ολόχληρη τη γραμμή των πραγματιχών αριθμών, ήτοι $-\infty < x < +\infty$. Επιπλέον, ως Γ ορίζεται η γνωστή μας Γάμμα συνάρτηση, όπου:

$$\Gamma(\alpha) = \int_0^\infty x^{a-1} e^{-x} dx \tag{2.5}$$

Η Student's t κατανομή είναι συμμετρική κοντά στο 0, που συμβαδίζει με το γεγονός ότι η μέση τιμή της είναι μηδέν και η τιμή ασυμμετρίας είναι, επίσης, μηδέν.

Η Student-t έχει αποδειχτεί κατάλληλη για την μοντελοποίηση δεδομένων, που έχουν παχιές (βαριές) ουρές (heavy tails). Το πεδίο μετασχηματισμού των κυματιδίων (wavelet domain) είναι βασικό παράδειγμα, καθώς υπάρχει μεγάλο πλήθος μικρών συντελεστών (ως προς την τιμή) και μικρό πλήθος πολύ μεγάλων συντελεστών.



Σχήμα 2.8 Ιστογράμματα των Συντελεστών προσέγγισης, των Οριζόντιων, Κάθετων και Διαγώνιων Συντελεστών που Προκύπτουν από το Μετασχηματισμό Κυματιδίων DWT.

Συνεπώς, η Student's t [30] κατανομή, ανάλογα με την επιλογή των παραμέτρων, μπορεί να έχει βαριές ουρές. Όσο οι βαθμοί ελευθερίας αυξάνονται, η συνάρτηση πυκνότητας πιθανότητας προσεγγίζει την κανονική κατανομή, ενώ, όσο πλησιάζουν το μηδέν, δεν έχει κάποια τιμή. Η κατανομή είναι συμμετρική, όσον αφορά την μέση τιμή της, και η διακύμανση είναι:

$$v/(v-2)$$
 µ ε v>2, (2.6)

διαφορετικά δεν ορίζεται. Γενικά, μόνο οι ροπές που είναι μικρότερες από το ν ορίζονται.

Η μορφή του προτεινόμενου παραμετριχού μοντέλου χαθορίζει, εν τέλει, τη λεπτομερή συμπεριφορά του μοντέλου αξιολόγησης χαι επιτρέπει έναν συγχεχριμένο βαθμό προσαρμοστιχότητας του μοντέλου σε διαφορετιχές πηγές δεδομένων (σε διαφορετιχές μορφές των ειχόνων που αξιολογούμε). Η Student's t είναι γνωστό ότι επιτρέπει μεγάλη ευελιξία στην μορφή της κατανομής της χαι, επιπλέον, μπορούμε να υπολογίσουμε την αβεβαιότητα, ως προς την φύση της ουράς της εμπειριχής κατανομής που μελετάμε. Κάτι τέτοιο χαθιστά τη προτεινόμενη κατανομή ένα ευέλιχτο μοντέλο, που μπορεί να καταγράψει το σχήμα της υποχείμενης συνάρτησης πυχνότητας πιθανότητας των δεδομένων που μας ενδιαφέρουν.

Για την περίπτωσή μας, είναι σημαντικό να μπορούμε να έχουμε καλές εκτιμήσεις των παραμέτρων που μας ενδιαφέρουν. Με στόχο την εύρεση των απαραίτητων παραμέτρων του προτεινόμενου μοντέλου, που πρόκειται να χρησιμοποιηθούν ως χαρακτηριστικά για τον αλγόριθμο αξιολόγησης, κάνουμε χρήση του επαναληπτικού αλγορίθμου ΕΜ (Expectation Maximization). Γενικά, πολλοί επαναληπτικοί αλγόριθμοι βασισμένοι στον ΕΜ έχουν αναπτυχθεί για να υπολογίσουν τις εκτιμήσεις της μέγιστης πιθανοφάνειας (ML-Maximum Likelihood) για τις παραμέτρους της t-κατανομής. Παρόλο που υπεισέρχεται μία μορφή υπολογιστικής πολυπλοκότητας, έχει αποδειχτεί ότι η ML εκτίμηση έχει πολύ μεγάλη ακρίβεια και, γι' αυτό το λόγο, έχει επιλεγεί και στη συγκεκριμένη μεθοδολογία.

Στην εργασία μας, μεταχειριζόμαστε τους συντελεστές ενδιαφέροντος ως τυχαίες μεταβλητές και υποθέτουμε ότι τα x_i^k για i = 1, ..., N είναι ανεξάρτητα και ομοιόμορφα κατανεμημένα δείγματα από την ίδια Student-t κατανομή (αλλά με διαφορετικό τελεστή k, ανάλογα με το επίπεδο wavelet ανάλυσης που χρησιμοποιείται) με μέσο μηδέν και παραμέτρους λ_k και v_k :

$$\boldsymbol{x}_{i}^{k} \sim St(\boldsymbol{x}_{i}^{k}; \boldsymbol{0}, \lambda_{k}, \boldsymbol{v}_{k}), \forall i, \forall k,$$

$$(2.7)$$

$$St(\mathbf{x}, 0, \lambda_k, \nu_k) = \frac{\Gamma(\nu_k/2 + 1/2)}{\Gamma(\nu_k/2)} \left(\frac{\lambda_k}{\pi\nu_k}\right)^{1/2} \left(1 + \frac{\lambda_k}{\nu_k} x^2\right)^{-\frac{\nu_k + 1}{2}}$$
(2.8)

Η υπόθεση της Student-t κατανομής συνεπάγεται μία διαδικασία δημιουργίας δεδομένων δύο επιπέδων [30]. Για να το αναλύσουμε περισσότερο, ας σκεφτεί κανείς ένα είδος ιεραρχικού μοντέλου δημιουργίας των συντελεστών, όπου αρχικά οι παράμετροι a_i^k δημιουργούνται, με ανεξάρτητο τρόπο η κάθε μία, από μία Gamma κατανομή. Σε ένα δεύτερο επίπεδο, κάθε συντελεστής x_i^k δημιουργείται από μία Κανονική κατανομή με μέσο μηδέν και ακρίβεια (precision) $\lambda_k a_i^k$.

Κατ' αρχάς, πρέπει να σημειωθεί ότι η συνάρτηση πυκνότητας πιθανότητας της εξίσωσης (2.7) μπορεί να γραφεί σαν το ολοκλήρωμα:

$$p(\boldsymbol{x}_{i}^{k}) = St(\boldsymbol{x}_{i}^{k}; \boldsymbol{0}, \lambda_{k}, \boldsymbol{v}_{k}) = \int_{0}^{+\infty} p(\boldsymbol{x}_{i}^{k} | \boldsymbol{a}_{i}^{k}) p(\boldsymbol{a}_{i}^{k}) d\boldsymbol{a}_{i}^{k}$$
(2.9)

όπου ισχύει:

$$p(\boldsymbol{x}_{i}^{k}|a_{i}^{k}) = N\left(0, \left(\lambda_{k}a_{i}^{k}\right)^{-1}\right), \ p(a_{i}^{k}) = Gamma\left(\frac{\nu_{k}}{2}, \frac{\nu_{k}}{2}\right)$$
(2.10)

Για να ερμηνεύσουμε τους πιο πάνω τύπους πιο διαισθητικά, θα πρέπει να δούμε τι γίνεται με τη χωρική μεταβλητότητα/προσαρμοστικότητα του

συγχεχριμένου μοντέλου. Σχεπτόμενοι τη χρησιμοποιούμενη μονοδιάστατη έχδοση των πιο πάνω τύπων, τότε σε χάθε θέση -*i* - παρατηρούμε ότι οι συντελεστές x_i^k μοντελοποιούνται ως μία Γχαουσσιανή χατανομή με μέσο μηδέν χαι διαχύμανση $(a_i^k)^{-1}$. Οι μεταβλητές x_i^k ονομάζονται «χρυφές1», λόγω του γεγονότος ότι δεν εμφανίζονται στην εξίσωση (εξάγονται λόγω της ολοχλήρωσης), αλλά παίζουν χρίσιμο ρόλο στο μοντέλο που δημιουργεί τους συντελεστές μας.

Επιπλέον, ιδιότητες, όπως χυρτότητα και λοξότητα, μπορούν να ληφθούν υπόψη από μία γενιχευμένη μορφή της προτεινόμενης κατανομής.

2.4 Η Οπτικά Εξέχουσα Περιοχή στο Πλαίσιο της Εκτίμησης της Ποιότητας της Εικόνας.

Τα τελευταία χρόνια [31], ο σχεδιασμός μετρικών ποιότητας της εικόνας βασισμένων στη μάθηση, σε συνδυασμό με την ανάπτυξη της τεχνητής νοημοσύνης, έχει γίνει πολύ δημοφιλής. Η βασική στρατηγική αυτού του είδους μετρικών είναι η μάθηση ενός μοντέλου παλινδρόμησης, το οποίο αντιστοιχίζει τα χαρακτηριστικά της εικόνας σε μια βαθμολογία ποιότητας. Υπάρχουν διάφορες τεχνικές παλινδρόμησης, όπως:

- μηχανές διανυσμάτων υποστήριξης (support vector machines-SVMs),
- νευρωνικά δίκτυα (neural networks-NNs),
- τυχαία δάση (random forests) και
- βαθιά μάθηση (deep learning),

που χρησιμοποιούνται ευρέως για την μάθηση ενός μοντέλου. Στην εργασία τους [32], οι K. Gu, G. Zhai et al εξάγουν δομικά χαρακτηριστικά και χαρακτηριστικά εμπνευσμένα από το ανθρώπινο οπτικό σύστημα (HVS). Έπειτα, χρησιμοποιώντας SVM παλινδρόμηση προκύπτει ένας εκτιμητής ποιότητας. Οι L. Kang, P. Ye et al, στην εργασία τους [33], πρότειναν ένα συμπαγές συνελικτικό νευρωνικό δίκτυο πολλαπλών διεργασιών για την ταυτόχρονη εκτίμηση της ποιότητας της εικόνας, αλλά και τον προσδιορισμό του τύπου παραμόρφωσης στην εικόνα. Στην εργασία τους [34], οι C. Li, A.C. Bovik et al ανέπτυξαν μια μετρική ποιότητας της εικόνας βασισμένη στην φάση, στην εντροπία και σε χαρακτηριστικά gradient, με τα οποία

 $^{^{1}}$ hidden $\acute{\eta}$ lattent

τροφοδότησαν ένα νευρωνικό δίκτυο παλινδρόμησης. Οι L. Zhang, Z. Gu et al παρουσίασαν την εικόνα ως ένα διάνυσμα χαρακτηριστικών και χρησιμοποίησαν ένα τυχαίο δάσος για την εκπαίδευση του μοντέλου παλινδρόμησης, το οποίο αντιστοιχούσε το διάνυσμα σε ένα υποκειμενικό score [35]. Οι W. Hou, X. Gao et al χρησιμοποίησαν ένα βαθύ δίκτυο μάθησης [36], το οποίο σχεδιάστηκε για να ταξινομήσει μια εικόνα σε 5 βαθμούς ποιότητας και, έπειτα, να μετατρέψει αυτές τις 5 κατηγορίες σε αριθμητικές βαθμολογίες. Οι συγγραφείς του [37] πρότειναν ένα συνελικτικό δίκτυο για εκτίμηση της ποιότητας της εικόνας, συνδυάζοντας την μάθηση και την παλινδρόμηση ως μια πλήρη βελτιστοποιημένη διαδικασία. Στην [38], προτάθηκε μια μέτρηση εκτίμησης ποιότητας των αραιών αναπαραστάσεων που μαθαίνεται από τις γενικές βάσεις δεδομένων εικόνας.

Όπως αναφέρθηκε στο πρώτο κεφάλαιο, στο τμήμα 1.7, η βελτίωση της αξιοπιστίας των μετρικών ποιότητας της εικόνας στηρίζεται στη κατανόηση του HVS, καθώς και στη μοντελοποίηση των πτυχών του που σχετίζονται με την αντίληψη της οπτικής ποιότητας. Για το σκοπό αυτό, μία από τις αυξανόμενες τάσεις στην έρευνα για την ποιότητα της εικόνας είναι η διεύρυνση του πώς η οπτική προσοχή (visual attention) επηρεάζει την αξιολόγηση της ποιότητας της εικόνας.

Η πιο χοινή μέθοδος, που υιοθετήθηχε, είναι η στάθμιση των τοπιχών παραμορφώσεων με χρήση οπτιχά εξέχουσας περιοχής, έχοντας σαν αποτέλεσμα την μετριχή σταθμισμένη εξέχουσα περιοχή. Για παράδειγμα, ο Moorthy [39] αναβάθμισε ένα υπάρχον μοντέλο οπτιχά εξέχουσας περιοχής, που ονομάζεται GAFFE, σε μια μετριχή SSIM, πετυχαίνοντας μια βελτίωση 1% εώς 4% όσον αφορά τη μετριχή απόδοση. Στο [40], σχεδιάστηχε μια non-reference μετριχή για την αξιολόγηση των τεχνουργημάτων συμπίεσης JPEG2000, χρησιμοποιώντας το μοντέλο εξέχουσας περιοχής του [41].

Πειραματικά αποτελέσματα έδειξαν, ότι οι πληροφορίες εξέχουσας περιοχής επέφεραν σημαντικές βελτιώσεις στην απόδοση της μετρικής ποιότητας χωρίς εξέχουσα περιοχή. Πρέπει να σημειωθεί, ότι η προσέγγιση στάθμισης της εξέχουσας περιοχής υποθέτει ότι ο μηχανισμός προσοχής του HVS λειτουργεί με έναν τρόπο μετά-επεξεργασίας, όταν εκτιμά την ποιότητα της εικόνας. Ωστόσο, η περίπλοκη αλληλεπίδραση μεταξύ οπτικής εξέχουσας περιοχής και ποιότητας εικόνας

αγνοείται. Η προστιθέμενη αξία της οπτικής εξέχουσας περιοχής στα IQMs εξαρτάται σε μεγάλο βαθμό από το μοντέλο εξέχουσας περιοχής, το IQM και τα χαρακτηριστικά της εικόνας ελέγχου. Οπότε, η προστιθέμενη αξία που αποκτάται είναι συχνά περιορισμένη και άλλοτε αρνητική. Επίσης, η προσέγγιση στάθμισης εξέχουσας περιοχής απαιτεί επιπλέον υπολογιστικό κόστος για την δημιουργία saliency maps και για την βελτίωση της σημασίας των τοπικών παραμορφώσεων. Μπορεί, επίσης, να περιορίσει περαιτέρω την ανάπτυξη αυτών των σταθμισμένων μετρικών σε εφαρμογές πραγματικού χρόνου. Συμπερασματικά, η διερεύνηση αντιληπτικά βελτιστο-ποιημένων τρόπων, για τη χρήση πληροφοριών εξέχουσας περιοχής σε IQMs, αξίζει περισσότερη μελέτη.

Ανάλογα με την επιλογή των μοντέλων εξεχουσών περιοχών ή εξέχουσας σημασίας και των IQM, ορισμένα ερευνητικά ευρήματα αποκάλυψαν ότι τα οφέλη από την προσθήχη αυτού του είδους της πληροφορίας στα IQM είναι οριαχά, ενώ ορισμένα ερευνητικά ευρήματα ανέφεραν ότι η εξέχουσα θέση θα μπορούσε να βελτιώσει σημαντικά τις μετρικές αξιολόγησης. Πολλά μοντέλα και αλλά και μετρικές καθίστανται διαθέσιμες, όμως ζητήματα, όπως το πώς ποικίλλουν τα οφέλη της συμπερίληψης της υπολογιστικής εξέχουσας περιοχής, αλλά και ποιες είναι οι αιτίες αυτής της παραλλαγής παραμένουν, αξίζουν περαιτέρω διερεύνησης. Όπως και να έχει, λόγω της περιορισμένης κατανόησής μας για τη σχέση μεταξύ οπτικής προσοχής και ποιότητας εικόνας, η εφαρμογή της οπτικής περιοχής σε διάφορες μετρικές εστιάζεται στη στάθμιση των τοπικών παραμορφώσεων (υπολογιζόμενες από μία μετρική) με την τοπική εξέχουσα περιοχή (που προέχυψε από ένα υπολογιστικό μοντέλο εξέχουσας περιοχής), αποδίδοντας ένα πιο εξελιγμένο μέσο πρόβλεψης ποιότητας εικόνας. Επιπλέον, για την απόκτηση περισσότερων γνώσεων για την ανθρώπινη όραση, έχουν πραγματοποιηθεί ψυχοφυσικές μελέτες για την καλύτερη κατανόηση των πτυχών της οπτικής προσοχής, σε σχέση με την αξιολόγηση της ποιότητας της εικόνας μέσω παρακολούθηση ματιών.

Επιπλέον των προαναφερομένων, ο μετασχηματισμός χυματιδίων έχει προσελχύσει το ενδιαφέρον των ερευνητών στη μοντελοποίηση οπτιχής προσοχής. Το πλεονέχτημα του WT είναι η ιχανότητά του να παρέχει ταυτόχρονα χωριχή ανάλυση και ανάλυση συχνοτήτων πολλαπλής χλίμαχας. Σ' αυτή την εργασία, προτείνουμε ένα νέο μοντέλο ανίχνευσης εξέχουσας περιοχής, με βάση τους

συντελεστές υψηλής διέλευσης της αποσύνθεσης κυματιδίων, αφού μειωθούν ή αλλαχθούν, ως προς την τιμή τους, κάποια υψηλής διέλευσης συστατικά της εικόνας. Η ιδέα είναι να δημιουργηθούν χάρτες χαρακτηριστικών (features maps) από τον αντίστροφο μετασχηματισμό κυματιδίων στην προτεινόμενη αποσύνθεση και αλλαγή των αντίστοιχων συντελεστών. Κάθε χάρτης χαρακτηριστικών περιέχει τοπικές παραλλαγές. Με άλλα λόγια, αντιπροσωπεύουν ζωνοπερατές (band-pass) τοπικές πληροφορίες με διαφορετικά εύρη ζώνης συχνοτήτων.

Η ανάλυση χυματιδίων είναι σε θέση να εχτελέσει χαλύτερη τοπιχή ανάλυση συχνότητας, καθώς εξετάζει το σήμα σε διαφορετιχές ζώνες και εύρη ζώνης. Αποτελεί διαδιχασία εφαρμογής μιας συστοιχίας φίλτρων, ιδιαίτερα όταν μιλούμε για πολλαπλή ανάλυση στο αρχικό σήμα. Η εφαρμογή του μετασχηματισμού έχει το πλεονέχτημα στην εξαγωγή λεπτομερειών προσανατολισμού (διαγώνιων, οριζοντίων, χάθετων) και επιτρέπει την υψηλή χωριχή ανάλυση με συστατιχά υψηλότερης συχνότητας και την χαμηλή χωριχή ανάλυση με συστατιχά χαμηλότερης ανάλυσης χωρίς απώλεια πληροφοριών σε λεπτομέρειες, κατά τη διάρχεια της διαδιχασίας αποσύνθεσης. Επομένως, ο μετασχηματισμός χυματιδίων με επιλεγμένα επίπεδα αποσύνθεσης μπορεί να παρέχει χάρτες χαραχτηριστιχών σε ζωνοπερατές περιοχές χωρίς περιεχόμενο, π.χ. χαμηλής συχνότητας μετά την αναχατασχευή, αγνοώντας τα σήματα προσέγγισης [42].

2.5 Ψεύδο-gradient

Παρόλο που ο μετασχηματισμός wavelet δεν παρέχει συντελεστές που θα μπορούσαν να ονομαστούν συντελεστές κλίσης, χρησιμοποιούμε σε κάθε επίπεδο τις οριζόντιες και κάθετες υπο-ζώνες συντελεστών για τον υπολογισμό των μέτρων και γωνιών, όπως στο [43], με τις ακόλουθες εξισώσεις:

$$mag = \sqrt{\mathbf{H}^2 + \mathbf{V}^2} \tag{2.11}$$

$$\phi = \begin{cases} \arctan\left(\frac{V}{H}\right), & \text{an } H \neq 0 \\ \frac{\pi}{2} , & \text{alling} \end{cases}$$
(2.12)

όπου τα Η και V αντιπροσωπεύουν τις υπο-ζώνες των οριζόντιων και κάθετων συντελεστών, αντίστοιχα. Η μεταβλητή mag είναι ο προκύπτων πίνακας με τα μέτρα που υπολογίζονται και φ είναι οι γωνίες του μετασχηματισμού κυματιδίων [44].

2.6 Μετρικές

2.6.1 Peak Signal-to-noise ratio PSNR

Αυτή η μετρική [45], [46], που προτάθηκε το 2004, λαμβάνει σαν σήμα τα αρχικά δεδομένα ή την αρχική εικόνα αναφοράς και ο θόρυβος είναι το σφάλμα που εισάγει κάποιο είδος επεξεργασίας. Λειτουργεί υπολογίζοντας πρώτα το μέσο τετραγωνικό σφάλμα και, στη συνέχεια, διαιρώντας το μέγιστο εύρος του τύπου των δεδομένων με το MSE. Εκφράζει την αναλογία μεταξύ της μέγιστης δυνατής ισχύος ενός σήματος και της ισχύος του θορύβου που επηρεάζει την πιστότητα της παράστασης και είναι λογαριθμικό μέγεθος. Ο τύπος του PSNR σε db είναι:

$$PSNR = 10 \cdot \log_{10} \left(\frac{MAX_{I}^{2}}{MSE} \right)$$
$$= 20 \cdot \log_{10} \left(\frac{MAX_{I}}{\sqrt{MSE}} \right)$$
$$= 20 \cdot \log_{10} \left(MAX_{I} \right) - 10 \cdot \log_{10} (MSE)$$
(2.13)

όπου MAX_I είναι η μέγιστη πιθανή τιμή που μπορεί να λάβει το pixel μια εικόνας. Το MSE ορίζεται ως το τετράγωνο των διαφορών των τιμών των αντίστοιχων pixel των δύο εικόνων. Το MSE [47] μιας εικόνας μεγέθους MxN δίνεται από τον τύπο:

$$MSE = \frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} \left(x(i, j) - y(i, j) \right)^2$$
(2.14)

όπου x(i,j) η τιμή του εικονοστοιχείου της αρχικής εικόνας μεγέθους MxN, και y(i,j) η τιμή του αντίστοιχου εικονοστοιχείου της παραμορφωμένης εικόνας μεγέθους MxN.

Όταν δύο εικόνες είναι πανομοιότυπες, τότε ο MSE ισούται με το μηδέν, καθώς ισχύει x(i,j)=y(i,j). Αν δεν υπάρχει κάποια παραμόρφωση, δηλαδή η τελική εικόνα είναι όμοια με την αρχική, το MSE είναι ίσο με το μηδέν, οπότε το PSNR τείνει στο άπειρο. Γενικά, το PSNR παρουσιάζει χαμηλές επιδόσεις σε σχέση με άλλες μετρικές, όσον αφορά την εκτίμηση της ποιότητας εικόνας και βίντεο.

2.6.2 Structural Similarity Index - SSIM

Ο δείκτης SSIM [47], [20], [48] είναι μια αντιληπτική μετρική, η οποία ποσοτικοποιεί την υποβάθμιση της ποιότητας της εικόνας που προκαλείται από την επεξεργασία, όπως η συμπίεση δεδομένων ή οι απώλειες κατά την μετάδοση δεδομένων. Είναι μια full reference μεθοδολογία, η οποία απαιτεί την αρχική και την παραμορφωμένη εικόνα, οπότε, ουσιαστικά, μετράει την ομοιότητα μεταξύ των δύο εικόνων. Χρησιμοποιεί τρία χαρακτηριστικά της εικόνας για να πραγματοποιήσει την σύγκριση των εικόνων: Την φωτεινότητα, την αντίθεση και την δομή.

1. Φωτεινότητα

Υπολογίζεται ως η μέση ένταση φωτεινότητας του χάθε σήματος από το τύπο:

$$\overline{x} = \frac{1}{N} \sum_{i=1}^{N} x_i \tag{2.15}$$

$$\overline{y} = \frac{1}{N} \sum_{i=1}^{N} y_i \tag{2.16}$$

Η σύγκριση αυτών των τιμών δίνεται από τον τύπο:

$$l(x, y) = (2\tilde{x}\tilde{y} + C1) / ((\tilde{x})^2 + (\tilde{y})^2 + C1)$$
(2.17)

2. Αντίθεση

Την υπολογίζουμε μέσω της τυπικής απόκλισης, που δίνεται από τον τύπο:

$$\sigma_x^2 = \frac{1}{N-1} \sum_{i=1}^{N} (x_i - \overline{x})^2$$
(2.18)

Η σύγκριση των αντιθέσεων δίνεται από το τύπο:

$$c(\mathbf{x}, \mathbf{y}) = \frac{2\sigma_x \sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2}$$
(2.19)

Οι C1 και C2 είναι σταθερές που προσθέτουμε, ώστε να αποφευχθεί η διαίρεση με το μηδέν.

3. Δομή

Για να κανονικοποιηθεί το σήμα, το διαιρούμε με την τυπική απόκλιση. Η σύγκριση γίνεται μεταξύ των κανονικοποιμένων σημάτων.

Η σύγχριση των δομών δίνεται από τον τύπο:

$$s(x, y) = \frac{\sigma_{xy} + c_3}{\sigma_x \sigma_y + c_3}$$
(2.20)

όπου:

$$c_3 = c_2 / 2 \tag{2.21}$$

Ο συνδυασμός των τριών παραπάνω χαραχτηριστιχών μας δίνει το μέτρο ομοιότητας. Ο τύπος του SSIM είναι ο εξής :

$$SSIM = \frac{(2 \times \overline{x} \times \overline{y} + C1) \times (2 \times \sigma_{xy} + C2)}{\left(\sigma_x^2 + \sigma_y^2 + C2\right) \times \left((\overline{x})^2 + (\overline{y})^2 + C1\right)}$$
(2.22)

Στο Σχήμα 2.9 βλέπουμε ένα διάγραμμα, που αναπαριστά τον αλγόριθμο SSIM.



Σχήμα 2.9 Σχηματική Αναπαράσταση της Λειτουργίας του Αλγορίθμου SSIM

Αν και πρωτοεμφανίστηκε το 2001, εμφανίζει μέχρι και σήμερα πολύ ανταγωνιστικά αποτελέσματα σε σχέση με νεότερες μεθόδους.

2.6.3 Visual Signal-to-Noise Ratio-VSNR

Η VSNR [49] είναι μια μέθοδος που αναπτύχθηκε το 2007 και λειτουργεί σε δύο στάδια. Στο πρώτο, τα κατώφλια αντίθεσης για ανίχνευση των παραμορφώσεων παρουσία φυσικών εικόνων υπολογίζονται μέσω μοντέλων οπτικής κάλυψης και οπτικής άθροισης και είναι βασισμένα σε κυματίδια. Αυτά τα μοντέλα χρησιμοποιούνται με σκοπό να προσδιοριστεί εάν οι στρεβλώσεις στην παραμορφωμένη εικόνα είναι ορατές. Εάν οι παραμορφώσεις είναι κάτω από το όριο ανίχνευσης, η παραμορφωμένη εικόνα θεωρείται ότι έχει τέλεια οπτική πιστότητα (VSNR = infin) και δεν απαιτείται περαιτέρω ανάλυση.

Εάν οι στρεβλώσεις είναι πάνω από το όριο ανίχνευσης, εφαρμόζεται ένα δεύτερο στάδιο, που λειτουργεί με βάση την οπτική ιδιότητα χαμηλού επιπέδου της αντιληπτής αντίθεσης και την οπτική ιδιότητα μεσαίου επιπέδου της καθολικής προτεραιότητας. Αυτές οι δύο ιδιότητες μοντελοποιούνται ως ευκλείδειες αποστάσεις στο διάστημα παραμόρφωσης-αντίθεσης μιας αποσύνθεσης κυματιδίων πολλαπλής κλίμακας και το VSNR υπολογίζεται με βάση ένα απλό γραμμικό άθροισμα αυτών των αποστάσεων.

Η προτεινόμενη μέτρηση VSNR είναι γενικά ανταγωνιστική με τις τρέχουσες μετρικές οπτικής πιστότητας. Είναι αποτελεσματική τόσο από την άποψη της χαμηλής υπολογιστικής πολυπλοκότητας όσο και από την άποψη των χαμηλών απαιτήσεων μνήμης.

2.6.4 Feature Similarity Index – FSIM

Η FSIM μετρική αφορά έναν full-reference αλγόριθμο εκτίμησης της ποιότητας που ανέπτυξαν οι Lin Zhang, Lei Zhang το 2011 [24]. Βασίζεται στο γεγονός ότι το ανθρώπινο οπτικό σύστημα αντιλαμβάνεται την εικόνα, κυρίως, σύμφωνα με τα χαρακτηριστικά χαμηλού επιπέδου της. Συγκεκριμένα, η συνάφεια φάσης (Phase Congruency-PC), η οποία είναι ένα αδιάστατο μέτρο της σημασίας μιας τοπικής δομής, χρησιμοποιείται ως το κύριο χαρακτηριστικό στο FSIM. Λαμβάνοντας υπόψη ότι η συνάφεια φάσης είναι αναξιόπιστη, ενώ οι πληροφορίες αντίθεσης επηρεάζουν την αντίληψη του HVS για την ποιότητα της εικόνας, το μέτρο της κλίσης (Gradient Magnitude - GM) της εικόνας χρησιμοποιείται ως δευτερεύον χαρακτηριστικό στο FSIM. Αυτά τα δύο χαρακτηριστικά διαδραματίζουν συμπληρωματικούς ρόλους στον χαρακτηρισμό της τοπικής ποιότητας της εικόνας.

Μετά την απόκτηση του τοπικού χάρτη ποιότητας (local quality map), χρησιμοποιούμε ξανά την συνάφεια φάσης ως λειτουργία στάθμισης, για να αντλήσουμε μια ενιαία βαθμολογία ποιότητας.

2.6.5 Visual Information Fidelity – VIF

Η VIF [50] μετρική είναι ένας Full-Reference αλγόριθμος του 2006, που χρησιμοποιεί μοντέλα στατιστικών φυσικών σκηνών (NSS) σε συνδυασμό με ένα μοντέλο παραμόρφωσης (καναλιού), για τον ποσοτικό προσδιορισμό των πληροφοριών που μοιράζονται μεταξύ της εικόνας αναφοράς και της εικόνας ελέγχου. Επιπλέον, ο δείκτης VIF βασίζεται στην υπόθεση ότι αυτές οι κοινές πληροφορίες είναι μια πτυχή της πιστότητας που σχετίζεται καλά με την οπτική ποιότητα. Πρόκειται για μια μέθοδο αρκετά ανταγωνιστική με τις μεθόδους QA.

Συγκεκριμένα, η εικόνα αναφοράς διαμορφώνεται ως η έξοδος μιας στοχαστικής «φυσικής» πηγής που διέρχεται από το κανάλι HVS και επεξεργάζεται αργότερα από τον εγκέφαλο. Το περιεχόμενο των πληροφοριών της εικόνας αναφοράς ποσοτικοποιείται ως η αμοιβαία πληροφορία μεταξύ της εισόδου και της εξόδου του καναλιού HVS. Το ίδιο μέτρο στη, συνέχεια, ποσοτικοποιείται παρουσία ενός καναλιού παραμόρφωσης εικόνας, το οποίο παραμορφώνει την έξοδο της φυσικής πηγής πριν περάσει μέσα από το κανάλι HVS, μετρώντας έτσι τις πληροφορίες που ο εγκέφαλος θα μπορούσε ιδανικά να εξάγει από την εικόνα ελέγχου.

Στη συνέχεια, τα δύο μέτρα πληροφόρησης συνδυάζονται για να σχηματίσουν ένα μέτρο πιστότητας οπτικής πληροφορίας, όπως παρουσιάζεται στην παρακάτω εικόνα.



Σχήμα 2.10 Αναπαράσταση Λειτουργίας Αλγορίθμου VIF

2.6.6 VCGS

Η μέθοδος VCGS, που πρότειναν οι Chenyang Shi και Yandan Lin [51] το 2020, είναι άξια αναφοράς όσον αφορά την υλοποίηση και τα αποτελέσματά της και αποτελεί την μέθοδο εκκίνησης για να προτείνουμε μια καινούρια, στην παρούσα διπλωματική εργασία, κάνοντας παραλλαγές σε διάφορα στάδιά της. Πρόκειται για μια full reference μεθοδολογία, που σημαίνει ότι δεν απαιτείται διαδικασία μάθησης, και βασίζεται στον συνδυασμό της ομοιότητας του visual saliency με εμφάνιση χρώματος, της gradient και chrominance ομοιότητας. Επιπλέον, γίνεται χρήση των δύο δεικτών εμφάνισης χρώματος, του vividness και του depth, με σκοπό την κατασκευή ενός χάρτη ομοιότητας βασισμένου στα εξέχοντα οπτικά χαρακτηριστικά μίας εικόνας. Αφού υπολογιστούν τα vividness και depth για την αρχική και για την παραμορφωμένη εικόνα, γίνεται χρήση ενός band-pass φίλτρου για την ανίχνευση saliency. Επιπλέον χρησιμοποιείται ένα Log-Gabor φίλτρο για την ποσοτικοποίηση του visual saliency. Η συνάρτηση μεταφοράς ενός Log-Gabor φίλτρου g(x) στο πεδίο συχνοτήτων εκφράζεται ως εξής:

$$G(\mathbf{u}) = \exp\left(-\left(\log\frac{||\mathbf{u}||_2}{\omega_0}\right)^2 / 2\sigma_F^2\right)$$
(2.23)

όπου $\mathbf{u} = (u, v) \in \mathbb{R}^2$ οι συντεταγμένες στο πεδίο συχνοτήτων, ω_0 το κέντρο συχνοτήτων του φίλτρου και σ_F η παράμετρος εύρους ζώνης του φίλτρου. Χρησιμοποιείται το G(u) αντί για το g(x), καθώς μπορεί να εκφραστεί αναλυτικά η συνάρτησή του, και για να πλησιάσουν προσεγγιστικά την τιμή του g(x) εφαρμόζουν αριθμητικό αντίστροφο μετασχηματισμό Fourier. Έπειτα υπολογίζουν το visual saliency με εμφάνιση χρώματος (visual saliency with color appearance-VC) υπολογίζεται από τον τύπο:

$$VC = \left((V * g)^{2} + (D * g)^{2} \right)^{1/2}$$
(2.24)

Όπου * δηλώνει την πράξη της συνέλιξης.



