

ΡΑΠΤΗΣ
ΕΥΑΓΓΕΛΟΣ
ΠΑΥΛΟΣ

ΕΦΑΡΜΟΓΕΣ ΤΗΣ ΘΕΩΡΙΑΣ ΤΩΝ
WAVELETS ΣΤΗΝ ΣΥΜΠΙΕΣΗ
ΣΗΜΑΤΟΣ ΚΑΙ ΣΤΙΣ
ΤΗΛΕΠΙΚΟΙΝΩΝΙΕΣ

**Εφαρμογές της Θεωρίας των Wavelets στην
Συμπίεση Σήματος
και στις Τηλεπικοινωνίες**

του

Ράπτη Ευάγγελου Παύλου

**Α.Τ.Ε.Ι. ΚΑΛΑΜΑΤΑΣ
ΠΑΡΑΡΤΗΜΑ ΣΠΑΡΤΗΣ
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ
ΤΗΛΕΠΙΚΟΙΝΩΝΙΩΝ**

2013

Επιβλέπων Καθηγητής: Μποζαντζής Βασίλειος

*στον πατέρα μου
που έφυγε τόσο νωρίς
και
στους ανθρώπους που
με στηρίζουν ...*

ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ	4
Πίνακας Σχημάτων	5
Πίνακας Εικόνων	6
Περίληψη	9
Κεφάλαιο 1 : Κλιμακοθετησιμότητα	10
1.1 Η Ανάγκη για Κλιμακοθετήσιμη Κωδικοποίηση Βίντεο	10
1.2 Χωρική Κλιμακοθετησιμότητα (Spatial Scalability)	11
1.3 Ποιοτική Κλιμακοθετησιμότητα) (Quality Scalability)	12
1.4 Χρονική Κλιμακοθετησιμότητα (Temporal Scalability)	13
1.5 Κλιμακοθετησιμότητα στη Πολυπλοκότητα (Complexity Scalability)	13
Κεφάλαιο 2 : Σήματα και Αναπαραστάσεις	15
2.1 Σήματα και Πληροφορία	15
2.2 Αναπαράσταση στο Χρόνο και τη Συχνότητα	16
2.3 Μια Εναλλακτική Αναπαράσταση	20
2.4 Ανάλυση	26
Κεφάλαιο 3 : Εισαγωγή στα Wavelets	27
3.1 Σύντομη Επισκόπηση Βασικών Εννοιών	27
3.2 Wavelets και Συστοιχίες Φίλτρων	32
3.3 Επιλογή Wavelet	36
Κεφάλαιο 4 : Εφαρμογές	40
Κεφάλαιο 5 : Εφαρμογή Matlab	41
5.1 Εισαγωγή	41
5.2 Λεπτομέρειες Χρήσης	41
Βιβλιογραφία	72

ΠΙΝΑΚΑΣ ΣΧΗΜΑΤΩΝ

Σχήμα 1.1: Αντιπροσωπευτικό σενάριο για video streaming μέσω του internet, όπου πολλοί πελάτες συνδέονται μέσω ετερογενών ασύρματων πρωτοκόλλων	11
Σχήμα 1.2: Προσαρμογή σε ετερογενή συσκευές και υποδομές δικτύων με το πρότυπο MPEG-21 DIA	12
Σχήμα 2.1: Ένα απλό διακριτό σήμα διακριτού χρόνου, αναπαράσταση στο χρόνο	16
Σχήμα 2.2: Αναπαράσταση του σήματος του σχ. 2.1 στο χώρο της συχνότητας (με χρήση ταχύ μετασχηματισμού Fourier-FFT)	17
Σχήμα 2.3: Το αρχικό σήμα συν ένας ισχυρός μοναδιαίος παλμός	17
Σχήμα 2.4: Εφαρμογή FFT στο σχ. 2.3	18
Σχήμα 2.5: Αρχική εικόνα (αριστερά) και εμφάνιση συνοριακών ασυνεχειών στα άκρα των παράθυρων ανάλυσης (δεξιά) σε υψηλή συμπίεση JPEG	18
Σχήμα 2.6: Επιλογή παραθύρου για το Μετασχηματισμό Fourier Βραχέως Χρόνου (STFT)	19
Σχήμα 2.7: Προσέγγιση της $a(n)$ με την $\varphi(n)$	20
Σχήμα 2.8: Προσέγγιση της $a(n)$ με την $\varphi(2n)$	21
Σχήμα 2.9: Προσέγγιση της $a(n)$ με την $\varphi(4n)$	22
Σχήμα 2.10: Διαδοχικές προσεγγίσεις του $a(n)$ με συνεισφορές από τα $\psi_k(2^m n)$	24
Σχήμα 2.11: Το $\psi(n)$ στο χρόνο και στη συχνότητα (με χρήση FFT)	24
Σχήμα 3.1: Γεωμετρικό ανάλογο μιας ορθογώνιας (a) και μιας διορθογώνιας (b) βάσης	31
Σχήμα 3.2: Υπολογισμός των σημάτων προσέγγισης και λεπτομερειών μέσω μιας συστοιχίας φίλτρων	35
Σχήμα 3.3: Μετασχηματισμός wavelet τριών σταδίων	36
Σχήμα 3.4: Συνάρτηση κλίμακας και wavelet Cohen-Daubechies-Feauveau 9/7	37
Σχήμα 3.5: Σταδιακή αποδόμηση του σήματος leleccum με χρήση του μετασχηματισμού wavelet CDF9/7	38
Σχήμα 3.6: (συνέχεια 3.5) Σταδιακή αποδόμηση του σήματος leleccum με χρήση του μετασχηματισμού wavelet CDF9/7	39

ΠΙΝΑΚΑΣ ΕΙΚΟΝΩΝ

Εικόνα 1: Matlab	41
Εικόνα 2: Matlab example	42
Εικόνα 3: Matlab run example	42
Εικόνα 4: Matlab Editor	43
Εικόνα 5: Matlab Editor Rere	44
Εικόνα 6: Matlab Renate-Demy	44
Εικόνα 7: mse , psnr	45
Εικόνα 8: Αυτοκίνητο με βήμα 2 (mse =5.7229e+003 και psnr = 10.5546)	45
Εικόνα 9: Αυτοκίνητο με βήμα 5 (mse =2.0796e+003 και psnr = 14.9509)	46
Εικόνα 10: Αυτοκίνητο με βήμα 8 (mse =592.6064 και psnr = 20.4031)	46
Εικόνα 11: Αυτοκίνητο με βήμα 10 (mse =249.4931 και psnr = 24.1602)	47
Εικόνα 12: Αυτοκίνητο με βήμα 15 (mse =11.1308 και psnr = 37.6655)	47
Εικόνα 13: Παρθενώνας με βήμα 2 (mse =3.2450 e+003 και psnr = 24.1602)	48
Εικόνα 14: Παρθενώνας με βήμα 5 (mse =1.0584 e+003 και psnr = 17.8841)	48
Εικόνα 15: Παρθενώνας με βήμα 10 (mse =115.2553 e+003 και psnr = 27.5142)	49
Εικόνα 16: Παρθενώνας με βήμα 15 (mse =2.7866 e+003 και psnr = 43.6801)	49
Εικόνα 17: Lasme με βήμα 2 (mse =1.9421 e+003 και psnr = 15.2482)	50
Εικόνα 18: Lasme με βήμα 5 (mse =1.4849 e+003 και psnr = 16.4139)	50
Εικόνα 19: Lasme με βήμα 12 (mse =32.2998 e+003 και psnr = 33.0388)	51

Εικόνα 20: Lasme με βήμα 15 (mse =3.8268 e+003 και psnr = 42.3025)	51
Εικόνα 21: Matlab Example 2	52
Εικόνα 22: Mask Original image - Compressed image	52
Εικόνα 23: Matlab mse,psnr	53
Εικόνα 24: Matlab example 3	54
Εικόνα 25: Matlab arms 1	54
Εικόνα 26: Matlab arms 2	55
Εικόνα 27: Matlab arms 3	55
Εικόνα 28: Matlab arms 4	56
Εικόνα 29: woodstatue loop 2	57
Εικόνα 30: 30 woodstatue loop 2	57
Εικόνα 31: 31 woodstatue loop 5	58
Εικόνα 32: woodstatue loop 2 new	58
Εικόνα 33: woodstatue loop 4	59
Εικόνα 34: : woodstatue loop 5	59
Εικόνα 35: woodstatue loop 9	60
Εικόνα 36: woodstatue compressed image	60
Εικόνα 37: Ολόκληρο το Ο.Α.Κ.Α. με βήμα 2	61
Εικόνα 38: Ολόκληρο το Ο.Α.Κ.Α. με βήμα 5	61
Εικόνα 39: Ολόκληρο το Ο.Α.Κ.Α. με βήμα 8	62
Εικόνα 40: Ολόκληρο το Ο.Α.Κ.Α. με βήμα 10	62

Εικόνα 41: Ολόκληρο το Ο.Α.Κ.Α. με βήμα 12	63
Εικόνα 42: Ολόκληρο το Ο.Α.Κ.Α. με βήμα 15	63
Εικόνα 43: Ο.Α.Κ.Α. με βήμα 2	64
Εικόνα 44: Ο.Α.Κ.Α. με βήμα 5	64
Εικόνα 45: Ο.Α.Κ.Α. με βήμα 8	65
Εικόνα 46: Ο.Α.Κ.Α. με βήμα 10	65
Εικόνα 47: Ο.Α.Κ.Α. με βήμα 12	66
Εικόνα 48: Ο.Α.Κ.Α. με βήμα 15	66

Περίληψη

Σκοπός της πτυχιακής αυτής εργασίας είναι η σχεδίαση, η ανάλυση και η εφαρμογή των wavelets στην συμπίεση σήματος και στις τηλεπικοινωνίες.

Αρχικά θα αναλύσω την κλιμακοθετησιμότητα και τους κλάδους της σε χωρική, ποιοτική και στην πολυπλοκότητα.

Στην συνέχεια θα κάνω μια σύντομη εισαγωγή στα wavelets και θα δείξω τις περιπτώσεις στις οποίες μας είναι απαραίτητα. Ταυτόχρονα, θα συνάγουμε (με μαθηματική χαλαρότητα) κάποιες σημαντικές έννοιες που τα διέπουν, και οι οποίες είναι αναγκαίες για τη συνέχεια της εργασίας. Επιπλέον, θα αναφερθούμε στις κυριότερες ως τώρα εφαρμογές τους και τη σημασία υλοποίησής τους σε υλικό.

Τέλος, στο πρακτικό μέρος και με την βοήθεια του προγράμματος Matlab θα γράψω έναν κώδικα ο οποίος διαβάζει την εικόνα, την συμπιέζει και εκτυπώνει την αρχική και την ανακατασκευασμένη εικόνα σε διάφορα βήματα υπολογίζοντας παράλληλα το μέσο τετραγωνικό λάθος (MSE) και τον μέγιστο σηματοθορυβικό λόγο (PSNR).

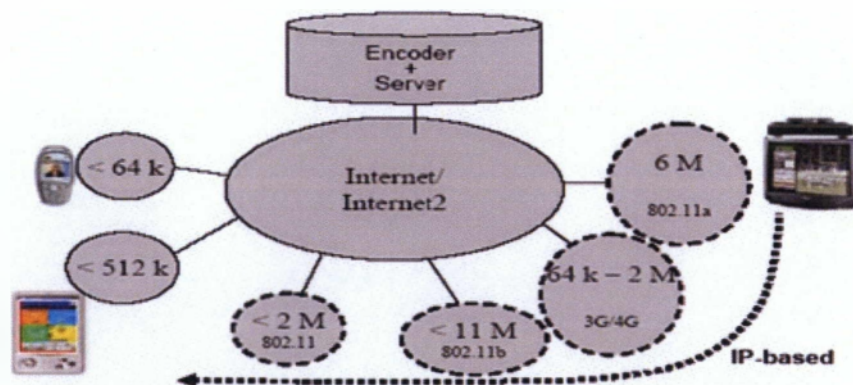
Κεφάλαιο 1 Κλιμακοθετησιμότητα

1.1 Ανάγκη για κλιμακοθετήσιμη κωδικοποίηση βίντεο

Στις μέρες μας ολοένα και περισσότερο αυξάνεται ο όγκος των δεδομένων βίντεο που χειριζόμαστε καθημερινά. Μια σημαντική πρόκληση είναι η διανομή του περιεχομένου βίντεο σε διαφορετικούς χρήστες. Το ψηφιακό βίντεο στη δεκαετία 1980 διανεμόταν μεταξύ εταιριών που διέθεταν ακριβά και αξιόπιστα κανάλια με μεγάλο εύρος ζώνης και δυνατότητα παροχής ποιότητας εξυπηρέτησης (QoS). Όμως, από το 1995 και έπειτα η εξάπλωση του Παγκόσμιου Ιστού άλλαξε τη φιλοσοφία της εμπορικής διανομής, με αποτέλεσμα σήμερα εκατομμύρια χρήστες να χρησιμοποιούν υπηρεσίες βίντεο ροοθήκευσης (video-streaming) για να δούνε ταινίες, μουσικά βίντεο και άλλα γεγονότα μέσω του Παγκόσμιου Ιστού. Επιπλέον, η ανάπτυξη της τεχνολογίας σύλληψης εικόνας προκάλεσε τη διεύρυνση καμερών εικόνας και βίντεο στις κινητές συσκευές, όπως τα κινητά τηλέφωνα και τα PDAs. Αυτό είχε ως αποτέλεσμα στις μέρες μας, ακόμα και οι φορητές συσκευές να λειτουργούν ως καταναλωτές ή και παραγωγοί ψηφιακού βίντεο.