Σχήμα 2.11 Το VC Μαρ της Αρχικής Εικόνας (αριστερά) και το VC Μαρ της Παραμορφωμένης Εικόνας (δεξιά).

Το VC map διαδραματίζει δύο ρόλους στο προτεινόμενο μοντέλο. Αρχικά, χρησιμοποιείται ως χαρακτηριστικό για τον υπολογισμό της τοπικής ποιότητας της παραμορφωμένης εικόνας και λειτουργεί και ως ένα μέσο στάθμισης, που αντικατοπτρίζει τη σημασία του τοπικού πεδίου. Τα χαρακτηριστικά δομής και χρωματισμού χαρακτηρίζονται από διαφορετικά κανάλια επιλεγμένου χρωματικού χώρου. Αρχικά, το μοντέλο λαμβάνει σαν είσοδο δύο RGB εικόνες X και Υ, με πανομοιότυπη χωρική ανάλυση, που αποτελούν την αρχική και την παραμορφωμένη εικόνα αντίστοιχα. Σαν έξοδο λαμβάνουμε μία βαθμολόγηση (score), που αφορά την αντιληπτή ποιότητα της παραμορφωμένης εικόνας και κινείται μεταξύ των τιμών 0 και 1, με το 1 να υποδηλώνει την καλύτερη δυνατή ποιότητα, που επιτυγχάνεται όταν η παραμορφωμένη εικόνα είναι πανομοιότυπη με την αρχική.

Στο Σχήμα 2.12, στην πρώτη σειρά, παρουσιάζονται η αρχική εικόνα και οι εικόνες που προκύπτουν από αυτή με 3 διαφορετικούς βαθμούς παραμόρφωσης θολώματος, τύπου Gaussian blur. Στην δεύτερη σειρά, παρουσιάζονται τα VC maps των εικόνων της πρώτης γραμμής έχοντας κανονικοποιηθεί. Παρατηρείται ότι όσο μεγαλύτερος είναι ο βαθμός παραμόρφωσης της εικόνας, τόσο χαμηλότερη είναι η ποιότητα του αντίστοιχου VC map. Αυτό οφείλεται στο γεγονός, ότι το σήμα υψηλής συχνότητας του VC map εξαρτάται από την ποιότητα της εικόνας εισόδου. Γι' αυτό το λόγο, το VC map χρησιμοποιείται για το σχεδιασμό του μοντέλου τους. Στην τρίτη σειρά, παρουσιάζονται τα gradient map των εικόνων της πρώτης σειράς. Εδώ παρατηρούμε, ότι όσο πιο παραμορφωμένη είναι η εικόνα, τόσο χαμηλότερη είναι η δομική πληροφορία του gradient map.



Α1. Αρχική Εικόνα



B1.

Gblur DMOS:1.0208 VCGS: 0.9914



C1. Gblur DMOS:1.9374 VCGS:0.9676



D1.Gblur DMOS:3.5416 V1CGS:0.9355



A2. VC map

B2. VC map





D2. VC map



A3. Gradient map

B3. Gradient map

C3. Gradient map

D4. Gradient map

Σχήμα 2.12 (A1) Αρχική Εικόνα από το LIVE DATASET, (B1-D1) Εικόνες με Παραμόρφωση Gblur σε Διαφορετικό Βαθμό Παραμόρφωσης η κάθε μία με το Αντίστοιχο MOS και το VCGS Score, (A2-D2) Τα VC Maps κάθε Εικόνας και (A3-D3) τα Gradient Maps κάθε Εικόνας.

ΚΕΦΑΛΑΙΟ 3

Προτεινομένη Full-Reference Μεθοδολογιά Βασισμένη στην Οπτικά Εξέχουσα Περιοχή, με Χρήση Στατιστικών Χαρακτηριστικών Μετασχηματισμού Κύματιδιών

- 3.1 Πρώτη Ομάδα Πειραμάτων
- 3.2 Δεύτερη Ομάδα Πειραμάτων
- 3.3 Τρίτη Ομάδα Πειραμάτων
- 3.4 Τέταρτη Ομάδα Πειραμάτων
- 3.5 Πέμπτη Ομάδα Πειραμάτων
- 3.6 Έχτη Ομάδα Πειραμάτων
- 3.7 Έβδομη Ομάδα Πειραμάτων
- 3.8 Όγδοη Ομάδα Πειραμάτων
- 3.9 Συγκεντρωτικά Αποτελέσματα

Τα τελευταία χρόνια, έχουν προταθεί πολλά μοντέλα εκτίμησης ποιότητας εικόνας με καλύτερη απόδοση, βασισμένα όχι μόνο στο κλασικό RGB χρωματικό μοντέλο αλλά εστιασμένο στο χρωματικό μοντέλο HVS. Αυτού του είδους τα μοντέλα χαρακτηρίζουν καλύτερα την δομική και χρωματική πληροφορία, την πληροφορία φωτεινότητας και αντίθεσης, τόσο στο χωρικό πεδίο όσο και στο πεδίο μετασχηματισμού. Ένα αντιπροσωπευτικό μοντέλο για τον υπολογισμό της δομικής πληροφορίας είναι ο γνωστός δείκτης Structural Similarity-SSIM, ο οποίος έχει πετύχει υψηλή απόδοση σε πολλές βάσεις δεδομένων. Τα γεωμετρικά χαρακτηριστικά είναι κατάλληλα για το πεδίο επιπέδων του γκρι (grayscale) με σχετικά ασήμαντη πληροφορία χρώματος και έχουν υψηλή υπολογιστική απόδοση. Παρόλα αυτά, δεν μπορούν να αντεπεξέλθουν το ίδιο καλά όταν έχουμε έγχρωμη εικόνα με χρωματικές αποκλίσεις, καθώς παραβλέπουν τον αντίκτυπο των χρωματικών πληροφοριών στην οπτική ποιότητα.

Συνήθως, για την εκτίμηση ποιότητας μιας έγχρωμης εικόνας χρησιμοποιείται ένα μοντέλο επιπέδων του γκρι (grayscale) σε μεμονωμένα RGB κανάλια και, έπειτα, συνδέονται οι βαθμολογίες των καναλιών. Αυτά τα μοντέλα θεωρούνται ανεπαρκή και καθιστούν απαραίτητη την πρόταση μοντέλων που εκμεταλλεύονται περισσότερο τις ιδιότητες της αντίληψης χρωμάτων. Έχοντας κατά νου τα προηγούμενα, χρησιμοποιήθηκαν ομοιόμορφοι χρωματικοί χώροι, όπως το CIELAB και το CIELUV, ώστε να χαρακτηρίσουν τις αντιληπτές χρωματικές διαφορές μεταξύ των ζευγών χρωμάτων. Οι κυλινδρικές πολικές συντεταγμένες στο CIELAB συσχετίζονται με τις χρωματικές ιδιότητες ενός ερεθίσματος. Κατά συνέπεια, η αντιληπτή παραμόρφωση των έγχρωμων εικόνων μπορεί να χαρακτηριστεί καλύτερα σε επίπεδα pixel χρησιμοποιώντας συνθέσεις CIE. Για υψηλότερη ακρίβεια στις εκτιμήσεις του μοντέλου, γίνεται χρήση δεικτών εμφάνισης χρωμάτων, ώστε να γίνει δυνατή η πρόσβαση στην χρωματική πληροφορία. Αυτοί οι δείκτες μπορούν να αναπαραστήσουν την χρωματική απόκλιση μιας έγχρωμης εικόνας με έναν HVS τρόπο αντίληψης.

Να σημειωθεί ότι, παρότι η χρωματική πληροφορία είναι ιδιαίτερα ελκυστική στο ανθρώπινο οπτικό σύστημα, σπάνια διερευνάται, τουλάχιστον μέχρι τη προηγούμενη δεκαετία. Η έρευνα στη γνωστική νευρο-επιστήμη, μας δίνει τη δυνατότητα να κατευθυνθούμε προς τη χρήση της χρωματικής πληροφορίας [52], [53].

Η χρήση του RGB έχει τα θετικά της, αλλά και τα αρνητικά της. Για παράδειγμα, οι εικόνες-χάρτες των τριών καναλιών του RGB μοντέλου είναι ιδιαίτερα συνεπείς με το περιεχόμενο της αρχικής εικόνας. Σε αυτή την εργασία, χρησιμοποιούνται RGB εικόνες ως είσοδοι (Σχήμα 3.1) και μεταφέρονται στο CIELAB χώρο (Σχήμα 3.2), καθώς πρόκειται για έναν χώρο ομοιόμορφο και πιο

συμβατό στην ανθρώπινη αντίληψη, με φωτεινότητα και δύο χρωματικά κανάλια (RG και BY). Το μοντέλο ακολουθεί μια full reference μεθοδολογία, που σημαίνει ότι δεν απαιτείται διαδικασία μάθησης και στηρίζεται στην μεθοδολογία VCGS που προτείνουν οι Chenyang Shi και Yandan Lin [51], αναφορά της οποίας έγινε στο προηγούμενο κεφάλαιο.



Σχήμα 3.1 Εικόνες RGB που Λαμβάνει ως Είσοδο η Προτεινόμενη Μέθόδος



Σχήμα 3.2 Μεταφορά των Εικόνων RGB που Λαμβάνει σαν Είσοδο η Προτεινόμενη Μέθοδος στο χώρο CIELAB

Αρχικά υπολογίζουμε τα εξής χαρακτηριστικά:

• Χαρακτηριστικό εξεχουσών περιοχών (Visual saliency feature)

Οι εξέχουσες περιοχές (Salient Regions) συνιστούν ένα καλό χαρακτηριστικό της αξιολόγησης μίας εικόνας, λόγω του ότι το Ανθρώπινο Οπτικό Σύστημα (HVS) είναι αρκετά ευαίσθητο σε αυτό. Επιπλέον, είναι πολύ σημαντικές για το ανθρώπινο οπτικό σύστημα, αφού τα ανθρώπινα όντα δίνουν μεγαλύτερη προσοχή σε αυτές τις περιοχές. Οι Zhang et al. [24] ερεύνησαν διεξοδικά τη σχέση αυτών των περιοχών με την αξιολόγηση της εικόνας και χρησιμοποίησαν τις πληροφορίες VS ως τη κύρια συνάρτηση στάθμισης για τη τελική συγκέντρωση της βαθμολογίας ποιότητας (σκορ ποιότητας).

Η σχέση μεταξύ οπτικής εξέχουσας θέσης (visual saliency) και ποιότητας εικόνας μπορεί να ενσωματωθεί σε ένα μοντέλο αξιολόγησης της εικόνας, ενώ η διερεύνησή τους μπορεί να ωθήσει σε ακόμη καλύτερα αποτελέσματα πρόβλεψης. Επομένως, στη βιβλιογραφία, μπορεί κανείς να δει προσπάθειες, που χρησιμοποιούν τη χρήση VS για τη βελτίωση της απόδοσης κατά το σχεδιασμό μοντέλων IQA [32]. Για αυτούς τους λόγους, ένα προτεινόμενο μοντέλο IQA σχεδιάζεται χρησιμοποιώντας είτε την αντίθεση είτε VS χαρακτηριστικά για την περιγραφή της ποιότητας.

Στην εργασία των Zhang et al. [24], έχει δειχθεί ότι η υποβάθμιση της αντιληπτικής ποιότητας εικόνας μπορεί να οδηγήσει σε αντιληπτικές αλλαγές σε χαμηλού-επιπέδου (Low-level) χαρακτηριστικά (features) της εικόνας. Από τη στιγμή που τα από κάτω προς τα πάνω (bottom-up) μοντέλα visual saliency βασίζονται, ουσιαστικά, σε χαμηλού επιπέδου (low-level) χαρακτηριστικά μίας εικόνας, οι ίδιες οι τιμές του Visual Saliency ποικίλλουν, ανάλογα με τις αλλαγές που υφίσταται η ποιότητα μίας εικόνας. Με μία τέτοιου είδους λογική, το visual saliency μπορεί να χρησιμοποιηθεί ως χαρακτηριστικό για να υπολογιστεί ένα τοπικός χάρτης ομοιότητας (Local similarity map). Όμως, τί κάνουμε στη παρούσα εργασία;

Υπολογίζουμε τα χαραχτηριστικά vividness και depth του χρωματικού μοντέλου CIE Lab με τους τύπους (2.1) και (2.2), αντίστοιχα. Έπειτα, αντί να χρησιμοποιήσουμε ζωνοπερατό (band pass) φιλτράρισμα με τη βοήθεια του Log-Gabor φίλτρου για την ανίχνευση του saliency, όπως κάνουν οι Chenyang Shi και Yandan Lin, εφαρμόζουμε wavelet μετασχηματισμό και διερευνούμε τους παραγόμενους συντελεστές. Λαμβάνοντας τις λεπτομέρειες κατεύθυνσης από τον μετασχηματισμό, αντικαθιστούμε την πληροφορία τους με παραμέτρους που χρησιμοποιούνται στην εκτίμηση των παραμέτρων της προσαρμοσμένης Student'st κατανομής. Στην συνέχεια, με αντίστροφο μετασχηματισμό επιστρέφουμε στο χωρικό πεδίο και υπολογίζεται το visual saliency, μέσω του τύπου :

$$VC = \sqrt{V'^2 + D'^2}$$
(3.1)



Σχήμα 3.3 Διάγραμμα Αναπαράστασης της Δημιουργίας των Χαρτών του Visual Saliency, Gradient, and Chrominance της Προτεινόμενης Μεθολογίας.

Αυτή, βέβαια, δεν είναι η μόνη λογική που ακολουθούμε, αλλά ελέγχουμε τί γίνεται με τη χρήση του προτεινόμενου δείκτη, χρησιμοποιώντας άμεσα στο πεδίο μετασχηματισμού τους προτεινόμενους συντελεστές λεπτομέρειας. Αλλαγές στους visual saliency χάρτες μίας εικόνας μπορεί να αποδειχτούν καλοί δείκτες του βαθμού της παραμόρφωσης που υφίσταται μία εικόνα, οπότε, όπως θα δούμε, ένας τέτοιου είδους χάρτης θα χρησιμοποιηθεί ως χαρακτηριστικό (feature) που μπορεί να χαρακτηρίσει την ποιότητα της εικόνας σε τοπικό επίπεδο.

Βέβαια, ο ρόλος ενός VS χάρτη δεν εξαντλείται μόνο στη χρήση του ως χάρτης που περιγράφει την «τοπική» ποιότητα μίας εικόνας, αλλά μπορεί να χρησιμοποιηθεί και σε επόμενο στάδιο ως μία συνάρτηση βάρους (weighting function), που δηλώνει την σημασία μίας τοπικής περιοχής για το pooling του τελικού σκορ.

• Δομικό χαρακτηριστικό (Structural feature)

Ωθούμενοι από το γεγονός ότι το ανθρώπινο οπτικό σύστημα είναι ευαίσθητο ως προς την πληροφορία ακμής, αρκετές μεθοδολογίες αξιολόγησης εμπλέκουν την έννοια των ακμών ως χαρακτηριστικό που σχετίζεται βάσιμα με την ποιότητα μίας εικόνας, με σκοπό να αξιολογηθεί η οπτική ποιότητα παραμορφωμένων εικόνων φυσικών σκηνών. Για αυτόν τον λόγο, ένας χάρτης κλίσης (gradient map) μίας εικόνας αποτελεί πλούσιο περιγραφέα της τοπικής δομής, αλλά και ποιότητας μίας εικόνας. Ο τρόπος που μεταλλάσσονται οι τιμές του χάρτη, ανταποκρίνεται στις διάφορες παραμορφώσεις. Το visual saliency μπορεί να χρησιμοποιηθεί ως χαρακτηριστικό για να υπολογιστεί ένας χάρτης ομοιότητας. Όμως, λόγω της αδυναμίας να χαρακτηριστεί με καλό τρόπο π.χ. ο τύπος της αλλαγής της αντίθεσης ως παραμόρφωση (Contrast Change distortion), μπορούμε να χρησιμοποιήσουμε επιπλέον χαρακτηριστικά, τα οποία «συμπληρώνουν» κατά μία έννοια την έλλειψη της ευαισθησίας αντίθεσης ενός VS χάρτη.

Έτσι, υπολογίζουμε την οριζόντια και κάθετη παράγωγο $G_x(i)$ και $G_y(i)$ των εικόνων εισόδου και μέσω του τύπου

$$G(i) = \sqrt{G_x^2(i) + G_y^2(i)}$$
(3.2)

παίρνουμε το μέτρο της πρώτης παραγώγου, που αποτελεί το δομικό χαρακτηριστικό.

Ως επιπλέον δομικό χαρακτηριστικό χρησιμοποιούμε, επίσης, ένα είδος gradient στο πεδίο μετασχηματισμού, χρησιμοποιώντας το μέτρο της πρώτης παραγώγου, βασισμένοι στους οριζόντιους και κάθετους wavelet συντελεστές, με το τύπο:

$$mag = \sqrt{\mathrm{H}^2 + \mathrm{V}^2} \tag{3.3}$$

• Χαρακτηριστικά χρωματικότητας (Chrominance features)

Στον CIELAB, το α* αντιπροσωπεύει το κόκκινο-πράσινο και το b* το μπλεκίτρινο και τα χρησιμοποιούμε ως χαρακτηριστικά.

Αφού υπολογίσουμε τα παραπάνω υπολογίζουμε τα εξής τρία similarity maps:

1) Visual Saliency Similarity Map

Έχοντας υπολογίσει το visual saliency χαρακτηριστικό για την αρχική και παραμορφωμένη εικόνα, τα συγκρίνουμε για να πάρουμε το Visual Saliency Similarity Map- S_{VS} .

$$S_{VS} = \frac{2VS_O \cdot VS_D + K_{VS}}{VS_O^2 + VS_D^2 + K_{VS}}$$
(3.4)

όπου Κ_{vc} μια σταθερά για την αποφυγή μηδενιχού παρονομαστή.

2) Gradient Similarity Map

Ομοίως, έχοντας πλέον το δομικό χαρακτηριστικό της αρχικής και παραμορφωμένης εικόνας μέσω της πρώτης παραγώγου, υπολογίζουμε την ομοιότητα τους, για να πάρουμε το Gradient Similarity Map- S_G , με τον τύπο:

$$S_{GS} = \frac{2GS_{O} \cdot GS_{D} + K_{GS}}{GS_{O}^{2} + GS_{D}^{2} + K_{GS}}$$
(3.5)

όπου K_{GS} μια σταθερά για την αποφυγή μηδενιχού παρονομαστή.

3) Chrominance Similarity Maps

Η ομοιότητα των συστατικών chrominance- S_c . στο CIELAB των καναλιών κόκκινου-πράσινου α^{*} και μπλε-κίτρινου b^{*} υπολογίζεται με τον τύπο:

$$S_{CHR} = \frac{2a_o \cdot a_D + K_{CHR}}{a_o^2 + a_D^2 + K_{CHR}} \cdot \frac{2b_o \cdot b_D + K_{CHR}}{b_o^2 + b_D^2 + K_{CHR}}$$
(3.6)

όπου Κς μια σταθερά για την αποφυγή μηδενιχού παρονομαστή

Τα τρία είδη similarity που αναφέραμε, δηλαδή το visual saliency similarity map- S_{VS} , το gradient similarity- S_G και το chrominance similarity- S_c , συνδυάζονται, ώστε να υπολογιστεί η τοπική ποιότητα της παραμορφωμένης εικόνας, μέσω του τύπου:

$$VSSC = \frac{\sum_{\Omega} S_{VS} \cdot S_{GS}^{\ a} \cdot S_{CHR}^{\lambda} \cdot VS_{max}}{\sum_{\Omega} VS_{max}}$$
(3.7)

όπου Ω ορίζει το χωρικό πεδίο και $VS_{max} = max(VS_o, VS_D)$ ορίζει την μέγιστη τιμή για κάθε pixel του map των εξεχουσών περιοχών μεταξύ των δύο saliency map. Το VS_{max} χρησιμοποιείται ως ένας σημαντικός συντελεστής στάθμισης μεταξύ αρχικής και παραμορφωμένης εικόνας.



Σχήμα 3.4 Διάγραμμα Αναπαράστασης της Δημιουργίας των Χαρτών Ομοιότητας του Visual Saliency, του Gradient και του Chrominance και του Τελικού Χάρτη Τοπικής Ποιότητας. Εφαρμογή Σταθμισμένου Pooling Βασισμένο στο Visual Saliency για την Παραγωγή της Συνολικής Βαθμολογία Ποιότητας.

Τα πειράματα δοχιμάζονται στο σύνολο δεδομένων LIVE IQA, το οποίο περιέχει ειχόνες που έχουν υποστεί ένα από τα πέντε είδη παραμορφώσεων, white noise, guassian blur, jp2k, jpeg, fast fading, ενώ θα εξετάσουμε χαι την επιρροή χάθε πειράματος χαι στον συνδυασμό όλων των παραμορφώσεων (All). Από το dataset αφαιρέθηχαν οι ειχόνες αναφοράς, έτσι ώστε οι αλγόριθμοι να εξετάζονται μόνο σε παραμορφωμένες ειχόνες. Επίσης, χρησιμοποιήθηχαν σαν υποχειμενιχά scores τα dmos, που έχουν υποστεί ένα είδος ευθυγράμμισης χαι διατίθενται επιπλέον στην βάση δεδομένων μαζί με αυτά που δεν έχουν υποστεί επεξεργασία. Η μέτρηση της απόδοσης τους γίνεται με :

A) τον συντελεστή γραμμική συσχέτισης του Pearson (Linear Correlation coefficient - LCC), που μετρά τη προβλεπτική γραμμικότητα των IQA μεθόδων, και

B) τον βαθμό κατάταξης του συντελεστή συσχέτισης Spearman (rank order correlation coefficient - SROCC) μεταξύ των εκτιμώμενων βαθμολογιών (scores), που υπολογίζει το εργαλείο ταξινόμησης που χρησιμοποιούμε, και των βαθμολογιών που έχουν αποδοθεί στις εικόνες από ανθρώπους (MOS, DMOS). Στην ουσία δηλαδή, μεταξύ των αντικειμενικών βαθμολογιών, που προβλέπονται από το προτεινόμενο μοντέλο, και των υποκειμενικών βαθμολογιών, που παρέχονται από το σύνολο δεδομένων.

Ο συντελεστής Pearson υπολογίζεται από τον τύπο:

$$rho(a,b) = \frac{\sum_{i=1}^{n} (X_{a,i} - \overline{X}_{a}) (Y_{b,i} - \overline{Y}_{b})}{\left\{ \sum_{i=1}^{n} (X_{a,i} - \overline{X}_{a})^{2} \sum_{j=1}^{n} (Y_{b,j} - \overline{Y}_{b})^{2} \right\}^{1/2}}$$
(3.8)

όπου X ο πίνακας που περιέχει τα MOS, DMOS, Y ο πίνακας με τις εκτιμώμενες βαθμολογίες που προέκυψαν από την προτεινόμενη μέθοδο, X_a μια στήλη στον πίνακα X και Y_b μια στήλη στον πίνακα Y, μέσες τιμές $\overline{X}_a = \sum_{i=1}^n (X_{a,i})/n$ και

 $\overline{Y}_{\!_{b}} = \sum_{j=1}^{n} \Bigl(X_{b,j}\Bigr) / n$ αντίστοιχα, και n το μέγεθος κάθε στήλης.

Για να μειώσουμε την μη γραμμικότητα στο μοντέλο παλινδρόμησης, εφαρμόζουμε, πριν το υπολογισμό της (3.7), μια λογιστική συνάρτηση πέντε παραμέτρων, που ορίζεται ως εξής:

$$s(p) = \beta_1 \cdot \left(\frac{1}{2} - \frac{1}{\exp(\beta_2 \cdot (p - \beta_3)) + 1}\right) + \beta_4 \cdot p + \beta_5$$
(3.9)

όπου s(p) είναι η προσαρμοσμένη (fitted) βαθμολογία ποιότητας, p είναι η αντιχειμενιχή βαθμολογία πρόβλεψης χαι β_i (i=1,2,3,4,5) ορίζουν τις παραμέτρους που πρέπει να προσαρμοστούν.

Οι τιμές του s(p) κυμαίνονται από το -1 στο +1. Όσο πιο κοντά στο +1 τόσο μεγαλύτερη η συσχέτιση της εκτιμώμενης βαθμολογίας με το ανθρώπινο κριτήριο, άρα τόσο καλύτερη και η απόδοση του αλγορίθμου.

Ομοίως ισχύει και για τις τιμές του συντελεστή Spearman, ο οποίος υπολογίζεται από τον τύπο:

$$rho = 1 - \frac{6 \cdot \sum_{i=1}^{v} \delta_i^2}{v \cdot (v^2 - 1)}$$
(3.10)

όπου δ η διαφορά των τάξεων των δυο στηλών.

Για λόγους συγκριτικής αξιολόγησης, η βασική μας αναφορά είναι η εργασία των Chenyang Shi και Yandan Lin [51], τα αποτελέσματα της οποίας εμφανίζονται στον Πίνακα 3.1 Η συγκεκριμένη εργασία αφορά αντίστοιχη εφαρμογή πρόβλεψης της ποιότητας σε έγχρωμες εικόνες, εκμεταλλευόμενη την πληροφορία που προκύπτει από υπομπάντες (subbands). Στο τέλος του κεφαλαίου θα γίνει σύγκριση της μεθόδου μας και με άλλες γνωστές μεθοδολογίες εκτίμησης αντιληπτής ποιότητας της εικόνας.

Vcgs-Αρχικό	Wn	Gblur	Jp2k	Jpeg	Fast fading	All
Pearson	0.9845	0.8206	0.824	0.716	0.7655	0.678
Spearman	0.9818	0.9688	0.965	-0.949	0.9623	0.947

Πίνακας 3.1 Συντελεστής Pearson και Spearman για την μέθοδο VCGS



Σχήμα 3.5 Οι Χάρτες που Προκύπτουν για την Μέθοδο VCGS. VSSImMatrix: Χάρτης Ομοιότητας μεταξύ των Δύο Εικόνων Βασισμένος στο Visual Saliency. Χρησιμοποιείται για τον Υπολογισμό του Τελικού Δείκτη. Weight: Πίνακας Στάθμισης Βασισμένος στην Σημείο-προς-Σημείο Μέγιστη Τιμή μεταξύ των Χαρτών Visual Saliency για τις Εικόνες που Συγκρίνονται. SimMatrixC: Τελικός Πίνακας Ομοιότητας πριν την Εφαρμοζόμενη Στάθμιση.

3.1 Πρώτη Ομάδα Πειραμάτων

Στην πρώτη ομάδα πειραμάτων, πραγματοποιούμε σε κάθε εικόνα μετασχηματισμό dwt2 ενός επιπέδου. Έπειτα προσαρμόζουμε την Student's t κατανομή στους οριζόντιους (cH), κάθετους (cV) και διαγώνιους συντελεστές (cD) λεπτομέρειας, που προκύπτουν από τον μετασχηματισμό, και με τη βοήθεια της ΕΜ μεθοοδολογίας εκτιμούμε τις παραμέτρους. Με αυτό τον τρόπο, αντικαθιστούμε τους συντελεστές λεπτομέρειας με την πληροφορία της βοηθητικής μεταβλητής που υπεισέρχεται στην ΕΜ επαναληπτική διαδικασία [30]. Στην συνέχεια, εφαρμόζουμε αντίστροφο μετασχηματισμό idwt2 με τα καινούρια cA', cV' και cD' που χρησιμοποιούν την βοηθητική μεταβλητή, ώστε να επιστρέψουμε στο χωρικό πεδίο και να υπολογίσουμε το νέο visual saliency.

Στο Final1, πραγματοποιούμε την παραπάνω διαδικασία χρησιμοποιώντας για τον αντίστροφο μετασχηματισμό την κρυμμένη πληροφορία των οριζόντιων, κάθετων και διαγώνιων συντελεστών. Επίσης, χρησιμοποιούμε τους συντελεστές προσέγγισης cA ως έχουν.

Στο Final2, χρησιμοποιούμε για τον αντίστροφο μετασχηματισμό την κρυμμένη πληροφορία των οριζόντιων, κάθετων και διαγώνιων συντελεστών, αλλά αυτή την φορά μηδενίζουμε τους συντελεστές προσέγγισης cA.

Στο Final3, δοχιμάσαμε να επιστρέψουμε την ειχόνα στο spatial domain χρησιμοποιώντας μόνο την χρυμμένη πληροφορία των διαγώνιων συντελεστών cD'. Οι χάθετοι cV, οριζόντιοι cH συντελεστές, χαθώς χαι οι συντελεστές προσέγγισης cA μηδενίζονται.

Στο Final4, χρησιμοποιούμε μόνο την πληροφορία των κάθετων συντελεστών cV' για τον αντίστροφο μετασχηματισμό κυματιδίων. Οπότε, μηδενίζουμε τους οριζόντιους και διαγώνιους συντελεστές και τους συντελεστές προσέγγισης.

Στο Final5, χρησιμοποιούμε μόνο την κρυμμένη πληροφορία των οριζόντιων συντελεστών cH', που σημαίνει ότι μηδενίζουμε τους διαγώνιους και κάθετους συντελεστές, αλλά και τους συντελεστές προσέγγισης.

Η απόδοση της πρώτης ομάδας πειραμάτων δίνεται από τους συντελεστές Pearson και Spearman, τα αποτελέσματα των οποίων παρουσιάζονται στον Πίνακα 3.2 και Πίνακα 3.3 αντίστοιχα.

PEARSON	WN	GBLUR	JP2K	JPEG	FF	ALL
VCGS	0.9814	0.9622	0.9738	0.9581	0.9490	0.9531
FINAL1	0.9783	0.9435	0.9584	0.9766	0.9471	0.9412
FINAL2	0.9784	0.9557	0.9605	0.9810	0.9514	0.9408
FINAL3	0.9674	0.9375	0.9437	0.9755	0.9573	0.9186
FINAL4	0.9672	0.9458	0.9519	0.9789	0.9560	0.9227
FINAL5	0.9682	0.9548	0.9566	0.9795	0.9601	0.9295

Πίναχας 3.2 Συντελεστής Pearson για τα Πειράματα της Πρώτης Ομάδας

Όσον αφορά τον συντελεστή Pearson, στο Final1, σημειώνεται αύξηση μόνο στην κατηγορία Jpeg σε σχέση με το πείραμα VCGS. Στα υπόλοιπα πειράματα έχουμε βελτιωμένες τιμές Pearson μόνο για τις κατηγορίες Jpeg και Fast Fading. Από τα πειράματα αυτής της ομάδας, αυτό που προσεγγίζει περισσότερο το αρχικό πείραμα VCGS είναι το πείραμα που χρησιμοποιεί την κρυμμένη πληροφορία όλων των συντελεστών λεπτομέρειας (Final2). Η μόνη κατηγορία, στην οποία δεν υπερτερεί έναντι των άλλων πειραμάτων της ομάδας, είναι η κατηγορία Fast Fading.

SPEARMAN	WN	GBLU R	JP2K	JPEG	FF	ALL
VCGS	0.9739	0.9488	0.9566	0.9742	0.9404	0.9564
FINAL1	0.9704	0.9305	0.9404	0.9583	0.9340	0.9424
FINAL2	0.9719	0.9453	0.9432	0.9683	0.9389	0.9404
FINAL3	0.9714	0.9340	0.9261	0.9566	0.9379	0.9169
FINAL4	0.9705	0.9350	0.9355	0.9591	0.9347	0.9211
FINAL5	0.9729	0.9458	0.9381	0,9630	0.9414	0.9290

Πίνακας 3.3 Συντελεστής Spearman για τα Πειράματα της Πρώτης Ομάδας

Όσον αφορά τον συντελεστή Spearman, τα πειράματα της πρώτης ομάδας παράγουν χειρότερα αποτελέσματα από το αρχικό πείραμα VCGS.



Σχήμα 3.6 VSMap Ref και VSMap Dis για το Πείραμα final2, όπου: VSMapRef: Visual Saliency Χάρτης για την Εικόνα Αναφοράς και VSMapDis: Visual Saliency Χάρτης για την Παραμορφωμένη Εικόνα



Σχήμα 3.7 Στην Πρώτη, Δεύτερη και Τρίτη Γραμμή Παρουσίαζονται το VSSimMatrix, το Weight και το SimMatrix του Πειράματος Final1, Final2 και Final5 αντίστοιχα, όπου: VSSImMatrix: Χάρτης Ομοιότητας μεταξύ των Δύο Εικόνων Βασισμένος στο Visual Saliency. Χρησιμοποιείται για τον Υπολογισμό του Τελικού Δείκτη. Weight: Πίνακας Στάθμισης Βασισμένος στην Σημείο-προς-Σημείο Μέγιστη Τιμή μεταξύ των Χαρτών Visual Saliency για τις Εικόνες που Συγκρίνονται. SimMatrixC: Τελικός Πίνακας Ομοιότητας πριν την Εφαρμοζόμενη Στάθμιση

3.2 Δεύτερη Ομάδα Πειραμάτων

Στην δεύτερη ομάδα πειραμάτων, επαναλαμβάνεται η διαδικασία της πρώτης ομάδας. Εφαρμόζουμε, δηλαδή, μετασχηματισμό dwt2 και κάνουμε fit την Student_t κατανομή στις οριζόντιες κάθετες και διαγώνιες λεπτομέρειες που προκύπτουν από τον μετασχηματισμό, ώστε να πάρουμε τις κρυμμένες μεταβλητές ai. Η μόνη διαφορά είναι, ότι στον αντίστροφο μετασχηματισμό idwt2 δεν χρησιμοποιούμε την κρυμμένη πληροφορία των εκάστοτε λεπτομερειών, αλλά υπολογίζουμε τον λογάριθμό της και χρησιμοποιούμε αυτή την τιμή για να επιστρέψουμε την εικόνα στο spatial πεδίο.