Στις περισσότερες περιπτώσεις η μετάδοση βίντεο πραγματοποιείται μέσω μη αξιόπιστων δικτύων, χρησιμοποιώντας πρωτόκολλα βασιζόμενα στα πακέτα όπως για παράδειγμα το TCP/IP το οποίο δεν υποστηρίζει λειτουργίες ποιότητας εξυπηρέτησης για ανθεκτική και χρονικά ορθή διανομή περιεχομένου πολυμέσων. Το πρόβλημα διογκώνεται με την χρήση ασύρματων υπηρεσιών στο περιβάλλον του χρήστη, αφού τα ασύρματα δίκτυα εμφανίζουν απώλειες πακέτων και μεγάλες διακυμάνσεις στο εύρος ζώνης. Για την τυπική περίπτωση της βίντεο ροοθήκευσης η σημερινή κατάσταση απεικονίζεται στην Εικόνα 1.1, όπου παρουσιάζονται διάφορες συσκευές και δίκτυα μαζί με το αντίστοιχο απαιτούμενο εύρος επικοινωνίας [1].

Motion Compensation – Scalable Video Coding



Σχήμα 1.1: Αντιπροσωπευτικό σενάριο για video streaming μέσω του internet, όπου πολλοί πελάτες συνδέονται μέσω ετερογενών ασύρματων πρωτοκόλλων

Για να αντιμετωπιστούν τα προβλήματα που δημιουργεί η ποικιλία των μέσων μετάδοσης βίντεο, η κοινότητα του MPEG αποφάσισε να διεξάγει έρευνα στις τεχνολογίες κωδικοποίησης βίντεο που μπορούν να χειριστούν τις απαιτήσεις μετάδοσης. Η έρευνα αυτή άρχισε το 2002 και συνεχίζεται ακόμα. Όσον αφορά το τεχνολογικό κομμάτι είναι γενικώς αποδεκτό ότι η κλιμακοθετήσιμη κωδικοποίηση (scalable coding) είναι πιο βιώσιμη τεχνικά λύση στα προβλήματα της μετάδοσης βίντεο σε ετερογενή δίκτυα και τερματικά χρηστών με διαφορετικές δυνατότητες. Η κλιμακοθετήσιμη κωδικοποίηση βίντεο αποτελεί ένα πλαίσιο εργασίας όπου από μια ενιαία συμπίεσμένη ακολουθία βίντεο μπορούν να προκύψουν αναπαραστάσεις του βίντεο με διαφορετική ποιότητα, ανάλυση (resolution) και χρονικό ρυθμό πλαισίων.

Ένα σημαντικό κομμάτι της επιτυχίας ενός συγκεκριμένου κλιμακοθετήσιμου αλγορίθμου κωδικοποίησης είναι το πόσο αποδοτική είναι η συμπίεση που δημιουργήσε. Συνεπώς, μια βασική απαίτηση κάθε αλγορίθμου κλιμακοθετήσιμης κωδικοποίησης είναι ότι πρέπει να αποδίδει το ίδιο καλά με τα υπάρχοντα μη κλιμακοθετήσιμα συστήματα κωδικοποίησης βίντεο όσον αφορά την οπτική ποιότητα. Αυτό θα πρέπει να συμβαίνει για την πλειοψηφία των απεικονίσεων, των επιπέδων ποιότητας και των χρονικών ρυθμών πλαισίων. Μια άλλη απαίτηση αφορά τον χρόνο αποκωδικοποίησης. Γενικά για βίντεο ροοθήκευσης (video streaming) ο χρόνος αποκωδικοποίησης είναι της τάξης δευτερολέπτων, ενώ για εφαρμογές όπως παρακολούθηση βίντεο (video monitoring) απαιτείται χρόνος της τάξης των 100 msec. Όσον αφορά τις απαιτήσεις για κλιμακοθετήσιμα (scalability) του περιεχομένου βίντεο, το MPEG ορίζει κάποιον ελάχιστο αριθμό χωρικών, χρονικών και ποιοτικών επιπέδων που πρέπει να διατίθενται. Τέλος, θα πρέπει να λάβουμε υπόψη και την πολυπλοκότητα της κωδικοποίησης και αποκωδικοποίησης. Οι απαιτήσεις αυτές αναλύονται περισσότερο στις ακόλουθες υποενότητες.

1.2 Χωρική Κλιμακοθετησιμότητα (Spatial Scalability)

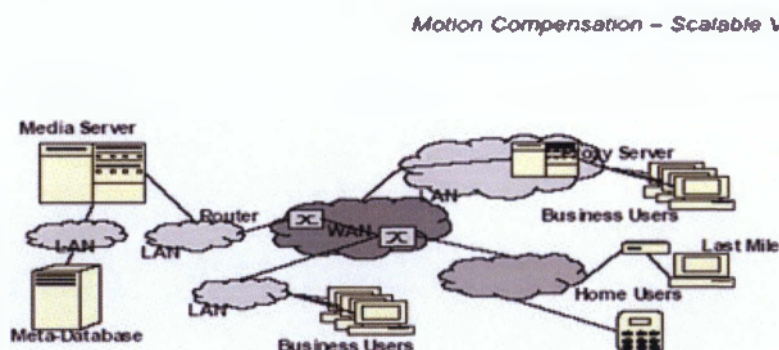
Η κλιμακοθετήσιμη κωδικοποίηση βίντεο πρέπει να υποστηρίζει το λιγότερο τέσσερα επίπεδα χωρικής ανάλυσης (ευκρίνειας) τα οποία κυμαίνονται από 1152 x 1408 έως 144 x 176 εικονοστοιχεία. Στο εύρος αυτό περιέχονται και τα δύο πολύ γνωστά πρότυπα για αναπαραστάση βίντεο: CIF (288 x 352) και QCIF (144 x 176). Η ικανότητα να παράγουμε πολλαπλές αναλύσεις από μία κοινή ακολουθία ψηφίων είναι σημαντική για αρκετές εφαρμογές. Μία από αυτές είναι η επιτήρηση χώρου μέσω βίντεο και τα συστήματα παρακολούθησης, όπου η λειτουργία χωρικής κλιμακοθετησιμότητας έχει δύο λόγους ύπαρξης [1].

Στο πρώτο σενάριο πολλές λήψεις από διαφορετικές θέσεις καταφτάνουν στον σταθμό επιτήρησης. Βασιζόμενοι στη δυνατότητα χωρικής κλιμακοθετησιμότητας κάθε λήψη μπορεί να μεγεθυνθεί. Έτσι μειώνεται το εύρος επικοινωνίας μεταξύ του δικτύου

καμερών και του σταθμού παρακολούθησης και το υψηλής ευκρίνειας περιεχόμενο μεταδίδεται μόνο έπειτα από συναγερμό ή από επιλογή του χρήστη.

Υποστήριξη αποθήκευσης βίντεο με λειτουργία διαγραφής. Στα συστήματα παρακολούθησης η σημαντικότητα των καταγεγραμμένων δεδομένων και κατ' επέκταση της απαιτούμενης ανάλυσης των ακολουθιών βίντεο, μειώνεται με το πέρασμα του χρόνου. Για παράδειγμα, το πλήρες ευκρίνειας βίντεο χρειάζεται να αποθηκευτεί μόνο για τρεις ημέρες. Οι ακολουθίες που είναι για χρονικό διάστημα μεγαλύτερο από μία εβδομάδα χρειάζονται μια ελάχιστη ευκρίνεια για την αρχειοθέτησή τους.

Τέλος, μία μεγάλη κατηγορία συστημάτων που απαιτούν κλιμακοθετησιμότητα στην ανάλυση είναι οι αρχιτεκτονικές επικοινωνίας που χρησιμοποιούν το πλαίσιο εργασίας MPEG-21 DIGITAL Item Adaptation (DIA), όπου ετερογενείς υποδομές επικοινωνίας χρησιμοποιούνται για video streaming προς ένα πλήθος συσκευών όπως απεικονίζεται στην Εικόνα 1.2. Σε αυτήν την περίπτωση, η προσαρμογή της ευκρίνειας βίντεο είναι σημαντική για να ικανοποιηθεί το εύρος ζώνης του καναλιού και η ευκρίνεια του αποκωδικοποιητή βίντεο.



Σχήμα 1.2: Προσαρμογή σε ετερογενή συσκευές και υποδομές δικτύων με το πρότυπο MPEG-21 DIA

1.3 Ποιοτική Κλιμακοθετησιμότητα (Quality Scalability)

Πρακτικά η ιδιότητα η ακολουθία που προκύπτει να είναι κλιμακοθετήσιμη στην ποιότητα σημαίνει να έχουμε κλιμακοθετησιμότητα στο ρυθμό δυαδικών ψηφίων (bit rate). Σύμφωνα με τις απαιτήσεις του MPEG τα κλιμακοθετήσιμα συστήματα κωδικοποίησης πρέπει να παρέχουν μία μεγάλη ποικιλία ρυθμών μετάδοσης δυαδικών ψηφίων που αντιστοιχούν στους ρυθμούς μετάδοσης που υποστηρίζουν τα σημερινά δίκτυα. Ένα πρακτικό παράδειγμα εύρους ρυθμών δυαδικών ψηφίων φαίνεται στην Εικόνα 1.1. Πρέπει επίσης να αναφερθεί ότι για συγκεκριμένες εφαρμογές αρχειοθέτησης βίντεο, όπως η βιομηχανία ψηφιακού σινεμά, η εξερεύνηση του διαστήματος ή κάποιες ιατρικές εφαρμογές, η δυνατότητα για κωδικοποίηση χωρίς απώλειες (lossless) πρέπει να υποστηρίζεται.

Η λειτουργία κλιμακοθετησιμότητας στην παρεχόμενη ποιότητα είναι πολύ χρήσιμη όταν η μετάδοση λαμβάνει χώρα σε μη αξιόπιστα δίκτυα, όπου οι αλλαγές στο εύρος ζώνης μπορούν να αντιμετωπιστούν αποτελεσματικά αλλάζοντας κατά τη μετάδοση (on the-fly) την ποιότητα του μεταδιδόμενου βίντεο. Μία άλλη εφαρμογή έχουμε στη διανομή περιεχομένου, όπου ο χρήστης χρεώνεται ανάλογα με την ποιότητα που επιθυμεί και έτσι με την κλιμακοθετήσιμη κωδικοποίηση από μία ακολουθία μπορούν να προκύψουν τα κατάλληλα επίπεδα ποιότητας που θα μεταδοθούν στους αντίστοιχους χρήστες [1].