Στο Final6, στον αντίστροφο μετασχηματισμό χρησιμοποιούμε τον λογάριθμο της κρυμμένης πληροφορίας των οριζόντιων, κάθετων και διαγώνιων συντελεστών. Επίσης, χρησιμοποιούμε τους συντελεστές προσέγγισης cA ως έχουν.

Στο Final7, χρησιμοποιούμε για τον αντίστροφο μετασχηματισμό τον λογάριθμο της κρυμμένης πληροφορίας των οριζόντιων, κάθετων και διαγώνιων συντελεστών, αλλά αυτή την φορά μηδενίζουμε τους συντελεστές προσέγγισης cA.

Στο Final8, επιστρέφουμε την εικόνα στο χωρικό πεδίο χρησιμοποιώντας μόνο τον λογάριθμο της κρυμμένης πληροφορίας των διαγώνιων συντελεστών. Οι κάθετοι cV, οριζόντιοι cH συντελεστές, καθώς και οι συντελεστές προσέγγισης cA μηδενίζονται.

Στο Final9, χρησιμοποιούμε μόνο τον λογάριθμο της κρυμμένης πληροφορίας των κάθετων συντελεστών για τον αντίστροφο μετασχηματισμό κυματιδίων. Οπότε, μηδενίζουμε τους οριζόντιους και διαγώνιους συντελεστές και τους συντελεστές προσέγγισης.

Στο Final10, χρησιμοποιούμε μόνο τον λογάριθμο της κρυμμένης πληροφορίας των οριζόντιων συντελεστών, που σημαίνει ότι μηδενίζουμε τους διαγώνιους και κάθετους συντελεστές, αλλά και τους συντελεστές προσέγγισης

Τα αποτελέσματα των πειραμάτων της δεύτερης ομάδας, που προχύπτουν από τον συντελεστή Pearson, παρουσιάζονται στον Πίναχα 3.4 και από τον συντελεστή Spearman στον Πίναχα 3.5.

PEARSON	WN	GBLUR	JP2K	JPEG	\mathbf{FF}	ALL
VCGS	0.9814	0.9622	0.9738	0.9581	0.9490	0.9531
FINAL6	0.9806	0.9497	0.9610	0.9763	0.9473	0.9443
FINAL7	0.9740	0.9623	0.9789	0.9876	0.9670	0.9501
FINAL8	0.9737	0.9622	0.9748	0.9851	0.9678	0.9472
FINAL9	0.9742	0.9641	0.9807	0.9881	0.9711	0.9528
FINAL10	0.9762	0.9700	0.9765	0.9883	0.9748	0.9533

Πίναχας 3.4 Συντελεστής Pearson για τα Πειράματα της Δεύτερης Ομάδας

Το Final6 καταφέρνει μεγαλύτερη απόδοση από το VCGS μόνο στην κατηγορία Jpeg. Τα Final7, Final8, Final9 πετυχαίνουν μεγαλύτερες τιμές Pearson στις κατηγορίες Gaussian Blur, Jp2k, Jpeg και Fast Fading, με το πείραμα Final9, που χρησιμοποιεί την κρυμμένη πληροφορία των κάθετων συντελεστών, να ξεχωρίζει περισσότερο σε σχέση με τα άλλα δύο. Τέλος, το πείραμα Final10, που χρησιμοποιεί την κρυμμένη πληροφορία των οριζόντιων συντελεστών, πετυχαίνει μεγαλύτερη απόδοση από το αρχικό VCGS σε όλες τις κατηγορίες εκτός από την κατηγορία White Noise. Επίσης, αποτελεί το πείραμα που πετυχαίνει τις υψηλότερες τιμές Pearson στην πλειοψηφία των κατηγοριών. Στην κατηγορία White Noise, κανένα πείραμα δεν καταφέρνει να ξεπεράσει την απόδοση που πετυχαίνει το VCGS, αν και το πείραμα Final6 προσεγγίζει πολύ την τιμή.

SPEARMAN	WN	GBLU R	JP2K	JPEG	FF	ALL
VCGS	0.9739	0.9488	0.9566	0.9742	0.9404	0.9564
FINAL6	0.9734	0.9355	0.9426	0.9588	0.9350	0.9454
FINAL7	0.9670	0.9532	0.9646	0.9796	0.9552	0.9528
FINAL8	0.9670	0.9532	0.9567	0.9734	0.9520	0.9474
FINAL9	0.9650	0.9527	$0.9\overline{658}$	0.9777	0.9542	$0.9\overline{544}$
FINAL10	0.9685	0.9601	0.9597	0.9783	0.9631	0.9562

Πίνακας 3.5 Συντελεστής Spearman για τα Πειράματα της Δεύτερης Ομάδας

Το Final6, που χρησιμοποιεί τους συντελεστές προσέγγισης και την κρυμμένη πληροφορία των συντελεστών λεπτομέρειας για τον αντίστροφο μετασχηματισμό κυματιδίων, δεν καταφέρνει να αυξήσει την απόδοση συγκριτικά με την μέθοδο VCGS. Ο μηδενισμός των συντελεστών προσέγγισης, στο πείραμα Final7, βοηθάει στην αύξηση της απόδοσης, καθώς πλέον πετυχαίνει μεγαλύτερη απόδοση από το VCGS στις κατηγορίες Gaussian Blur, Jp2k, Jpeg και Fast Fading. Το Final8 πετυχαίνει μεγαλύτερες τιμές Spearman στις κατηγορίες Gaussian Blur, Jp2k και Fast fading. Το Final9 δεν ξεπερνάει σε απόδοση το VCGS στις κατηγορίες White noise και all. Πετυχαίνει, όμως, την καλύτερη απόδοση στην κατηγορία Jp2k. Τέλος, το Final10, όπως και με τον συντελεστή Pearson έτσι και με τον Spearman,
εμφανίζει τις καλύτερες τιμές στην πλειοψηφία των κατηγοριών σε σύγκριση με όλα τα προηγούμενα πειράματα, συμπεριλαμβανομένου και του VCGS. Η κατηγορία στην οποία είναι πιο αδύναμο, σε σχέση με το VCGS, είναι για μια ακόμα φορά η White noise. Το πείραμα που καταφέρνει να προσεγγίσει την υψηλή απόδοση της VCGS στο White noise είναι το Final 6, στο οποίο χρησιμοποιούνται οι συντελεστές προσέγγισης και η κρυμμένη πληροφορία των συντελεστών λεπτομέρειας.



Σχήμα 3.8 VSMap Ref και VSMap Dis για το Πείραμα final10, όπου: VSMapRef: Visual Saliency Χάρτης για την Εικόνα Αναφοράς και VSMapDis: Visual Saliency Χάρτης για την Παραμορφωμένη Εικόνα



Σχήμα 3.9 Στην Πρώτη, Δεύτερη και Τρίτη Γραμμή Παρουσίαζονται το VSSimMatrix, το Weight και το SimMatrix του Πειράματος Final6, Final7 και Final10 αντίστοιχα, όπου: VSSImMatrix: Χάρτης Ομοιότητας μεταξύ των Δύο Εικόνων Βασισμένος στο Visual Saliency. Χρησιμοποιείται για τον Υπολογισμό του Τελικού Δείκτη. Weight: Πίνακας Στάθμισης Βασισμένος στην Σημείο-προς-Σημείο Μέγιστη Τιμή μεταξύ των Χαρτών Visual Saliency για τις Εικόνες που Συγκρίνονται. SimMatrixC: Τελικός Πίνακας Ομοιότητας πριν την Εφαρμοζόμενη Στάθμιση

3.3 Τρίτη Ομάδα Πειραμάτων

Η διαδικασία παραμένει η ίδια με τις προηγούμενες ομάδες πειραμάτων. Αυτή τη φορά, στον αντίστροφο μετασχηματισμό idwt2 δεν χρησιμοποιούμε την κρυμμένη πληροφορία των εκάστοτε λεπτομερειών, ούτε τον λογάριθμό της. Υπολογίζουμε τον λογάριθμο της αντίστροφης κρυμμένης πληροφορίας των εκάστοτε λεπτομερειών και χρησιμοποιούμε αυτή την τιμή για να επιστρέψουμε την εικόνα στο χωρικό πεδίο.

Στο final11, για τον αντίστροφο μετασχηματισμό χρησιμοποιούμε τον λογάριθμο της αντίστροφης τιμής της κρυμμένης πληροφορίας των οριζόντιων, κάθετων και διαγώνιων συντελεστών. Επίσης χρησιμοποιούμε τους συντελεστές προσέγγισης ως έχουν.

Στο final12, χρησιμοποιούμε για τον αντίστροφο μετασχηματισμό τον λογάριθμο της αντίστροφης τιμής της κρυμμένης πληροφορίας των οριζόντιων, κάθετων και διαγώνιων συντελεστών, αλλά αυτή την φορά μηδενίζουμε τους συντελεστές προσέγγισης cA.

Στο final13, επιστρέφουμε την εικόνα στο χωρικό πεδίο χρησιμοποιώντας μόνο τον λογάριθμο της αντίστροφης τιμής της κρυμμένης πληροφορίας των διαγώνιων συντελεστών. Οι κάθετοι, οριζόντιοι συντελεστές, καθώς και οι συντελεστές προσέγγισης μηδενίζονται.

Στο final14, χρησιμοποιούμε μόνο τον λογάριθμο της αντίστροφης τιμής της κρυμμένης πληροφορίας των κάθετων συντελεστών για τον αντίστροφο μετασχηματισμό κυματιδίων. Οπότε, μηδενίζουμε τους οριζόντιους και διαγώνιους συντελεστές και τους συντελεστές προσέγγισης.

Στο final15, χρησιμοποιούμε μόνο τον λογάριθμο της αντίστροφης τιμής της κρυμμένης πληροφορίας των οριζόντιων συντελεστών, που σημαίνει ότι μηδενίζουμε τους διαγώνιους και κάθετους συντελεστές, αλλά και τους συντελεστές προσέγγισης.

Τα αποτελέσματα των πειραμάτων της τρίτης ομάδας, που προχύπτουν από τον συντελεστή Pearson, παρουσιάζονται στον Πίναχα 3.6 και από τον συντελεστή Spearman στον Πίναχα 3.7.

PEARSON	WN	GBLUR	JP2K	JPEG	FF	ALL
VCGS	0.9814	0.9622	0.9738	0.9581	0.9490	0.9531
FINAL11	0.9797	0.9485	0.9608	0.9749	0.9490	0.9430
FINAL12	0.9736	0.9621	0.9791	0.9880	0.9667	0.9503
FINAL13	0.9737	0.9614	0.9739	0.9854	0.9684	0.9471
FINAL14	0.9742	0.9642	0.9811	0.9881	0.9721	0.9524
FINAL15	0.9766	0.9703	0.9765	0.9880	0.9759	0.9528

Πίνακας 3.6 Συντελεστής Pearson για τα Πειράματα της Τρίτης Ομάδας

SPEARMAN	WN	GBLUR	JP2K	JPEG	FF	ALL
VCGS	0.9739	0.9488	0.9566	0.9742	0.9404	0.9564
FINAL11	0.9734	0.9352	0.9422	0.9544	0.9360	0.9440
FINAL12	0.9680	0.9547	0.9649	0.9778	0.9537	0.9532
FINAL13	0.9655	0.9527	0.9578	0.9738	0.9517	0.9478
FINAL14	0.9655	0.9532	0.9661	0.9782	0.9554	0.9550
FINAL15	0.9692	0.9596	0.9597	0.9779	0.9631	0.9549

Πίναχας 3.7 Συντελεστής Spearman για τα Πειράματα της Τρίτης Ομάδας

Τα πειράματα Final12, Final14 και Final15 στον Πίνακα 3.6 πετυχαίνουν μεγαλύτερη απόδοση από το VCGS στην κατηγορία Jpeg με σχεδόν ίδια τιμή. Επιπλέον, το Final15, που χρησιμοποιεί την κρυμμένη πληροφορία των οριζόντιων συντελεστών, εκτός από την κατηγορία Jpeg καταφέρνει μεγαλύτερη απόδοση από το VCGS, αλλά και από τα άλλα πειράματα της ομάδας, και στις κατηγορίες Gaussian Blur xai Fast Fading. To Final14, που χρησιμοποιεί την κρυμμένη πληροφορία των χάθετων συντελεστών, ξεχωρίζει από τα υπόλοιπα πειράματα της ομάδας, πλην του Final15, καθώς εκτός από την κατηγορία Jpeg καταφέρνει μεγαλύτερη απόδοση από το VCGS και στην κατηγορία Jp2k, στην οποία πετυχαίνει την μεγαλύτερη τιμή από όλα τα πειράματα της ομάδας. Το πείραμα Final13 έχει καλύτερη απόδοση από το VCGS στις κατηγορίες Jp2k, Jpeg και Fast Fading. Στην κατηγορία White Noise, κανένα πείραμα δεν καταφέρνει να ξεπεράσει την απόδοση που πετυχαίνει το VCGS, αν και το πείραμα Final11 προσεγγίζει πολύ την τιμή. Συγκεντρωτικά, το πείραμα της ομάδας που ξεχωρίζει είναι το Final15, καθώς και έχει την καλύτερη απόδοση και από τα υπόλοιπα πειράματα της ομάδας, αλλά και από το VCGS στις κατηγορίες Gaussian Blur, Jpeg, και Fast Fading, και προσεγγίζει, χωρίς ωστόσο να προσπερνάει, το VCGS στην κατηγορία All.

Το Final11 στο Πίναχα 3.7, που χρησιμοποιεί τους συντελεστές προσέγγισης χαι την χρυμμένη πληροφορία των συντελεστών λεπτομέρειας για τον αντίστροφο μετασχηματισμό χυματιδίων, δεν χαταφέρνει να αυξήσει την απόδοση συγχριτιχά με την μέθοδο VCGS. Ο μηδενισμός των συντελεστών προσέγγισης, στο πείραμα Final12, βοηθάει στην αύξηση της απόδοσης, χαθώς πλέον πετυχαίνει μεγαλύτερη απόδοση από το VCGS στις κατηγορίες Gaussian Blur, Jp2k, Jpeg και Fast Fading. Το Final13 πετυχαίνει μεγαλύτερες τιμές Spearman στις κατηγορίες Gaussian Blur, Jp2k και Fast fading. Το Final14 δεν ξεπερνάει σε απόδοση το VCGS μόνο στην κατηγορία White noise, πετυχαίνει, όμως, την καλύτερη απόδοση στις κατηγορίες Jp2k και Jpeg. Τέλος, το Final15 εμφανίζει τις καλύτερες τιμές στις κατηγορίες Gaussian Blur και Fast Fading σε σύγκριση με όλα τα προηγούμενα πειράματα, συμπεριλαμβανομένου και του VCGS, ενώ στις κατηγορίες Jp2k και Jpeg πετυχαίνει καλύτερη απόδοση από το VCGS. Η κατηγορία στην οποία υπερτερεί το VCGS είναι για μια αχόμα φορά η White noise.



Σχήμα 3.10 VSMap Ref και VSMap Dis για το Πείραμα final15, όπου: VSMapRef: Visual Saliency Χάρτης για την Εικόνα Αναφοράς και VSMapDis: Visual Saliency Χάρτης για την Παραμορφωμένη Εικόνα



Σχήμα 3.11 Στην Πρώτη, Δεύτερη και Τρίτη Γραμμή Παρουσίαζονται το VSSimMatrix, το Weight και το SimMatrix του Πειράματος Final11, Final12 και Final15 αντίστοιχα, όπου: VSSImMatrix: Χάρτης Ομοιότητας μεταξύ των Δύο Εικόνων Βασισμένος στο Visual Saliency. Χρησιμοποιείται για τον Υπολογισμό του Τελικού Δείκτη. Weight: Πίνακας Στάθμισης Βασισμένος στην Σημείο-προς-Σημείο Μέγιστη Τιμή μεταξύ των Χαρτών Visual Saliency για τις Εικόνες που Συγκρίνονται. SimMatrixC: Τελικός Πίνακας Ομοιότητας πριν την Εφαρμοζόμενη Στάθμιση

3.4 Τέταρτη Ομάδα Πειραμάτων

Η διαδικασία παραμένει η ίδια με τις προηγούμενες ομάδες πειραμάτων. Για τον αντίστροφο μετασχηματισμό κυματιδίων, υπολογίζουμε την αντίστροφη τιμή της κρυμμένης πληροφορίας των εκάστοτε λεπτομερειών.

Στο Final16, πραγματοποιούμε την γνωστή διαδικασία και χρησιμοποιούμε για τον αντίστροφο μετασχηματισμό την αντίστροφη τιμή της κρυμμένης πληροφορίας των οριζόντιων, κάθετων και διαγώνιων συντελεστών. Χρησιμοποιούμε τους συντελεστές προσέγγισης cA ως έχουν.

Στο Final17, χρησιμοποιούμε για τον αντίστροφο μετασχηματισμό την αντίστροφη τιμή της κρυμμένης πληροφορίας των οριζόντιων, κάθετων και διαγώνιων συντελεστών, αλλά αυτή την φορά μηδενίζουμε τους συντελεστές προσέγγισης.

Στο Final18, επιστρέφουμε την εικόνα στο χωρικό πεδίο χρησιμοποιώντας μόνο την αντίστροφη τιμή της κρυμμένης πληροφορίας των διαγώνιων συντελεστών. Οι κάθετοι, οριζόντιοι συντελεστές, καθώς και οι συντελεστές προσέγγισης μηδενίζονται.

Στο πείραμα Final19, χρησιμοποιούμε μόνο την αντίστροφη τιμή της κρυμμένης πληροφορίας των κάθετων συντελεστών για τον αντίστροφο μετασχηματισμό κυματιδίων. Οπότε, μηδενίζουμε τους οριζόντιους και διαγώνιους συντελεστές και τους συντελεστές προσέγγισης.

Στο πείραμα Final20, χρησιμοποιούμε μόνο την αντίστροφη τιμή της κρυμμένης πληροφορίας των οριζόντιων συντελεστών, που σημαίνει ότι μηδενίζουμε τους διαγώνιους και κάθετους συντελεστές, αλλά και τους συντελεστές προσέγγισης.

Τα αποτελέσματα των πειραμάτων της τέταρτης ομάδας, που προχύπτουν από τον συντελεστή Pearson, παρουσιάζονται στον Πίναχα 3.8 και από τον συντελεστή Spearman στον Πίναχα 3.9.

PEARSON	WN	GBLUR	JP2K	JPEG	FF	ALL
VCGS	0.9814	0.9622	0.9738	0.9581	0.9490	0.9531
FINAL16	0.8812	0.9074	0.9223	0.9522	0.9350	0.8626
FINAL17	0.9763	0.9294	0.9774	0.9872	0.9637	0.9369
FINAL18	0.9762	0.9577	0.9775	0.9882	0.9621	0.9533
FINAL19	0.9763	0.9260	0.9788	0.9889	0.9627	0.9361
FINAL20	0.9754	0.9322	0.9745	0.9866	0.9656	0.9348

Πίναχας 3.8 Συντελεστής Pearson για τα Πειράματα της Τέταρτης Ομάδας

SPEARMAN	WN	GBLU R	JP2K	JPEG	FF	ALL
VCGS	0.9739	0.9488	0.9566	0.9742	0.9404	0.9564
FINAL16	0.8350	0.8953	0.9011	0.9345	0.9207	0.8551
FINAL17	0.9677	0.9271	0.9612	0.9757	0.9507	0.9395
FINAL18	0.9660	0.9463	0.9633	0.9766	0.9478	0.9552
FINAL19	0.9670	0.9227	0.9639	0.9795	0.9468	0.9396
FINAL20	0.9660	0.9310	0.9581	0.9761	0.9532	0.9384

Πίνακας 3.9 Συντελεστής Spearman για τα Πειράματα της τέταρτης ομάδας

Με αυτή την μορφή της κρυμμένης πληροφορίας, δεν πετυχαίνουμε τόσο καλά αποτελέσματα όσο σε άλλες ομάδες πειραμάτων. Για μια ακόμα φορά, αποδεικνύεται ότι η χρήση των συντελεστών προσέγγισης δεν προσδίδει στις μεθόδους μας, καθώς ο μηδενισμός τους οδηγεί πάντα σε καλύτερα αποτελέσματα. Με τα πειράματα Final17, Final18, Final19, Final 20 έχουμε μεγαλύτερες τιμές Pearson στις κατηγορίες Jp2k, Jpeg και FF, σε σχέση με την VCGS μέθοδο. Επίσης, στο πείραμα που χρησιμοποιούμε την κρυμμένη πληροφορία των διαγώνιων συντελεστών (Final 18), έχουμε και καλύτερη απόδοση στην κατηγορία All.

Παρομοίως, για τον συντελεστή Spearman, όταν δεν χρησιμοποιούμε τους συντελεστές προσέγγισης, έχουμε καλύτερα αποτελέσματα. Με το Final17, Final20, πετυχαίνουμε καλύτερη απόδοση στις κατηγορίες Jp2k, Jpeg και Fast Fading. Όταν χρησιμοποιούμε την κρυμμένη πληροφορία των διαγώνιων συντελεστών, προσεγγίζουμε σε μεγάλο βαθμό και την τιμή Spearman της κατηγορίας All που πετυχαίνει το VCGS.



Σχήμα 3.12 VSMap Ref και VSMap Dis για το Πείραμα final19, όπου: VSMapRef: Visual Saliency Χάρτης για την Εικόνα Αναφοράς και VSMapDis: Visual Saliency Χάρτης για την Παραμορφωμένη Εικόνα



Σχήμα 3.13 Στην Πρώτη, Δεύτερη και Τρίτη Γραμμή Παρουσίαζονται το VSSimMatrix, το Weight και το SimMatrix του Πειράματος Final16, Final17 και Final20 αντίστοιχα, όπου: VSSImMatrix: Χάρτης Ομοιότητας μεταξύ των Δύο Εικόνων Βασισμένος στο Visual Saliency. Χρησιμοποιείται για τον Υπολογισμό του Τελικού Δείκτη. Weight: Πίνακας Στάθμισης Βασισμένος στην Σημείο-προς-Σημείο Μέγιστη Τιμή μεταξύ των Χαρτών Visual Saliency για τις Εικόνες που Συγκρίνονται. SimMatrixC: Τελικός Πίνακας Ομοιότητας πριν την Εφαρμοζόμενη Στάθμιση

3.5 Πέμπτη Ομάδα Πειραμάτων

 Σ ' αυτή την ομάδα, έγινε επιλογή δύο πειραμάτων από αυτά που έδωσαν καλύτερα αποτελέσματα στις παραπάνω ομάδες και γίνεται μετασχηματισμός κυματιδίων δύο επιπέδων. Έπειτα, εφαρμόζουμε την Student's t κατανομή στους οριζόντιους, κάθετους και διαγώνιους συντελεστές λεπτομέρειας κάθε επιπέδου, που προχύπτουν από τον αντίστοιχο μετασχηματισμό, χαι παίρνουμε τις αντίστοιχες ai. Στην συνέχεια, κρυμμένες μεταβλητές εφαρμόζουμε αντίστροφο μετασχηματισμό idwt2, χρησιμοποιώντας τις χρυμμένες πληροφορίες ai που προέχυψαν από τους συντελεστές λεπτομέρειας δευτέρου επίπεδου. Το αποτέλεσμα, που προχύπτει από αυτή την διαδιχασία, χρησιμοποιείται μαζί με τις χρυμμένες πληροφορίες ai των συντελεστών λεπτομέρειας πρώτου επιπέδου για να υπολογίσουμε τον αντίστροφο μετασχηματισμό idwt2, ώστε να γίνει επιστροφή της εικόνας στο χωρικό πεδίο.

Για το FinalA, για να υπολογίσουμε τον αντίστροφο μετασχηματισμό στο δεύτερο επίπεδο, χρησιμοποιούμε μόνο τον λογάριθμο της χρυμμένης πληροφορίας των οριζόντιων συντελεστών του δεύτερου επιπέδου, ενώ οι κάθετοι, διαγώνιοι και οι προσεγγιστικοί συντελεστές του δεύτερου επιπέδου μηδενίζονται. Τέλος, για να επιστρέψουμε την εικόνα στο χωρικό πεδίο, μηδενίζουμε την κάθετη και την διαγώνια πληροφορία στο πρώτο επίπεδο και χρησιμοποιούμε τον λογάριθμο της κρυμμένης πληροφορίας των οριζόντιων συντελεστών του πρώτου επιπέδου, καθώς και το αποτέλεσμα του αντίστροφου μετασχηματισμού στο δεύτερο επίπεδο, στην θέση του συντελεστή προσέγγισης του πρώτου επιπέδου.

Τα αποτελέσματα του συντελεστή Pearson για το πείραμα Final Α παρουσιάζονται στον Πίναχα 3.10 μαζί με το πείραμα Final 10 χαι την μέθοδο VCGS, ώστε να γίνει σύγχριση της απόδοσης των μεθόδων. Ομοίως χαι για τον συντελεστή Spearman, στον Πίναχα 3.11.

PEARSON	WN	GBLUR	JP2K	JPEG	FF	ALL
VCGS	0.9814	0.9622	0.9738	0.9581	0.9490	0.9531
FINAL10	0.9762	0.9700	0.9765	0.9883	0.9748	0.9533
FINAL_A	0.9746	0.9729	0.9768	0.9877	0.9739	0.9522

Πίναχας 3.10 Συντελεστής Pearson για το Πείραμα FinalA

SPEARMAN	WN	GBLUR	JP2K	JPEG	FF	ALL
VCGS	0.9739	0.9488	0.9566	0.9742	0.9404	0.9564
FINAL10	0.9685	0.9601	0.9597	0.9783	0.9631	0.9562
FINAL_A	0.9665	0.9606	0.9599	0.9773	0.9611	0.9544

Πίναχας 3.11 Συντελεστής Spearman για το Πείραμα FinalA

Στο Final B, υλοποιείται η ίδια διαδικασία με το πείραμα Final A με την διαφορά ότι, αντί για τον λογάριθμο της κρυμμένης πληροφορίας των οριζόντιων συντελεστών, χρησιμοποιείται ο λογάριθμος της αντίστροφης τιμής της κρυμμένης πληροφορίας τους.

Τα αποτελέσματα του συντελεστή Pearson για το πείραμα Final B παρουσιάζονται στον Πίναχα 3.12 μαζί με το πείραμα Final 10 χαι την μέθοδο VCGS, ώστε να γίνει σύγχριση της απόδοσης των μεθόδων. Ομοίως χαι για τον συντελεστή Spearman στον Πίναχα 3.13.

Πίναχας 3.12 Συντελεστής Pearson για το Πείραμα FinalB

PEARSON	WN	GBLUR	JP2K	JPEG	FF	ALL
VCGS	0.9814	0.9622	0.9738	0.9581	0.9490	0.9531
FINAL15	0.9766	0.9703	0.9765	0.9880	0.9759	0.9528
FINAL_B	0.9754	0.9724	0.9768	0.9872	0.9747	0.9522

Πίναχας 3.13 Συντελεστής Spearman για το Πείραμα FinalB

SPEARMAN	WN	GBLUR	JP2K	JPEG	FF	ALL
VCGS	0.9739	0.9488	0.9566	0.9742	0.9404	0.9564
FINAL15	0.9692	0.9596	0.9597	0.9779	0.9631	0.9549
FINAL_B	0.9680	0.9606	0.9600	0.9778	0.9616	0.9548

Τόσο για τον συντελεστή Pearson όσο και για τον συντελεστή Spearman, του πειράματος Final A, ισχύουν τα εξής: Με την χρήση δύο επιπέδων μετασχηματισμού wavelet έχουμε αύξηση των συντελεστών μόνο στην κατηγορία Gaussian Blur και Jp2k, ενώ στις υπόλοιπες παρατηρείται μείωση. Παρόλη την μείωση, οι συντελεστές στις κατηγορίες Gaussian Blur, Jp2k, Jpeg και Fast Fading διατηρούνται σε αρκετά υψηλό επίπεδο, ώστε να ξεπερνάνε ακόμα την VCGS μέθοδο.

Το ίδιο ισχύει και όταν χρησιμοποιούμε τον λογάριθμο της αντίστροφης τιμής της κρυμμένης πληροφορίας των οριζόντιων συντελεστών στο πείραμα Final B.



Σχήμα 3.14 VSMap Ref και VSMap Dis για το Πείραμα finalB, όπου: VSMapRef: Visual Saliency Χάρτης για την Εικόνα Αναφοράς και VSMapDis: Visual Saliency Χάρτης για την Παραμορφωμένη Εικόνα



Σχήμα 3.15 Οι Χάρτες που Προκύπτουν για το Πειραμα finalB. όπου VSSImMatrix: Χάρτης Ομοιότητας μεταξύ των Δύο Εικόνων Βασισμένος στο Visual Saliency. Χρησιμοποιείται για τον Υπολογισμό του Τελικού Δείκτη. Weight: Πίνακας Στάθμισης Βασισμένος στην Σημείο-προς-Σημείο Μέγιστη Τιμή μεταξύ των Χαρτών Visual Saliency για τις Εικόνες που Συγκρίνονται. SimMatrixC: Τελικός Πίνακας Ομοιότητας πριν την Εφαρμοζόμενη Στάθμιση

3.6 Έκτη Ομάδα Πειραμάτων

Σε αυτήν την ομάδα επιλέχθηκαν δύο από τα καλύτερα περάματα των παραπάνω ομάδων και γίνεται μετασγηματισμός κυματιδίων τριών επιπέδων. Έπειτα, για κάθε ένα επίπεδο υπολογίζουμε την Student's t κατανομή των οριζόντιων κάθετων διαγώνιων συντελεστών λεπτομέρειας, που προχύπτουν από και τον μετασχηματισμό χάθε επιπέδου, για να πάρουμε τις αντίστοιχες χρυμμένες πληροφορίες ai. Στην συνέχεια, εφαρμόζουμε αντίστροφο μετασχηματισμό idwt2, χρησιμοποιώντας τις κρυμμένες πληροφορίες ai που προέκυψαν από τους συντελεστές λεπτομέρειας τρίτου επίπεδου. Αχολουθεί ένας αχόμα αντίστροφος μετασχηματισμός idwt2, όπου αυτήν την φορά χρησιμοποιούνται οι κρυμμένες πληροφορίες ai που προέχυψαν από τους συντελεστές λεπτομέρειας δευτέρου επίπεδου, αλλά στην θέση του συντελεστή προσέγγισης δευτέρου επιπέδου χρησιμοποιείται το αποτέλεσμα του αντίστροφο μετασχηματισμό idwt2 του τρίτου επιπέδου. Το αποτέλεσμα, που προχύπτει από αυτή την διαδιχασία, χρησιμοποιείται μαζί με τις κρυμμένες πληροφορίες ai των συντελεστών λεπτομέρειας πρώτου επιπέδου, για να υπολογίσουμε τον αντίστροφο μετασχηματισμό idwt2, ώστε να γίνει επιστροφή της εικόνας στο χωρικό πεδίο.

Στο Final i, για να υπολογίσουμε τον αντίστροφο μετασχηματισμό στο τρίτο επίπεδο, χρησιμοποιούμε μόνο τον λογάριθμο της κρυμμένης πληροφορίας των οριζόντιων συντελεστών του δεύτερου επιπέδου, ενώ οι κάθετοι, διαγώνιοι και οι προσεγγιστικοί συντελεστές του τρίτου επιπέδου μηδενίζονται. Έπειτα, για να υπολογίσουμε τον αντίστροφο μετασχηματισμό στο δεύτερο επίπεδο, μηδενίζουμε τους κάθετους και την διαγώνιους συντελεστές στο δεύτερο επίπεδο και χρησιμοποιούμε τον λογάριθμο της κρυμμένης πληροφορίας των οριζόντιων συντελεστών του δευτέρου επιπέδου, καθώς στο δεύτερο επίπεδο και χρησιμοποιούμε τον λογάριθμο της κρυμμένης πληροφορίας των οριζόντιων συντελεστών του δευτέρου επιπέδου, καθώς και το αποτέλεσμα του αντίστροφου μετασχηματισμού στο τρίτο επίπεδο, στην θέση του συντελεστή προσέγγισης του δευτέρου επιπέδου, μηδενίζουμε τον χωρικό πεδίο, χρησιμοποιούμε τον λογάριθμο της κρυμμένης πληροφορίας των οριζόντιων συντελεστών του δευτέρου επιπέδου, καθώς και το αποτέλεσμα του αντίστροφου μετασχηματισμού στο τρίτο επίπεδο, στην θέση του συντελεστή προσέγγισης του διαγώνιους συντελεστών του δευτέρου επιπέδου, μηδενίζουμε τοις κάθετους και του αντίστροφου μετασχηματισμού στο τρίτο επίπεδο, στην θέση του συντελεστή προσέγγισης του διαγώνιους συντελεστώς των οριζόντιων συντελεστών του πρώτου επιπέδου, μηδενίζουμε τους κάθετους και τους λογάριθμο της κρυμμένης πληροφορίας των οριζόντιων συντελεστών του πρώτου επιπέδου, μηδενίζουμε του αντίστροφομαι του αντίστροφομαι του αντισιούμε τον λογάριθμο της κρυμμένης πληροφορίας των οριζώντων συντελεστών του πρώτου επιπέδου, μηδενίζουμε τους κάθετους και τους διαγώνιους συντελεστές στο πρώτο επίπεδο και στην θέση του συντελεστή προσέργισης του αντελεστή προσέργισης του αντελεστή προσέργισης του αντελεστά του αντιστοι επίπεδο και στην θέση του συντελεστή προσέγρισης του αντίστου επίπεδου χρησιμοποιούμε το αποτέλεσμα του αντίστροφορίας των αντίστου επίπεδου χρησιμοποιούμε το αποτέλεσμα του αντίστου επίπεδου και στην θέση του συντελεστή προσέρου μετασχηματισμού στο δεύτερο επίπεδο.

63

Τα αποτελέσματα του συντελεστή Pearson για το πείραμα Final i παρουσιάζονται στον Πίναχα 3.14 μαζί με το πείραμα Final 10 χαι την μέθοδο VCGS, ώστε να γίνει σύγχριση της απόδοσης των μεθόδων. Ομοίως χαι για τον συντελεστή Spearman στον Πίναχα 3.15.

PEARSON	WN	GBLUR	JP2K	JPEG	\mathbf{FF}	ALL
VCGS	0.9814	0.9622	0.9738	0.9581	0.9490	0.9531
FINAL10	0.9762	0.9700	0.9765	0.9883	0.9748	0.9533
FINAL_I	0.9742	0.9716	0.9764	0.9876	0.9738	0.9523

Πίναχας 3.14 Συντελεστής Pearson για το Πείραμα Final i

Πίναχας 3.15 Συντελεστής Spearman για το Πείραμα Final i

SPEARMAN	WN	GBLUR	JP2K	JPEG	FF	ALL
VCGS	0.9739	0.9488	0.9566	0.9742	0.9404	0.9564
FINAL10	0.9685	0.9601	0.9597	0.9783	0.9631	0.9562
FINAL_I	0.9640	0.9586	0.9596	0.9779	0.9594	0.9550

Στο Final ii υλοποιείται η ίδια διαδικασία με το πείραμα Final i με την διαφορά ότι, αντί για τον λογάριθμο της κρυμμένης πληροφορίας των οριζόντιων συντελεστών, χρησιμοποιείται ο λογάριθμος της αντίστροφης τιμής της κρυμμένης πληροφορίας τους.

Τα αποτελέσματα του συντελεστή Pearson για το πείραμα Final ii παρουσιάζονται στον Πίναχα 3.16 μαζί με το πείραμα Final 10 χαι την μέθοδο VCGS, ώστε να γίνει σύγχριση της απόδοσης των μεθόδων. Ομοίως χαι για τον συντελεστή Spearman στον Πίναχα 3.17.

Πίνακας 3.16 Συντελεστής Pearson για το Πείραμα Final ii

PEARSON	WN	GBLUR	JP2K	JPEG	FF	ALL
VCGS	0.9814	0.9622	0.9738	0.9581	0.9490	0.9531
FINAL15	0.9766	0.9703	0.9765	0.9880	0.9759	0.9528
FINAL_II	0.9739	0.9725	0.9772	0.9875	0.9725	0.9522

Πίναχας 3.17 Συντελεστής Spearman για το Πείραμα Final ii

SPEARMAN	WN	GBLUR	JP2K	JPEG	${ m FF}$	ALL
VCGS	0.9739	0.9488	0.9566	0.9742	0.9404	0.9564
FINAL15	0.9692	0.9596	0.9597	0.9779	0.9631	0.9549
FINAL_II	0.9653	0.9601	0.9606	0.9775	0.9596	0.9547

Το τρίτο επίπεδο επιφέρει περαιτέρω μείωση των δύο συντελεστών, σε σχέση με τα αντίστοιχα πειράματα της 5ης ομάδας στα οποία χρησιμοποιούνταν δύο επίπεδα.



Σχήμα 3.16 VSMap Ref και VSMap Dis για το Πείραμα final i, όπου: VSMapRef: Visual Saliency Χάρτης για την Εικόνα Αναφοράς και VSMapDis: Visual Saliency Χάρτης για την Παραμορφωμένη Εικόνα.



Σχήμα 3.17 Οι Χάρτες που Προχύπτουν για το Πείραμα final i. όπου VSSImMatrix: Χάρτης Ομοιότητας μεταξύ των Δύο Ειχόνων Βασισμένος στο Visual Saliency. Χρησιμοποιείται για τον Υπολογισμό του Τελιχού Δείχτη. Weight: Πίναχας Στάθμισης Βασισμένος στην Σημείο-προς-Σημείο Μέγιστη Τιμή μεταξύ των Χαρτών Visual Saliency για τις Ειχόνες που Συγχρίνονται. SimMatrixC: Τελιχός Πίναχας Ομοιότητας πριν την Εφαρμοζόμενη Στάθμιση

3.7 Έβδομη Ομάδα Πειραμάτων

Σε αυτήν την ομάδα πειραμάτων εφαρμόζουμε έναν νέο τρόπο υπολογισμού του gradient map, μια "ψεύδο-gradient". Αρχικά, χρησιμοποιούμε μετασχηματισμό κυματιδίων ενός επιπέδου στο L χρωματικό κανάλι της αρχικής και της παραμορφωμένης εικόνας, αφού βρίσκονται στον Lab χρωματικό χώρο. Έτσι, υπολογίζονται οι οριζόντιοι (cH), κάθετοι (cV) και διαγώνιοι συντελεστές (cD) λεπτομέρειας για κάθε μια εικόνα. Έπειτα, για τον υπολογισμό της ψεύδο-gradient κάθε εικόνας, χρησιμοποιούμε την πληροφορία του πλάτους (magnitude) mag= $\sqrt{(H^2+V^2)}$, με H, V να είναι οι οριζόντιοι και οι κάθετοι συντελεστές λεπτομέρειας, αντίστοιχα. Οι δύο αυτές "ψεύδο-gradient" χρησιμοποιούνται για να παράγουν έναν νέο gradient similarity matrix, ο οποίος χρησιμοποιείται στον υπολογισμό του VCGS, είτε αντικαθιστώντας το παλιό gradient similarity matrix.

Στο VCGS_ADD-GRADIENT, ο υπολογισμός των visual saliency γίνεται όπως στη VCGS. Επιπλέον, υπολογίζεται ο νέος gradient similarity matrix που χρησιμοποιεί τις "ψεύδο-gradient". Έπειτα, για τον υπολογισμό του VCGS, χρησιμοποιείται και ο παλιός και ο νέος gradient similarity matrix, χωρίς να γίνει κάποια αντικατάσταση.

Στο VCGS_REPLACE-GRADIENT, αντικαθιστούμε στο gradient similarity matrix, που χρησιμοποιεί η μέθοδος VCGS, την "ψεύδο-gradient".

66

Τα αποτελέσματα του συντελεστή Pearson, για το πείραμα της επιπλέον χρήσης της "ψεύδο-gradient" στην μέθοδο VCGS, αλλά και με την αντικατάστασή της στο δικό τους τρόπο υπολογισμού της gradient, παρουσιάζονται στον Πίνακα 3.18. Ομοίως και για τον συντελεστή Spearman στον Πίνακα 3.19.

Πίνακας	3.18	Συ	ντελ	εστής	Pearson	ι για	την	M	Ιέθοδα	o VC	CGS µε	Επιπλέον	Ψεί	όδο-
Gradient	και	με	την	Αντικ	ατάστα	ση τη	ς σ	το	διχό	της	Τρόπο	Υπολογισ	μού	της
Gradient														

PEARSON	WN	GBLUR	JP2K	JPEG	FF	ALL
VCGS	0.9814	0.9622	0.9738	0.9581	0.9490	0.9531
VCGS ADD-	0.9743	0.9850	0.9732	0.9423	0.9538	0.9327
GRADIENT						
VCGS						
REPLACE-	0.9732	0.9840	0.9727	0.9311	0.9533	0.9048
GRADIENT						

Με την προσθήκη της επιπλέον "ψεύδο-gradient", πετυχαίνουμε καλύτερο συντελεστή pearson στις κατηγορίες Gaussian Blur και Fast Fading. Στις υπόλοιπες κατηγορίες, παρατηρούμε μειωμένη απόδοση συγκριτικά με την μέθοδο VCGS. Όταν αντικαθιστούμε την gradient με την "ψεύδο-gradient", έχουμε μειωμένα αποτελέσματα σε σχέση με όταν την προσθέταμε επιπλέον. Παρόλα αυτά, η μείωση είναι μικρού βαθμού, οπότε η απόδοση στις κατηγορίες Gaussian Blur και Fast Fading παραμένει υψηλότερη της μεθόδου VCGS.

Πίνακας 3.19 Συντελεστής Spearman για την Μέθοδο VCGS με με Επιπλέον Ψεύδο-Gradient και με την Αντικατάσταση της στο δικό της Τρόπο Υπολογισμού της Gradient

SPEARMAN	WN	GBLUR	JP2K	JPEG	FF	ALL
VCGS	0.9739	0.9488	0.9566	0.9742	0.9404	0.9564
VCGS ADD-	0 9690	0 9734	0 9578	0 9404	0 9438	0 9376
GRADIENT	0.0000	0.0701	0.0070	0.0101	0.0100	0.0070
VCGS						
REPLACE-	0.9655	0.9692	0.9575	0.9330	0.9419	0.9051
GRADIENT						

Όσον αφορά τον συντελεστή spearman, η προσθήκη της "ψεύδο-gradient" οδηγεί στην αύξηση της απόδοσης στις κατηγορίες Gaussian Blur, Fast Fading και Jp2k, ενώ στις υπόλοιπες παρατηρούμε μείωση. Με την αντικατάσταση της gradient με την "ψεύδο-gradient", έχουμε μειωμένα αποτελέσματα σε σχέση με όταν την προσθέταμε επιπλέον. Η μείωση είναι πάλι μικρού βαθμού, οπότε η απόδοση στις κατηγορίες Gaussian Blur, Fast Fading και Jp2k παραμένει υψηλότερη της μεθόδου VCGS.

Τα δύο παραπάνω πειράματα επαναλαμβάνονται και για το πείραμα final10, στο οποίο χρησιμοποιούσαμε τον λογάριθμο της κρυμμένης πληροφορίας των οριζόντιων συντελεστών για τον υπολογισμό του visual saliency. Τα αποτελέσματα του συντελεστή Pearson και Spearman παρουσιάζονται στον Πίνακα 3.20 και Πίνακα 3.21 αντίστοιχα.

Πίνακας 3.20 Συντελεστής Pearson για την Μέθοδο Final 10 με Επιπλέον Ψεύδο-Gradient και με την Αντικατάσταση της στο δικό της Τρόπο Υπολογισμού της Gradient

PEARSON	WN	GBLUR	JP2K	JPEG	FF	ALL
VCGS	0.9814	0.9622	0.9738	0.9581	0.9490	0.9531
FINAL10	0.9762	0.9700	0.9765	0.9883	0.9748	0.9533
FINAL10						
ADD-	0.9611	0.9481	0.9767	0.9878	0.9624	0.9308
GRADIENT						
FINAL10						
REPLACE-	0.9678	0.9347	0.9745	0.9866	0.9549	0.9084
GRADIENT						

Όσον αφορά τον συντελεστή Pearson, με την επιπλέον προσθήκη της "ψεύδοgradient" έχουμε καλύτερη απόδοση στην κατηγορία Jp2k από το final10, ενώ στις υπόλοιπες παρατηρείται μείωση. Με την αντικατάσταση έχουμε περαιτέρω μείωση. Συγκριτικά με την μέθοδο VCGS, η επιπλέον προσθήκη της "ψεύδο-gradient" έχει υψηλότερες τιμές pearson στις κατηγορίες Jp2k, Jpeg και Fast Fading. Ομοίως και με την αντικατάσταση.

Πίνακας 3.21 Συντελεστής Spearman για την Μέθοδο Final 10 με Επιπλέον Ψεύδο-Gradient και με την Αντικατάσταση της στο δικό της Τρόπο Υπολογισμού της Gradient

SPEARMAN	WN	GBLUR	JP2K	JPEG	FF	ALL
VCGS	0.9739	0.9488	0.9566	0.9742	0.9404	0.9564
FINAL10	0.9685	0.9601	0.9597	0.9783	0.9631	0.9562
FINAL10						
ADD-	0.9645	0.9475	0.9609	0.9762	0.9542	0.9354
GRADIENT						
FINAL10						
REPLACE-	0.9621	0.9389	0.9578	0.9730	0.9404	0.9086
GRADIENT						

Σχετικά με τον συντελεστή Spearman, η επιπλέον προσθήκη της "ψεύδοgradient" οδηγεί στην αύξηση της τιμής στην κατηγορία Jp2k σε σχέση με το final10, ενώ στις υπόλοιπες παρατηρείται μείωση. Με την αντικατάσταση έχουμε περαιτέρω μείωση. Συγκριτικά με την μέθοδο VCGS, η επιπλέον προσθήκη της "ψεύδο-gradient" έχει υψηλότερες τιμές Spearman στις κατηγορίες Jp2k, Jpeg και Fast Fading. Με την αντικατάσταση έχουμε υψηλότερες τιμές Spearman μόνο στις κατηγορίες Jp2k και Fast Fading.

3.8 Όγδοη Ομάδα Πειραμάτων

Στις παραπάνω ομάδες πειραμάτων για τον μετασχηματισμό κυματιδίων χρησιμοποιούμε το φίλτρο Daubechies τάξης 4 (db4). Επιλέξαμε τα δύο καλύτερα από τα πειράματα και σε αυτά εφαρμόσαμε στον μετασχηματισμό κυματιδίων τα φίλτρα Daubechies τάξης 1(db1) ή, αλλιώς, φίλτρο Haar και Daubechies τάξης 8(db8). Αρχικά, εφαρμόσαμε τα φίλτρα στο πείραμα Final 10, στο οποίο χρησιμοποιούσαμε τον λογάριθμο της κρυμμένης πληροφορίας των οριζόντιων συντελεστών για τον υπολογισμό του visual saliency.

Στον Πίνακα 3.22 παρουσιάζονται και συγκρίνονται τα αποτελέσματα της μεθόδου Final 10 με το φίλτρο db4, db1 και db8, όπως προκύπτουν από τον συντελεστή Pearson. Τα αποτελέσματα του συντελεστή Spearman εμφανίζονται στον Πίνακα 3.23.

Πίνα
κας 3.22 Συντελεστής Pearson για το Final 10 με Εφαρμογή των Φίλτρων db
4, db1 και db8

PEARSON	WN	GBLUR	JP2K	JPEG	FF	ALL
VCGS	0.9814	0.9622	0.9738	0.9581	0.9490	0.9531
FINAL10	0.9762	0.9700	0.9765	0.9883	0.9748	0.9533
FINAL10 db1	0.9727	0.9692	0.9733	0.9855	0.9768	0.9484
FINAL10 db8	0.9738	0.9548	0.9746	0.9863	0.9757	0.9434

Η εφαρμογή των φίλτρων db1 και db8 επιφέρει αύξηση της τιμής Pearson μόνο στην κατηγορία Fast Fading σε σχέση με όταν χρησιμοποιούμε το φίλτρο db4. Σε σχέση με το πείραμα VCGS η εφαρμογή του φίλτρου db1 βελτιώνει την απόδοση στις κατηγορίες Jp2k, Jpeg και Fast Fading. Ομοίως και το db8.

Πίνα
κας 3.23 Συντελεστής Spearman για το Final 10 μ με Εφαρμογή των Φίλτρων db
4, db1 και db8

SPEARMAN	WN	GBLUR	JP2K	JPEG	FF	ALL
VCGS	0.9739	0.9488	0.9566	0.9742	0.9404	0.9564
FINAL10	0.9685	0.9601	0.9597	0.9783	0.9631	0.9562
FINAL10 db1	0.9655	0.9512	0.9563	0.9750	0.9621	0.9503
FINAL10 db8	0.9650	0.9419	0.9572	0.9748	0.9621	0.9454

Όσον αφορά τον συντελεστή Spearman, η χρήση του φίλτρου db4 υπερτερεί έναντι των db1 και db8 σε όλες τις κατηγορίες. Όταν χρησιμοποιούμε db1, πετυχαίνει καλύτερες τιμές στις κατηγορίες Gaussian Blur, Jpeg και Fast Fading συγκριτικά με το VCGS, ενώ το db8 στις κατηγορίες Jp2k, Jpeg και Fast Fading.

Η ίδια διαδικασία εφαρμόστηκε και στο πείραμα Final15, στο οποίο χρησιμοποιούσαμε τον λογάριθμο της αντίστροφης κρυμμένης πληροφορίας των οριζόντιων συντελεστών. Τα αποτελέσματα του συντελεστή Pearson και Spearman εμφανίζονται στον Πίνακα 3.24

Πίνα
χας 3.24 Συντελεστής Pearson για το Final 15 με Εφαρμογή των Φίλτρων db
4, db1 και db8

PEARSON	WN	GBLUR	JP2K	JPEG	FF	ALL
VCGS	0.9814	0.9622	0.9738	0.9581	0.9490	0.9531
FINAL15	0.9766	0.9703	0.9765	0.9880	0.9759	0.9528
FINAL15 db1	0.9720	0.9694	0.9739	0.9853	0.9757	0.9482
FINAL15 db8	0.9741	0.9558	0.9738	0.9751	0.9751	0.9434

Και αυτή την φορά παρατηρούμε, ότι τα καλύτερα δυνατά αποτελέσματα τα έχουμε με το φίλτρο Daubechies τάξης 4. Τα αμέσως καλύτερα αποτελέσματα τα έχει το φίλτρο Daubechies τάξης 1. Το ίδιο ισχύει και για τον συντελεστή Spearman.

Πίνα
κας 3.25 Συντελεστής Spearman για το Final 15 με Εφαρμογή των Φίλτρων db
4, db1 και db8

SPEARMAN	WN	GBLUR	JP2K	JPEG	FF	ALL
VCGS	0.9739	0.9488	0.9566	0.9742	0.9404	0.9564
FINAL15	0.9692	0.9596	0.9597	0.9779	0.9631	0.9549
FINAL15 db1	0.9650	0.9522	0.9569	0.9742	0.9621	0.9504
FINAL15 db8	0.9655	0.9414	0.9560	0.9751	0.9616	0.9444



Σχήμα 3.18 VSMap Ref και VSMap Dis για το Πείραμα final10 db1, όπου: VSMapRef: Visual Saliency Χάρτης για την Εικόνα Αναφοράς και VSMapDis: Visual Saliency Χάρτης για την Παραμορφωμένη Εικόνα.



Σχήμα 3.19 Οι Χάρτες που Προχύπτουν για το Πείραμα final 10 db1, όπου VSSImMatrix: Χάρτης Ομοιότητας μεταξύ των Δύο Ειχόνων Βασισμένος στο Visual Saliency. Χρησιμοποιείται για τον Υπολογισμό του Τελιχού Δείχτη. Weight: Πίναχας Στάθμισης Βασισμένος στην Σημείο-προς-Σημείο Μέγιστη Τιμή μεταξύ των Χαρτών Visual Saliency για τις Ειχόνες που Συγχρίνονται. SimMatrixC: Τελιχός Πίναχας Ομοιότητας πριν την Εφαρμοζόμενη Στάθμιση



Σχήμα 3.20 VSMap Ref και VSMap Dis για το Πείραμα final10 db8, όπου: VSMapRef: Visual Saliency Χάρτης για την Εικόνα Αναφοράς και VSMapDis: Visual Saliency Χάρτης για την Παραμορφωμένη Εικόνα.



Σχήμα 3.21 Οι Χάρτες που Προκύπτουν για το Πείραμα final 10 db8, όπου VSSImMatrix: Χάρτης Ομοιότητας μεταξύ των Δύο Εικόνων Βασισμένος στο Visual Saliency. Χρησιμοποιείται για τον Υπολογισμό του Τελικού Δείκτη. Weight: Πίνακας Στάθμισης Βασισμένος στην Σημείο-προς-Σημείο Μέγιστη Τιμή μεταξύ των Χαρτών Visual Saliency για τις Εικόνες που Συγκρίνονται. SimMatrixC: Τελικός Πίνακας Ομοιότητας πριν την Εφαρμοζόμενη Στάθμιση

3.9 Συγκεντρωτικά Αποτελέσματα

Η καλύτερη μέθοδος είναι το Final 10 της δεύτερης ομάδας πειραμάτων, στην οποία χρησιμοποιούμε μόνο την κρυμμένη πληροφορία των οριζόντιων συντελεστών του μετασχηματισμού wavelet για την παραγωγή του visual saliency. Η μέθοδος αυτή παρουσιάζεται στους πίνακες 3.26 και 3.27 ως VSSC και συγκρίνεται με πέντε γνωστές μετρικές εκτίμησης της ποιότητας της εικόνας. Στο Σχήμα 3.22 παρουσιάζεται και το scatter plot των υποκειμενικών και αντικειμενικών scores αυτής της μεθόδου.

ALL	PSNR	SSIM	MS-SSIM	VSI	VCGS	VSSC
SROCC	0.8739	0.9462	0.9498	0.9512	0.9564	0.9562
PLCC	0.8742	0.9444	0.9499	0.9492	0.9531	0.9533

Πίναχας 3.26 SROCC και PLCC για 6 Full Eeference Μεθόδους για το Dasatet IQA

Πίνα
χας 3.27 SROCC για 6 Reference Μεθόδους για κάθε Είδος Παραμόρφωσης του Dataset IQA

Spearman	PSNR	SSIM	MSSIM	VSI	VCGS	VSSC
WN	0.9793	0.9640	0.9680	0.9773	0.9739	0.9685
GB	0.7746	0.9429	0.9461	0.9433	0.9488	0.9601
JP2K	0.8872	0.9540	0.9551	0.9536	0.9566	0.9598
JPEG	0.8736	0.9700	0.9757	0.9702	0.9742	0.9783
FF	0.8833	0.9492	0.9429	0.9350	0.9404	0.9631
ALL	0.8739	0.9462	0.9498	0.9512	0.9564	0.9562



Σχήμα 3.22 Scatter Συσχέτισης μεταξύ των Υποκειμενικών και Αντικειμενικών Scores για τη Μέθοδο VSSC στο οποίο Χρησιμοποιούμε για τον Αντίστροφο Μετασχηματισμό Wavelet τον Λογάριθμο της Κρυμμένης Πληροφορίας των Οριζόντιων Συντελεστών

Από τους πίναχες παρατηρούμε, ότι η απόδοση της μεθόδου μας την χατατάσσει μεταξύ των τριών χαλύτερων μεθόδων.

ΚΕΦΑΛΑΙΟ 4

ΠΡΟΤΕΙΝΟΜΕΝΕΣ NO-REFERENCE ΜΕΘΟΔΟΛΟΓΙΕΣ

- 4.1 DATASET IQA
- 4.2 DATASET C
- 4.3 Σύγχριση Αποτελεσμάτων με Μοντέλα Βασισμένα στους MCSN Συντελεστές
- 4.4 Συμπληρωματικά Πειράματα

Όπως αναφέραμε, μία NSS μέθοδος χρησιμοποιεί χαρακτηριστικά, που ανήκουν είτε στο χωρικό πεδίο είτε στο πεδίο των κυματιδίων είτε, γενικά, στον τομέα του πεδίου των συχνοτήτων. Στη BRISQUE μεθοδολογία [6], οι ερευνητές έστρεψαν το ενδιαφέρον τους στα χαρακτηριστικά από το χωρικό πεδίο. Συγκεκριμένα, κάνουν χρήση κανονικοποιημένων συντελεστών φωτεινότητας (mean subtracted contrast normalized -MCSN) στο πεδίο του χώρου, καθώς και των κατά ζεύγη γινομένων των γειτονικών κανονικοποιημένων συντελεστών φωτεινότητας, χρησιμοποιώντας για την προσαρμογή τους την Γενιχευμένη Γκαουσιανή κατανομή (Generalized Gaussian Distribution-GGD) και την Ασύμμετρη Γενικευμένη Γκαουσιανή κατανομή (Asymmetric Generalized Gaussian Distribution-AGGD), αντίστοιγα. Η συγχεχριμένη μεθοδολογία είναι γεγονός ότι επιτυγχάνει υψηλή επίδοση, όσον αφορά την πρόβλεψη. Επιπλέον, λόγω του γεγονότος ότι όλα τα χαραχτηριστικά που χρησιμοποιεί εξάγονται στο χωρικό πεδίο, η χρονική πολυπλοκότητά της είναι χαμηλή.

Στην εργασία τους [54], οι Fu-Zhao Ou et al κάνουν εξαγωγή χαρακτηριστικών από το χωρικό πεδίο με τον ίδιο τρόπο, αλλά χρησιμοποιούν και χαρακτηριστικά

από το πεδίο μετασχηματισμού και συγκεκριμένα από το Μετασχηματισμό Συνημιτόνου (DCT), δημιουργώντας έτσι ένα υβριδικό μοντέλο. Η επίδοση αυξήθηκε σε σχέση με το χωρικό μοντέλο BRISQUE [6], όπως, όμως, και τα χαρακτηριστικά και η πολυπλοκότητά του.

Βασιζόμενοι στο γεγονός ότι μία τέτοιου είδους μεθοδολογία αξιολόγησης της ποιότητας, που βασίζεται σε διαδικασίες εκπαίδευσης – μάθησης, συνιστά μία αναδυόμενη -εξελισσόμενη τεχνολογία, επιλέξαμε να μελετήσουμε τις διαφορετικές στατιστικές κατανομές στο πλαίσιό της. Σε προηγούμενη εργασία [55], αναπτύξαμε μεθόδους αξιολόγησης της αντιληπτής ποιότητας εικόνων μέσω χρήσης εναλλακτικών στατιστικών κατανομών. Κάναμε συγκριτική μελέτη των ήδη υπαρχουσών μεθόδων, που χρησιμοποιούν Γενικευμένη Γκαουσιανή Κατανομή, με μεθόδους που χρησιμοποιούν Student's t κατανομή, αλλά και μιας πιο ειδικής περίπτωσής της, της Cauchy κατανομής.

Συγκεκριμένα, αναπτύξαμε μοντέλα που χρησιμοποιούν κανονικοποιημένους συντελεστές φωτεινότητας (MSCN) στο πεδίο του χώρου, καθώς και τα κατά ζεύγη γινόμενα των γειτονικών κανονικοποιημένων συντελεστών φωτεινότητας. Το συγκεκριμένο μοντέλο βασίζεται στην παραδοχή ότι οι MSCN συντελεστές έχουν χαρακτηριστικές στατιστικές ιδιότητες, που διαταράσσονται με την παρουσία των παραμορφώσεων και η μοντελοποίηση αυτών των συντελεστών δίνει τη δυνατότητα σε μία μέθοδο να προβλέψει το είδος και τον βαθμό της παραμόρφωσης. Οι MCSN συντελεστές υπολογίζονται από τον τύπο:

$$L(i, j) = \frac{I(i, j) - \mu(i, j)}{\sigma(i, j) + C}$$
(4.1)

όπου i, j είναι δείκτες χωρικού πεδίου, L είναι μια εικόνα φωτεινότητας και C μια σταθερά για να αποφύγουμε τον μηδενικό παρονομαστή. Με το $\mu(i,j)$ και το $\sigma(i,j)$ υπολογίζουμε το τοπικό μέσο και την τυπική απόκλιση.

Για τη ποσοτιχοποίηση της ποιότητας μιας ειχόνας παρουσία χάποιας παραμόρφωσης, συνήθως χρησιμοποιούμε μια παραμετριχή στατιστιχή χατανομή, όπου περιγράφει το σχήμα των δεδομένων (πχ. Ιστόγραμμα) σε μια συμπαγή αναπαράσταση μέσω των παραμέτρων της. Η αρχιχή μας σχέψη αφορούσε το γεγονός, ότι η Student's t μπορεί να προσαρμοστεί σε ένα μεγάλο εύρος μορφών της χατανομής των δεδομένων που έχουμε. Γενιχά, η μορφή των MSCN δεδομένων είναι Γχαουσιανής μορφής.



Σχήμα 4.1 Εικόνα από το Dataset C.



Σχήμα 4.2 Πραγματική Κατανομή των MCSN Συντελεστών συγκριτικά με την Κανονική Κατανομή (Κόκκινη Καμπύλη).



Σχήμα 4.3 Οι Πραγματικές Κατανομές των Κατά Ζεύγη Γειτονικών MSCN του Αρχικού Σχήματος 4.1, για τη Οριζόντια (Πάνω Αριστερά), Κάθετη (Πάνω Δεξιά), Κύρια Διαγώνιο (Κάτω Αριστερά) και Δευτερεύουσα Διαγώνιο (Κάτω Δεξιά) Κατεύθυνση σε Σχέση με την Κανονική Κατανομή (Κόκκινη Καμπύλη)

Ιδιαίτερη σημασία έχει, επίσης, η διερεύνηση της στατιστικής σχέσης στο χωρικό πεδίο μεταξύ των γειτονικών κανονικοποιημένων συντελεστών. Οπότε, γίνεται εξαγωγή και επιπλέον χαρακτηριστικών από προσαρμογή της κατανομής Studentt στα κατά ζεύγη γινόμενα των οριζόντιων, κάθετων και διαγώνιων κατευθύνσεων. Συνήθως, χρησιμοποιείται ο ΕΜ αλγόριθμος για την εκτίμηση παραμέτρων της κατανομής. Όμως, έπειτα από πολλά πειράματα, και για να αυξήσουμε την απόδοση των μεθόδων μας, ακολουθήσαμε μια εναλλακτική προσέγγιση εκτίμησης, κατά την οποία υπολογίζουμε δύο παραμέτρους, ήτοι την παράμετρο κλίμακας σ² και τους βαθμούς ελευθερίας α. Η παράμετρος κλίμακας και οι βαθμοί ελευθερίας υπολογίζονται από το ιστόγραμμα των MCSN συντελεστών και των κατά ζεύγη γινομένων τους μέσω των τύπων:

$$\sigma_{\chi}^2 = \frac{\alpha}{\alpha - 2} \sigma^2 \tag{4.2}$$

$$\kappa_{\chi} = \frac{6}{\alpha - 4} \tag{4.3}$$

όπου σ_{χ}^2 η διαχύμανση και κ_{χ} η χύρτωση και εχτιμώνται από την δεύτερη και τέταρτη ροπή των δεδομένων.

Μέσω αυτών των παραμέτρων εξάγουμε 2 χαρακτηριστικά από τους MCSN συντελεστές και 8 από τα κατά ζεύγη γινόμενα τους, ήτοι σύνολο 10 χαρακτηριστικά. Έπειτα, υποκλιμακώνουμε την εικόνα κατά ½ και κάνουμε εξαγωγή άλλων 10 χαρακτηριστικών. Κάναμε χρήση χαρακτηριστικών και από το πεδίο μετασχηματισμού, δημιουργώντας έτσι ένα υβριδικό μοντέλο.



Σχήμα 4.4 Αναπαράσταση Λειτουργίας ενός Υβριδικού Μοντέλου

Διαπιστώσαμε, ότι η χρήση μιας διαφορετικής κατανομής από αυτήν της ευρέως χρησιμοποιούμενης Γενικευμένης Γκαουσιανής μπορεί να οδηγήσει σε καλύτερα αποτελέσματα, είτε σημαίνει μεγαλύτερη μετρήσιμη απόδοση είτε παρόμοια είτε ακόμα και μικρότερη, κάνοντας χρήση, όμως, λιγότερων χαρακτηριστικών. Με λιγότερα χαρακτηριστικά εξασφαλίζουμε μείωση της χρονικής πολυπλοκότητας, καθώς και της υπολογιστικής πολυπλοκότητας.

Όμως, όπως αντιληφθήκαμε από την λεπτομερή εξέταση των κατανομών των δεδομένων των εικόνων, η μορφή τους σε διάφορες «ακραίες» περιπτώσεις προσαρμόζεται καλύτερα από τη Γενικευμένη Γκαουσιανή και όχι από τη Student's t. Αυτού του είδους οι «ακραίες» περιπτώσεις καθιστούν τα δεδομένα μας περισσότερο κοντά στην ομοιόμορφη κατανομή, μορφή στην οποία δε μπορεί να προσαρμοστεί η Student's t, ενώ η GGD εμφανίζει σαφές συγκριτικό πλεονέκτημα.

Το συγχεχριμένο μειονέχτημα της πρότασής μας, μας οδήγησε στην ιδέα εφαρμογής αυτής χαι της περαιτέρω διερεύνησης σε δεδομένα, που δεν θα εχφυλίζεται η μορφή τους σε ομοιόμορφη χατανομή. Τέτοιου είδους χατανομές εμπειριχών δεδομένων είναι τα δεδομένα των μετασχηματισμών χυματιδίων. Έτσι, εφαρμόζουμε WT στην ειχόνα χαι παίρνουμε τους συντελεστές διαγώνιας, χάθετης και οριζόντιας λεπτομέρειας, cD, cV και cH, αντίστοιχα, σε διάφορα επίπεδα του μετασχηματισμού. Αυτοί οι συντελεστές προσαρμόζονται με μια από τις εξής κατανομές:

- Student's t κατανομή με χρήση του επαναληπτικού αλγορίθμου EM (Expectation Maximization), όπως περιγράφηκε στην ενότητα 2.3 για την εκτίμηση των παραμέτρων της κατανομής.
- Student's t κατανομή με δύο παραμέτρους: Η παράμετρος κλίμακας σ² και οι βαθμοί ελευθερίας α υπολογίζονται από το ιστόγραμμα ενδιαφέροντος κάθε φορά, στην συγκεκριμένη περίπτωση από το ιστόγραμμα των συντελεστών κυματιδίων, και υπολογίζονται από τον τύπους (4.2) και (4.3).
- Students't κατανομή με τέσσερις παραμέτρους: Εκτός από την παράμετρο κλίμακας σ² και τους βαθμούς ελευθερίας α χρησιμοποιούνται σαν χαρακτηριστικά και η διακύμανση σ_x² και η κύρτωση k_x.
- Cauchy κατανομή: Όταν ένα στατιστικό μοντέλο μπορεί να μετατραπεί σε ένα άλλο, επιβάλλοντας περιορισμούς στις παραμέτρους του πρώτου, τότε ορίζουμε αυτά τα μοντέλα ως φωλιασμένα (nested). Για παράδειγμα, αυτό συμβαίνει για την Student- t κατανομή, καθώς έχει φωλιασμένο το σύνολο των Cauchy και των κανονικών κατανομών. Το πιο σημαντικό, όμως, είναι ότι υπάρχουν περιπτώσεις, όπου αναμένουμε το πιο γενικό μοντέλο να έχει καλύτερη απόδοση, αφού περιέχει το λιγότερο γενικό μοντέλο. Άλλες φορές, η εξάλειψη της ανάγκης να βρεθούν όλες εκείνες οι παράμετροι του λιγότερου γενικού μοντέλου, καθιστά τη χρήση του αποτελεσματική με κριτήριο την χρονική πολυπλοκότητα, χωρίς να θυσιάζεται η επίδοση.