1.4 Χρονική κλιμακοθετησιμότητα (Temporal Scalability)

Η χρονική κλιμακοθετησιμότητα της ακολουθίας βίντεο ισοδυναμεί με τη δυνατότητα κλιμακοθετησιμότητας του ρυθμού πλαισίων (frame rate). Ο ρυθμός πλαισίων του βίντεο αντιστοιχεί στον αριθμό των πλαισίων που προβάλλονται στη διάρκεια ενός δευτερολέπτου. Έτσι αλλάζοντας το ρυθμό των πλαισίων στην πραγματικότητα αλλάζουμε την ποιότητα του βίντεο που αντιλαμβάνεται το ανθρώπινο μάτι. Ο πιο συνηθισμένος ρυθμός πλαισίων είναι 30 Hz. Η κλιμακοθετήσιμη κωδικοποίηση βίντεο πρέπει να παρέχει τουλάχιστον τέσσερα επίπεδα χρονικής κλιμακοθετησιμότητας με ρυθμό αποκωδικοποίησης πλαισίων ως και 60 Hz [1],

Η κύρια εφαρμογή της χρονικής κλιμακοθετησιμότητας είναι η παραγωγή και διανομή πολυκαναλικού περιεχομένου, όπου η ίδια ακολουθία βίντεο προορίζεται για ένα πλήθος συσκευών που υποστηρίζουν διαφορετικές χρονικές αναλύσεις. Για παράδειγμα 7.5 Hz, 15 Hz, 30 Hz και 60 Hz πρέπει να υποστηρίζονται ώστε να εξυπηρετηθούν ποικίλοι χρήστες, από συσκευές στούντιο μέχρι βίντεο κατά απαίτηση (video-on-demand) σε κινητά τηλέφωνα με μικρές επεξεργαστικές ικανότητες. Επιπλέον, η χρονική κλιμακοθετησιμότητα μπορεί να χρησιμοποιηθεί για να εξομοιώσει λειτουργίες fast forward/backward που συναντάμε στις αναλογικές συσκευές, όπως τα βίντεο. Τέλος η χρονική κλιμακοθετησιμότητα βρίσκει εφαρμογή σε εφαρμογές παρακολούθησης, όπου οι κάμερες καταγράφουν στατικές εικόνες και έτσι βίντεο υψηλού χρονικού πλαισίων είναι απαραίτητα μόνο έπειτα από κάποιο συναγερμό.

1.5 Κλιμακοθετησιμότητα στη Πολυπλοκότητα (Complexity scalability)

Μια άλλη σημαντική λειτουργία είναι η κλιμακοθετησιμότητα στην πολυπλοκότητα. Είναι επιθυμητό η πολυπλοκότητα αποκωδικοποίησης να μεταβάλλεται ανάλογα με την χρονική και χωρική ανάλυση. Επίσης, ανάλογα με τους διαθέσιμους πόρους στην πλατφόρμα κωδικοποίησης θα πρέπει να συμβαίνει κατά τη μετάδοση (on-the-fly) προσαρμογή της ακολουθίας δυαδικών ψηφίων, ώστε να επιλέγεται η υποακολουθία που οδηγεί σε μικρής πολυπλοκότητας αποκωδικοποίηση. Η πολυπλοκότητα σχετίζεται με αλγοριθμικά χαρακτηριστικά και με την πλατφόρμα υλοποίησης, οπότε για να έχουμε αξιόπιστα αποτελέσματα θα πρέπει να χρησιμοποιηθούν σωστά μοντέλα πολυπλοκότητας.

Σημαντικές εφαρμογές της κλιμακοθετησιμότητας στην πολυπλοκότητα έχουμε στην περιοχή της ασύρματης βίντεο ροοθήκευσης (video streaming) σε κινητές συσκευές, όπως κινητά τηλέφωνα και PDAs. Επιπλέον, ανάλογα με την εφαρμογή, μπορεί να είναι απαραίτητο η κωδικοποίηση να πληρεί ορισμένα όρια πολυπλοκότητας. Αυτό είναι σημαντικό σε σενάρια όπου οι συσκευές κωδικοποίησης διαμοιράζονται στο δίκτυο, όπως στις εφαρμογές παρακολούθησης.

Κεφάλαιο 2 Σήματα και Αναπαραστάσεις

2.1 Σήματα και πληροφορία

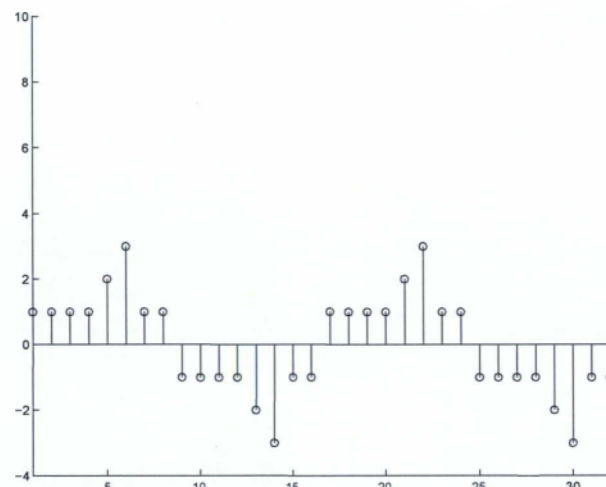
Στην ενότητα αυτή θα κάνουμε μια σύντομη, περισσότερο ποιοτική, εισαγωγή σε βασικές έννοιες σημάτων, με σκοπό κυρίως την αυτάρκεια του κειμένου της εργασίας.

Τόσο ο φυσικός όσο και ο τεχνητός κόσμος αποτελούνται από συστήματα, τα οποία επικοινωνούν μεταξύ τους μέσω σημάτων. Ένα σήμα μπορεί να είναι συνεχούς (π.χ. ηλεκτρική τάση) ή διακριτού χρόνου (π.χ. η έξοδος ενός ψηφιακού κυκλώματος) και να λαμβάνει συνεχείς ή διακριτές τιμές. Επίσης, μπορεί να είναι μιας διάστασης (π.χ. ήχος, τάση ρευματοδότη, καρδιογράφημα), δύο διαστάσεων (π.χ. εικόνα), τριών (π.χ. βίντεο) ή και περισσότερων.

Εμείς εδώ θα ασχοληθούμε με σήματα μιας διάστασης, καθώς τα συμπεράσματα (αλλά και οι μετασχηματισμοί) που προκύπτουν μπορούν εύκολα να μεταφερθούν σε περισσότερες διαστάσεις. Θα θεωρήσουμε ότι η μια αυτή διάσταση είναι ο **χρόνος**. Στην πραγματικότητα, αυτό γίνεται για λόγους διευκόλυνσης και μόνο, καθώς υπάρχει μια πληθώρα σημάτων με μονοδιάστατο πεδίο ορισμού (π.χ. η τάση κατά μήκος ενός αγωγού, μια χρονική στιγμή) που δεν ορίζονται με βάση το χρόνο.

2.2 Αναπαράσταση στο χρόνο και τη συχνότητα

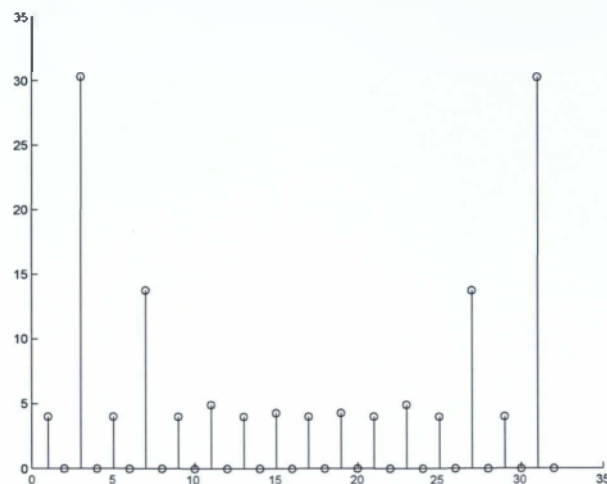
Η πιο συνηθισμένη αναπαράσταση ενός σήματος είναι η αναπαράστασή του στο χρόνο, γνωστή και ως αναπαράσταση πλάτους-χρόνου. Αυτή προκύπτει άμεσα από τις τιμές του σήματος και μας δίνει μια άμεση πληροφορία για τη χρονική του εξέλιξη.



Σχήμα 2.1 Ένα απλό διακριτό σήμα διακριτού χρόνου, αναπαράσταση στο χρόνο

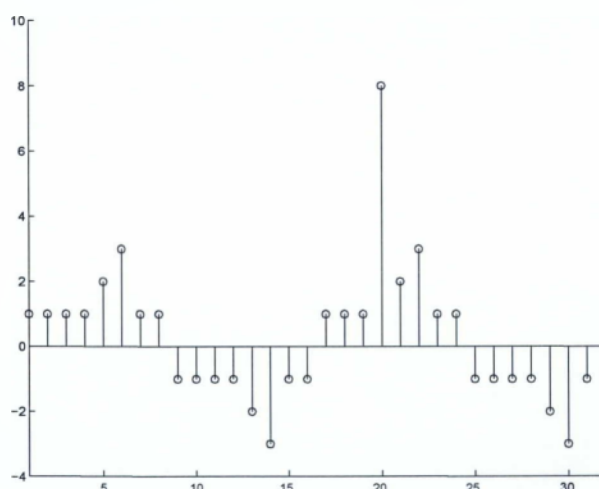
Ωστόσο, είναι αρκετές οι περιπτώσεις που η κύρια πληροφορία για το σύστημα που παράγει το σήμα, ή εν γένει οι κύριες πληροφορίες που μας ενδιαφέρουν, δεν προέρχονται από τη μορφή του σήματος στο χρόνο, αλλά από τη μορφή του σήματος στο χώρο της συχνότητας.

Ένα πολύ διαδεδομένο εργαλείο που αναπαριστά σχεδόν οποιοδήποτε σήμα από το χώρο του χρόνου στο χώρο της συχνότητας είναι ο μετασχηματισμός Fourier και οι «συγγενείς» μετασχηματισμοί συνημιτόνου (DCT) ή τροποποιημένου συνημιτόνου (MDCT). Για παράδειγμα, το σήμα του σχ. 1.1 έχει την εξής μορφή στο χώρο της συχνότητας:



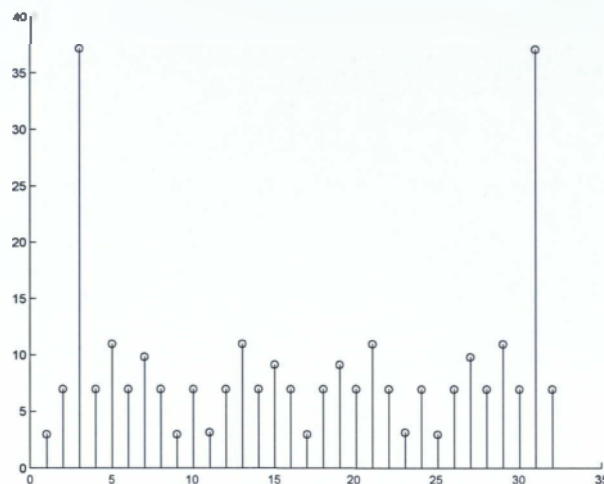
Σχήμα 2.2 Αναπαράσταση του σήματος του σχ. 1.1 στο χώρο της συχνότητας (με χρήση ταχύ μετασχηματισμού Fourier-FFT)

Παρατηρείστε ότι το σχ. 1.2 μας δείχνει *ποιές* συχνότητες απαντώνται στο αρχικό σήμα, αλλά δε μας δίνει καμία πληροφορία για το *που* στο σήμα μπορούμε να τις βρούμε. Για να κάνουμε τα πράγματα χειρότερα, θα προσθέσουμε ένα ισχυρό μοναδιαίο παλμό στο αρχικό μας σήμα:



Σχήμα 2.3 Το αρχικό σήμα συν ένας ισχυρός μοναδιαίος παλμός

Ο παλμός που προστέθηκε στη χρονική στιγμή $t=20$, προφανώς θα αλλάξει αρκετά το φασματικό περιεχόμενο του σήματος, καθώς λόγω της απότομης ανόδου/καθόδου που εισάγει, θα προκαλέσει μια σειρά από καινούριες αρμονικές. Πράγματι, ο γρήγορος μετασχηματισμός Fourier μας δίνει:

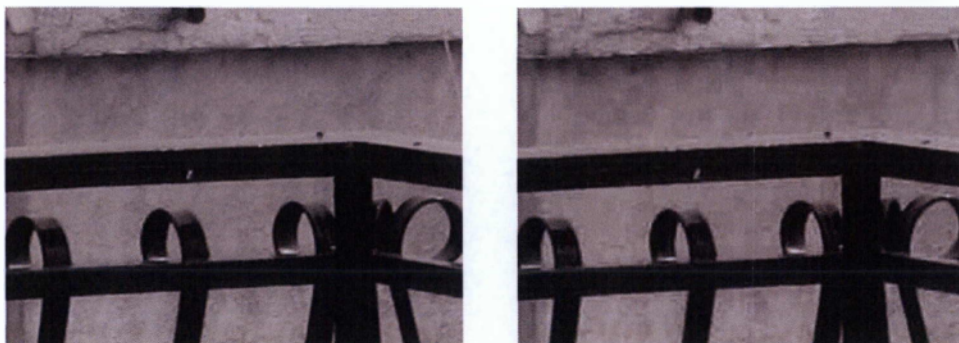


Σχήμα 2.4 Εφαρμογή FFT στο σχ. 2.3

Αν το σχ. 1.2 έδινε το περιεχόμενο συχνότητας ενός σήματος που παράγεται κατά την κανονική λειτουργία ενός μηχανήματος, και το σχ. 1.4 κατά την ανώμαλη λειτουργία, συγκρίνοντάς τα βλέπουμε ότι στη δεύτερη περίπτωση υπήρξε κάποιο απότομο μεταβατικό φαινόμενο. Δεν μπορούμε όμως να διακρίνουμε πότε αυτό έλαβε χώρα. Επίσης, παρατηρούμε ότι παρά την ένταση του παλμού στο σχ. 1.3, οι φασματικές συνιστώσες που αντιστοιχούν σε αυτόν είναι σχετικά μικρές, καθώς η ενέργειά του έχει διαμοιραστεί σε όλο το μήκος του σήματος.

Η ικανότητα συσχέτισης ενός φαινομένου με το χρόνο είναι πολύ σημαντική, συνεπώς πρέπει να βρεθεί κάποιος τρόπος να είναι δυνατό να γνωρίζουμε ταυτόχρονα το περιεχόμενο ενός σήματος στο χώρο της συχνότητας και τη χρονική στιγμή που αυτό μεταβάλλεται. Μια λύση στο πρόβλημα έχει δώσει ο μετασχηματισμός fourier βραχέως χρόνου: χωρίζουμε το αρχικό σήμα σε μια σειρά από μικρότερα σήματα ίσου μήκους, στο κάθε ένα από τα οποία κάνουμε μετασχηματισμό fourier. Το μέγεθος καθενός από αυτά τα μικρότερα σήματα λέγεται *παράθυρο ανάλυσης*.

Ωστόσο, αυτή η τεχνική του διαχωρισμού του αρχικού σήματος σε μικρότερα δημιουργεί τεχνητές συνοριακές συνθήκες στα άκρα κάθε ενός από τα παράθυρα ανάλυσης και επιπλέον, λόγω του ότι το καθένα από τα δύο στοιχεία εκατέρωθεν του συνόρου δέχεται ξεχωριστή επεξεργασία, είναι πολύ πιθανό, αν επιχειρηθεί ανασύνθεση του αρχικού σήματος μέσω των αποτελεσμάτων του μετασχηματισμού fourier

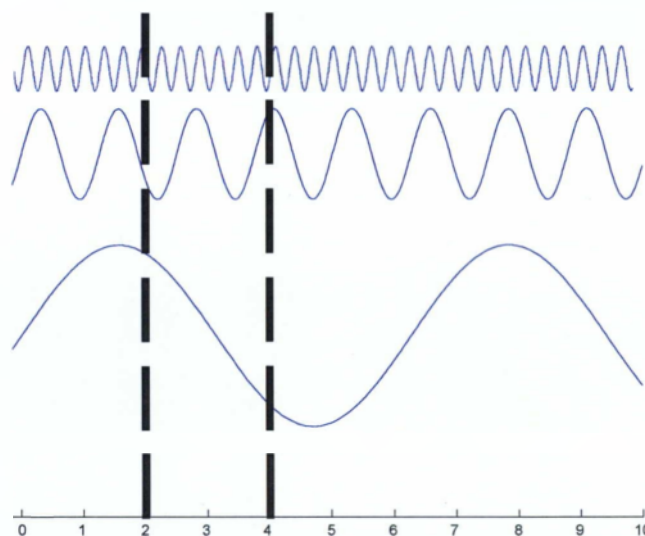


Σχήμα 2.5 Αρχική εικόνα (αριστερά) και εμφάνιση συνοριακών ασυνεχειών στα άκρα των παραθύρων ανάλυσης (δεξιά) σε υψηλή συμπίεση JPEG

βραχέως χρόνου να δημιουργηθούν τεχνητές ασυνέχειες στα σημεία τομής. Είναι για

παράδειγμα πολύ γνωστές οι ασυνέχειες που εμφανίζει σε μια εικόνα υπό υψηλή συμπίεση ο αλγόριθμος JPEG (ο οποίος χρησιμοποιεί το (διακριτό) μετασχηματισμό συνημιτόνου σε ομάδες 8x8 pixel, πολύ όμοιο με το μετασχηματισμό fourier).

Το πρόβλημα των τεχνητών συνόρων μπορεί να λυθεί με διάφορους τρόπους, για παράδειγμα με επιλογή παραθύρων που αλληλοεπικαλύπτονται και απόρριψη των οριακών σημείων κατά την ανασύνθεση. Αυτό το τέχνασμα χρησιμοποιείται από τον τροποποιημένο διακριτό μετασχηματισμό συνημιτόνου (MDCT-Modified Discrete Cosine Transform) που χρησιμοποιούν οι αλγόριθμοι συμπίεσης ήχου MP3, AAC και AC3. Ωστόσο, ένα πρόβλημα που παραμένει είναι ότι και πάλι, η επιλογή του μεγέθους του παραθύρου αποτελεί έναν συμβιβασμό.



Σχήμα 2.6 Επιλογή παραθύρου για το Μετασχηματισμό Fourier Βραχέως Χρόνου (STFT)

Για παράδειγμα, βλέπουμε στο σχ. 1.5 ότι το μέγεθος του παραθύρου που επιλέξαμε είναι αρκετά καλό για τη μεσαία ημιτονική συνάρτηση. Ωστόσο, το παράθυρο είναι πολύ μικρό σε μήκος για να μπορούμε να έχουμε πληροφορία για μικρότερες συχνότητες (αφού η περιόδός τους ολοκληρώνεται σε πολλά μήκη του) ενώ για τις υψηλότερες συχνότητες, εμφανίζεται πάλι το πρόβλημα της αδυναμίας εντοπισμού.

Το τελευταίο πρόβλημα είναι επακόλουθο μιας πολύ γενικότερης αρχής που ονομάζεται **αρχή της αβεβαιότητας**. Η αρχή αυτή μας λέει ότι είναι αδύνατο να έχουμε ταυτόχρονα πλήρη πληροφορία και στο πεδίο του χρόνου και στο πεδίο της συχνότητας. Αν έχουμε όλη την πληροφορία στο πεδίο του χρόνου (σχ. 1.1, 1.3) δεν έχουμε καμία πληροφορία για το πεδίο της συχνότητας. Αντίθετα, αν έχουμε ακριβείς πληροφορίες για το σύνολο του χώρου της συχνότητας (σχ. 1.2, 1.4) δεν έχουμε καμία πληροφορία για το πεδίο του χρόνου. Θα συναντήσουμε την αρχή της αβεβαιότητας και παρακάτω όταν θα αναφερθούμε στον εντοπισμό των συναρτήσεων βάσης.

2.3 Μια εναλλακτική αναπαράσταση

Είδαμε πριν ότι οι κλαστικοί μετασχηματισμοί σημάτων (fourier, συνημιτόνου) ενώ μας δείχνουν ποιες αρμονικές υπάρχουν σε ένα σήμα, δε μας δείχνουν πότε αυτές εμφανίζονται. Επομένως, είναι ακατάλληλοι για σήματα με μεγάλες και μη-κανονικές ασυνέχεις (που ωστόσο, παρουσιάζουν μεγάλο πρακτικό ενδιαφέρον). Εδώ θα προσπαθήσουμε να εισάγουμε μια εναλλακτική αναπαράσταση η οποία είναι κατάλληλη για τέτοια σήματα. Αρχικά, ας ορίσουμε τη διακριτή συνάρτηση διακριτού χρόνου (ακολουθία)

$$\varphi(n) = \begin{cases} 0, n < 0 \\ 1, n(\text{modulo})8=0 \\ 0, \text{αλλού} \end{cases} \quad (1.1)$$

και βάσει αυτής ας ορίσουμε την οικογένεια συναρτήσεων

$$\varphi_k(n) = \varphi(n - 8 * k) \quad (1.2)$$

Είναι φανερό ότι οι διάφορες ακολουθίες που ορίζονται από την (1.2) είναι γραμμικά ανεξάρτητες μεταξύ τους, ενώ μπορούν να αναπαραστήσουν (μέχρι κάποια ανάλυση, βέβαια) κάθε διακριτό σήμα ως γραμμικό συνδυασμό τους.

Για παράδειγμα, το σήμα του σχ. 1.3 μπορεί να προσεγγιστεί από αυτές ως εξής:



Σχήμα 1.7 Προσέγγιση της $a(n)$ με την $\varphi(n)$

Με βάση το παραπάνω σχήμα, αν $a(n)$ το αρχικό σήμα του σχ. 1.3 μπορούμε να γράψουμε ότι

$$a(n) \approx \varphi_0(n) - \varphi_1(n) + \varphi_2(n) - \varphi_3(n) = (1, -1, 1, -1) \quad (1.3)$$

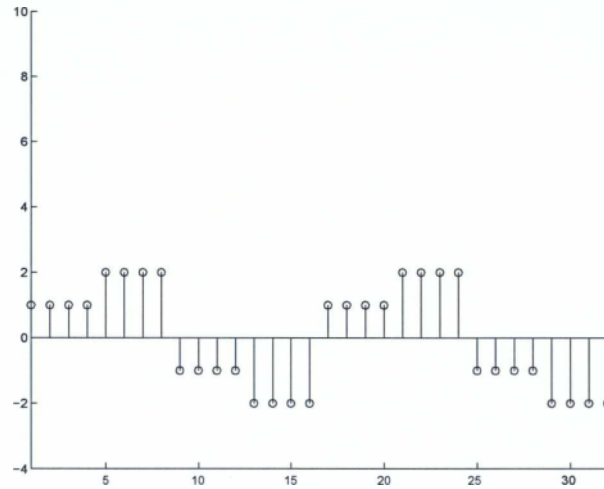
Στη συνέχεια θα χρησιμοποιήσουμε κυρίως τον τελευταίο συμβολισμό. Ας θεωρήσουμε τώρα την $\varphi(2n)$. Τότε, σύμφωνα με την (1.1) έχω ότι

$$\varphi(2n) = \begin{cases} 0, n < 0 \\ 1, 2n(\text{modulo})8 = 0 \Leftrightarrow n(\text{modulo})4 = 0 \\ 0, \text{αλλού} \end{cases} \quad (1.4)$$

και επίσης, από την (1.2):

$$\varphi_k(2n) = \varphi(2n - 8 * k) = \varphi(2(n - 4 * k)) \quad (1.5)$$

Όπως πριν, το $a(n)$ μπορεί να αναπαρασταθεί προσεγγιστικά, αλλά με περισσότερες λεποτιμέρεις ως γραμμικός συνδυασμός των $\varphi_{2k}(n)$:



Σχήμα 1.8 Προσέγγιση της $a(n)$ με την $\varphi(2n)$

Οπότε, ισχύει για την $a(n)$ σε σχέση με την $\varphi(2n)$:

$$a(n) \approx (1, 2, -1, -2, 1, 2, -1, -2) \quad (1.6)$$

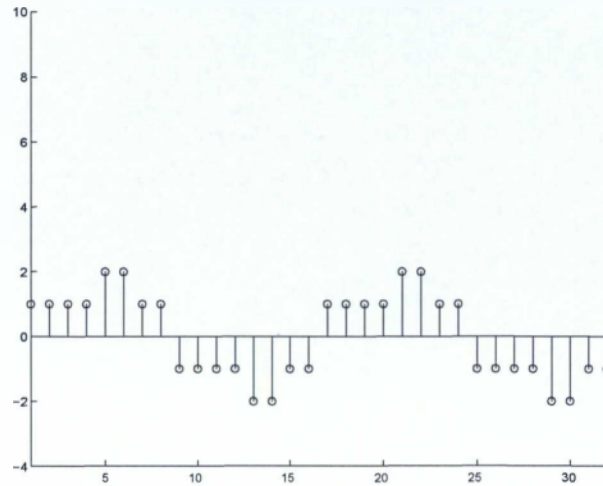
Συνεχίζοντας παρόμοια, έχουμε για την $\varphi(4n)$:

$$\varphi(4n) = \begin{cases} 0, n < 0 \\ 1, 4n \pmod{8} = 0 \Leftrightarrow n \pmod{2} = 0 \\ 0, \text{αλλού} \end{cases} \quad (1.7)$$

και επίσης, από την (1.2):

$$\varphi_k(4n) = \varphi(4n - 8 * k) = \varphi(4(n - 2 * k)) \quad (1.8)$$

Οπότε, και το $a(n)$ προσεγγίζεται ως εξής:



Σχήμα 1.9 Προσέγγιση της $a(n)$ με την $\varphi(4n)$

Και ισχύει για την $a(n)$ σε σχέση με την $\varphi(4n)$

$$a(n) \approx (1, 1, 2, 1, -1, -1, -2, -1, 1, 1, 2, 1, -1, -1, -2, -1) \quad (1.9)$$

Τέλος, αν προχωρήσουμε στην $\varphi(8n)$ θα διαπιστώσουμε ότι

$$\varphi(8n) = \begin{cases} 0, n < 0 \\ 1, 8n \pmod{ulo} = 0 \Leftrightarrow n \pmod{ulo} = 0 \Leftrightarrow n = 0 \\ 0, \text{αλλιώς} \end{cases} \quad (1.10)$$

και επίσης, από την (1.2):

$$\varphi_k(8n) = \varphi(8n - 8 * k) = \varphi(8(n - k)) \quad (1.11)$$

δηλαδή, κατ'ουσίαν, η $\varphi(8n)$ είναι ο μοναδιαίος παλμός, και άρα, οι συντελεστές της αναπαράστασης του $a(n)$ ως γραμμικού συνδυασμού των $\varphi_k(8n)$ είναι η χρονική αναπαράστασή του.