Επομένως, καθώς η t-κατανομή περιέχει την Cauchy και την Κανονική κατανομή ως ειδικές περιπτώσεις, η χρήση της είναι σημαντική για γενικά δεδομένα μοντελοποίησης δεδομένων, ως εναλλακτικό και πιο ισχυρό μοντέλο πιθανότητας σε σύγκριση π.χ. με την Κανονική κατανομή.

Το μοντέλο Cauchy έχει χρησιμοποιηθεί επιτυχώς σε διάφορες εφαρμογές, όπως για την υδατογράφηση (watermarking) στο πεδίο μετασχηματισμού, κάνοντας χρήση των συντελεστών DCT χαμηλής έως μεσαίας συχνότητας και των συντελεστών λεπτομέρειας του μετασχηματισμού DWT. Η συνάρτηση πυκνότητας πιθανότητας της Cauchy κατανομής, με παράμετρο θέσης $-\infty < \delta < \infty$ και παράμετρο σχήματος $\gamma > 0$, είναι :

80

$$p(x|\gamma,\delta) = \frac{1}{\pi} \frac{\gamma}{\gamma^2 + (x-\delta)^2}$$
(4.4)

με $-\infty < x < \infty$. Η Cauchy κατανομή με $\delta = 0$ (η οποία είναι συμμετρική κοντά στο μηδέν) δίνεται από την σχέση:

$$p(x|\gamma) \coloneqq p(x|\gamma,0) \tag{4.5}$$

Σε αντίθεση με την Gaussian κατανομή, οι ουρές της Cauchy κατανομής εξασθενούν με ρυθμό βραδύτερο από την εκθετική. Ως εκ τούτου παρατηρούμε παχιές ουρές στη συνάρτηση πυκνότητας πιθανότητας.



Σχήμα 4.5 Πραγματική Κατανομή των Οριζόντιων Συντελεστών, που Προκύπτουν από το Μετασχηματισμό Κυματιδίων μιας Εικόνας, συγκριτικά με την Κανονική (Κόκκινη Καμπύλη) και Student-t Κατανομή (Μπλέ Καμπύλη)

Εκτίμηση Παραμέτρων της Cauchy

Όσον αφορά την εκτίμηση της παραμέτρου γ της Cauchy κατανομής, καταφεύγουμε στην προσέγγιση της εκτίμησης Μέγιστης Πιθανοφάνειας. Αν θεωρήσουμε ότι έχουμε ένα πλήθος N δεδομένων x[1],...,x[N], τότε αυτά υποδηλώνουν την πραγματοποίηση N τυχαίων ανεξάρτητων και ομοιόμορφα κατανεμημένων μεταβλητών που ακολουθούν μια Cauchy κατανομή με $\delta = 0$. Στην περίπτωση αυτή, η ML εκτίμηση του γ δίνεται από την ρίζα της εξίσωσης:

$$\frac{1}{N} \sum_{t=1}^{N} \frac{2}{1 + (x[t]/\gamma)^2} - 1 =: h(\gamma)$$
(4.6)

Για την εύρεση ρίζας με τη μέθοδο του Newton–Raphson χρειαζόμαστε το $h'(\gamma)$. Η πρώτη παράγωγος του $h(\gamma)$ είναι:

$$h'(\gamma) \coloneqq \frac{\partial h}{\partial \gamma} = \frac{4\gamma}{N} \sum_{t=1}^{N} \frac{x[t]^2}{\left(\gamma^2 + x[t]^2\right)^2}$$
(4.7)

όπου οδηγεί στην σχέση:

$$\hat{\gamma}_{n+1} = \hat{\gamma}_n - h(\hat{\gamma}_n) / h'(\hat{\gamma}_n)$$
 (4.8)

για την n-οστή επανάληψη (με n > 1). Ως αρχική τιμή $\hat{\gamma}_1$ χρησιμοποιούμε μια εκτίμηση βασισμένη στην μέθοδο του δείγματος κβάντων x_p :

$$\hat{\gamma}_1 = 0.5 (x_p - x_{1-p}) \tan(\pi(1-p)), \quad \mu \in 0.5 (4.9)$$

Σ' αυτή την μεταπτυχιαχή εργασία, όποτε χρειάζεται να χρησιμοποιήσουμε κάποιον αλγόριθμο εχτίμησης, επιλέγουμε έναν γρήγορο αλγόριθμο, που παρουσιάστηχε αρχιχά από τους Trihrintzis και Nikias [56] στο πλαίσιο της εχτίμησης των παραμέτρων μίας άλλης οιχογένειας, της οιχογένειας των Symmetric a Stable Κατανομών. Δεδομένου ότι υποθέτουμε ότι $\delta = 0$ και $\alpha = 1$, ο προτεινόμενος εχτιμητής για το γδίνεται από το τύπο:

$$\hat{\gamma} = \left[\frac{\sqrt{3}}{2N} \sum_{t=1}^{N} |x[t]|^{1/3}\right]^3$$
(4.10)

ο οποίος μπορεί να υπολογιστεί με γραμμική πολυπλοκότητα. Τέλος, παρατηρούμε ότι η εκτίμηση ροπής δεν είναι δυνατή στην περίπτωση της Cauchy κατανομής, καθώς δεν υπάρχουν ροπές.

Εκτός από τα χαρακτηριστικά του πεδίου κυματιδίων χρησιμοποιούνται και χαρακτηριστικά από το Μετασχηματισμό Συνημιτόνου, τα οποία προσθέτουμε σταδιακά, ώστε να παρατηρήσουμε την επιρροή τους στην απόδοση των μοντέλων και να καταλήξουμε σε ένα μοντέλο με υψηλή απόδοση. Η αύξηση του πλήθους των χαρακτηριστικών οδηγεί σε αύξηση της υπολογιστικής αλλά και χρονικής πολυπλοκότητας και, γι' αυτό, στόχος μας είναι να βρούμε το μοντέλο που κατορθώνει μια ισοσταθμία μεταξύ της απόδοσης που επιτυγχάνει και του πλήθους των χαρακτηριστικών που χρησιμοποιεί. Τα χαρακτηριστικά που προσθέτουμε από το Μετασχηματισμό Συνημιτόνου είναι τα εξής:

• 2 χαρακτηριστικά βασισμένα στο νόμο του Benford (Bendford's law)

Πρόχειται για χαραχτηριστικά ιδιαίτερα ευαίσθητα στον λευκό θόρυβο, στο Γκαουσιανό θόρυβο και στο Fast fading. Ο νόμος του Benford υποστηρίζει ότι, για ένα σύνολο δεδομένων, η κατανομή των πιο σημαντικών ψηφίων ικανοποιεί την σχέση:

$$p(n) = \log_m \left(1 + \frac{1}{n}\right), n = 1, 2, \dots, m - 1$$
 (4.11)

όπου m ένας συγκεκριμένος αριθμός για διαφορετικές βάσεις του λογαρίθμου. Έχει αναχαλυφθεί ότι η χατανομή των σημαντιχών ψηφίων των συντελεστών DCT μια πλήρους εικόνας είναι σύμφωνη με την με την υποκειμενική βαθμολογία της εικόνας. Εκτεταμένα πειράματα δείχνουν ότι για τις παραμορφωμένες εικόνες, ιδίως όσον αφορά τους τρεις τύπους παραμόρφωσης, WN, GB και FF, η κατανομή θα αλλάξει σημαντικά. Για παράδειγμα, η εικόνα αναφοράς και οι παραμορφωμένες εικόνες θόλωσης διαφορετικού επιπέδου παραμόρφωσης εμφανίζονται στην Εικόνα 4.6. Οι κατανομές των κορυφαίων ψηφίων των συντελεστών DCT πλήρους εικόνας σχεδιάζονται στο Σχήμα 4.7. Μπορούμε να δούμε, από το Σχήμα. 4.6, ότι η παραμόρφωση θόλωσης υποβαθμίζει σημαντικά την οπτική ποιότητα της εικόνας. Αντίθετα, μπορεί να παρατηρηθεί, από την Εικόνα 4.7, ότι η κατανομή της εικόνας αναφοράς (δηλαδή, αρχική εικόνα, DMOS=0) είναι σχεδόν η ίδια με εκείνη του νόμου του Benford, αλλά η κατανομή της θολής παραμορφωμένης εικόνας δεν το κάνει. Μπορούμε, επίσης, να διαπιστώσουμε, ότι η κατανομή της περισσότερο παραμορφωμένης εικόνας (DMOS=82.3) είναι η πιο μαχριά από εχείνη του νόμου του Benford.



Σχήμα 4.6 Η Οπτική Ποιότητα μιας Εικόνας Αναφοράς και Δυο Παραμορφωμένων Εικόνων με Φίλτρο Θολώματος. Με Σειρά Εμφάνισης Παρουσιάζονται οι εξής Εικόνες: Εικόνα Αναφοράς με DMOS=0 (Εικόνα Αριστερά), η Εικόνα Παραμόρφωσης με Θόλωση με DMOS=55.1 (Εικόνα Μέση) και η Εικόνα Παραμόρφωσης με DMOS (82.3).



Σχήμα 4.7 Ιστόγραμμα των κορυφαίων ψηφίων των DCT συντελεστών

Με βάση αυτό το νόμο και το γεγονός ότι η κατανομή του είναι διακριτή και έχει μικρό χώρο δείγματος, οι Fu-Zhao Ou et al [54] παρουσίασαν μια ευκλείδεια απόσταση για την μέτρηση της απόκλισης μεταξύ δύο αποκλίσεων. Έτσι, το πρώτο από τα δύο χαρακτηριστικά εξάγεται με τον τύπο:

$$\mathbf{x} = \sqrt{\sum_{i=1}^{9} \left[p^{d}(i) - p(i) \right]^{2}}$$
(4.12)

όπου p^d δηλώνει τη κατανομή του νόμου Benford μιας παραμορφωμένης εικόνας.

Το δεύτερο χαρακτηριστικό εξάγεται με το ίδιο τρόπο, με την διαφορά ότι η εικόνα έχει υποβληθεί σε υποκλιμάκωση ½ και σε επεξεργασία με Γκαουσιανό χαμηλοπερατό φίλτρο μεγέθους 3X3, με μηδενική μέση τιμή και τυπική απόκλιση ίση με 1.

6 χαρακτηριστικά βασισμένα στην ενεργειακή αναλογία υποζώνης

Η έρευνα των Saad et al [57] απέδειξε ότι η παραμόρφωση αλλάζει το τοπικό περίγραμμα των εικόνων με μη φυσικό τρόπο και, επιπλέον, ότι μπορεί να φανερωθεί στην ενεργειακή αλλαγή ενός DCT block. Για να συλλεχθούν κατευθυντικές πληροφορίες από τις τοπικές επιδιορθώσεις εικόνων, το μπλοκ DCT χωρίζεται κατευθυντικά σε τρεις προσανατολισμένες υπο-περιοχές. Αυτός ο διαχωρισμός εκφράζει τρεις υπό-ζώνες συχνοτήτων στο DCT block. Τα ανώτερα, μεσαία και χαμηλότερα χωρίσματα αντιστοιχούν στις υπό-ζώνες χαμηλών συχνοτήτων (n=1), μεσαίων συχνοτήτων (n=2) και υψηλών συχνοτήτων (n=3), αντίστοιχα.

DC	C ₁₂	C ₁₃	C ₁₄	C ₁₅	DC	C ₁₂	C ₁₃	C ₁₄	C ₁₅
C ₂₁	C ₂₂	C ₂₃	C ₂₄	C ₂₅	C ₂₁	C ₂₂	C ₂₃	C ₂₄	C ₂₅
C ₃₁	C ₃₂	C ₃₃	C ₃₄	C ₃₅	C ₃₁	C ₃₂	C ₃₃	C ₃₄	C ₃₅
C ₄₁	C ₄₂	C ₄₃	C ₄₄	C ₄₅	C ₄₁	C ₄₂	C ₄₃	C ₄₄	C ₄₅
C ₅₁	C ₅₂	C ₅₃	C ₅₄	C ₅₅	C ₅₁	C ₅₂	C ₅₃	C ₅₄	C ₅₅

Σχήμα 4.8 Οι DCT Συντελεστές κατά Μήκος Τριών Ζωνών Συχνοτήτων (αριστερά) και Τριών Προσανατολισμών (δεξιά).

Ο ενεργειαχός λόγος υποζώνης των συντελεστών DCT για χάθε block δίνεται από το τύπο:

$$R_{n} = \frac{\left|\sigma_{n}^{2} - \frac{1}{n-1}\sum_{j < n}\sigma_{j}^{2}\right|}{\sigma_{n}^{2} + \frac{1}{n-1}\sum_{j < n}\sigma_{j}^{2}}$$
(4.13)

όπου n=2,3 και σ_j^2 δηλώνει την ενέργεια της j-οστής υποζώνης. Έπειτα, υπολογίζεται το $\overline{R} = (R_2 + R_3)/2$. Η μέση τιμή του 100-οστού και του 10-ου υψηλότερου \overline{R} προσδιορίζουν τα δύο από τα 6 χαρακτηριστικά. Τα υπόλοιπα 6 εξάγονται με την ίδια διαδικασία, αλλά για την εικόνα που έχει υποστεί Γκαουσιανό χαμηλοπερατό φίλτρο δύο φορές.

• 6 χαρακτηριστικά βασισμένα στην διακύμανση συχνότητας.

Έστω ζ_j η διαχύμανση συχνότητας για χάθε DCT block , που υπολογίζεται για τρείς χατευθύνσεις από τον τύπο:

$$\zeta_j = \frac{\sigma_j}{\mu_j} \tag{4.14}$$

όπου j=1,2,3 αναπαριστούν τρεις υπο-περιοχές, μ_j η μέση τιμή και σ_j η τυπική απόκλιση των συντελεστών DCT. Υπολογίζουμε την διακύμανση για όλα τα block και ο μέσος όρος του 100-στου και του 10-ου υψηλότερου εκατοστημορίου αποτελούν τα έξι χαρακτηριστικά. Έπειτα από την εφαρμογή Γκαουσιανού χαμηλοπερατού φίλτρου στην αρχική εικόνα, κάνουμε εξαγωγή άλλων 6 χαρακτηριστικών.

Αφού υπολογιστούν τα απαραίτητα χαραχτηριστικά, ακολουθεί ένα μοντέλο μάθησης. Αρχικά, χωρίζουμε το σύνολο δεδομένων σε δύο υποσύνολα. Το μοντέλο εκπαιδεύεται αντιστοιχίζοντας τα παραπάνω χαρακτηριστικά ενός τυχαίου υποσυνόλου εικόνων, το οποίο αποτελεί το 80% του αρχικού συνόλου και ονομάζεται σύνολο εκπαίδευσης (train), στις υποκειμενικές βαθμολογίες (scores) των εικόνων. Αυτή η αντιστοιχία μεταξύ των χαρακτηριστικών και των τιμών MOS (ή DMOS) έγινε με χρήση μηχανών διανυσμάτων στήριξης (SVM). Η ποιότητα προβλέπεται από αυτό το μοντέλο εισάγοντας τα χαρακτηριστικά του εναπομείναντος μη επικαλυπτόμενου υποσυνόλου, που αποτελεί το 20% του αρχικού και ονομάζεται υποσύνολο ελέγχου (test). Επαναλαμβάνουμε αυτή την διαδικασία 1000 φορές και ως μέτρο επίδοσης κρατάμε το μέσο όρο των αποδόσεων, που παίρνουμε με τους συντελεστές Spearman και Pearson. Αυτή η διαδικασία εφαρμόζεται σε δύο σύνολα δεδομένων, το LIVE IQA και το LIVE C.

86
4.1 DATASET IQA

To Live IQA, όπως αναφέραμε στην ενότητα 1.7, είναι ένα σύνολο 982 εικόνων, οι οποίες έχουν υποστεί ένα είδος συνθετικής παραμόρφωσης από τα εξής: white noise, Gaussian Blur, Jp2k, jpeg και Fast Fading.

4.1.1 Student's t κατανομή με ΕΜ εκτίμηση παραμέτρων

• Logistic Pearson

Πίνακας 4.1 Συντελεστής Pearson Μοντέλων που Έχουν σαν Βάση Χαρακτηριστικά που Εξάγονται μέσω της Student-t Κατανομής με ΕΜ Εκτίμηση Παραμέτρων.

V	Πλήθος	White	Gaussian	T 91-	in a T	Fast	A 11
κατανομη	features	noise	Blur	јрźк	Jpeg	Fading	All
Student_t_1	6 wavelets	0.9205	0.9013	0.6697	0.6620	0.8019	0.4553
Student_t_2	6 wav+2 benford	0.9314	0.9086	0.8627	0.7555	0.8131	0.6740
Student_t_3	6 wav+6 enr	0.9554	0.9337	0.9083	0.6551	0.8653	0.6066
Student_t_4	6 wav+6 freq	0.9496	0.9373	0.9337	0.9259	0.8832	0.7142
Student_t_5	6 wav+6 enr+2 benford	0.9571	0.9348	0.9195	0.7771	0.8698	0.7412
Student_t_6	6 wav+6 freq +2 benford	0.9424	0.9414	0.9351	0.9247	0.8646	0.8780
Student_t_7	6 wav+ 6 enr+ 6 frq	0.9566	0.9437	0.9338	0.9277	0.8796	0.8486
Student_t_8	6 wav + 14 DCT	0.9571	0.9476	0.9419	0.9310	0.8897	0.8885

Αρχικά, σε κάθε ένα από τα 8 πειράματα, χρησιμοποιείται ο τύπος (2.4) για τις εικόνες που προκύπτουν για τους διαγώνιους, οριζόντιους και κάθετους συντελεστές, που δημιουργούνται από το μετασχηματισμό κυματιδίων. Υπολογίζονται οι δύο παράμετροι της κατανομής για κάθε εικόνα, με αποτέλεσμα την εξαγωγή 6 χαρακτηριστικών.

Στο Student_t_1, χρησιμοποιούμε μόνο αυτά τα 6 χαρακτηριστικά για την εκπαίδευση του μοντέλου μάθησης.

Στο Student_t_2, χρησιμοποιούμε τα 6 χαρακτηριστικά που βασίζονται στα wavelets, προσθέτοντας άλλα δύο χαρακτηριστικά Benford. Οπότε, έχουμε συνολικά 6+2=8 χαρακτηριστικά. Παρατηρούμε αύξηση του συντελεστή Pearson σε κάθε κατηγορία παραμόρφωσης, διατηρώντας το μικρό πλήθος χαρακτηριστικών.

Στο Student_t_3, χρησιμοποιούμε τα 6 χαρακτηριστικά που βασίζονται στα wavelets, προσθέτοντας αυτή την φορά 6 χαρακτηριστικά βασισμένα στην ενεργειαχή αναλογία υποζώνης. Συνολιχά έχουμε 6+6=12 χαραχτηριστιχά. Παρατηρούμε αύξηση συντελεστή Pearson, του σχέση όταν σε με χρησιμοποιούσαμε μόνο τα χαρακτηριστικά που στηρίζονταν στα wavelets, στο πείραμα Student_t_1. Συγκριτικά με τα αποτελέσματα του πειράματος Student_t_2, που χρησιμοποιούσαμε 8 χαρακτηριστικά, παρατηρούμε χαμηλότερη τιμή του συντελεστή στις κατηγορίες jpeg και all, παρόλο που έχουμε 4 χαρακτηριστικά παραπάνω. Στις υπόλοιπες κατηγορίες έχουμε αυξημένες τιμές του συντελεστή.

Στο Student_t_4, χρησιμοποιούμε τα 6 χαρακτηριστικά που βασίζονται στα wavelets, προσθέτοντας αυτή την φορά 6 χαρακτηριστικά βασισμένα στην διακύμανση συχνοτήτων. Συγκριτικά με το πείραμα Student_t_1, έχουμε υψηλότερες τιμές του συντελεστή σε όλες τις κατηγορίες. Παρομοίως, και αν το συγκρίνουμε με το Student_t_2. Σε σχέση με το Student_t_3, με το οποίο έχουν ίδιο πλήθος χαρακτηριστικών, μπορούμε να συμπεράνουμε ότι η χρήση χαρακτηριστικών βασισμένων στην διαχύμανση συχνοτήτων είναι πιο αποδοτική από αυτά που βασίζονται στην ενεργειακή αναλογία υποζώνης. Στην μόνη κατηγορία που παρουσιάζει χαμηλότερη απόδοση είναι η white noise.

Στο Student_t_5, χρησιμοποιούμε τα 12 χαρακτηριστικά του πειράματος Student_t_3 και προσθέτουμε και τα 2 Benford χαρακτηριστικά. Αυτή η προσθήκη προκαλεί αύξηση των τιμών του συντελεστή Pearson σε όλες τις κατηγορίες παραμόρφωσης. Το μόνο πείραμα, που δεν καταφέρνει να ανταγωνιστεί από τα

προαναφερθέντα, είναι το Student_t_4, καθώς παρουσιάζει αύξηση μόνο στις κατηγορίες white noise και all και σε όλες τις άλλες κατηγορίες παρουσιάζει μείωση.

Στο Student_t_6, χρησιμοποιούμε τα 12 χαραχτηριστικά του πειράματος Student_t_4 και προσθέτουμε και τα 2 Benford χαραχτηριστικά. Η προσθήκη αυτή δεν έχει ανάλογη αύξηση, όπως σε αυτήν της περίπτωσης Student_t_5. Στις κατηγορίες white noise, jpeg και fast fading, φαίνεται ότι τα Benford χαραχτηριστικά μείωσαν την απόδοση που είχαν πετύχει τα χαραχτηριστικά βασισμένα στην διακύμανση συχνοτήτων. Συγκριτικά με το πείραμα Student_t_5, που χρησιμοποιούσε 6 χαραχτηριστικά βασισμένα στην ενεργειακή αναλογία υποζώνης, αντί των 6 της διακύμανσης συχνοτήτων, έχουμε καλύτερα αποτελέσματα σε όλες τις κατηγορίες εκτός από την white noise και fast fading.

Στο πείραμα Student_t_7, χρησιμοποιούμε τα 6 χαρακτηριστικά που βασίζονται στα wavelets και προσθέτουμε 6 χαρακτηριστικά βασισμένα στην διακύμανση συχνοτήτων και 6 χαρακτηριστικά βασισμένα στην ενεργειακή αναλογία υποζώνης. Συνολικά έχουμε 6+6+6=18 χαρακτηριστικά. Παρά την σημαντική αύξηση χαρακτηριστικών, δεν καταφέρνει σημαντική αύξηση της απόδοσης σε σχέση με τα πειράματα Student_t_5 και Student_t_6, ενώ στην κατηγορία jp2k και all παρουσιάζει και μείωση συγκριτικά με το Student_t_6.

Στο Student_t_8, χρησιμοποιούμε τα 6 χαρακτηριστικά που βασίζονται στα wavelets με προσθήκη των 2 Benford χαρακτηριστικών, των 6 χαρακτηριστικών βασισμένων στην διακύμανση συχνοτήτων και των 6 χαρακτηριστικών βασισμένων στην ενεργειακή αναλογία υποζώνης. Συνολικά έχουμε 6+6+6+2= 20 χαρακτηριστικά. Το πλήθος χαρακτηριστικών φαίνεται να επιδρά θετικά στην απόδοση, καθώς πετυχαίνει υψηλότερη απόδοση συγκριτικά με τα προηγούμενα πειράματα σε όλες τις κατηγορίες.

• Spearman

Για τα ίδια πειράματα υπολογίστηκε και ο συντελεστής Spearman.

Πίνακας 4.2 Συντελεστής Spearman Μοντέλων που Έχουν σαν Βάση Χαρακτηριστικά που Εξάγονται μέσω της Student-t Κατανομής με ΕΜ Εκτίμηση Παραμέτρων.

Varannamá	Πλήθος	White	Gaussian	T 91-	Inco	Fast	A 11
κατανομη	features	noise	blur	јрźк	Jpeg	Fading	All
Student_t_1	6 wavelets	0.9047	0.9057	0.6143	0.6129	0.7653	0.2154
Student_t_2	6 wav+2 benford	0.8866	0.9204	0.8328	0.7417	0.7719	0.6618
Student_t_3	6 wav+6 enr	0.9540	0.9299	0.8890	0.6218	0.8385	0.5855
Student_t_4	6 wav+6 freq	0.9498	0.9273	0.9107	0.8913	0.8462	0.5782
Student_t_5	6 wav+6 enr+2 benford	0.9568	0.9365	0.8994	0.7620	0.8482	0.7182
Student_t_6	6 wav+6 freq +2 benford	0.9296	0.9384	0.9196	0.8921	0.8374	0.8651
Student_t_7	6 wav+ 6 enr+ 6 frq	0.9528	0.9320	0.9183	0.8977	0.8539	0.8484
Student_t_8	6 wav + 14 DCT	0.9572	0.9421	0.9274	0.8991	0.8658	0.8827

Στην Student_t_1, παρατηρούμε πολύ χαμηλό Spearman στην κατηγορία all. Στην Student_t_2, η προσθήκη των 2 χαρακτηριστικών Benford αυξάνει ραγδαία αυτή την τιμή, αλλά αυξάνει τις τιμές και στις υπόλοιπες κατηγορίες. Στην κατηγορία white noise η προσθήκη των Benford προκαλεί μείωση. Στην Student_t_3, η προσθήκη των 6 χαρακτηριστικών, βασισμένων στην ενεργειακή αναλογία υποζώνης, στα χαρακτηριστικά του πειράματος Student_t_1 προκαλεί αύξηση του συντελεστή σε όλες τις κατηγορίες, ενώ, σε σχέση με το πείραμα που προσθέταμε τα 2 benford χαρακτηριστικά, έχουμε καλύτερη απόδοση στις κατηγορίες white noise, gaussian blur, jp2k και fast fading και χαμηλότερη στις jpeg και all.

Στην Student_t_4, η προσθήκη 6 χαρακτηριστικών, βασισμένων στην διακύμανση συχνοτήτων, στα χαρακτηριστικά του πειράματος Student_t_1 πετυχαίνει αύξηση της απόδοσης. Συγκριτικά με το πείραμα που προσθέταμε τα 2 benford χαρακτηριστικά, έχουμε αύξηση των τιμών στις σε όλες τις κατηγορίες εκτός από την κατηγορία all. Σε σχέση με το πείραμα που προσθέταμε στην θέση τους 6 χαρακτηριστικά βασισμένα στην ενεργειακή αναλογία υποζώνης, πετυχαίνουμε καλύτερα αποτελέσματα στις κατηγορίες jp2k, jpeg και fast fading.

Στην Student_t_5, αν στο πείραμα Student_t_3 προσθέσουμε και τα 2 χαρακτηριστικά Benford, πετυχαίνουμε ακόμα μεγαλύτερη βελτίωση από ότι είχαμε χωρίς την προσθήκη τους. Η αύξηση στην κατηγορία all είναι μεγάλου βαθμού. Συγκριτικά με τα υπόλοιπα πειράματα, τις μόνες τιμές που δεν καταφέρνει να ξεπεράσει, είναι αυτές στις κατηγορίες jp2k και jpeg του πειράματος Student_t_4.

Στην Student_t_6, αν προσθέσουμε τα 2 χαρακτηριστικά Benford στο πείραμα Student_t_4, παρατηρούμε αύξηση στις τιμές του συντελεστή Spearman στις κατηγορίες Gaussian blur, jp2k, jpeg και all. Η αύξηση στην κατηγορία all είναι μεγάλου βαθμού. Σε σχέση με τα υπόλοιπα πειράματα, υπερτερεί σε όλες τις κατηγορίες εκτός από το white noise και fast fading των πειραμάτων Student_t_3 και Student_t_5.

Στο πείραμα Student_t_7, έχουμε αύξηση τιμών στις κατηγορίες jpeg και fast fading σε σχέση με όλα τα άλλα πειράματα. Παρουσιάζει χαμηλότερη τιμή στην κατηγορία all και Gaussian blur σε σχέση με το Student_t_6, καθώς και μικρότερο Gaussian blur και white noise σε σχέση με το Student_t_5.

Στο πείραμα Student_t_8, που προστίθενται και τα 14 DCT χαρακτηριστικά, επιτυγχάνονται τα υψηλότερα αποτελέσματα σε όλες τις κατηγορίες.

Στο Σχήμα 4.9 εμφανίζουμε ένα scatter, που δείχνει την συσχέτιση των υποκειμενικών scores με τα αντίστοιχα αντικειμενικά scores που προέκυψαν από το μοντέλο με την υψηλότερη απόδοση (Student_t_8).



Σχήμα 4.9 Scatter Συσχέτισης μεταξύ των Υποκειμενικών και Αντικειμενικών Scores του Μοντέλου Student_t_8

4.1.2 Student's t με 2 παραμέτρους

• Logistic Pearson

Πίνακας 4.3 Συντελεστής Pearson Μοντέλων που Έχουν σαν Βάση Χαρακτηριστικά που Εξάγονται μέσω της Student-t Κατανομής με Εκτίμηση Δύο Παραμέτρων.

Κατανομή	Πλήθος	White	Gaussian	Infle	inor	Fast	A 11
Κατανομη	features	noise	Blur	јрак	jpeg	Fading	7111
Student t2 1	6	0 9514	0 9271	0 6414	0 6999	0 7407	0 4489
otudent_t2_1	wavelets	0.0014	0.0271	010111	0.0000	0.7107	0.1100
	6						
Student_t2_2	wav+2	0.9755	0.9322	0.8514	0.7978	0.7782	0.6841
	benford						
	6						
Student_t2_3	wav+6	0.9797	0.9118	0.8647	0.7212	0.8242	0.6191
	enr						
	6						
Student_t2_4	wav+6	0.9794	0.9448	0.9278	0.9193	0.8444	0.8526
	freq						
	6					0 8400	0.7096
Student +2 5	wav+6	0 9807	0 9258	0 8986	0.8079		
Student_t2_5	enr+2	0.3007	0.0200	0.0500	0.0075	0.0400	0.7030
	benford						
	6						
Student +2 6	wav+6	0 9823	0 9449	0 9320	0 9179	0.8476	0 8885
Student_t2_0	freq +2	0.0020	0.0442	0.0020	0.0170	0.0470	0.0000
	benford						
	6 wav+						
Student_t2_7	6 enr+	0.9814	0.9364	0.9305	0.9264	0.8738	0.8672
	6 frq						
Student +9 8	6 wav +	0 9827	0 9394	0 9373	0 9266	0.8800	0 8921
	14 DCT	0.0021	0.0004	0.0010	0.0200	0.0000	0.0021

Στο πείραμα Student_t2_1, χρησιμοποιούμε τους τύπους (4.2), (4.3) για να εξάγουμε τα χαρακτηριστικά εκπαίδευσης του SVM μοντέλου, υπολογίζοντας τις 2 παραμέτρους της κατανομής για τις εικόνες διαγώνιων, οριζόντιων και κάθετων συντελεστών, που προκύπτουν από τον μετασχηματισμό κυματιδίων. Έτσι εξάγουμε συνολικά 6 χαρακτηριστικά.

Στο Student_t2_2, προσθέτουμε στα χαρακτηριστικά του πρώτου πειράματος 2 χαρακτηριστικά Benford, τα οποία οδηγούν σε αύξηση του συντελεστή Pearson σε όλες τις κατηγορίες.

Στο πείραμα Student_t2_3, προσθέτουμε 6 χαραχτηριστικά της ενεργειαχής αναλογίας υποζώνης στα 6 χαραχτηριστικά που βασίζονται στα wavelets του πειράματος Student_t2_1. Η προσθήκη αυτή οδηγεί σε αύξηση της απόδοσης σε όλες τις κατηγορίες παραμορφώσεων εκτός της Gaussian Blur. Σε σχέση με το πείραμα Student_t2_2, έχει μεγαλύτερη τιμή Pearson στις κατηγορίες white noise, jp2k και fast fading, ενώ στις υπόλοιπες κατηγορίες έχει μικρότερη, παρά το αυξημένο πλήθος features.

Στο πείραμα Student_t2_4, προσθέτουμε αυτή την φορά 6 χαραχτηριστικά, βασισμένα στην διαχύμανση αναλογίας, στα 6 χαραχτηριστικά που βασίζονται στα wavelets του πειράματος Student_t2_1. Αυτή η προσθήκη επιφέρει αύξηση του συντελεστή σε όλες τις κατηγορίες, όπως αυξημένη απόδοση παρουσιάζει και συγκριτικά με το πείραμα Student_t2_2. Επίσης, συγκριτικά με τα αποτελέσματα του πειράματος Student_t2_3, μπορούμε να συμπεράνουμε, ότι τα χαρακτηριστικά που βασίζονται στην διακύμανση αναλογίας αποφέρουν καλύτερη απόδοση από αυτά που βασίζονται στην ενεργειακή αναλογία υποζώνης, αφού για ίδιο πλήθος χαρακτηριστικών παρουσιάζουν μεγαλύτερη τιμή Pearson σε όλες τις κατηγορίες.

Στο Student_t2_5, προσθέτουμε στα χαρακτηριστικά του πειράματος Student_t2_3 επιπλέον 2 χαρακτηριστικά Benford, τα οποία βοηθούν στην απόδοση, καθώς πετυχαίνουν αύξηση του συντελεστή σε όλες τις κατηγορίες και καταφέρνει να ξεπεράσει την απόδοση του πειράματος Student_t2_2, που χωρίς τα 2 χαρακτηριστικά δεν το κατάφερνε. Σε σχέση με το Student_t2_4, παρουσιάζει μικρότερη απόδοση, παρά το μεγαλύτερο πλήθος χαρακτηριστικών.