Παρατηρούμε από τα παραπάνω ότι κάθε φορά που διπλασιάζαμε το χρόνο στη $\varphi(n)$ αυτή «λέπταινε» και ήταν σε θέση να αναπαραστήσει περισσότερες λεπτομέρειες από το αρχικό σήμα $a(n)$. Δηλαδή, κάθε νέα βάση από $\varphi_k(n)$ είχε πρόσβαση σε περισσότερη φασματική πληροφορία και υψηλότερες συχνότητες.

Το ενδιαφέρον θα ήταν να μπορούσαμε να δημιουργήσουμε μια τέτοια βάση από ακολουθίες, στην οποία διαφορετικά μέλη να ήταν μονοσήμαντα υπεύθυνα για διαφορετικό επίπεδο λεπτομερειών. Αυτό προφανώς δε μπορεί να γίνει αν ενώσουμε π.χ. τις $\varphi_k(n)$ με τις $\varphi_k(2n)$: τα δύο σύνολα δεν είναι γραμμικά ανεξάρτητα και συνεπώς δε μπορούμε να ορίσουμε μια βάση με την ένωσή τους. Αυτό, όμως, που μπορούμε να κάνουμε είναι να βρούμε μια νέα οικογένεια ακολουθιών, οι οποίες να είναι γραμμικά ανεξάρτητες από τις $\varphi_k(n)$ και, εάν το σύνολό τους ενωθεί με το σύνολο των $\varphi_k(n)$ το σύνολο επέκτασης της ένωσης να ισούται με το σύνολο επέκτασης των $\varphi(2n)$.

Μπορούμε εύκολα να δούμε ότι μια ακολουθία που ικανοποιεί την παραπάνω απαίτηση είναι η εξής (βλ. σχ. 1.10):

$$\psi(n) = \begin{cases} 0, n < 0 \\ 1, n \bmod 4 = 0 \\ -1, (n-4) \bmod 4 = 0 \\ 0, \text{αλλιώς} \end{cases} \quad (1.12)$$

και κατ'αναλογία με πριν, ορίζουμε ότι:

$$\psi_k(n) = \psi(n - 8 \cdot k) \quad (1.13)$$

Τότε, μπορούμε εύκολα να δούμε ότι το $\{\psi_k(n)\} \cup \{\phi_k(n)\}$ έχουν το ίδιο σύνολο επέκτασης με το $\phi_k(2n)$. Για παράδειγμα, το σήμα του σχήματος 1.7 μπορεί να γραφεί ως:

$$a(n) = 1.5\phi_0(n) - 1.5\phi_1(n) + 1.5\phi_2(n) - 1.5\phi_3(n) - 0.5\psi_0(n) + 0.5\psi_1(n) - 0.5\psi_2(n) + 0.5\psi_3(n) \quad (1.14)$$

Κατά παρόμοιο τρόπο, μπορούμε να ορίσουμε τα $\psi_k(2n)$ για τα $\phi_k(2n)$ και τα $\psi_k(4n)$ για τα $\phi_k(4n)$. Συνεπώς, αφού το $a(n)$ μπορεί να αναπαρασταθεί επακριβώς από το $\{\phi_k(8n)\}$, θα μπορεί να αναπαρασταθεί από το $\{\psi_k(4n)\} \cup \{\phi_k(4n)\}$, αλλά και το $\{\phi_k(4n)\}$ έχει το ίδιο σύνολο επέκτασης με το $\{\psi_k(2n)\} \cup \{\phi_k(2n)\}$ κ.ο.κ, άρα το $a(n)$ μπορεί να αναπαρασταθεί επακριβώς και με μοναδικό τρόπο ως γραμμικός συνδυασμός των στοιχείων της βάσης

$$\{\psi_k(4n)\} \cup \{\psi_k(2n)\} \cup \{\psi_k(n)\} \cup \{\phi_k(n)\} \quad (1.15)$$

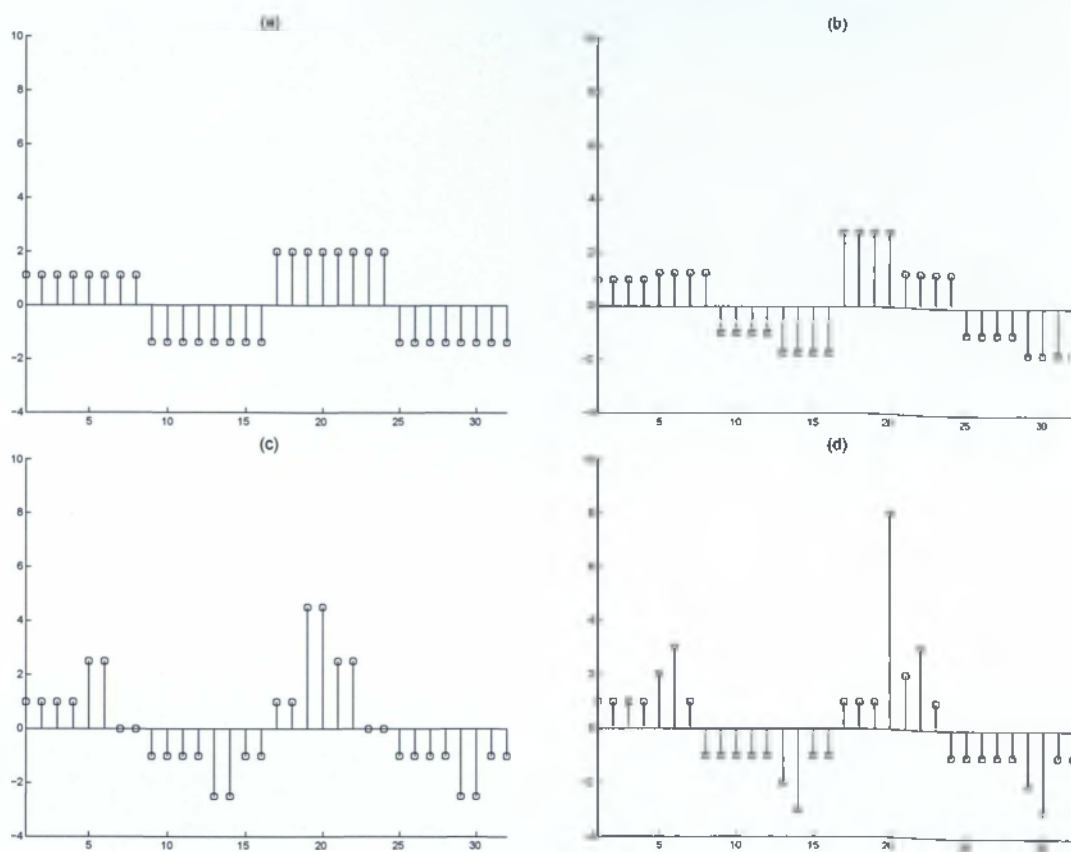
Κάθε ένα από τα σύνολα που αποτελούν τη βάση στην (1.15) αντιστοιχεί (και μόνο αυτό) σε ένα διαφορετικό επίπεδο ανάλυσης, και συνεπώς σε κάποιες υψηλότερες συχνότητες. Για παράδειγμα, το $a(n)$ μπορεί να γραφεί ως εξής:

$$\text{Συντελεστές για τα } \phi_k(n) \text{ (σχ. 1.9a): } (1.125, -1.375, 2, -1.375) \quad (1.16a)$$

$$\text{Συντελεστές για τα } \psi_k(n) \text{ (σχ. 1.9b): } (-0.125, 0.375, 0.75, 0.375) \quad (1.16b)$$

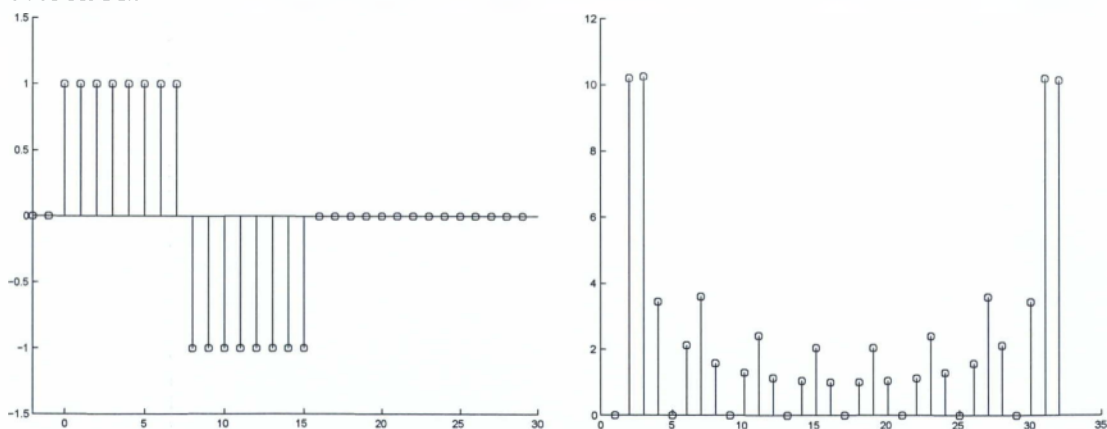
$$\text{Συντελεστές για τα } \psi_k(2n) \text{ (σχ. 1.9c): } (0, 1.25, 0, -0.75, -1.75, 1.25, 0, -0.75) \quad (1.16c)$$

$$\text{Συντελεστές για τα } \psi_k(4n) \text{ (σχ. 1.9d): } (0, 0, -0.5, 1, 0, 0, 0.5, 0, 0, -3.5, -0.5, 1, 0, 0, 0.5, 0) \quad (1.16d)$$



Σχήμα 1.10 Διαδοχικές προσεγγίσεις του $a(n)$ με συνεισφορές από τα $\psi_k(2^m n)$

Μια άμεση παρατήρηση είναι ότι λόγω όλων των προηγούμενων, τα διανύσματα στις (1.16) περιέχουν πληροφορίες για το περιεχόμενο συχνότητας του $a(n)$ ως το συγκεκριμένο επίπεδο λεπτομερειών: για παράδειγμα, στην (1.16a) μας δίνονται οι «θεμελιώδεις» ταλαντώσεις του σήματος, και **ταυτόχρονα** στην (1.16d) βλέπουμε ότι κάπου ανάμεσα στις χρονικές στιγμές 19 και 20 υπήρξε μια ισχυρή υψίσυχη συνιστώσα.