Στο πείραμα Student_t2_6, προσθέτουμε τα 2 χαρακτηριστικά Benford στα χαρακτηριστικά του πειράματος Student_t2_4. Η προσθήκη οδηγεί, κυρίως, σε αύξηση του συντελεστή Pearson. Στις κατηγορίες Gaussian blur και Jpeg έχουμε

μείωση, αν και μικρού βαθμού. Παρόλα αυτά, έχει καλύτερη απόδοση από το Student_t2_5, σαν συνέχεια του προηγούμενος συμπεράσματος ότι τα χαρακτηριστικά που βασίζονται στην διακύμανση αναλογίας αποφέρουν καλύτερη απόδοση από αυτά που βασίζονται στην ενεργειακή αναλογία υποζώνης.

Στο Student_t2_7, χρησιμοποιούμε, επιπλέον των χαρακτηριστικών που βασίζονται στα wavelets, τα 6 χαρακτηριστικά που βασίζονται στην ενεργειακή αναλογία υποζώνης και 6 που βασίζονται στην διακύμανση συχνοτήτων. Η αύξηση του πλήθους των χαρακτηριστικών δεν προκαλεί ανάλογη αύξηση του συντελεστή Pearson, ενώ στην περίπτωση του πειράματος Student_t2_6 έχει μικρότερη απόδοση στις κατηγορίες Gaussian blur, jpeg και all, παρόλο που έχει 4 χαρακτηριστικά περισσότερα από αυτό.

Στο Student_t2_8, προσθέτουμε και τα 14 DCT στα χαρακτηριστικά που βασίζονται στα wavelets του πειράματος Student_t2_1. Το μεγάλο πλήθος χαρακτηριστικών έχει αντίκτυπο στην απόδοση, καθώς ο συντελεστής Pearson είναι ο μεγαλύτερος από όλα τα προαναφερθέντα πειράματα σε όλες τις κατηγορίες εκτός από την κατηγορία Gaussian Blur του Student_t2_4.

Παρόλο που το Student_t2_8 παρουσιάζει καλύτερη απόδοση στην πλειοψηφία των κατηγοριών, το Student_t2_6 μπορεί να θεωρηθεί ένα πολύ ανταγωνιστικό μοντέλο, καθώς προσφέρει έναν συμβιβασμό μεταξύ του πλήθους των χαρακτηριστικών και της απόδοσης του μοντέλου. Με 6 χαρακτηριστικά λιγότερα από το Student_t2_8, πετυχαίνει παρόμοια απόδοση και μάλιστα στην κατηγορία Gaussian blur μεγαλύτερη.

• Spearman

Πίνακας 4.4 Συντελεστής Spearman Μοντέλων που Έχουν σαν Βάση Χαρακτηριστικά που Εξάγονται μέσω της Student-t Κατανομής με Εκτίμηση 2 Παραμέτρων.

Κατανομή	Πλήθος features	White noise	Gaussian Blur	Jp2k	Jpeg	Fast Fading	All
Student_t2_1	6	0.9589	0.9134	0.6016	0.6761	0.7060	0.0210
Student_t2_2	8	0.9787	0.9373	0.8093	0.7710	0.7395	0.6618
Student_t2_3	12	0.9784	0.9014	0.8413	0.7133	0.7974	0.5794
Student_t2_4	12	0.9790	0.9348	0.9031	0.8696	0.8158	0.8508
Student_t2_5	14	0.9790	0.9251	0.8803	0.7863	0.8161	0.6918
Student_t2_6	14	0.9813	0.9425	0.9140	0.8708	0.8205	0.8782
Student_t2_7	18	0.9806	0.9228	0.9125	0.8888	0.8457	0.8681
Student_t2_8	20	0.9808	0.9313	0.9220	0.8886	0.8524	0.8880

Στο Student_t2_1, χρησιμοποιούμε μόνο τα χαρακτηριστικά που εξάγονται μέσω των τύπων (4.2), (4.3), για τις εικόνες συντελεστών λεπτομέρειας που προκύπτουν από το μετασχηματισμό κυματιδίων.

Στο Student_t2_2, προστίθενται στα παραπάνω χαρακτηριστικά 2 χαρακτηριστικά Benford, τα οποία οδηγούν σε αύξηση του συντελεστή Spearman σε όλες τις κατηγορίες παραμορφώσεων, κυρίως στην κατηγορία all.

Στο Student_t3_3, προστίθενται στα χαρακτηριστικά του πρώτου πειράματος 6 χαρακτηριστικά, βασισμένα στην ενεργειακή αναλογία υποζώνης. Ο συντελεστής Spearman αυξάνεται, αλλά όχι σε τόσο μεγάλο βαθμό όσο στο Student_t2_2, καθώς μόνο στις κατηγορίες jpek και fast fading καταφέρνει να το ξεπεράσει.

Στο πείραμα Student_t2_4, προσθέτουμε στα χαρακτηριστικά του Student_t2_1 6 χαρακτηριστικά διακύμανσης συχνοτήτων. Ο συντελεστής Spearman αυξάνεται σε όλες τις κατηγορίες, συγκριτικά με όλα τα προαναφερθέντα πειράματα. Επιπλέον, μπορούμε να εξάγουμε σαν συμπέρασμα, ότι τα χαρακτηριστικά διαχύμανσης συχνοτήτων συνεισφέρουν περισσότερο στην βελτίωση της απόδοσης, από ότι συνεισφέρουν τα χαραχτηριστιχά ενεργειαχής αναλογίας υποζώνης που δοχιμάσαμε στο πείραμα Student_t2_3.

Στο πείραμα Student_t2_5, προσθέσαμε 2 χαρακτηριστικά Benford στο πείραμα Student_t2_3, τα οποία αυξάνουν την απόδοση του μοντέλου, αλλά δεν καταφέρνει να ξεπεράσει αυτό του πειράματος Student_t2_4.

Στο πείραμα Student_t2_6, προσθέτουμε 2 χαρακτηριστικά Benford στο πείραμα Student_t2_4, τα οποία βελτιώνουν ακόμα περισσότερο την ήδη υψηλή απόδοσή του.

Στο Student_t2_7, προσθέτουμε 6 στα χαρακτηριστικά του πειράματος Student_t2_1, 6 χαρακτηριστικά ενεργειακής αναλογίας υποζώνης και 6 χαρακτηριστικά διακύμανσης συχνοτήτων. Η αύξηση της απόδοσης είναι αποτέλεσμα, σε αυτήν την περίπτωση, της αύξησης του πλήθους των χαρακτηριστικών και έτσι ξεπερνά τα προαναφερθέντα πειράματα που χρησιμοποιούσαν λιγότερα χαρακτηριστικά. Εξαίρεση αποτελεί το πείραμα Student_t2_7, που με 4 λιγότερα fearures παρουσιάζει μεγαλύτερη απόδοση στις κατηγορίες white noise, Gaussian blur και all.

Στο Student_t2_8, προσθέτουμε και τα 14 DCT χαρακτηριστικά στα 6 χαρακτηριστικά που βασίζονται στα wavelets. Αυτή η προσθήκη αυξάνει τον συντελεστή Spearman των κατηγοριών jp2k, jpeg, fast fading και all, σε βαθμό που να ξεπερνά όλα τα προαναφερθέντα πειράματα. Δεν καταφέρνει να ξεπεράσει την απόδοση που επιτυγχάνεται στις κατηγορίες white noise και gaussian blur του πειράματος Student_t2_6.

Το πείραμα Student_t2_8, είναι αρχετά αποδοτικό αν έχουμε σαν κριτήριο τις υψηλές τιμές Spearman. Όμως, αν λάβουμε σαν συμπληρωματικό κριτήριο και το πλήθος των χαρακτηριστικών, τότε πολύ ανταγωνιστικό μοντέλο είναι και το Student_t2_6, το οποίο με 6 λιγότερα χαρακτηριστικά προσεγγίζει, και σε ορισμένες παραμορφώσεις ξεπερνά, τις τιμές του Student_t2_8.

Στο Σχήμα 4.10 εμφανίζουμε ένα scatter, που δείχνει την συσχέτιση των υποχειμενιχών scores με τα αντίστοιχα αντιχειμενιχά scores που προέχυψαν από το μοντέλο με την υψηλότερη απόδοση (Student_t2_8).



Σχήμα 4.10 Scatter Συσχέτισης μεταξύ των Υποκειμενικών και Αντικειμενικών Scores του Μοντέλου Student_t2_8

4.1.3 Student's t με 4 παραμέτρους

Logistic Pearson

Πίνακας 4.5 Συντελεστής Pearson Μοντέλων που Έχουν σαν Βάση Χαρακτηριστικά που Εξάγονται μέσω της Student-t Κατανομής με Εκτίμηση Τεσσάρων Παραμέτρων.

Κατανομή	Πλήθος	White	Gaussian	In9k	Inor	Fast	A 11
Κατανομη	features	noise	Blur	јрак	Jpeg	Fading	All
Student t4 1	12	0 9498	0 9231	0 6998	0.8170	0 7721	0 4533
orddent_t1_1	wavelets	0.0100	0.0201	0.0000	0.0170	0.7721	0.1000
	12						
Student_t4_2	wav+2	0.9709	0.9358	0.8588	0.8273	0.7871	0.7082
	benford						
	12						
Student_t4_3	wav+6	0.9809	0.9157	0.8708	0.7976	0.8318	0.6329
	enr						
	12						
Student_t4_4	wav+6	0.9807	0.9448	0.9276	0.9289	0.8416	0.8582
	freq						
	12						
Student t4 5	wav+6	0.9823	0 9292	0 9019	0.8355	0.8413	0 7157
orddent_t1_0	enr+2	0.0020	0.0202	0.0010	0.0000	0.0110	0.7107
	benford						
	12						
Student t4 6	wav+6	0 9829	0 9461	0 9324	0 9258	0.8452	0 8904
otudent_t1_0	freq +2	0.0020	0.0101	0.0021	0.0200	0.0102	0.0001
	benford						
	12 wav+						
Student_t4_7	6 enr+	0.9823	0.9395	0.9306	0.9323	0.8714	0.8704
	6 frq						
	12 wav						
Student_t4_8	+ 14	0.9839	0.9408	0.9366	0.9305	0.8742	0.8933
	DCT						

Στο πείραμα Student_t4_1, χρησιμοποιούμε τους τύπους (4.2), (4.3) για να εξάγουμε τα χαρακτηριστικά για την μάθηση του SVM μοντέλου, υπολογίζοντας τις 4 παραμέτρους της κατανομής για τις εικόνες διαγώνιων, οριζόντιων και κάθετων συντελεστών, που προκύπτουν από τον μετασχηματισμό κυματιδίων. Με αυτό τον τρόπο εξάγουμε 12 χαρακτηριστικά.

Στο πείραμα Student_t4_2, προσθέτουμε επιπλέον στα 12 χαρακτηριστικά 2 Benford, τα οποία προκαλούν αύξηση του συντελεστή σε όλες τις κατηγορίες.

Στο πείραμα Student_t4_3, στα χαρακτηριστικά του πειράματος Student_t4_1 γίνεται προσθήκη 6 χαρακτηριστικών βασισμένων στην ενεργειακή αναλογία υποζώνης. Αυτή η προσθήκη αυξάνει την απόδοση, αλλά όχι σε τόσο μεγάλο βαθμό όσο κατάφεραν τα χαρακτηριστικά Benford στο πείραμα Student_t4_2. Οι κατηγορίες στις οποίες πέτυχε μεγαλύτερη απόδοση είναι η white noise, jp2k και fast fading.

Στο Student_t4_4, προσθέτουμε 6 χαρακτηριστικά βασισμένα στην διακύμανση συχνοτήτων. Αυτή η προσθήκη επιφέρει μεγαλύτερη απόδοση από ότι η προσθήκη των χαρακτηριστικών που βασίζονται στην ενεργειακή αναλογία υποζώνης, καθώς η αύξηση του συντελεστή είναι τέτοιου βαθμού, που ξεπερνάει τις τιμές του πειράματος Student_t4_2, αυτήν την φορά, αλλά και του πειράματος Student_t4_3

Στο πείραμα Student_t4_5, προσθέτουμε στα χαρακτηριστικά του πειράματος Student_t4_3 επιπλέον 2 χαρακτηριστικά Benford, η οποία προσθήκη αυξάνει το συντελεστή Pearson σε όλες τις κατηγορίες, σε σημείο να ξεπερνούν, αυτή την φορά, τις αντίστοιχες τιμές του πειράματος Student_t4_2. Πάλι δεν φτάνει την τιμή της κατηγορίας Gaussian Blur. Συγκριτικά με το πείραμα Student_t4_4 παρουσιάζει χαμηλότερη απόδοση, παρά το γεγονός ότι ξεπερνάει το πλήθος των χαρακτηριστικών του κατά δύο.

Στο πείραμα Student_t4_6, προσθέτουμε στα χαρακτηριστικά του πειράματος Student_t4_4 επιπλέον 2 χαρακτηριστικά Benford. Η προσθήκη αυτή αυξάνει τις τιμές του συντελεστή όλων των κατηγοριών εκτός της Jpeg. Σε σχέση με το Student_t4_5, που έχει το ίδιο πλήθος χαρακτηριστικών, παρουσιάζει μεγαλύτερη απόδοση, καθιστώντας τα χαρακτηριστικά διακύμανσης συχνοτήτων καλύτερα αυτών της ενεργειακής υποζώνης συχνοτήτων.

Στο πείραμα Student_t4_7, προσθέτουμε τα 6 χαρακτηριστικά βασισμένα στην διακύμανση συχνοτήτων και τα 6 χαρακτηριστικά βασισμένα στην ενεργειακή αναλογία υποζώνης, επί πλέον των χαρακτηριστικών του πειράματος Student_t4_1. Το πείραμα είναι καλύτερο από προαναφερθέντα πειράματα, καθώς παρουσιάζει υψηλότερες τιμές του συντελεστή Pearson, αν και με αρκετά περισσότερα χαρακτηριστικά. Το Student_t4_6 παραμένει, όμως, αρκετά ανταγωνιστικό, καθώς με 4 λιγότερα χαρακτηριστικά παρουσιάζει υψηλότερες τιμές στην πλειοψηφία των κατηγοριών. Στις κατηγορίες jpeg και fast fading έχει χαμηλότερες τιμές.

Στο πείραμα Student_t4_8, προσθέτουμε στα χαρακτηριστικά του Student_t4_1 και τα 14 DCT χαρακτηριστικά. Αυτή η αύξηση στο πλήθος των χαρακτηριστικών επιφέρει αύξηση της απόδοσης στις περισσότερες κατηγορίες με εξαίρεση τις Gaussian Blur και jpeg.

Αν και καλύτερο μοντέλο, ως προς το πλήθος των κατηγοριών που πετυχαίνει μεγαλύτερο Pearson, είναι το Student_t4_8, αν συνυπολογίσουμε για την επιλογή και το πλήθος των χαρακτηριστικών, τα μοντέλα Student_t4_6 και Student_t4_7 είναι επίσης πολύ ανταγωνιστικά.

• Spearman

Πίνακας 4.6 Συντελεστής Speaman Μοντέλων που Έχουν σαν Βάση Χαρακτηριστικά που Εξάγονται μέσω της Student-t Κατανομής με Εκτίμηση Τεσσάρων Παραμέτρων.

Κατανομή	Πλήθος features	White noise	Gaussian Blur	Jp2k	jpeg	Fast Fading	All
Student_t4_1	12	0.9567	0.9109	0.6683	0.7879	0.7394	0.1892
Student_t4_2	14	0.9749	0.9363	0.8240	0.7964	0.7515	0.6896
Student_t4_3	18	0.9808	0.9053	0.8508	0.7806	0.8031	0.6059
Student_t4_4	18	0.9794	0.9328	0.9042	0.8821	0.8173	0.8594
Student_t4_5	20	0.9821	0.9281	0.8840	0.8090	0.8168	0.7028
Student_t4_6	20	0.9818	0.9391	0.9142	0.8823	0.8197	0.8840
Student_t4_7	24	0.9813	0.9266	0.9147	0.8971	0.8461	0.8735
Student_t4_8	26	0.9822	0.9324	0.9219	0.8979	0.8521	0.8903

Στο Student_t4_1, έχουμε τα 12 χαρακτηριστικά που εξάγουμε από τους τύπυς (4.2) και (4.3) για τους συντελεστές λεπτομερειών του μετασχηματισμού κυματιδίων.

Στο Student_t4_2, με την προσθήκη 2 χαρακτηριστικών Benford, ο συντελεστής Spearman αυξάνεται σε όλες τις κατηγορίες.

Στο Student_t4_3, με την προσθήκη 6 χαρακτηριστικών ενεργειακής αναλογίας υποζώνης, αυξάνει επίσης τον συντελεστή, σε σχέση με όταν δεν χρησιμοποιούνταν. Στις κατηγορίες Gaussian blur, jpeg και all τα Benford χαρακτηριστικά απέδωσαν καλύτερα, αν και λιγότερα σε πλήθος.

Στο Student_t4_4, με την προσθήκη 6 χαρακτηριστικών διακύμανσης συχνοτήτων, έχουμε μεγαλύτερη αύξηση του συντελεστή, από ότι είχαμε με τις προηγούμενες προσθήκες, εκτός από το white noise που είχαμε με την προσθήκη των χαρακτηριστικών ενεργειακής υποζώνης και το Gaussian Blur με την προσθήκη των 2 benford.

Στο Student_t4_5, προσθέτουμε τα 2 Benford χαρακτηριστικά στα χαρακτηριστικά του πειράματος Student_t4_3, τα οποία βελτιώνουν την απόδοσή του, αλλά συγκριτικά με το πείραμα Student_t4_4 έχει χειρότερα αποτελέσματα, εκτός από την κατηγορία white noise που έχει υψηλότερη τιμή Spearman.

Στο Student_t4_6, προσθέτουμε τα 2 Benford χαρακτηριστικά στα χαρακτηριστικά του πειράματος Student_t4_4, τα οποία αυξάνουν τις τιμές Spearman σε όλες τις κατηγορίες. Σημειώνει τη υψηλότερη απόδοση από τα προηγούμενα μοντέλα, εκτός από την κατηγορία white noise που δεν καταφέρνει να ξεπεράσει αυτή του Student_t4_5, αν και για μικρή διαφορά.

Στο Student_t4_7, προσθέτουμε στα χαρακτηριστικά του πειράματος Student_t2_1 6 χαρακτηριστικά ενεργειακής αναλογίας υποζώνης και 6 χαρακτηριστικά διακύμανσης συχνοτήτων. Η αύξηση των features αποδίδει συγκριτικά με την πλειοψηφία των πειραμάτων. Το Student_t4_6, με 4 λιγότερα χαρακτηριστικά, προσεγγίζει πολύ τις τιμές του Student_t2_7 και στις κατηγορίες Gaussian blur και all τις ξεπερνά.

Στο Student_t4_8, προσθέτουμε στα χαρακτηριστικά που βασίζονται στα wavelets και τα 14 DCT χαρακτηριστικά. Αυτή η προσθήκη αποφέρει στην απόδοση του μοντέλου, καθώς η μόνη κατηγορία στην οποία δεν επιτυγχάνει την υψηλότερη

απόδοση είναι η Gaussian Blur, στην οποία χαλύτερη απόδοση επιτυγχάνει το πείραμα Student_t2_6.

Στο Σχήμα 4.11 εμφανίζουμε ένα scatter, που δείχνει την συσχέτιση των υποκειμενικών scores με τα αντίστοιχα αντικειμενικά scores που προέκυψαν από το μοντέλο με την υψηλότερη απόδοση (Student_t4_8).



Σχήμα 4.11 Scatter Συσχέτισης μεταξύ των Υποκειμενικών και Αντικειμενικών Scores (με Χρήση Κατανομής Student-t με Εκτίμηση 4 Παραμέτρων)

4.1.4 Cauchy

• Logistic Pearson Cauchy

Πίνακας 4.7 Συντελεστής Pearson Μοντέλων που Έχουν σαν Βάση Χαρακτηριστικά που Εξάγονται μέσω της Cauchy Κατανομής.

V	Πλήθος	White	Gaussian	L- 91-	·	Fast	A 11
κατανομη	features	noise	Blur	Јр2к	jpeg	Fading	All
Cauchy _1	3 wavelets	0.9375	0.8759	0.7616	0.7096	0.7846	0.3469
Cauchy _2	3 wav+2 benford	0.9762	0.9033	0.8679	0.8168	0.7792	0.6412
Cauchy _3	3 wav+6 enr	0.9745	0.8734	0.8772	0.7515	0.8044	0.6191
Cauchy _4	3 wav+6 freq	0.9735	0.9316	0.9285	0.9176	0.8509	0.8271
Cauchy _5	3 wav+6 enr+2 benford	0.9773	0.9098	0.9036	0.8228	0.8221	0.6612
Cauchy_6	3 wav+6 freq +2 benford	0.9813	0.9365	0.9324	0.9144	0.8521	0.8894
Cauchy _7	3 wav+ 6 enr+ 6 frq	0.9775	0.9235	0.9312	0.9249	0.8682	0.8554
Cauchy _8	3 wav + 14 DCT	0.9819	0.9359	0.9368	0.9246	0.8798	0.8957

Στο πείραμα Cauchy_1, κάνοντας χρήση του τύπου (4.5), υπολογίζουμε για κάθε εικόνα διαγώνιων, οριζόντιων και κάθετων συντελεστών, που προκύπτουν από το μετασχηματισμό κυματιδίων, την παράμετρο της κατανομής Cauchy. Με αυτό τον τρόπο εξάγουμε 1 χαρακτηριστικό για κάθε μια από τις εικόνες λεπτομερειών, δηλαδή 3 χαρακτηριστικά. Παρατηρούμε ότι στην κατηγορία all ο συντελεστής

Pearson είναι πολύ χαμηλός, ενώ στις άλλες κατηγορίες ο συντελεστής είναι αρκετά υψηλός.

Έπειτα, στο Cauchy_2, προσθέτουμε 2 Benford χαρακτηριστικά σε αυτά του πειράματος Cauchy_1, το οποίο προκαλεί αύξηση του συντελεστή, εκτός της κατηγορίας fast fading.

Στο πείραμα Cauchy_3, προσθέτουμε στα χαρακτηριστικά του Cauchy_1 6 χαρακτηριστικά βασισμένα στην ενεργειακή αναλογία υποζώνης. Έχουμε συνολικά 3+6=9 χαρακτηριστικά. Αυτή η προσθήκη επιφέρει αύξηση του συντελεστή σε όλες τις κατηγορίες, εκτός της Gaussian Blur. Συγκριτικά με το Cauchy_2, που προσθέταμε λιγότερα features, δεν έχουμε ανάλογη αύξηση, καθώς με 4 χαρακτηριστικά παραπάνω από αυτή ο συντελεστής είναι υψηλότερος μόνο στις κατηγορίες jp2k και fast fading.

Αντιθέτως, αν στα χαραχτηριστικά του Cauchy_1 προσθέσουμε 6 χαραχτηριστικά, αυτή την φορά βασισμένα στην διαχύμανση συχνοτήτων, έχουμε καλύτερα αποτελέσματα, καθώς τα 4 επιπλέον features επιφέρουν αύξηση στο πείραμα Cauchy_2, εκτός από την κατηγορία white noise. Οπότε, και συγκριτικά με το Cauchy_3, με το οποίο χρησιμοποιούν ίδιο πλήθος χαρακτηριστικών, παρουσιάζει καλύτερη απόδοση σε όλες τις κατηγορίες, εκτός από την κατηγορία white noise.

Στο Cauchy_5, προσθέτουμε 2 χαρακτηριστικά Benford στα 9 χαρακτηριστικά του πειράματος Cauchy_3. Αυτή η προσθήκη βοηθάει πολύ στην απόδοση, καθώς έχουμε αύξηση του συντελεστή σε όλες τις κατηγορίες. Το Cauchy_4 παραμένει καλύτερο, παρά το μικρότερο πλήθος features, εκτός από την κατηγορία white noise στην οποία παρουσιάζει χαμηλότερη τιμή Pearson.

Στο Cauchy_6, προσθέτουμε 2 χαρακτηριστικά Benford στα 9 χαρακτηριστικά του πειράματος Cauchy_4. Στην πλειοψηφία των κατηγοριών παρουσιάζει αύξηση του συντελεστή. Στην μόνη που μειώνεται ο συντελεστής, είναι η κατηγορία jpeg.

Στο πείραμα Cauchy_7, χρησιμοποιούμε τα 3 χαρακτηριστικά που εξάγουμε μέσω των συντελεστών λεπτομέρειας του μετασχηματισμού κυματιδίων, μαζί με τα 6 χαρακτηριστικά που βασίζονται στην ενεργειακή αναλογία υποζώνης και τα 6 που βασίζονται στην διακύμανση συχνοτήτων. Συνολικά έχουμε 3+6+6=15 χαρακτηριστικά. Έχει καλύτερη απόδοση από τα προαναφερθέντα πειράματα,

αλλά δεν φτάνει αυτή του πειράματος Cauchy_6. Η μόνη κατηγορία, στην οποία το ξεπερνάει, είναι του fast fading, αν και έχει 6 χαρακτηριστικά περισσότερα.

Τέλος, στο Cauchy_8 προσθέτουμε και τα 14 DCT χαρακτηριστικά στα 3 που βασίζονται στα wavelets. Η προσθήκη αυτή επιφέρει αύξηση του συντελεστή στην πλειοψηφία των κατηγοριών, αν και όχι ανάλογη του πλήθους, και επιτυγχάνει μικρότερη απόδοση από αυτή του πειράματος Cauchy_6 στην κατηγορία Gaussian blur και παρόμοια με αυτή του πειράματος Cauchy_7 στην κατηγορία Jpeg.

Συμπερασματικά, ναι μεν στο πείραμα Cauchy_8 έχουμε την καλύτερη απόδοση, καθώς έχουμε την μεγαλύτερη τιμή Pearson στην πλειοψηφία των κατηγοριών, αλλά αν λάβουμε υπόψη και το πλήθος των χαρακτηριστικών, τότε το Cauchy_6 έχει πλεονέκτημα, καθώς με τα μισά σχεδόν χαρακτηριστικά πετυχαίνει απόδοση πολύ κοντά σε αυτή του Cauchy_8, ενώ στην κατηγορία Gaussian blur την ξεπερνάει.

• Spearman

Πίνακας 4.8 Συντελεστής Spearman Μοντέλων που Έχουν σαν Βάση Χαρακτηριστικά που Εξάγονται μέσω της Cauchy Κατανομής.

Κατανοικά	Πλήθος	White	Gaussian	Infle	inor	Fast	A 11
κατανομη	features	noise	Blur	јрак	Jpcg	Fading	All
Cauchy_1	3	0.9458	0.8755	0.7222	0.6758	0.7523	0.4776
Cauchy_2	5	0.9810	0.9127	0.8359	0.7780	0.7446	0.6206
Cauchy_3	9	0.9742	0.8624	0.8617	0.7132	0.7711	0.5663
Cauchy_4	9	0.9758	0.9191	0.9074	0.8607	0.8212	0.8264
Cauchy_5	11	0.9775	0.9147	0.8871	0.7857	0.8027	0.6343
Cauchy_6	11	0.9806	0.9338	0.9153	0.8658	0.8214	0.8821
Cauchy_7	15	0.9775	0.9039	0.9139	0.8878	0.8434	0.8612
Cauchy_8	17	0.9806	0.9300	0.9204	0.8894	0.8545	0.8897

Στο Cauchy_1, χρησιμοποιούμε τα 3 χαρακτηριστικά που βασίζονται στα wavelets.

Στο Cauchy_2, με την προσθήκη 2 χαρακτηριστικών Benford, αυξάνεται ο συντελεστής Spearman σε όλες τις κατηγορίες.

Στο Cauchy_3, με την προσθήκη των 6 χαρακτηριστικών ενεργειακής αναλογίας υποζώνης, έχουμε επίσης αύξηση. Συγκριτικά με το Cauchy_2, παρουσιάζει μικρότερες τιμές Spearman, εκτός από τις κατηγορίες jp2k και fast fading.

Στο Cauchy_4, με την προσθήκη των 6 χαρακτηριστικών διακύμανσης συχνοτήτων, η αύξηση της απόδοσης είναι μεγαλύτερη από ό,τι κατάφεραν οι προηγούμενες προσθήκες. Στη μόνη κατηγορία που παρουσιάζει μικρότερη απόδοση, είναι η white noise του πειράματος Cauchy_2.

Στο Cauchy_5, προσθέτουμε 2 χαρακτηριστικά Benford στα 9 χαρακτηριστικά του πειράματος Cauchy_3, βελτιώνοντας έτσι την τιμή Spearman σε κάθε κατηγορία. Το μόνο πείραμα που δεν ξεπερνά είναι το Cauchy_4, παρόλο που χρησιμοποιεί 2 χαρακτηριστικά παραπάνω.

Στο Cauchy_6, προσθέτουμε 2 χαρακτηριστικά Benford στα 9 χαρακτηριστικά του πειράματος Cauchy_4, βελτιώνοντας ακόμα περισσότερα την υψηλή απόδοση του μοντέλου.

Στο Cauchy_7, χρησιμοποιούμε τα 3 χαραχτηριστικά που εξάγουμε μέσω των συντελεστών λεπτομέρειας του μετασχηματισμού κυματιδίων, μαζί με τα 6 χαραχτηριστικά που βασίζονται στην ενεργειακή αναλογία υποζώνης και τα 6 που βασίζονται στην διακύμανση συχνοτήτων. Οι τιμές Spearman αυξάνονται και ειδικά στις κατηγορίες jpeg, fast fading και all είναι τέτοια η αύξηση, που ξεπερνάει όλα τα προαναφερθέντα πειράματα.

Στο Cauchy_8, προσθέτουμε και τα 14 χαρακτηριστικά DCT. Στις κατηγορίες jp2k, jpeg, fast fading και all πετυχαίνει την υψηλότερη απόδοση από όλα τα προηγούμενα πειράματα. Συνεχίζει να ισχύει το συμπέρασμα που εξήγαμε από τον συντελεστή Pearson, δηλαδή μπορεί το Cauchy_8 να πετυχαίνει την υψηλότερη απόδοση στην πλειοψηφία των παραμορφώσεων, αλλά το Cauchy_6 έχει πλεονέκτημα, καθώς με τα μισά σχεδόν χαρακτηριστικά πετυχαίνει απόδοση πολύ κοντά σε αυτή του Cauchy_8.

Στο Σχήμα 4.12 εμφανίζουμε ένα scatter, που δείχνει την συσχέτιση των υποχειμενιχών scores με τα αντίστοιχα αντιχειμενιχά scores που προέχυψαν από το μοντέλο με την υψηλότερη απόδοση (Cauchy_8).



Σχήμα 4.12 Scatter Συσχέτισης μεταξύ των Υποκειμενικών και Αντικειμενικών Scores (με χρήση Κατανομής Cauchy)

4.2 DATASET C

Στην βάση δεδομένων C έχουμε 1069 εικόνες που έχουν υποστεί αυθεντικές παραμορφώσεις, δηλαδή παραμορφώσεις οι οποίες προκλήθηκαν κατά την λήψη και αποθήκευση φωτογραφιών από κινητά τηλέφωνα και προκλήθηκαν είτε από περιβαλλοντικούς παράγοντες (π.χ. ακτινοβολία ήλιου) είτε από την ποιότητα της συσκευής.

4.2.1 Student's t κατανομή με EM εκτίμηση παραμέτρων

Πίνακας 4.9 Συντελεστής Pearson και Spearman Μοντέλων που Έχουν σαν Βάση Χαρακτηριστικά που Εξάγονται μέσω της Student-t Κατανομής ΕΜ Εκτίμηση Παραμέτρων.

Κατανομή	Πλήθος features	Pearson	spearman
Student_t_1	6	0.4730	0.4822
Student_t_2	8	0.4783	0.5001
Student_t_3	12	0.4830	0.5019
Student_t_4	12	0.4585	0.4729
Student_t_5	14	0.5045	0.5214
Student_t_6	14	0.4731	0.4989
Student_t_7	18	0.4778	0.4933
Student_t_8	20	0.5066	0.5185

Στην Student_t_1, με 6 μόλις χαρακτηριστικά επιτυγχάνεται αρκετά υψηλή απόδοση.

Στην Student_t_2, η προσθήκη 2 επίπλέον features Benford αυξάνει τις τιμές και των δύο συντελεστών.

Στην Student_t_3, αν στα χαραχτηριστικά που βασίζονται στα wavelets του πειράματος Student_t_1 προσθέσουμε 6 χαραχτηριστικά βασισμένα σε ενεργειακή αναλογία υποζώνης, έχουμε αύξηση των συντελεστών, αλλά όχι ανάλογη του πλήθους των features που χρησιμοποιούνται. Αυτό φαίνεται και αν το συγκρίνουμε με την Student_t_2, που με την προσθήκη μόλις 2 χαρακτηριστικών οι τιμές της απόδοσης είναι πολύ κοντά.

Στο Student_t_4, προσθέτουμε στα χαρακτηριστικά της Student_t_1 πάλι 6 χαρακτηριστικά, αυτή την φορά βασισμένα στη διακύμανση συχνοτήτων. Η προσθήκη τους προκαλεί μείωση της απόδοσης, παρά την αύξηση του πλήθους των χαρακτηριστικών. Οπότε, παρατηρούμε ότι τα 6 χαρακτηριστικά ενεργειακής αναλογίας υποζώνης αποδίδουν καλύτερα από τα 6 της διακύμανσης συχνοτήτων.

Στο Student_t_5, προσθέτουμε στα χαρακτηριστικά του πειράματος Student_t_3 2 χαρακτηριστικά Benford. Και οι δύο συντελεστές αυξάνονται σημαντικά.

Στο Student_t_6, προσθέτουμε στα χαρακτηριστικά του πειράματος Student_t_4 πάλι 2 χαρακτηριστικά Benford. Τα 2 χαρακτηριστικά Benford αυξάνουν την απόδοση του πειράματος. Συγκριτικά με το Student_t_5, επιτυγχάνει αρκετά μειωμένη απόδοση, παρά το ίδιο πλήθος χαρακτηριστικών.