Σχήμα 1.11 Το $\psi(n)$ στο χρόνο και στη συχνότητα (με χρήση FFT)

Επίσης, ο συνολικός αριθμός των συντελεστών των (1.16) είναι 32, δηλαδή, όσο και ο αριθμός των δειγμάτων του σήματος, ή των δειγμάτων στο χώρο της συχνότητας στην αναπαράσταση που μας δίνει ο μετασχηματισμός fourier. Έχουμε, δηλαδή, μια **συμπαγή** αναπαράσταση του σήματος στην οποία έχουμε πληροφορία **και για**

το φασματικό του περιεχόμενο και για τις χρονικές στιγμές στις οποίες αυτό μεταβάλλεται. Επιτύχαμε, λοιπόν, μια αναπαράσταση **χρόνου-συχνότητας**.

Παρατηρήστε στο σχ. 1.10 ότι το φάσμα του $\psi(t)$ δεν είναι συγκεντρωμένο σε δύο σημεία, όπως του ημιτόνου, αλλά παρουσιάζει μεγαλύτερη διασπορά. Έτσι, η αναπαράσταση χρόνου-συχνότητας που πετύχαμε παρουσιάζει αβεβαιότητα στη συχνότητα (δηλαδή, δεν είναι δυνατό να αντιστοιχίσουμε π.χ. τους συντελεστές της (1.16d) σε μια συγκεκριμένη συχνότητα, όπως κάνει ο μετασχηματισμός fourier) αλλά και στο χρόνο (π.χ. στο 1.16d δεν μπορούμε να κρίνουμε αν η υψίσυχη συνιστώσα παρουσιάζεται για $n=19$ ή για $n=20$). Αυτό δεν είναι παρά φυσική συνέπεια της αρχής της απροσδιοριστίας που αναφέραμε πριν. Ωστόσο, αυτή η αβεβαιότητα είναι πολύ μικρή για τη συγκεκριμένη εφαρμογή, και δε μας απασχολεί

2.4 Ανάλυση

Είδαμε πριν πως μια αναπαράσταση του σήματος χρησιμοποιώντας ως ακολουθίες βάσης μια συγκεκριμένη ακολουθία και τις χρονικές της συμπίεσεις μπορεί και αποδίδει σωστά και με σαφήνεια τα χαρακτηριστικά σημάτων με απότομες μεταβολές ενώ παράλληλα δίδει και πληροφορία ταυτόχρονα στα πεδία χρόνου και συχνότητας, δύο ρόλοι στους οποίους ο μετασχηματισμός fourier αποτυγχάνει να ανταποκριθεί σωστά. Ο λόγος για τον οποίον αυτό συμβαίνει έχει να κάνει (και μόνο) με τις συναρτήσεις-ακολουθίες βάσης που χρησιμοποιεί ο κάθε μετασχηματισμός:

Ο μετασχηματισμός fourier και οι «συγγενικοί» σε αυτόν μετασχηματισμοί (συνημιτόνου, διακριτοί κλπ) χρησιμοποιεί ως συναρτήσεις βάσης τις τριγωνομετρικές συναρτήσεις. Οι συναρτήσεις αυτές είναι *εντοπισμένες στη συχνότητα* (δηλαδή είναι παντού μηδέν εκτός από ένα φραγμένο διάστημα) αλλά δεν είναι *εντοπισμένες στο χρόνο*. Έτσι, αν αναλύσουμε ένα σήμα με βάση αυτές παίρνουμε *ακριβή* πληροφορία στη συχνότητα και *καθόλου* πληροφορία στο χρόνο

Από την άλλη, ο μετασχηματισμός που κάναμε στο $a(n)$ πριν έχει ως ακολουθίες βάσης την $\phi(n)$ και τις $\psi_k(n)$ οι οποίες είναι εντοπισμένες στο χρόνο αλλά δεν είναι εντοπισμένες στη συχνότητα. Ωστόσο, σε αντίθεση με το ημίτονο που έχει την ίδια ενέργεια σε κάθε χρονική στιγμή, οι $\psi_k(n)$ έχουν την *περισσότερη* ενέργειά τους συγκεντρωμένη σε λίγες αρμονικές (βλ. 1.10). Έτσι, μπορούν να μας δώσουν *προσεγγιστική* πληροφορία **και** στο χρόνο **και** στη συχνότητα

Οι τριγωνομετρικές συναρτήσεις (ημίτονα, συνημίτονα) είναι συναρτήσεις των οποίων οι τιμές «ταλαντώνονται» επ'άπειρον και έχουν άπειρη ενέργεια. Είναι λοιπόν, **κύματα**. Από την άλλη, οι $\psi_k(n)$ είναι ακολουθίες που αν και ταλαντώνονται οι τιμές τους, είναι εντοπισμένες χρονικά και έχουν πεπερασμένη ενέργεια. Συνεπώς, θα μπορούσαμε να τις ονομάσουμε **κυματίδια ή wavelets**. Στην πραγματικότητα, οι ακολουθίες $\psi(n)$ και $\phi(n)$ που παρουσιάσαμε πριν είναι οι «διακριτοποιημένες» εκδοχές της πρώτης συνάρτησης wavelet και της πρώτης συνάρτησης κλίμακας που ανακαλύφθηκαν ποτέ και σήμερα γνωστές με το όνομα “Haar Wavelet”. Περισσότερα γι'αυτές τις έννοιες, στην επόμενη ενότητα.

Κεφάλαιο 3 Εισαγωγή στα Wavelets

3.1 Σύντομη επισκόπηση βασικών εννοιών

Είδαμε πριν πως η αναπαράσταση ενός σήματος χρησιμοποιώντας κάποιες (μάλλον αυθαίρετες) συναρτήσεις βάσης οι οποίες είναι σχεδιασμένες έτσι ώστε η κάθε μια οικογένεια από αυτές να αντιστοιχεί σε συγκεκριμένο επίπεδο ανάλυσης/λεπτομέρειας, άρα και σε κάποιες συγκεκριμένες συχνότητες, μας δίνει πολύ καλύτερες πληροφορίες για ένα μη-περιοδικό σήμα με απότομες μεταβολές απ'ότι ο μετασχηματισμός Fourier. Είμαστε σε θέση πλέον, έχοντας ως πρακτικό παράδειγμα την ανάλυση της προηγούμενης ενότητας, να προχωρήσουμε στους ορισμούς των εννοιών που αφορούν τα wavelets.

Τα wavelets είναι ένας μάλλον καινούριος (οι σημαντικές ανακαλύψεις έγιναν μετά το 1985) και ραγδαία αναπτυσσόμενος κλάδος των μαθηματικών και της επεξεργασίας σήματος. Αυτό αποτελεί εν μέρει ένα μειονέκτημα καθώς πολλά βιβλία που θεωρούνταν κάποτε «κλασσικά» σήμερα στερούνται των τελευταίων εξελίξεων με πολύ σημαντικές εφαρμογές ή είναι γραμμένα σε γλώσσα για μαθηματικούς και όχι για μηχανικούς! Εμείς εδώ θα κάνουμε μια αρκετά χαλαρή μαθηματικά και «πρακτική» εισαγωγή σε όσες βασικές έννοιες χρειάζονται για την κατανόηση αυτών που θα ακολουθήσουν. Στη βιβλιογραφία παραθέτουμε πλήθος συγγραμμάτων για τον ενδιαφερόμενο αναγνώστη. Ενδεικτικά, το [A2] αποτελεί μάλλον την καλύτερη πρακτική εισαγωγή στο αντικείμενο, προσφέροντας καλή ενημέρωση πάνω σε καινούριες σχετικά εξελίξεις όπως τα διορθογώνια wavelets. Το [A6] έχει το πλεονέκτημα της αυστηρής μαθηματικής θεμελίωσης που δεν χρειάζεται βαρύ μαθηματικό υπόβαθρο για να γίνει κατανοητό, όπως το [A7], το οποίο όμως είναι ένα αδιαμφισβήτητο κλασσικό σύγγραμμα από την πρωτοπόρο του κλάδου. Ωστόσο, το [A6] παρουσιάζει πολύ στοιχειώδεις έννοιες, με αποτέλεσμα να μην είναι επαρκές για την κατανόηση σύγχρονων εφαρμογών. Τέλος, το [A4] θεωρείται από πολλούς ως το καλύτερο και πιο περιεκτικό βιβλίο για τα wavelets και τις εφαρμογές τους, ωστόσο απαιτεί κάποιο παραπάνω μαθηματικό υπόβαθρο.

Ας θεωρήσουμε όλες τις (γενικευμένες) πραγματικές συναρτήσεις που ανήκουν στον απειροδιάστατο διανυσματικό χώρο $L^2(\mathbb{R})$, δηλαδή, αν $f(t)$ μια από αυτές, ισχύει πάντοτε

$$\sqrt{\int_{-\infty}^{\infty} |f(t)|^2 dt} < \infty \quad (1.17)$$

Η συνθήκη αυτή σημαίνει πρακτικά ότι η αν η $f(t)$ αναπαριστά ένα σήμα, το σήμα αυτό έχει πεπερασμένη ενέργεια. Αυτό ισχύει για όλα σχεδόν τα σήματα που συναντώνται, και απαιτείται και από άλλα είδη μετασχηματισμών.

Ας θεωρήσουμε, λοιπόν, αυθαίρετα, μια κατάλληλη συνάρτηση $\varphi(t)$ που να ανήκει στο $L^2(\mathbb{R})$ έτσι ώστε όλες οι συναρτήσεις που ορίζονται ως

$$\varphi_k(t) = \varphi(t-k) \quad (1.18)$$

όπου k ακέραιος να είναι γραμμικά ανεξάρτητες πολύ από τα πλαίσια αυτής.

Το υποσύνολο του το $L^2(\mathcal{H})$ που αποτελεί το σύνολο επέκτασης των $\{\varphi_k(t)\}$ ορίζουμε ως

$$V_0 = \overline{\text{Span}_k \{\varphi_k(t)\}} \quad (1.19)$$

για όλους τους ακεραίους k από το μείον ως το συν άπειρο. Η μπάρα από πάνω δείχνει ότι πρόκειται για κλειστό σύνολο. Είναι φανερό ότι το V_0 δε θα είναι ποτέ κενό. Ας θεωρήσουμε μια $f(t) \in V_0$. Τότε, αυτή θα μπορεί να γραφεί πάντοτε ως:

$$f(t) = \sum_k a_k \varphi_k(t) \quad (1.20)$$

Μια ακόμα προϋπόθεση για την επιλογή της $\varphi(t)$ είναι να είναι τέτοια ώστε αυξάνοντας τη χρονική ακρίβεια της αρχικής συνάρτησης, το σύνολο επέκτασης των $\varphi_k(mt)$ που προκύπτουν να είναι γνήσιο υπερσύνολο του V_0 . Έτσι, ορίζουμε μια οικογένεια από $\varphi(t)$ με δύο παραμέτρους, τη χρονική μετατόπιση και τη χρονική συμπίεση. Ένας συχνά χρησιμοποιούμενος σύμβαση και συμβολισμός για μια τέτοια οικογένεια είναι ο

$$\varphi_{j,k}(t) = 2^{j/2} \varphi(2^j t - k) \quad (1.21)$$

από τις οποίες, ο διανυσματικός χώρος που καλύπτει η κάθε ομάδα $\varphi_j(t)$ είναι ο

$$V_j = \overline{\text{Span}_k \{\varphi_k(2^j t)\}} = \overline{\text{Span}_k \{\varphi_{j,k}(t)\}} \quad (1.22)$$

για όλους τους ακεραίους k . Αυτό σημαίνει ότι αν η $f(t) \in V_j$, τότε μπορεί να γραφτεί $f(t) = \sum_k a_k \varphi_k(2^j t + k)$

(1.23)

Λόγω των προηγούμενων, έχουμε ότι ισχύει για τα υποσύνολα V_j του $L^2(\mathcal{H})$ που προκύπτουν από κάθε $\{\varphi_j(t)\}$:

$$\dots \subset V_{-1} \subset V_0 \subset V_1 \subset V_2 \subset \dots \subset L^2 \quad (1.24a)$$

ή, εναλλακτικά,

$$V_j \subset V_{j+1}, \text{ με } V_{-\infty} = \{0\} \text{ και } V_{\infty} = L^2 \quad (1.24b)$$

Μια $\varphi(t)$ που ικανοποιεί τις συνθήκες (1.20)...(1.24) ονομάζεται **συνάρτηση κλίμακας (scaling function)**. Μια «συνάρτηση» (ακολουθία) κλίμακας υπό τον ορισμό αυτό ήταν και η $\varphi(t)$ που θέσαμε στην (1.1), αφού είδαμε πως οι επόμενες χρονικές συμπίεσεις της μπορούσαν να αναπαραστήσουν σήματα με περισσότερες λεπτομέρειες. Φυσικά, όπως και στο παράδειγμά μας πριν, ενώ οι $\varphi(t)$ που ορίζονται από την (1.21) και έχουν το ίδιο j και διαφορετικά k είναι γραμμικά ανεξάρτητες, αυτές που έχουν διαφορετικά j δεν είναι κατ'ανάγκη (και ούτε θα πρέπει να είναι, όλες, για να ικανοποιείται η (1.24a)) γραμμικά ανεξάρτητες.