Στο Student_t_7, προσθέτουμε στα χαρακτηριστικά του πειράματος Student_t_1 6 χαρακτηριστικά βασισμένα σε ενεργειακή αναλογία υποζώνης και 6 χαρακτηριστικά βασισμένα στη διακύμανση συχνοτήτων. Η απόδοση δεν φτάνει στο επίπεδο προηγουμένων πειραμάτων, παρά το αυξημένο πλήθος χαρακτηριστικών που χρησιμοποιεί.

Τέλος, στο Student_t_8, χρησιμοποιούνται επιπλέον των χαρακτηριστικών που βασίζονται στα wavelets και τα 14 DCT χαρακτηριστικά. Η απόδοση αυξάνεται σε σχέση με τα υπόλοιπα πειράματα, όσον αφορά τον συντελεστή Pearson, αλλά ο συντελεστής Spearman είναι μικρότερος αυτού του πειράματος Student_t_5.

Μεταξύ του πειράματος Student_t_5 και του πειράματος Student_t_8 καλύτερο μοντέλο είναι το Student_t_5, καθώς πετυχαίνει υψηλή απόδοση με μικρό πλήθος χαρακτηριστικών.

4.2.2 Student's t με 2 παραμέτρους

Πίνακας 4.10 Συντελεστής Pearson και Spearman Μοντέλων που Έχουν σαν Βάση Χαρακτηριστικά που Εξάγονται μέσω της Student-t Κατανομής με Εκτίμηση Δύο Παραμέτρων.

Κατανομή	Πλήθος features	Pearson	Spearman
Student_t2_1	6	0.4760	0.4454
Student_t2_2	8	0.4611	0.4150
Student_t2_3	12	0.4904	0.4467
Student_t2_4	12	0.4372	0.4026
Student_t2_5	14	0.4830	0.4347
Student_t2_6	14	0.4533	0.4160
Student_t2_7	18	0.4636	0.4264
Student_t2_8	20	0.4884	0.4451

Στο Student_t2_1, χρησιμοποιούμε για χαρακτηριστικά μάθησης 6 χαρακτηριστικά που εξάγονται μέσω των τύπων (4.2), (4.3) για τις εικόνες διαγώνιων, οριζοντίων και κάθετων συντελεστών, που προκύπτουν από τον μετασχηματισμό κυματιδίων.

Στο Student_t2_2, προσθέτουμε 2 χαρακτηριστικά Benford, τα οποία μειώνουν την απόδοση του Student_t2_1.

Στο πείραμα Student_t2_3, προσθέτουμε 6 χαρακτηριστικά ενεργειακής αναλογίας υποζώνης στα χαρακτηριστικά του πρώτου πειράματος. Αυτή η προσθήκη επιφέρει αύξηση της απόδοσης, καθώς και οι δύο συντελεστές αυξάνονται.

Στο Student_t2_4, η προσθήκη 6 χαρακτηριστικών διακύμανσης συχνοτήτων δεν έχει την ίδια επιρροή στους συντελεστές, καθώς μειώνουν την απόδοση του πειράματος Student_t2_1.

Στο Student_t2_5, προσθέτουμε στα χαρακτηριστικά του πειράματος Student_t2_3 2 χαρακτηριστικά Benford, τα οποία ρίχνουν την υψηλή απόδοση που

πέτυχε το μοντέλο χωρίς αυτά. Συγκριτικά με τα υπόλοιπα πειράματα, παρουσιάζει καλύτερα αποτελέσματα.

Στο Student_t2_6, προσθέτουμε στα χαρακτηριστικά του πειράματος Student_t2_4 2 χαρακτηριστικά Benford, τα οποία αυξάνουν την απόδοσή του, αλλά όχι σε τόσο μεγάλο βαθμό, ώστε να θεωρηθεί ανταγωνιστικό ως προς άλλα πειράματα.

Στο πείραμα Student_t2_7, προσθέτουμε στα χαρακτηριστικά του Student_t2_3 6 χαρακτηριστικά διακύμανσης συχνοτήτων, τα οποία φαίνεται να επηρεάζουν αρνητικά την επίδοση που πετύχαινε χωρίς αυτά.

Στο πείραμα Student_t2_8, προσθέτουμε στο πείραμα Student_t2_7 άλλα 2 χαρακτηριστικά Benford, τα οποία αυξάνουν την απόδοση. Το πείραμα, αν και χρησιμοποιεί μεγάλος πλήθος χαρακτηριστικών, δεν πετυχαίνει αρκετά υψηλή απόδοση, ειδικά αν το συγκρίνουμε με το πείραμα Student_t2_3, που χρησιμοποιεί 8 χαρακτηριστικά λιγότερα και επιτυγχάνει υψηλότερη απόδοση.

4.2.3 Student's t με 4 παραμέτρους

Πίνακας 4.11 Συντελεστής Pearson και Spearman Μοντέλων που έχουν σαν Βάση Χαρακτηριστικά που Εξάγονται μέσω της Student-t Κατανομής με Εκτίμηση Τεσσάρων Παραμέτρων.

Κατανομή	Πλήθος features	Pearson	Spearman
Student_t4_1	12	0.5173	0.4819
Student_t4_2	14	0.5103	0.4685
Student_t4_3	18	0.5178	0.4728
Student_t4_4	18	0.4896	0.4597
Student_t4_5	20	0.5201	0.4759
Student_t4_6	20	0.5037	0.4763
Student_t4_7	24	0.4961	0.4647
Student_t4_8	26	0.5206	0.4854

Στο Student_t4_1, παρουσιάζονται οι τιμές των συντελεστών Pearson και Spearman για 12 χαρακτηριστικά που εξάγονται μέσω των τύπων (4.2), (4.3) για τους συντελεστές λεπτομέρειας, που προκύπτουν από το μετασχηματισμό κυματιδίων της εικόνας. Στο πείραμα Student_t4_2, γίνεται προσθήκη 2 χαρακτηριστικών Benford, η οποία προσθήκη δεν βοηθά στη απόδοση του μοντέλου, καθώς μειώνονται και οι δύο συντελεστές μέτρησης.

Στο πείραμα Student_t4_3, προσθέτουμε στα χαρακτηριστικά του πειράματος Student_t4_1 6 χαρακτηριστικά που βασίζονται στην ενεργειακή αναλογία υποζώνης. Ούτε αυτή η προσθήκη έχει καλά αποτελέσματα, καθώς ο συντελεστής Pearson δεν δείχνει να μεταβάλλεται, παρά την αύξηση των χαρακτηριστικών, και ο συντελεστής Spearman μειώνεται.

Στο πείραμα Student_t4_4, προσθέτουμε στα χαρακτηριστικά του πειράματος Student_t4_1 6 χαρακτηριστικά, που αυτή την φορά βασίζονται στην διακύμανση συχνοτήτων. Η μείωση της απόδοσης είναι ακόμα μικρότερη από αυτή που προκαλούσαν οι προσθήκες άλλων χαρακτηριστικών.

Στο πείραμα Student_t4_5, προσθέτουμε στα χαρακτηριστικά του πειράματος Student_t4_3 2 χαρακτηριστικά Benford. Η προσθήκη βελτιώνει την απόδοση του Student_t4_3. Στην περίπτωση του συντελεστή Pearson, μάλιστα, η βελτίωση είναι τέτοιου βαθμού, που ξεπερνάει την υψηλή απόδοση του πειράματος Student_t4_1.

Στο πείραμα Student_t4_6, προσθέτουμε στα χαρακτηριστικά του πειράματος Student_t4_4 2 χαρακτηριστικά Benford. Η προσθήκη δεν αποδίδει τόσο, όσο στο προηγούμενο πείραμα.

Στο Student_t4_7, χρησιμοποιούμε, επιπλέον των χαρακτηριστικών που στηρίζονται στα wavelet, 6 χαρακτηριστικά που βασίζονται στην ενεργειακή αναλογία υποζώνης και 6 χαρακτηριστικά που βασίζονται στην διακύμανση συχνοτήτων. Η αύξηση των χαρακτηριστικών δεν συνεπάγεται, σε αυτή την περίπτωση, την αύξηση της απόδοσης καθώς, δεν πετυχαίνει υψηλές τιμές στους συντελεστές συγκριτικά με άλλα πειράματα.

Στο Student_t4_8, αν προσθέσουμε αχόμα δύο χαραχτηριστιχά Benford σε αυτά του Student_t4_7, τότε έχουμε αύξηση των τιμών χαι των δύο συντελεστών, σε βαθμό που το χαθιστά χαλύτερο μοντέλο από όλα τα προαναφερθέντα.

Αν συνυπολογίσουμε για την επιλογή του κατάλληλου μοντέλου και το πλήθος των χαρακτηριστικών, τότε αρκετά ανταγωνιστικά μοντέλα είναι και το Student_t4_1, που με 14 χαρακτηριστικά λιγότερα πλησιάζει πολύ την απόδοση του Student_t4_8, αλλά και το Student_t4_5.

4.2.4 Cauchy

Πίναχας 4.12 Συ	ντελεστής Pears	on και Spearman	Μοντέλων που	Έχουν σαι	, Βάση
Χαρακτηριστικά	που Εξάγονται	μέσω της Cauchy	/ Κατανομής.		

Κατανομή	Πλήθος features	Pearson	spearman
Cauchy_1	3	0.4393	0.4063
Cauchy_2	5	0.4349	0.3794
Cauchy_3	9	0.4564	0.4177
Cauchy_4	9	0.4490	0.4261
Cauchy_5	11	0.4563	0.4175
Cauchy_6	11	0.4430	0.4197
Cauchy_7	15	0.4807	0.4550
Cauchy_8	17	0.4706	0.4477

Στο Cauchy_1 παρουσιάζονται οι τιμές των συντελεστών Pearson και Spearman για 3 χαρακτηριστικά που υπολογίζονται μέσω του τύπου (4.5) για τους συντελεστές λεπτομέρειας, που προκύπτουν από το μετασχηματισμό κυματιδίων της εικόνας.

Στο Cauchy_2, προσθέτουμε επιπλέον 2 χαρακτηριστικά Benford, τα οποία μειώνουν και τους δύο συντελεστές.

Στο Cauchy_3, προσθέτουμε στα χαρακτηριστικά του Cauchy_1 6 χαρακτηριστικά βασισμένα στην ενεργειακή αναλογία υποζώνης. Η προσθήκη αυτή αυξάνει τις τιμές και των δύο συντελεστών. Επίσης, έχει καλύτερα αποτελέσματα και από το Cauchy_2, καθώς αυξάνει αρκετά τον συντελεστή Spearman, χρησιμοποιώντας 4 χαρακτηριστικά παραπάνω.

Στο Cauchy_4, προσθέτουμε στα χαρακτηριστικά του Cauchy_1 6 χαρακτηριστικά βασισμένα στην διακύμανση συχνοτήτων, το οποίο οδηγεί σε αύξηση και των δύο συντελεστών. Επίσης, έχει καλύτερα αποτελέσματα και από το Cauchy_2. Έχουμε μικρότερο Pearson, σε σχέση με όταν χρησιμοποιούσαμε ως επιπλέον χαρακτηριστικά τα έξι βασισμένα στην ενεργειακή αναλογία υποζώνης, αλλά αυξημένο συντελεστή Spearman.

Στο Cauchy_5, προσθέτουμε στα χαρακτηριστικά του Cauchy_3 και 2 Benford χαρακτηριστικά. Αυτή η προσθήκη δεν επηρεάζει τους δύο συντελεστές.

Στο Cauchy_6, προσθέτουμε τα 2 χαρακτηριστικά Benford στα χαρακτηριστικά του Cauchy_4. Η προσθήκη των 2 χαρακτηριστικών μειώνει και τους δύο συντελεστές. Τα αποτελέσματα, συγκριτικά με το Cauchy_5 το οποίο χρησιμοποιεί ίδιο πλήθος χαρακτηριστικών, είναι παρόμοια.

Στο Cauchy_7, προσθέτουμε στα χαρακτηριστικά του Cauchy_1 6 χαρακτηριστικά βασισμένα στην ενεργειακή αναλογία υποζώνης και 6 χαρακτηριστικά βασισμένα στην διακύμανση συχνοτήτων. Το αυξημένο πλήθος χαρακτηριστικών, σε αυτήν την περίπτωση, βοηθά στην αύξηση και των δύο συντελεστών.

Τέλος, στο Cauchy_8, προσθέτουμε στα χαρακτηριστικά του Cauchy_1 και τα 14 χαρακτηριστικά DCT. Το υψηλό πλήθος χαρακτηριστικών αποδίδει σε σχέση με τα προαναφερθέντα πειράματα, αλλά δεν βοηθά συγκριτικά με το Cauchy_7. Για μια ακόμα φορά, συμπεραίνουμε ότι η χρήση των χαρακτηριστικών Benford ρίχνει την απόδοση.

4.3 Σύγκριση αποτελεσμάτων με μοντέλα βασισμένα στους MCSN συντελεστές

 Σ ' αυτή την ενότητα θα εξετάσουμε την επιρροή των κατανομών Student's t και Cauchy στην απόδοση του μοντέλου εκτίμησης της ποιότητας της εικόνας, ανάλογα με τον τύπο δεδομένων που θέλουμε να μοντελοποιήσουμε. Κάθε σύγκριση θα γίνεται μεταξύ ενός χωρικού μοντέλου (δηλαδή ενός μοντέλου που χρησιμοποιεί μόνο χαρακτηριστικά από το χωρικό πεδίο) και ενός μοντέλου που χρησιμοποιεί χαρακτηριστικά μόνο από το πεδίο μετασχηματισμού (κυματιδίων).

Στον πίναχα 4.13, γίνεται σύγχριση της απόδοσης, που προχύπτει από τον συντελεστή Spearman, μεταξύ δύο μεθόδων που χρησιμοποιούν Student's t κατανομή, η μια μέθοδος για την προσαρμογή των MCSN συντελεστών (FeaturesSttMSCN) και η άλλη για την προσαρμογή των συντελεστών λεπτομέρειας του wavelet μετασχηματισμού (Student_t_1).

Για τους MCSN συντελεστές (4.1), υπολογίζονται οι δύο παράμετροι της student's t (2.4) κατανομής, εξάγοντας έτσι 2 χαρακτηριστικά. Έπειτα, υποκλιμακώνουμε την εικόνα κατά ½ και κάνουμε με το ίδιο τρόπο εξαγωγή άλλων δύο χαρακτηριστικών. Στη συνέχεια, παίρνουμε και τα κατά ζεύγη γινόμενα των

γειτονικών MCSN συντελεστών, με τον τρόπο που περιγράφηκε στην αρχή του κεφαλαίου 4 και υπολογίζονται οι 2 παράμετροι της κατανομής για κάθε μια από τις 4 κατευθύνσεις, οριζόντια, κάθετη, πρώτη διαγώνια και δευτερεύουσα διαγώνια, υπολογίζοντας έτσι 8 χαρακτηριστικά. Υποκλιμακώνουμε και πάλι την εικόνα κατά ¹/₂ και κάνουμε εξαγωγή άλλων 8 ίδιων χαρακτηριστικών. Έτσι, το χωρικό μοντέλο που δημιουργείται έχει συνολικά 20 χαρακτηριστικά.

Για τους wavelet συντελεστές, υπολογίζονται οι δύο παράμετροι της κατανομής για τις εικόνες οριζόντιων, κάθετων και διαγώνιων συντελεστών, με αποτέλεσμα την εξαγωγή 6 χαρακτηριστικών.

Πίνακας 4.13 Σύγκριση Μεθόδων με Κατανομή Student's t για το Σύνολο Δεδομένων LIVE IQA

Μέθοδος	White	Gaussian	Jp2k	jpeg	Fast Fading	All	Πλήθος features
	noise	Diui			Fauling		leatures
FeaturesStt MSCN	0.9748	0.8556	0.8585	0.8531	0.8790	0.7905	20
Student_t_1	0.9205	0.9013	0.6697	0.6620	0.8019	0.4553	6

Παρατηρούμε, ότι με τους wavelets συντελεστές πετυχαίνουμε καλύτερη απόδοση μόνο στην κατηγορία Gaussian Blur, και ότι στις κατηγορίες White noise και fast fading, αν και υπάρχει μείωση, δεν είναι τόσο μεγάλου βαθμού. Στις κατηγορίες jp2k, Jpeg και all έχουμε μεγάλη μείωση σε σχέση με όταν χρησιμοποιούμε MCSN συντελεστές, που μπορεί να θεωρηθεί λογική συνέπεια της μεγάλης μείωσης του πλήθους των χαρακτηριστικών.

Στον Πίνακα 4.14, γίνεται σύγκριση της απόδοσης, που προκύπτει από τον ,συντελεστή Spearman, μεταξύ δύο μεθόδων που χρησιμοποιούν Cauchy κατανομή, η μια μέθοδος για την προσαρμογή των MCSN συντελεστών (FeaturesCauchyMSCN) και η άλλη για την προσαρμογή των συντελεστών λεπτομέρειας του wavelet μετασχηματισμού (Cauchy_1).

Για τους MCSN συντελεστές (4.1), υπολογίζεται η μια παράμετρος της Cauchy κατανομής (4.5), εξάγοντας έτσι 1 χαρακτηριστικό. Έπειτα, υποκλιμακώνουμε την εικόνα κατά ½ και κάνουμε εξαγωγή ενός ακόμα χαρακτηριστικού. Στη συνέχεια, παίρνουμε και τα κατά ζεύγη γινόμενα των γειτονικών MCSN συντελεστών και

υπολογίζεται η παράμετρος της Cauchy για την οριζόντια, κάθετη, πρώτη διαγώνια και δευτερεύουσα διαγώνια κατεύθυνση, εξάγοντας έτσι 4 χαρακτηριστικά. Υποκλιμακώνουμε και πάλι την εικόνα κατά ½ και κάνουμε εξαγωγή άλλων 4 χαρακτηριστικών. Οπότε, το χωρικό μοντέλο που δημιουργείται χρησιμοποιεί 10 χαρακτηριστικά.

Για τους wavelet συντελεστές, υπολογίζουμε την παράμετρο της Cauchy κατανομής για τις εικόνες οριζόντιων, κάθετων και διαγώνιων συντελεστών, με αποτέλεσμα την εξαγωγή 3 χαρακτηριστικών.

Πίνακας 4.14 Σύγκριση Μεθόδων με Κατανομή Cauchy για το Σύνολο Δεδομένων LIVE IQA

Μέθοδος	White noise	Gaussia n Blur	Jp2k	jpeg	Fast Fading	All	Πλήθος features
FeaturesCauch	0.9821	0.8856	0.869	0.889	0.8254	0.8365	10
yMSCN	0.001	0.0000	4	8		0.0000	10
Cauchy 1	0 9375	0.8759	0.761	0.709	0 7846	0.3469	3
Caucity_1	0.0070	0.0700	6	6	0.7040	0.0400	0

Παρατηρώντας τον Πίναχα 4.14, διαχρίνουμε μειωμένη απόδοση σε όλες τις κατηγορίες, όταν χρησιμοποιούμε την Cauchy χατανομή για τους wavelets συντελεστές, σε σχέση με όταν χρησιμοποιούμε MCSN συντελεστές. Πρέπει να σημειωθεί, όμως, ότι χρησιμοποιούνται μόλις 3 χαραχτηριστιχά, τα οποία φαίνεται να πιάνουν μεγάλο βαθμό της απόδοσης, αν τα συγχρίνουμε με τα αντίστοιχα του πειράματος FeaturesCauchyMSCN, που εξάγει 7 χαραχτηριστιχά παραπάνω, αλλά χωρίς να επηρεάζεται ανάλογα χαι η απόδοση. Η μόνη χατηγορία, που πετυχαίνει ανάλογη απόδοση του πλήθους των features που χρησιμοποιεί, είναι η χατηγορία all.

Η ίδια σύγκριση μεθόδων γίνεται και για το σύνολο δεδομένων LIVE C. Εξετάζουμε και την τιμή του συντελεστή Pearson σε αυτή την βάση.

ΜέθοδοςPearsonSpearmanNumber of
featuresFeaturesSttSym0.24170.277020Student_t_10.47300.48226

Πίνα
κας 4.15 Σύγκριση Μεθόδων με Κατανομή Student's
t για το Σύνολο Δεδομένων LIVE C

Τα συμπεράσματα σε αυτή την βάση είναι πολύ διαφορετικά. Παρατηρούμε στον Πίνακα 4.15, ότι το μοντέλο με τα 6 χαρακτηριστικά από το πεδίο μετασχηματισμού (Student_t_1) πετυχαίνει διπλάσια απόδοση σε σχέση με το χωρικό μοντέλο που χρησιμοποιεί 20 χαρακτηριστικά.

Πίνα
κας 4.16 Σύγκριση Μεθόδων με Κατανομή Cauchy για το Σύνολο Δεδομένων LIVE C

Μέθοδος	Doorcon	Speerman	Number of
Μευσσος	rearson	Spearman	features
FeaturesCCauchySym	0.3680	0.3519	10
Cauchy_1	0.4393	0.4063	3

Ομοίως, το μοντέλο, που χρησιμοποιεί 3 χαρακτηριστικά από την προσαρμογή της Cauchy κατανομής (Πίνακας 4.16) στους wavelet συντελεστές, πετυχαίνει μεγαλύτερη απόδοση από το μοντέλο που εξάγει 10 χωρικά χαρακτηριστικά.

4.4 Συμπληρωματικά πειράματα

Για την κατανομή Student's t και Cauchy κατανομή, πραγματοποιήθηκε μια ακόμα σειρά πειραμάτων, κατά την οποία εφαρμόζουμε πάλι μετασχηματισμό κυματιδίων και χρησιμοποιούμε τις εικόνες διαγώνιων, οριζόντιων και κάθετων συντελεστών, για να υπολογίσουμε τις παραμέτρους κάθε κατανομής και να πάρουμε τα αντίστοιχα χαρακτηριστικά. Έπειτα, χρησιμοποιούμε για πρώτη φορά και την εικόνα των συντελεστών προσέγγισης, στην οποία εφαρμόζουμε DCT μετασχηματισμό. Κρατάμε ένα πλήθος των DCT συντελεστών που προκύπτουν και εφαρμόζουμε σε αυτούς την κατανομή που επιθυμούμε να εξετάσουμε.

4.4.1 Student's t

Σε αυτή την ενότητα, εφαρμόζουμε μετασχηματισμό χυματιδίων και εφαρμόζουμε DCT μετασχηματισμό στους συντελεστές προσέγγισης. Στους συντελεστές που προχύπτουν, κάνουμε προσαρμογή καμπύλης (fit) την Student's t κατανομή και εξάγουμε δύο χαραχτηριστικά. Επιπλέον, υπολογίζουμε τις παραμέτρους της Student's t κατανομής για τις εικόνες διαγώνιων, οριζόντιων και κάθετων συντελεστών, εξάγοντας και με αυτό τον τρόπο άλλα 6 χαραχτηριστικά. Εξετάζουμε, σε πρώτη φάση, την απόδοση που επιτυγχάνουν μόνο τα 2 χαραχτηριστικά που εξάγουμε από τους συντελεστές DCT της εικόνας προσέγγισης. Χρησιμοποιούμε διαφορετικά πλήθη συντελεστών για την ίδια μέτρηση απόδοση. Έπειτα, προσθέτουμε σε αυτά τα 2, τα 6 χαραχτηριστικά που εξάγουμε από τους συντελεστές λεπτομέρειας και επανεξετάζουμε την απόδοση για τα διαφορετικά πλήθη συντελεστών DCT. Η μέτρηση γίνεται, και πάλι, με την βοήθεια του συντελεστή Pearson και Spearman.

- Spearman
- •

Πίνακας 4.17 Συντελεστής Spearman Μοντέλων που Έχουν σαν Βάση 2 Χαρακτηριστικά που Εξάγονται από τους Συντελεστές DCT της Εικόνας Προσέγγισης μέσω της Student-t Κατανομής με ΕΜ Εκτίμηση Παραμέτρων.

Student	Πλήθος	wn	gblur	Jp2k	jpeg	Fast fading	All
10.000	2	0.4968	0.6065	0.3086	0.1774	0.4731	0.1501
20.000	2	0.6978	0.6415	0.3372	0.2525	0.5250	0.1559
30.000	2	0.7224	0.6752	0.3845	0.3116	0.5628	0.1606
40.000	2	0.7264	0.7071	0.3359	0.3469	0.5957	0.1724

Όταν για την εκπαίδευση του μοντέλου μάθησης χρησιμοποιούμε μόνο τα 2 χαρακτηριστικά, που προκύπτουν από το fit της κατανομής Student's t στους συντελεστές DCT της εικόνας προσέγγισης, η απόδοση κινείται σε χαμηλά επίπεδα στην πλειοψηφία των παραμορφώσεων. Όσο αυξάνουμε τους συντελεστές DCT, παρατηρούμε αύξηση των τιμών του συντελεστή. Πίνακας 4.18 Συντελεστής Spearman Μοντέλων που Έχουν σαν Βάση 8 Χαρακτηριστικά που Εξάγονται από τους Συντελεστές DCT της Εικόνας Προσέγγισης και από τις Εικόνες Διαγώνιων, Οριζόντιων και Κάθετων Συντελεστών μέσω της Student-t Κατανομής με ΕΜ Εκτίμηση Παραμέτρων.

Student	Πλήθος	Wn	oblur	In9k	ineg	Fast	Δ11
Student	11/1/005	**11	goiui	JPZK	Jpeg	fading	7 111
10.000	8	0.8881	0.9042	0.6091	0.6167	0.7714	0.2209
20.000	8	0.7852	0.9042	0.5999	0.6213	0.7728	0.2279
30.000	8	0.7923	0.9059	0.5982	0.6206	0.7779	0.2231
40.000	8	0.7963	0.9055	0.6425	0.6307	0.7740	0.2232

Με την προσθήκη των 6 χαρακτηριστικών που προκύπτουν από το fit της Student's t στις εικόνες των συντελεστών λεπτομέρειας, έχουμε αύξηση της απόδοσης. Για την παραμόρφωση White noise, παρατηρείται ότι όσο αυξάνονται οι συντελεστές DCT, μειώνεται η απόδοση.

• Pearson(Logistic)

Πίνακας 4.19 Συντελεστής Pearson Μοντέλων που Έχουν σαν Βάση 2 Χαρακτηριστικά που Εξάγονται από τους Συντελεστές DCT της Εικόνας Προσέγγισης μέσω της Student-t Κατανομής με ΕΜ Εκτίμηση Παραμέτρων.

Student	Πλήθος	wn	gblur	Jp2k	jpeg	Fast fading	all
10.000	2	0.5476	0.6513	0.4201	0.3049	0.5683	0.3095
20.000	2	0.7346	0.6608	0.4540	0.3909	0.6238	0.3131
30.000	2	0.7511	0.6804	0.4735	0.4370	0.6545	0.3303
40.000	2	0.7533	0.7200	0.4377	0.4580	0.6821	0.3558

Αντίστοιχα με τον συντελεστή Pearson, συνήθως η αύξηση συντελεστών DCT αυξάνει την απόδοση (Πίναχας 4.19). Υπάρχουν εξαιρέσεις, όπως αυτή στην κατηγορία jp2k, που για 20.000 και 30.000 συντελεστές εμφανίζει υψηλότερες τιμές Pearson από ότι για 40.000 συντελεστές.

Πίνακας 4.20 Συντελεστής Pearson Μοντέλων που Έχουν σαν Βάση 8 Χαρακτηριστικά που Εξάγονται από τους Συντελεστές DCT της Εικόνας Προσέγγισης και από τις Εικόνες Διαγώνιων, Οριζόντιων και Κάθετων Συντελεστών μέσω της Student-t Κατανομής με ΕΜ Εκτίμηση Παραμέτρων.

Student	Πλήθος	wn	gblur	Jp2k	jpeg	Fast fading	all
10.000	8	0.9066	0.9053	0.6704	0.6687	0.8083	0.4581
20.000	8	0.8003	0.9060	0.6480	0.6625	0.8102	0.4655
30.000	8	0.8053	0.9087	0.6407	0.6636	0.8131	0.4634
40.000	8	0.8071	0.9074	0.6921	0.6771	0.8114	0.4638

Με την προσθήκη των 6 χαρακτηριστικών έχουμε βελτίωση της απόδοσης (Πίνακας 4.20). Για την παραμόρφωση White noise, παρατηρείται ότι όσο αυξάνονται οι συντελεστές DCT, μειώνεται η απόδοση.

Επαναλαμβάνουμε τα παραπάνω πειράματα και για το σύνολο δεδομένων dataset C.

Πίνακας 4.21 Συντελεστές Pearson και Spearman Μοντέλων που Έχουν σαν Βάση 2 Χαρακτηριστικά που Εξάγονται από τους Συντελεστές DCT της Εικόνας Προσέγγισης μέσω της Student-t Κατανομής με ΕΜ Εκτίμηση Παραμέτρων.

Student	Πλώθος	Choormon	Pearson-	
Student	πληθος	Spearman	(logistic)	
10.000	2	0.2535	0.4076	
20.000	2	0.3046	0.4130	
30.000	2	0.4220	0.4487	
40.000	2	0.4361	0.4880	

Έχουμε αύξηση της απόδοσης όσο αυξάνουμε τους συντελεστές DCT τόσο για τον συντελεστής Spearman όσο και για τον συντελεστή Pearson (Πίνακας 4.21)

Πίνακας 4.22 Συντελεστές Pearson και Spearman Μοντέλων που Έχουν σαν Βάση 8 Χαρακτηριστικά που Εξάγονται από τους Συντελεστές DCT της Εικόνας Προσέγγισης και από τις Εικόνες Διαγώνιων, Οριζόντιων και Κάθετων Συντελεστών μέσω της Student-t Κατανομής με ΕΜ Εκτίμηση Παραμέτρων.

Curdent	Πλάθοια	Cue a a runne a ru	Pearson-
Student	Πληθος	Spearman	(logistic)
10.000	8	0.4793	0.4677
20.000	8	0.4810	0.4722
30.000	8	0.4822	0.4762
40.000	8	0.4803	0.4766

Παρατηρώντας τον πίνακα 4.22, παρατηρούμε ότι έχουμε αύξηση της απόδοσης, όσο αυξάνουμε τους συντελεστές DCT.

4.4.2 Cauchy

Σε αυτή την ενότητα, επαναλαμβάνουμε την διαδικασία που αναφέραμε στην ενότητα 4.3.1. Εφαρμόζουμε μετασχηματισμό κυματιδίων και εφαρμόζουμε DCT μετασχηματισμό στην εικόνα προσέγγισης. Στους συντελεστές που προκύπτουν, κάνουμε fit την Cauchy κατανομή (4.5) και εξάγουμε 1 χαρακτηριστικό. Έπειτα, υπολογίζεται η παράμετρος της Cauchy για κάθε μια από τις εικόνες διαγώνιων, κάθετων και οριζόντιων συντελεστών, εξάγοντας άλλα 3 χαρακτηριστικά.

• Spearman

Πίνακας 4.23 Συντελεστής Spearman Μοντέλων που Έχουν σαν Βάση 1 Χαρακτηριστικό που Εξάγεται από τους Συντελεστές DCT της Εικόνας Προσέγγισης μέσω της Cauchy Κατανομής.

Cauchy	Πλήθος	wn	gblur	Jp2k	jpeg	Fast fading	all
10.000	1	0.5581	0.8131	0.2577	0.0205	0.6985	0.2282
20.000	1	0.6066	0.8239	0.2792	0.0547	0.7099	0.2265
30.000	1	0.6425	0.8364	0.3127	0.0821	0.7091	0.2398
40.000	1	0.6691	0.8498	0.3377	0.1287	0.7079	0.2488

Με το 1 χαρακτηριστικό, από το fit της Cauchy στους DCT συντελεστές, έχουμε χαμηλή απόδοση στις κατηγορίες Jp2k, jpeg και all και αρκετά υψηλή στις υπόλοιπες κατηγορίες, αν αναλογιστούμε ότι χρησιμοποιείται μόνο 1 χαρακτηριστικό. Με αύξηση των συντελεστών DCT, έχουμε αύξηση και στη απόδοση του μοντέλου μάθησης (Πίνακας 4.23).

Πίνακας 4.24 Συντελεστής Spearman Μοντέλων που Έχουν σαν Βάση 4 Χαρακτηριστικά που Εξάγονται από τους Συντελεστές DCT της Εικόνας Προσέγγισης και από τις Εικόνες Διαγώνιων, Οριζόντιων και Κάθετων Συντελεστών μέσω της Cauchy Κατανομής.

Cauchy	Πλήθος	wn	gblur	Jp2k	jpeg	Fast fading	all
10.000	4	0.9600	0.8648	0.7386	0.7613	0.7597	0.3891
20.000	4	0.9565	0.8721	0.7303	0.7602	0.7663	0.3938
30.000	4	0.9552	0.8776	0.7294	0.7538	0.7654	0.4016
40.000	4	0.9521	0.8822	0.7258	0.7452	0.7646	0.4192

Με προσθήκη των 3 χαρακτηριστικών, από το fit της Cauchy στους συντελεστές λεπτομέρειας, έχουμε αύξηση σε όλες τις κατηγορίες παραμόρφωσης και, μάλιστα, μεγάλου βαθμού στις κατηγορίες που είχαν πολύ χαμηλές τιμές χωρίς αυτά τα 3 χαρακτηριστικά. Στις κατηγορίες white noise, jp2k και jpeg έχουμε μείωση της απόδοσης, όσο αυξάνονται οι συντελεστές DCT (Πίνακας 4.24).

• Pearson(Logistic)

Πίνακας 4.25 Συντελεστής Pearson Μοντέλων που Έχουν σαν Βάση 1 Χαρακτηριστικό που Εξάγεται από τους Συντελεστές DCT της Εικόνας Προσέγγισης μέσω της Cauchy Κατανομής.