Ας θεωρήσουμε τώρα μια κατάλληλη $\varphi(t)$ με σύνολο επέκτασης το V_0 . Τότε, προφανώς, η $\varphi(t)$ περιέχεται στο V_0 αλλά και (λόγω της (1.24a)) σε κάθε άλλο

υπερσύνολό του, π.χ. το V_1 . Συνεπώς, θα μπορεί να γραφεί ως γραμμικός συνδυασμός των στοιχείων της βάσης του V_1 , δηλαδή, τα $\varphi_k(2t)$ ως εξής:

$$\varphi(t) = \sum_n h(n) \sqrt{2} \varphi(2t-n), \text{ όπου } n \text{ ακέραιος} \quad (1.25)$$

Το $h(n)$ είναι μια ακολουθία (πραγματικών) συντελεστών που ονομάζονται **συντελεστές της συνάρτησης κλίμακας** (scaling function coefficients), **φίλτρο κλίμακας** (scaling filter) ή **διάνυσμα κλίμακας** (scaling vector). Η τετραγωνική ρίζα του 2 υπάρχει στο γινόμενο ως επακόλουθο της (1.23) για να είναι κανονικοποιημένα τα $h(n)$. Για παράδειγμα, στο παράδειγμά μας στην ενότητα B, από τις (1.1), (1.4), (1.7), (1.10) μπορούμε εύκολα να δούμε ότι $h(n) = (2^{-1}, 2^{-1})$.

Παρόμοια με την ενότητα B, η επιθυμία μας να έχουμε μια ενιαία βάση για το $L^2(\mathbb{R})$ της οποίας μερικά μέλη να ευθύνονται αποκλειστικά για κάποιο επίπεδο λεπτομερειών μας οδηγεί να ορίσουμε κάποιες νέες συναρτήσεις $\psi(n)$, το σύνολο επέκτασης των οποίων θα είναι η **διαφορά** μεταξύ των διαδοχικών V_j και θα διέπονται από αρχές παραπλήσιες της (1.21)

Δηλαδή, αν $\psi(t)$ μια κατάλληλη τέτοια συνάρτηση που αντιστοιχεί στην $\varphi(t)$ που αναλύσαμε στις (1.20)-(1.24), τότε, θα πρέπει να μπορώ να ορίσω τις:

$$\psi_{j,k}(t) = 2^{j/2} \psi(2^j t - k) \quad (1.26)$$

τέτοιες ώστε, κάθε υποσύνολο από αυτές ψ_j να έχει ως σύνολο επέκτασης το:

$$\text{Span}_k \{ \psi_{j,k}(t) \} = W_j = V_{j+1} - V_j \quad (1.27a)$$

ή, εναλλακτικά,

$$\text{Span}_k \{ \{ \psi_{j,k}(t) \} \cup \{ \varphi_{j,k}(t) \} \} = V_{j+1} \quad (1.27)$$

Η (1.27) σημαίνει, αναδρομικά, ότι θα πρέπει όχι μόνο τα $\psi_{j,k}(t)$ με το ίδιο j να είναι γραμμικά ανεξάρτητα μεταξύ τους, αλλά όλα τα $\psi_{j,k}(t)$ να είναι γραμμικά ανεξάρτητα (πάλι, δε χρειάζεται τόση αυστηρότητα αλλά αυτό είναι πέραν των σκοπών αυτής της εργασίας). Οι συναρτήσεις $\psi_{j,k}(t)$ ονομάζονται συναρτήσεις κυματιδίου ή συναρτήσεις *wavelet*. Για παράδειγμα, η $\psi(n)$ που ορίσαμε στην (1.12) αποτελεί μια ακολουθία *wavelet*. Το όνομα “*wavelet*” («κυματίδιο», αν και εμείς εδώ θα επιμένουμε στον αγγλικό όρο) προέρχεται από την μορφή μικρού, εντοπισμένου στο χρόνο κύματος (δηλαδή, λαμβάνουν τιμές και πάνω και κάτω από το μηδέν) που έχουν οι περισσότερες από αυτές.

Από τη στιγμή που τα $\{\psi_j\}$ είναι βάση για το V_{j+1} , τότε ανήκουν σε αυτό. Αυτό σημαίνει, μπορούν να γραφούν ως γραμμικός συνδυασμός της βάσης του όπως πριν, ότι $\{\varphi_{j+1}\}$, οπότε, κατ’αναλογία με την (1.25) ισχύει για το $\psi(t)$

$$\psi(t) = \sum h_1(n) \sqrt{2} \varphi(2t-n), \text{ όπου } n \text{ ακέραιος} \quad (1.28)$$

Εδώ, το $h_1(n)$ είναι πάλι μια ακολουθία πραγματικών συντελεστών που ονομάζονται **συντελεστές της συνάρτησης wavelet** (wavelet function coefficients), **φίλτρο wavelet** (wavelet filter) ή **διάνυσμα κλίμακας** (wavelet vector).

Συνοψίζοντας, από τις (1.24) και (1.27) αν επιλέξουμε μια τυχαίας ακρίβειας συνάρτηση κλίμακας $\varphi(t)$ και ορίσουμε την κατάλληλη γι' αυτή συνάρτηση wavelet $\psi(t)$, τότε, σύμφωνα με τα παραπάνω, κάθε πραγματικό σήμα που ανήκει στο $L^2(\mathbb{R})$ μπορεί να γραφεί ως

$$g(t) = \sum_{k=-\infty}^{\infty} c_{j0}(k) * \varphi_k(t) + \sum_{j=0}^{\infty} \sum_{k=-\infty}^{\infty} d_j(k) * \psi_{j,k}(t) \quad (1.29)$$

Η ανάλυση ενός σήματος στα $c_{j0}(k)$ και $d_j(k)$ σύμφωνα με κάποια δεδομένα $\varphi(t)$ και $\psi(t)$ λέγεται **διακριτός μετασχηματισμός wavelet (discrete wavelet transform)** του σήματος αυτού, και, όπως είδαμε στην ενότητα Β, τα $d_j(k)$ μπορεί να έχουν αξία ανάλογη των συντελεστών που επιστέφει ο μετασχηματισμός fourier. Τα $c_{j0}(k)$ (και εν γένει, οι συντελεστές της συνάρτησης κλίμακας) θα ονομάζονται από εδώ και εμπρός **σήμα προσέγγισης** ή συντελεστές προσέγγισης, και τα $d_j(k)$ (για ένα συγκεκριμένο j) **σήμα λεπτομερειών**, ή συντελεστές wavelet.

Ουσιαστικά, έχουμε ορίσει όλες σχεδόν τις παραμέτρους που απαιτούνται για την επιλογή των συναρτήσεων $\psi(t)$, $\varphi(t)$ (υπάρχουν και μερικές άλλες απαιτήσεις, μαθηματικής φύσεως που δε μας αφορούν εδώ). Πράγματι, σε αντίθεση με τους μετασχηματισμούς fourier, συνημιτόνου, czt κλπ. οι υποψήφιες συναρτήσεις κλίμακας και wavelet για το διακριτό μετασχηματισμό wavelet είναι *άπειρες*. Ωστόσο, υπάρχουν διάφορα κριτήρια που ευνοούν κάποιες συναρτήσεις έναντι κάποιων άλλων. Πριν αναφερθούμε όμως στα κριτήρια αυτά, θα προσπαθήσουμε να εξάγουμε τα $c_{j0}(k)$, $d_j(k)$ από την (1.29).

Μέχρι τώρα, το μόνο κριτήριο που θέσαμε για τη σχέση μεταξύ των $\varphi_k(t)$, $\psi_{j,k}(t)$ είναι να είναι γραμμικά ανεξάρτητα μεταξύ τους έτσι ώστε να μπορούν να αποτελούν μια βάση. Ένας τρόπος να εξάγουμε τα $c_{j0}(k)$, $d_j(k)$ από την (1.29) θα ήταν να απαιτούσαμε τα $\varphi_k(t)$, $\psi_{j,k}(t)$ να είναι εκτός από γραμμικά ανεξάρτητα να είναι και *ορθογώνια* μεταξύ τους, δηλαδή, να ισχύει πάντοτε:

$$\int_{-\infty}^{\infty} \psi_{j,k}(t) \psi_{l,m}(t) dt = \delta(j-l, m-k) \quad (1.30)$$

και παρόμοια και για τα $\varphi_k(t)$. Σημείωση: στην ουσία η 1.30 απαιτεί οι συναρτήσεις να είναι ορθοκανονικές και όχι απλά ορθογώνιες, ωστόσο, σε περίπτωση που είναι ορθογώνιες και όχι ορθοκανονικές, μπορούμε να τις μετατρέψουμε σε ορθοκανονικές με κατάλληλους συντελεστές)

Στην περίπτωση που οι συναρτήσεις είναι ορθογώνιες, τότε οι συντελεστές μπορούν να ληφθούν από τις σχέσεις:

$$c_{j0}(k) = \int_{-\infty}^{\infty} g(t) \varphi_k'(t) dt \quad (1.31a)$$

$$d_j(k) = \int_{-\infty}^{\infty} g(t) * \psi_{j,k}'(t) dt \quad (1.31b)$$

Ωστόσο, υπάρχουν περιπτώσεις που οι συναρτήσεις κλίμακας και wavelet δε θέλουμε να είναι ορθογώνιες ή δεν *πρέπει* να είναι ορθογώνιες: στην πράξη οι ορθογώνιες συναρτήσεις έχουν μια σειρά από περιορισμούς που τις κάνουν μη επιθυμητές για τις περισσότερες εφαρμογές.

Τότε, μπορούμε να απαιτήσουμε τα $\varphi_k(t)$, $\psi_{j,k}(t)$ να είναι *διορθογώνια*.

Τι σημαίνει διορθογώνια; Θα δώσουμε ένα γεωμετρικό ανάλογο. Στο σχ. 3.1(a), τα διανύσματα i και j είναι ορθογώνια. Αυτό σημαίνει ότι αν εκφράσουμε οποιοδήποτε διάνυσμα του επιπέδου ως γραμμικό συνδυασμό των i , j τότε ο πολλαπλασιασμός του (εσωτερικό γινόμενο) με π.χ. το i θα μας δώσει τον συντελεστή του i στο διάνυσμα αυτό.



Σχήμα 3.1 Γεωμετρικό ανάλογο μιας ορθογώνιας (a) και μιας διορθογώνιας (b) βάσης

Στο σχ. 3.1b αυτό δε μπορεί να γίνει με τα i, j , καθώς το εσωτερικό τους γινόμενο είναι διαφορετικό του μηδενός. Ωστόσο, μπορούμε να ορίσουμε δυο νέα διανύσματα, τα k και l , τα οποία είναι ορθογώνια ως προς το j και το i αντίστοιχα. Έτσι, αν έχουμε εκφράσει ένα οποιοδήποτε διάνυσμα του επιπέδου ως γραμμικό συνδυασμό των i, j τότε, το εσωτερικό γινόμενό του π.χ. με το k θα μας δώσει το συντελεστή του j . Έτσι, τα (i, j) σχηματίζουν μαζί με τα *δυναδικά* τους (k, l) μια *διορθογώνια* βάση στο επίπεδο.