Cauchy	Πλήθος	wn	gblur	Jp2k	jpeg	Fast fading	all
10.000	1	0.6038	0.8243	0.3888	0.2828	0.7623	0.4394
20.000	1	0.6575	0.8312	0.3997	0.2808	0.7670	0.4554
30.000	1	0.6893	0.8399	0.4252	0.2820	0.7672	0.4744
40.000	1	0.7084	0.8506	0.4402	0.2948	0.7696	0.4872
Ο συντελεστής Pearson επηρεάζεται με τον ίδιο τρόπο, όπως ο συντελεστής Spearman (Πίναχας 4.25).

Πίναχας 4.26 Συντελεστής Pearson Μοντέλων που Έχουν σαν Βάση 4 Χαραχτηριστικά που Εξάγονται από τους Συντελεστές DCT της Εικόνας Προσέγγισης και από τις Εικόνες Διαγώνιων, Οριζόντιων και Κάθετων Συντελεστών μέσω της Cauchy Κατανομής.

Cauchy	Πλήθος	wn	gblur	Jp2k	jpeg	Fast fading	all
10.000	4	0.9533	0.8717	0.7816	0.8018	0.7989	0.5188
20.000	4	0.9497	0.8754	0.7772	0.8032	0.8027	0.5201
30.000	4	0.9478	0.8803	0.7743	0.7980	0.8037	0.5247
40.000	4	0.9457	0.8829	0.7689	0.7905	0.8054	0.5306

Ο συντελεστής Pearson επηρεάζεται με τον ίδιο τρόπο, όπως ο συντελεστής Spearman (Πίναχας 4.26).

Επαναλαμβάνουμε τα παραπάνω πειράματα και για το σύνολο δεδομένων dataset C.

Πίνακας 4.27 Συντελεστές Pearson και Spearman Μοντέλων που Έχουν σαν Βάση 1 Χαρακτηριστικό που Εξάγεται από τους Συντελεστές DCT της Εικόνας Προσέγγισης μέσω της Cauchy Κατανομής.

Cauchy	Πλήθος	Spearman	Pearson-	
Cauchy	πηθος	Spearman	(logistic)	
10.000	1	0.4229	0.4773	
20.000	1	0.4439	0.5002	
30.000	1	0.4581	0.5134	
40.000	1	0.4703	0.5257	

Έχουμε αύξηση της απόδοσης, όσο αυξάνουμε τους συντελεστές DCT (Πίνακας 4.27).

Πίνακας 4.28 Συντελεστές Spearman Pearson και Μοντέλων που Έχουν σαν Βάση 4 Χαρακτηριστικά που Εξάγονται από τους Συντελεστές DCT της Εικόνας Προσέγγισης και από τις Εικόνες Διαγώνιων, Οριζόντιων και Κάθετων Συντελεστών μέσω της Cauchy κατανομής.

Cauchy	Πλήθος	Spearman	Pearson- (logistic)
10.000	4	0.4428	0.4929
20.000	4	0.4480	0.5017
30.000	4	0.4639	0.5167
40.000	4	0.4745	0.5282

Με προσθήκη των 3 χαρακτηριστικών, από το fit της Cauchy στους συντελεστές λεπτομέρειας, έχουμε αύξηση σε όλες τις κατηγορίες παραμόρφωσης. Έχουμε αύξηση της απόδοσης, όσο αυξάνουμε τους συντελεστές DCT (Πίνακας 4.28).

ΚΕΦΑΛΑΙΟ 5

Σ ΥΜΠΕΡΑΣΜΑΤΑ

Όπως είδαμε στα προηγούμενα χεφάλαια, η ανάπτυξη αλγορίθμων εκτίμησης της αντιληπτής ποιότητας της εικόνας αποτελεί ένα σημαντικά αναπτυσσόμενο ερευνητικό πεδίο, καθώς η εικόνα είναι ένα μέσο που χρησιμοποιείται ευρέως σε όλους τους τομείς της καθημερινότητας μας. Τα μέσα που χρησιμοποιούνται για την μετάδοση της εικόνας, λόγω του διαθέσιμου εύρους ζώνης, αλλά και των φυσικών ιδιοτήτων τους, θέτουν περιορισμούς στην ποιότητά της, υποβαθμίζοντάς την. Έτσι, δημιουργείται η ανάγκη ανάπτυξης αλγορίθμων, που θα υπολογίζουν την υποβάθμιση αυτή και θα εκτιμούν με ακρίβεια την ποιότητά της, με τέτοιο τρόπο ώστε να συνάδει με την υποκειμενική ανθρώπινη αξιολόγηση. Τέτοιου είδους αλγόριθμοι μπορούν να ενσωματωθούν σε διάφορα αυτόματα συστήματα, τα οποία πλέον θα είναι ικανά να προσαρμόζονται σε πραγματικό χρόνο, με βάση κάποια ανάδραση (feedback) της ποιότητας εξόδου, και να βελτιώνουν αυτόματα την ποιότητα της εικόνας εισόδου. Είναι, επίσης, σημαντικοί για εφαρμογές ελέγχου, benchmarking, παρακολούθησης κ.α.

Πολύ σημαντικό στοιχείο των μετρικών αξιολόγησης της ποιότητας της εικόνας αποτελούν τα διαφορετικά αντιληπτικά χαρακτηριστικά, που χρησιμοποιούνται για την εκτίμησή της. Τα χαρακτηριστικά που βασίζονται στη στατιστική της φυσικής σκηνής έχουν επιδείξει πολύ καλά αποτελέσματα. Ένα άλλο στοιχείο, που έχει προσελκύσει το ενδιαφέρον των ερευνητών στο πλαίσιο των IQA αλγορίθμων, είναι

η οπτική προσοχή, δηλαδή ο μηχανισμός του ανθρώπινου οπτικού συστήματος που καθοδηγεί την επιλογή της οπτικής πληροφορίας σε μια οπτική σκηνή και πιο συγκεκριμένα η bottom up εκδοχή του, που είναι γνωστή ως η οπτικά εξέχουσα περιοχή. Είναι βασικό χαρακτηριστικό, που μας βοηθά στην μείωση της πλεονάζουσας πληροφορίας και στην εστίαση της προσοχή μας στις οντότητες που εμφανίζονται στις οπτικά εξέχουσες περιοχές, που θεωρούνται ότι είναι αυτές που ασκούν την πιο σημαντική επίδραση στην αντίληψη της οπτικής ποιότητας.

Κύριος στόχος αυτής της εργασίας ήταν η ανάπτυξη και η συγκριτική μελέτη μεθόδων αξιολόγησης της αντιληπτής ποιότητας εικόνων, μέσω χρήσης εναλλακτικών στατιστικών κατανομών. Τα δεδομένα που επιλέξαμε να μοντελοποιήσουμε είναι οι συντελεστές του μετασχηματισμού κυματιδίων. Στο πλαίσιο του συγκεκριμένου μετασχηματισμού, εκμεταλλευτήκαμε εκείνους του συντελεστές που θα μπορούσαν να αποτελέσουν τα πιο χρήσιμα αντιληπτικά χαρακτηριστικά. Αυτό επιτεύχθηκε είτε με την εκμετάλλευση των συντελεστών κατεύθυνσης (στοιχείο που επηρεάζει το ανθρώπινο οπτικό σύστημα) είτε δημιουργώντας νέους χάρτες οπτικά εξεχουσών περιοχών (visual saliency), κατάλληλους να συλλαμβάνουν τις αλλαγές που προκαλούν οι διάφορες παραμορφώσεις.

Τα πειράματα που διεξήχθησαν, αφορούσαν είτε τους συντελεστές προσέγγισης είτε τους συντελεστές οριζόντιων, κάθετων και διαγώνιων λεπτομερειών, που προκύπτουν από την εφαρμογή του wavelet μετασχηματισμού. Υποθέσαμε, ότι οι συντελεστές έχουν στατιστικές ιδιότητες, που μεταβάλλονται ανάλογα από τον τύπο και τον βαθμό παραμόρφωσης μια εικόνας, και η ποσοτικοποίηση αυτών των μεταβολών μπορεί να καταστήσει δυνατή την εκτίμηση του τύπου παραμόρφωσης που στρεβλώνει την εικόνα, αλλά και της αντίληψης της ποιότητάς της.

Ένας αχόμα λόγος που επιλέξαμε σαν δεδομένα μοντελοποίησης τους wavelet συντελεστές, είναι οι κατανομές που επιλέξαμε να εξετάσουμε. Οι κατανομές αυτές είναι η Student-t κατανομή και η πιο ειδική μορφή της, η Cauchy. Σε προηγούμενη εργασία είχαμε προσπαθήσει να μοντελοποιήσουμε μέσω αυτών των κατανομών τους MCSN συντελεστές και ενώ τα αποτελέσματα ήταν αρχετά ικανοποιητικά, παρατηρήσαμε ότι, σε κάποιες «ακραίες» περιπτώσεις, η μορφή τους εκφυλιζόταν σε ομοιόμορφη κατανομή, με αποτέλεσμα να προσαρμόζονται καλύτερα από την ευρέως χρησιμοποιούμενη Γενικευμένη Γκαουσιανή κατανομή.

Αρχικά, στο 3ο κεφάλαιο, αναπτύξαμε Full-Reference αλγορίθμους. Σε αυτούς τους αλγορίθμους, βασικά χαρακτηριστικά αποτέλεσαν δύο νέοι δείκτες, όπως η ζωντάνια (vividness) και το βάθος (depth) του χρωματικού μοντέλου CIELAB. Στο συγχεκριμένο χρωματικό χώρο γινόταν η μετατροπή των RGB εικόνων του συνόλου δεδομένων LIVE IQA. Σε αυτά τα χαρακτηριστικά εφαρμόσαμε wavelet λαμβάνοντας λεπτομέρειες κατεύθυνσης, μετασχηματισμό και, τις αντικαταστήσαμε την πληροφορία τους με παραμέτρους που χρησιμοποιούνται στην εκτίμηση των παραμέτρων της προσαρμοσμένης Student's t κατανομής. Στην συνέχεια, με αντίστροφο μετασχηματισμό, επιστρέφαμε στο χωρικό πεδίο και υπολογίζαμε την οπτικά εξέχουσα περιοχή. Έπειτα από πολλά πειράματα, στα οποία συνδυάζαμε συντελεστές λεπτομέρειας ή τους χρησιμοποιούσαμε μεμονωμένα, αγνοώντας ή όχι τους συντελεστές προσέγγισης, καταλήξαμε σε μεθόδους με αρχετά υψηλά αποτελέσματα, με μία μέθοδο να ξεχωρίζει, χαθώς πέτυχε αρχετά υψηλή απόδοση, σε σημείο να μπορεί να ανταγωνιστεί άλλες ευρέως γνωστές μεθόδους. Σε αυτή την μέθοδο χρησιμοποιούνταν μόνο οι οριζόντιοι συντελεστές λεπτομέρειας.

Παρατήρηση 1: Η μορφή με την οποία χρησιμοποιούσαμε τους συντελεστές λεπτομέρειας για τον αντίστροφο μετασχηματισμό επηρέαζε σημαντικά την απόδοση της μεθόδου. Για το λόγο αυτό, γινόταν αντικατάσταση των συντελεστών λεπτομέρειας στις διάφορες κατευθύνσεις (οριζόντια, κάθετη ή διαγώνια) από τις αντιληπτικές μάσκες, όπως αυτές προέκυπταν από την εκμετάλλευση των κρυμμένων πληροφοριών, που χρησιμοποιούνταν κατά την εκτίμηση των παραμέτρων της προτεινόμενης κατανομής. Λέγοντας εκμετάλλευση, εννοούμε την αντικατάστασή τους με τις αντίστοιχες τιμές θέση-προς-θέση στις διάφορες κατευθύνσεις. Στο σημείο αυτό, πραγματοποιήθηκε μία πληθώρα πειραμάτων με διάφορες «διαμορφώσεις» τους, όπως φαίνεται και στο τέταρτο κεφάλαιο. Τα καλύτερα αποτελέσματα προέκυψαν, όταν χρησιμοποιήσαμε τον λογάριθμο της κρυμμένης πληροφορίας των εκάστοτε συντελεστών. Τα αμέσως καλύτερα ήταν με τον λογάριθμο της αντίστροφης κρυμμένης πληροφορίας. Αντιθέτως, όταν δεν χρησιμοποι-ούσαμε λογάριθμο, τα αποτελέσματα ήταν αρκετά χαμηλά.

Παρατήρηση 2: Οι συντελεστές προσέγγισης δεν φάνηκε να συνεισφέρουν στις μεθόδους. Κάθε άλλο, μηδενίζοντάς τους παρατηρήσαμε αύξηση της απόδοσης της μεθόδου, συγκριτικά με όταν χρησιμοποιούνταν. Προφανώς, η πληροφορία που

σχετίζεται με τις υψηλές συχνότητες μίας εικόνας (αφού τα καλύτερα αποτελέσματα μας τα έδιναν οι συντελεστές πρώτου επιπέδου) είναι πιο ευμετάβλητη στις διάφορες παραμορφώσεις, καθιστώντας την εκμετάλλευσή τους από κάποια μετρική πιο αποδοτική.

Παρατήρηση 3: Εξετάσαμε τρία διαφορετικά φίλτρα για τον μετασχηματισμό wavelet, όλα της οικογένειας Daubechies, αλλά διαφορετικής τάξης. Τα φίλτρα ήταν τάξης 1(Haar wavelets), 4 και 8. Τα καλύτερα αποτελέσματα επιτυγχάνονταν με χρήση του φίλτρου Daubechies τάξης 4. Ουσιαστικά, δηλαδή, φίλτρα μεσαίου μεγέθους της οικογένειας Daubechies, υποδηλώνοντας ότι ένα μέσο μέγεθος μπορεί να «συλλάβει» τις δομές μίας εικόνας, ικανές να συσχετίζονται με αντιληπτικά χαρακτηριστικά, σχετικά με παραμορφώ-σεις λόγω διαφόρων επεξεργασιών της εικόνας.

Παρατήρηση 4: Τα πειράματα που διεξήχθησαν ήταν με μετασχηματισμό ενός επιπέδου. Κάθε φορά που προσθέταμε ένα επίπεδο, παρατηρούνταν μείωση μικρού βαθμού. Πειράματα με χρήση δύο και τριών επιπέδων του μετασχηματισμού wavelet αποτελούν την πειραματική απόδειξη για την εξαγωγή του συγκεκριμένου συμπεράσματος.

Παρατήρηση 5: Για την παραγωγή του χάρτη κλίσης (gradient map), εκτός από τον τρόπο στον οποίο χρησιμοποιούσαμε το μέτρο της πρώτης παραγώγου σαν δομικό χαρακτηριστικό, εξετάσαμε και την χρήση μιας «ψεύδο-gradient». Δηλαδή, ουσιαστικά, χρησιμοποιήσαμε το μέτρο της πρώτης παραγώγου, βασισμένοι στους οριζόντιους και κάθετους wavelet συντελεστές. Η εναλλακτική αυτή μορφή της gradient δεν είχε το ίδιο καλά αποτελέσματα με την αρχική. Αντιθέτως, ακόμα και η επιπλέον προσθήκη της μείωνε την απόδοση που επιτύγχανε από μόνη της η αρχική μορφή gradient.

Στο 4ο κεφάλαιο, αναπτύξαμε No-Reference αλγορίθμους και δοκιμάσαμε περισσότερες κατανομές για την προσαρμογή των συντελεστών wavelets. Αφού κάναμε εξαγωγή των βασικών χαρακτηριστικών, μέσω προσαρμογής της εκάστοτε κατανομής στους συντελεστές λεπτομέρειας του μετασχηματισμού κυματιδίων, προσθέταμε σταδιακά DCT χαρακτηριστικά, ώστε να καταλήξουμε στο κατάλληλο μοντέλο το οποίο πετυχαίνει μια ισορροπία μεταξύ της απόδοσης και του πλήθους των χαρακτηριστικών που χρησιμοποιεί. Οι αλγόριθμοι εφαρμόστηκαν σε δύο

σύνολα δεδομένων και τα συμπεράσματα διαφέρουν ανάλογα το σύνολο που χρησιμοποιήθηκε.

Παρατήρηση 6: Στο σύνολο IQA, την υψηλότερη απόδοση ανεξαρτήτου κατανομής την επιτύγχανε το μοντέλο, στο οποίο χρησιμοποιούσαμε ολόκληρο το σύνολο των DCT χαρακτηριστικών. Στο σύνολο C δεν ισχύει ο ίδιος κανόνας.

Παρατήρηση 7: Τα χαρακτηριστικά Benford είχαν μεγάλη συνεισφορά στην απόδοση, όσον αφορά το σύνολο IQA, ενώ αντίθετα στο σύνολο C ήταν μικρότερη και σε πολλές περιπτώσεις προκαλούσε ακόμα και μείωση, όταν τα χρησιμοποιούσαμε.

Παρατήρηση 8: Επίσης, παρατηρήθηκε ότι στο σύνολο δεδομένων IQA είχαμε καλύτερα αποτελέσματα, όταν χρησιμοποιούσαμε τα χαρακτηριστικά που βασίζονται στην διακύμανση συχνοτήτων, ενώ στο dataset C είχαμε υψηλότερη απόδοση, όταν χρησιμοποιού-σαμε χαρακτηριστικά βασισμένα στην ενεργειακή αναλογία υποζώνης.

Παρατήρηση 9: Στο dataset IQA δεν είναι ξεχάθαρο ποια χατανομή μας δίνει καλύτερα αποτελέσματα, χαθώς χάθε μία δείχνει να επιδρά διαφορετικά, ανάλογα το είδος της παραμόρφωσης που έχει υποστεί η ειχόνα. Πολύ χαλά αποτελέσματα δίνει η Student't με εχτίμηση 4 παραμέτρων, χάτι το οποίο μπορεί να θεωρηθεί λογικό συμπέρασμα, χαθώς είναι η χατανομή με το μεγαλύτερο πλήθος παραμέτρων. Αλλά, πολύ χαλά αποτελέσματα έχει χαι η Cauchy χατανομή, που, παρά το μιχρό πλήθος παραμέτρων της, πλησιάζει πολύ τις τιμές των συντελεστών χατανομών με περισσότερες παραμέτρους. Στο dataset C είναι πιο ξεχάθαρο το συμπέρασμα, ότι η υψηλότερη απόδοση επετεύχθη με χρήση της χατανομής Student's t με τον ΕΜ αλγόριθμο για την εχτίμηση των παραμέτρων. Πιθανόν χάτι τέτοιο να συμβαίνει λόγω μεγαλύτερης αχρίβειας στην εχτίμησή τους. Στο σημείο αυτό, πραγματοποι-ήθηχαν περισσότερα πειράματα, με αποτέλεσμα να αντιληφθούμε ότι η τελιχή αντιχειμενιχή πρόβλεψη του αλγορίθμου δεν παρουσιάζει μεγάλες μεταβολές λόγω της αχρίβειας εχτίμησης των παραμέτρων.

Παρατήρηση 10: Μέσω της σύγκρισης μεθόδων του πεδίου μετασχηματισμού wavelet και του χωρικού πεδίου παρατηρήσαμε τα εξής. Στο σύνολο δεδομένων LIVE C, οι μέθοδοι με τα wavelet χαρακτηριστικά είχαν πολύ καλύτερα αποτελέσματα από αυτές με τα χωρικά, παρά το μικρότερο πλήθος των χαρακτηριστικών τους. Αντιθέτως, στο LIVE IQA, οι μέθοδοι με τα χωρικά

χαρακτηριστικά, που ήταν και τα περισσότερα, είχαν καλύτερη απόδοση από τις μεθόδους του πεδίου μετασχηματισμού, αν και η αύξηση αυτή δεν ήταν ανάλογη του πλήθους των χαρακτηριστικών τους.

Παρατήρηση 11: Από την διεξαγωγή όλων των πειραμάτων προέχυψε το εξής συμπέρασμα, όσον αφορά τις βάσεις δεδομένων που χρησιμοποιούμε: Στο LIVE IQA παρατηρήθηχε, ότι όσο αυξάναμε το πλήθος των χαραχτηριστιχών, αυξανόταν και η απόδοση. Στο LIVE C δεν ισχύει το ίδιο, χαθώς μια αύξηση χαραχτηριστιχών μπορεί να μην είχε χαμία επιρροή στην απόδοση ή, αχόμα, μπορούσε να προχαλέσει χαι μείωση. Στο LIVE IQA, το πλήθος των χαραχτηριστιχών φαίνεται να επηρεάζει την απόδοση, ενώ στο LIVE C επηρεάζει περισσότερο το τί χαραχτηριστιχά χρησιμοποιούμε, παρά το πλήθος τους.

Τέλος, χρησιμοποιήσαμε και τους συντελεστές προσέγγισης για να εξάγουμε χαρακτηριστικά και να τα χρησιμοποιήσουμε επιπλέον των χαρακτηριστικών που είχαμε εξάγει από τους συντελεστές λεπτομέρειας. Κρατούσαμε διαφορετικά πλήθη DCT συντελεστών κάθε φορά και εφαρμόσαμε σε αυτούς την Student's t και Cauchy κατανομή. Όσο αυξάναμε τους συντελεστές DCT, παρατηρούσαμε συνήθως αύξηση των τιμών των συντελεστών, με λίγες εξαιρέσεις που αφορούσαν συγκεκριμένες παραμορφώσεις εικόνων.

Συμπερασματικά, ο συνδυασμός του Διακριτού Μετασχηματισμού Wavelet (DWT) σε συνδυασμό με την προσαρμογή με μία κατάλληλη στατιστική κατανομή, όπως η Student's t, μας οδηγεί σε βελτιωμένα αποτελέσματα, σε σχέση με αντίστοιχες μεθοδολογίες που κάνουν χρήση των Log-Gabor ή Gabor φίλτρων. Κάτι τέτοιο επιτυγχάνεται ακόμη περισσότερο με την προτεινόμενη εκμετάλλευση της κρυφής πληροφορίας, η οποία εμπλέκεται κατά την εκτίμηση των παραμέτρων της Student's t. Το σημαντικό είναι, ότι μας δίνεται η δυνατότητα να υλοποιήσουμε διάφορες εκδοχές χειρισμού των συντελεστών ενδιαφέροντος. Επίσης, τα χαρακτηριστικά που προκύπτουν από τον μετασχηματισμό κυματιδίων, μέσω της κατανομής Student't και των πιο ειδικών μορφών της, φαίνεται να είναι πιο ικανά να αντιληφθούν της διάφορες στρεβλώσεις της εικόνας από ότι τα χαρακτηριστικά πού πεδίο, και είναι ακόμα πιο αποτελεσματικά, αν συνδυαστούν με χαρακτηριστικά από το Μετασχηματισμό Συνημιτόνου.

Μελλοντικά, μπορεί να εξεταστεί, αν η χρήση επιπλέον αντιληπτικών χαρακτηριστικών, π.χ. χαρακτηριστικά υφής, μπορούν να βελτιώσουν την επίδοση

πρόβλεψης της προτεινόμενης μεθοδολογίας είτε από τους Full-reference αλγορίθμους είτε από τους No-reference. Επιπλέον, μπορεί να διερευνηθεί η χρήση των contourlets, ως μετασχηματισμού που ενδυναμώνει την visual saliency πληροφορία και την σχέση της με την αξιολόγηση της ποιότητας μια εικόνας. Τέλος, θα είχε αξία η διερεύνηση της προσέγγισης που αναπτύξαμε, με χαμηλότερη υπολογιστική πολυπλοκότητα, αλλά διατηρώντας τα επίπεδα πρόβλεψης της ποιότητας της εικόνας.

ΒΙΒΛΙΟΓΡΑΦΙΑ

- [1] Moorthy, A. K., & Bovik, A. C. (2011). Blind image quality assessment: From natural scene statistics to perceptual quality. IEEE transactions on Image Processing, 20(12), 3350-3364.
- [2] Gao, F., Wang, Y., Li, P., Tan, M., Yu, J., & Zhu, Y. (2017). Deepsim: Deep similarity for image quality assessment. Neurocomputing, 257, 104-114.
- [3] Wasson, V., & Kaur, B. (2019, March). Full Reference Image Quality Assessment from IQA Datasets: A Review. In 2019 6th International Conference on Computing for Sustainable Global Development (INDIACom) (pp. 735-738). IEEE.
- [4] Ma, K., & Fang, Y. (2021, October). Image Quality Assessment in the Modern Age. In Proceedings of the 29th ACM International Conference on Multimedia (pp. 5664-5666).
- [5] Mohammadi, P. Ebrahimi-Moghadam, A. Shirani, S. (2014). "Subjective and Objective Quality Assessment of Image: A Survey". Majlesi Journal of Electrical Engineering, 9, 55-83.
- [6] Mittal, A., Moorthy, A. K., & Bovik, A. C. (2012). No-reference image quality assessment in the spatial domain. IEEE Transactions on image processing, 21(12), 4695-4708.
- [7] Saad, M. A., Bovik, A. C., & Charrier, C. (2010). A DCT statistics-based blind image quality index. IEEE Signal Processing Letters, 17(6), 583-586.
- [8] Wang, Z., & Bovik, A. C. (2011). Reduced-and no-reference image quality assessment. IEEE Signal Processing Magazine, 28(6), 29-40.
- [9] Ghadiyaram, D., & Bovik, A. C. (2015). Massive online crowdsourced study of subjective and objective picture quality. IEEE Transactions on Image Processing, 25(1), 372-387.
- Borji, A., Sihite, D. N., & Itti, L. (2012). Quantitative analysis of human-model agreement in visual saliency modeling: A comparative study. IEEE Transactions on Image Processing, 22(1), 55-69.

- [11] Fecteau, J. H., & Munoz, D. P. (2006). Salience, relevance, and firing: a priority map for target selection. Trends in cognitive sciences, 10(8), 382-390.
- [12] Itti, L., Koch, C., & Niebur, E. (1998). A model of saliency-based visual attention for rapid scene analysis. IEEE Transactions on pattern analysis and machine intelligence, 20(11), 1254-1259.
- [13] Engelke, U., Kaprykowsky, H., Zepernick, H. J., & Ndjiki-Nya, P. (2011).
 Visual attention in quality assessment. IEEE Signal Processing Magazine, 28(6), 50-59.
- [14] Ma, Q., & Zhang, L. (2008, December). Image quality assessment with visual attention. In 2008 19th International Conference on Pattern Recognition (pp. 1-4). IEEE.
- [15] Farias, M. C., & Akamine, W. Y. (2012). On performance of image quality metrics enhanced with visual attention computational models. Electronics letters, 48(11), 631-633.
- [16] Sheikh, H. R. (2005). LIVE image quality assessment database release2. http://live. ece. utexas. edu/research/quality.
- [17] Sheikh, H. R., Sabir, M. F., & Bovik, A. C. (2006). A statistical evaluation of recent full reference image quality assessment algorithms. IEEE Transactions on image processing, 15(11), 3440-3451.
- [18] Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). Image quality assessment: from error visibility to structural similarity. IEEE transactions on image processing, 13(4), 600-612.
- [19] Ghadiyaram, D., & Bovik, A. (2017). Live in the Wild Image Quality Challenge Database.
 2015. Online: http://live.ece.utexas.edu/research/ChallengeDB/index.html.
- [20] Rehman, A., & Wang, Z. (2012). Reduced-reference image quality assessment by structural similarity estimation. IEEE transactions on image processing, 21(8), 3378-3389.
- [21] Luo, M. R., Cui, G., & Rigg, B. (2001). The development of the CIE 2000 colour-difference formula: CIEDE2000. Color Research & Application 26(5), 340-350.

- [22] Toet, A., & Lucassen, M. P. (2003). A new universal colour image fidelity metric. Displays, 24(4-5), 197-207.
- [23] Wang, Z., Lu, L., & Bovik, A. C. (2004). Video quality assessment based on structural distortion measurement. Signal processing: Image communication, 19(2), 121-132.
- [24] Zhang, L., Zhang, L., Mou, X., & Zhang, D. (2011). FSIM: A feature similarity index for image quality assessment. IEEE transactions on Image Processing, 20(8), 2378-2386.
- [25] Shi, Y., Ding, Y., Zhang, R., & Li, J. (2009, February). Structure and hue similarity for color image quality assessment. In 2009 International Conference on Electronic Computer Technology (pp. 329-333). IEEE.
- [26] Kikuchi, H., Huttunen, H., Hwang, J., Yukawa, M., Muramatsu, S., & Shin, J. (2012, December). Color-tone similarity on digital images. In Proceedings of The 2012 Asia Pacific Signal and Information Processing Association Annual Summit and Conference (pp. 1-4). IEEE.
- [27] Wang, Q., Chu, J., Xu, L., & Chen, Q. (2016). A new blind image quality framework based on natural color statistic. Neurocomputing, 173, 1798-1810.
- [28] https://www.geeksforgeeks.org/difference-between-rgb-cmyk-hsv-and-yiqcolor-models/
- [29] Παναγοπούλου, Δ. Ε. (2017). Απεικόνιση εικόνων υψηλής δυναμικής περιοχής (Doctoral dissertation).
- [30] Bishop, C. M., & Nasrabadi, N. M. (2006). Pattern recognition and machine learning (Vol. 4, No. 4, p. 738). New York: springer.
- [31] Zhang, W., Zou, W., & Yang, F. (2019). Linking visual saliency deviation to image quality degradation: A saliency deviation-based image quality index. Signal Processing: Image Communication, 75, 168-177.
- [32] Gu, K., Zhai, G., Yang, X., & Zhang, W. (2014). Using free energy principle for blind image quality assessment. IEEE Transactions on Multimedia, 17(1), 50-63.
- [33] Kang, L., Ye, P., Li, Y., & Doermann, D. (2015, September). Simultaneous estimation of image quality and distortion via multi-task convolutional neural

networks. In 2015 IEEE international conference on image processing (ICIP) (pp. 2791-2795). IEEE.

- [34] Li, C., Bovik, A. C., & Wu, X. (2011). Blind image quality assessment using a general regression neural network. IEEE Transactions on neural networks, 22(5), 793-799.
- [35] L. Zhang, Z. Gu, X. Liu, H. Li, J. Lu, Training quality-aware filters for noreference image quality assessment, IEEE MultiMedia 21 (4) (2014) 67–75
- [36] Hou, W., Gao, X., Tao, D., & Li, X. (2014). Blind image quality assessment via deep learning. IEEE transactions on neural networks and learning systems, 26(6), 1275-1286.
- [37] Kang, L., Ye, P., Li, Y., & Doermann, D. (2014). Convolutional neural networks for no-reference image quality assessment. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1733-1740).
- [38] Temel, D., Prabhushankar, M., & AlRegib, G. (2016). UNIQUE: Unsupervised image quality estimation. IEEE signal processing letters, 23(10), 1414-1418.
- [39] Moorthy, A. K., & Bovik, A. C. (2009). Visual importance pooling for image quality assessment. IEEE journal of selected topics in signal processing, 3(2), 193-201.
- [40] Barland, R., & Saadane, A. (2006, October). Blind quality metric using a perceptual importance map for JPEG-20000 compressed images. In 2006 International Conference on Image Processing (pp. 2941-2944). IEEE.
- [41] Osberger, W. M., & Rohaly, A. M. (2001, June). Automatic detection of regions of interest in complex video sequences. In Human vision and electronic imaging VI (Vol. 4299, pp. 361-372). SPIE.
- [42] Imamoglu, N., Lin, W., & Fang, Y. (2012). A saliency detection model using low-level features based on wavelet transform. IEEE transactions on multimedia, 15(1), 96-105.
- [43] Li, H. C., Krylov, V. A., Fan, P. Z., Zerubia, J., & Emery, W. J. (2015). Unsupervised learning of generalized gamma mixture model with application in statistical modeling of high-resolution SAR images. IEEE Transactions on Geoscience and Remote Sensing, 54(4), 2153-2170.

- [44] Rami, H., El Maliani, A. D., & El Hassouni, M. (2019). A finite mixture of weibull-based statistical model for texture retrieval in the complex wavelet domain. IEEE Access, 7, 130144-130155.
- [45] https://www.mathworks.com/help/vision/ref/psnr.html
- [46] https://en.wikipedia.org/wiki/Peak_signal-to-noise_ratio
- [47] Τολιόπουλος, Ι. (2013) Ανάπτυξη και συγκριτική μελέτη μεθόδων αποτίμησης ποιότητας ψηφιακών εικόνων.
- [48] Λάλος, Β. (2012). Αξιολόγηση ποιότητας εικόνας.
- [49] Chandler, D. M., & Hemami, S. S. (2007). VSNR: A wavelet-based visual signal-to-noise ratio for natural images. IEEE transactions on image processing, 16(9), 2284-2298.
- [50] https://en.wikipedia.org/wiki/Visual_information_fidelity
- [51] Shi, C., & Lin, Y. (2020). Full reference image quality assessment based on visual salience with color appearance and gradient similarity. IEEE Access, 8, 97310-97320.
- [52] de Weijer, J. V., & Schmid, C. (2006, May). Coloring local feature extraction. In European conference on computer vision (pp. 334-348). Springer, Berlin, Heidelberg.
- [53] Gevers, T., Gijsenij, A., Van de Weijer, J., & Geusebroek, J. M. (2012). Color in computer vision: fundamentals and applications (Vol. 23). John Wiley & Sons.
- [54] Ou, F. Z., Wang, Y. G., & Zhu, G. (2019, September). A novel blind image quality assessment method based on refined natural scene statistics. In 2019 IEEE International Conference on Image Processing (ICIP) (pp. 1004-1008). IEEE.
- [55] Mairgiotis, A., Tsampra, D., & Kondi, L. P. (2021, June). Improved hybrid blind IQA using alternative NSS characterization in the spatial domain. In 2021 Picture Coding Symposium (PCS) (pp. 1-5). IEEE.
- [56] Tsihrintzis, G. A., & Nikias, C. L. (1996). Fast estimation of the parameters of alpha-stable impulsive interference. IEEE transactions on signal processing, 44(6), 1492-1503.

[57] Saad, M. A., Bovik, A. C., & Charrier, C. (2012). Blind image quality assessment: A natural scene statistics approach in the DCT domain. IEEE transactions on Image Processing, 21(8), 3339-3352.