Ας επιστρέψουμε, τώρα, στο διακριτό μετασχηματισμό wavelet. Αν λοιπόν απαιτήσουμε τα $\varphi_k(t)$, $\psi_{j,k}(t)$ να είναι διορθογώνια, τότε θα πρέπει να είμαστε σε θέση να ορίσουμε και τα *δυναδικά* τους $\varphi'_k(t)$, $\psi'_{j,k}(t)$, παρόμοια με το γεωμετρικό ανάλογο, έτσι ώστε, π.χ. για τα $\psi_{j,k}(t)$ να ισχύει

$$\int_{-\infty}^{\infty} \psi_{j,k}(t) \psi'_{l,m}(t) dt = \delta(j-l, m-k) \quad (1.32)$$

Από τη στιγμή που τα $\varphi'_k(t)$, $\psi'_{j,k}(t)$ είναι *δυναδικά* των $\varphi_k(t)$ και $\psi_{j,k}(t)$, μπορούν και αυτά να αποτελέσουν βάσεις των αντίστοιχων υποσυνόλων του $L^2(\mathbb{R})$ και επιπλέον γι'αυτά θα ισχύουν (αντίστοιχα) οι (1.21) και (1.26). Από το γεγονός ότι αποτελούν (και αυτά) βάση για τους αντίστοιχους διανυσματικούς υποχώρους, σημαίνει ότι θα πρέπει να ορίζονται τα $h'(t)$, $h_1'(t)$ έτσι ώστε και εδώ έχουμε:

$$\varphi'(t) = \sum_n h'(n) \sqrt{2} \varphi'(2t-n) \quad (1.33)$$

$$\text{και } \psi'(t) = \sum_n h_1'(n) \sqrt{2} \varphi'(2t-n) \quad (1.34)$$

όπου n ακέραιος

Τότε τα σήματα προσέγγισης και λεπτομερειών θα μπορούν να υπολογιστούν:

$$c_{j0}(k) = \int_{-\infty}^{\infty} g(t) \varphi_k'(t) dt \quad (1.35a)$$

$$\text{και } d_j(k) = \int_{-\infty}^{\infty} g(t) \cdot \psi_{jk}'(t) dt \quad (1.35b)$$

Οι διορθογώνιες συναρτήσεις κλίμακας και wavelet (που συχνά θα αποκαλούμε συλλογικά ως «διορθογώνια wavelets») έχουν πολύ περισσότερους βαθμούς ελευθερίας και συνεπώς πολύ μεγαλύτερη ευελιξία από τα ορθογώνια. Πράγματι, οι περισσότεροι μετασχηματισμοί wavelet που συμβαίνουν στην πράξη έχουν να κάνουν με διορθογώνια wavelets.

Ωστόσο, ούτε οι (1.31), ούτε οι (1.33) μας δίνουν κάποιον αποδοτικό και αποτελεσματικό τρόπο να υπολογίσουμε τα σήματα προσέγγισης και λεπτομερειών. Ένα ολοκλήρωμα είναι δύσκολο να υπολογιστεί γρήγορα με κάποιον αλγόριθμο και επιπλέον, πριν προχωρήσουμε στην ολοκλήρωση θα πρέπει αν ξέρουμε ποια σήματα ακριβώς θέλουμε, καθώς και τα $c_{j0}(k)$ και τα $d_j(k)$ είναι θεωρητικά άπειρα. Επιπλέον, οι (1.31), (1.33) αφορούν μόνο πραγματικά σήματα, και όχι σήματα ψηφιακά που έχουν προέλθει από δειγματοληψία, όπως αυτά που υπάρχουν μέσα σε ψηφιακά κυκλώματα. Έτσι, θα πρέπει κανονικά όλη η παραπάνω θεωρία να αναδιατυπωθεί για σήματα διακριτού χρόνου, κατά τρόπο ανάλογο με την εξαγωγή του μετασχηματισμού Fourier διακριτού χρόνου από το μετασχηματισμό Fourier.

Ευτυχώς, κάτι τέτοιο δε χρειάζεται να γίνει. Όπως θα δείξουμε στην αμέσως επόμενη υποενότητα, τα σήματα προσέγγισης και λεπτομερειών μπορούν να υπολογιστούν πολύ πιο εύκολα και αποδοτικά από μια διάταξη που λέγεται συστοιχία φίλτρων. Στην πραγματικότητα, στον υπολογισμό ενός μετασχηματισμού wavelet, δε χρειάζεται καν να γνωρίζουμε την ακριβή μορφή της συνάρτησης wavelet και της συνάρτησης κλιμάκωσης που σχετίζονται με αυτόν.

3.2 Wavelets και Συστοιχίες Φίλτρων

Θα δείξουμε εδώ συνοπτικά πως ο μετασχηματισμός wavelet μπορεί να πραγματοποιηθεί όχι από (1.31), (1.35) αλλά από μια πολύ πιο απλή διάταξη που ονομάζεται **συστοιχία φίλτρων (filter bank)**. Η «απόδειξη» που θα κάνουμε αφορά κυρίως ορθογώνιες συναρτήσεις κλίμακας και wavelet, αλλά είναι πολύ παρόμοια για την περίπτωση των διορθογωνίων συναρτήσεων βάσης.

Ας θεωρήσουμε μια ορθογώνια βάση από $\{\varphi_k, \psi_{j,k}\}$, όπου η φ_k έχει επιλεγεί αυθαίρετα έτσι ώστε να μπορεί να αναπαριστά σήματα μέχρι ένα επίπεδο λεπτομερειών. Τότε, επεκτείνοντας την (1.25) μπορούμε να γράψουμε:

$$\varphi(2^j t - k) = \sum_n h(n) \sqrt{2} \varphi(2(2^j t - k) - n) = \sum_n h(n) \sqrt{2} \varphi(2^{j+1} t - 2k - n) \quad (1.36)$$

Αν στην (1.36) αντικαταστήσουμε το $2k+n$ με m , παίρνουμε:

$$\varphi(2^j t - k) = \sum_m h(m - 2k) \sqrt{2} \varphi(2^{j+1} t - m) \quad (1.37)$$

$$\text{Ας θεωρήσουμε τώρα, έναν από τους υποχώρους } V_j \text{ που ορίσαμε στην (1.22), δηλαδή} \\ V_j = \text{Span}_k \{\varphi_{j,k}(t)\} = \text{Span}_k \{\varphi(2^j t - k)\} \quad (1.38)$$

Τότε, αν μια $f(t)$ ανήκει στο V_{j+1} θα μπορεί να γραφεί, σύμφωνα με την (1.23) και χρησιμοποιώντας ορολογία της (1.29) ως

$$f(t) = \sum_k c_{j+1}(k) \sqrt{2} \varphi_{j+1,k}(t) = \sum_k c_{j+1}(k) 2^{(j+1)/2} \varphi(2^{j+1}t - k) \quad (1.39)$$

Αν ωστόσο θέλουμε να εκφράσουμε την $f(t)$ ως προς τις $\varphi_{j,k}(t)$ θα πρέπει αναγκαστικά να χρησιμοποιήσουμε και τις συναρτήσεις wavelet $\psi_{j,k}(t)$, οπότε, από την (1.29), θα έχουμε:

$$f(t) = \sum_k c_j(k) \sqrt{2} \varphi_{j,k}(t) + \sum_k d_j(k) \psi_{j,k}(t) \quad (1.40)$$

Επειδή τα $\{\varphi_k(t)\}$, $\{\psi_{j,k}(t)\}$ είναι ορθογώνια, τότε, από την (1.40) θα μπορούμε να λάβουμε το σήμα λεπτομερειών $c_j(k)$ από το ολοκλήρωμα

$$c_j(k) = \int_{-\infty}^{\infty} f(t) \psi_{j,k}(t) dt = \int_{-\infty}^{\infty} f(t) 2^{j/2} \varphi(2^j t - k) dt \quad (1.41)$$

Από την (1.37) ισχύει όμως $\varphi(2^j t - k) = \sum_m h(m - 2k) \sqrt{2} \varphi(2^{j+1} t - m)$, άρα, αντικαθιστώντας στην (1.41) παίρνουμε:

$$\begin{aligned} c_j(k) &= \int_{-\infty}^{\infty} f(t) 2^{j/2} \varphi(2^j t - k) dt = \\ &= \int_{-\infty}^{\infty} f(t) 2^{j/2} \left[\sum_m h(m - 2k) 2^{1/2} \varphi(2^{j+1} t - m) \right] dt = \\ &= \sum_m h(m - 2k) \int_{-\infty}^{\infty} f(t) 2^{(j+1)/2} \varphi(2^{j+1} t - m) dt \end{aligned} \quad (1.42)$$

αλλά από την (1.41) έχουμε ότι

$$\int_{-\infty}^{\infty} f(t) 2^{(j+1)/2} \varphi(2^{j+1} t - m) dt = c_{j+1}(m) \quad (1.43)$$

άρα, η (1.42) μετατρέπεται στην:

$$c_j(k) = \sum_m h(m - 2k) c_{j+1}(m) \quad (1.44a)$$

Η (1.44) μας δείχνει τον τρόπο να λάβουμε το σήμα προσέγγισης ενός μετασχηματισμού wavelet από το σήμα προσέγγισης της αμέσως μεγαλύτερης

ακρίβειας μέσω ενός απλού γραμμικού φίλτρου με υποδειματοληψία ανά δύο στην έξοδο. Παρόμοια, για το σήμα λεπτομερειών (συντελεστές wavelet) ισχύει

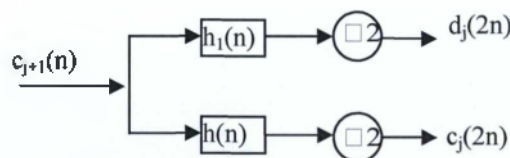
$$d_j(k) = \sum_m h_1(m-2k) c_{j+1}(m) \quad (1.44b)$$

Ενώ με τον ίδιο τρόπο αποδεικνύεται ότι αν τα $\{\varphi_k(n)\}$, $\{\psi_{j,k}(n)\}$ διορθογώνια, οι (1.44) έχουν τη μορφή:

$$c_j(k) = \sum_m h'(m-2k) c_{j+1}(m) \quad (1.45a)$$

$$d_j(k) = \sum_m h_1'(m-2k) c_{j+1}(m) \quad (1.45b)$$

Η σημασία των σχέσεων (1.44) και (1.45) είναι θεμελιώδης, αφού μας επιτρέπει να εκτελέσουμε το μετασχηματισμό wavelet ενός σήματος με τη χρήση ενός ζεύγους φίλτρων (εξ' ου και ο όρος filter bank) χωρίς να χρειάζεται καν να ξέρουμε τη μορφή των συναρτήσεων wavelet και κλίμακας που ορίζουν το μετασχηματισμό. Υπάρχει ένα



Σχήμα 3.2 Υπολογισμός των σημάτων προσέγγισης και λεπτομερειών μέσω μιας συστοιχίας φίλτρων

πρόβλημα, ωστόσο. Πως θα υπολογίσουμε τα c_{j+1} για κάποιο j ; Και επιπλέον, αφού τα j είναι άπειρα για ένα σήμα για το οποίο δε γνωρίζουμε σε ποιο υποσύνολο V_j του $L^2(\mathcal{R})$ ανήκει, δε θα πρέπει αυτό να γίνεται συνέχεια; Η απάντηση είναι όχι, και μας δίνει τη λαβή για την εφαρμογή των παραπάνω στα σήματα διακριτού χρόνου. Έχουμε ότι

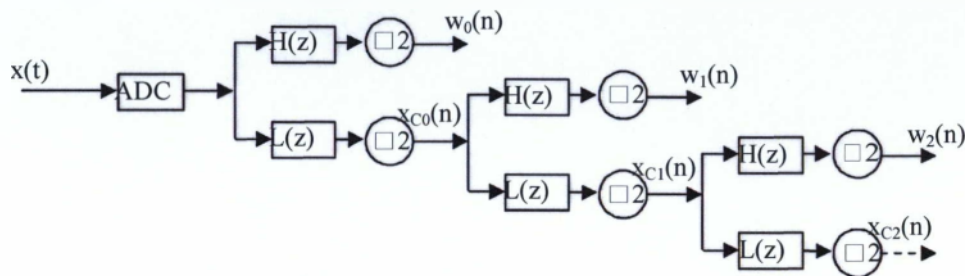
$$c_j(k) = \int f(t) \varphi_{j,k}(t) dt. \text{ Αν η } \varphi(t) \text{ είναι σωστά επιλεγμένη (κάτι που ισχύει για όλες}$$

τις συναρτήσεις κλίμακας που χρησιμοποιούνται στην πράξη), τότε για μεγάλα j η $\varphi_{j,k}(t)$ δρα σαν τη συνάρτηση dirac, η οποία είναι γνωστή και ως συνάρτηση δειγματοληψίας (ως ανάλογο παράδειγμα, στην (3.1) η $\varphi(n)$ τελικά έγινε ο μοναδιαίος παλμός). Συνεπώς, με σεβασμό του θεωρήματος δειγματοληψίας του Shannon (και σε περίπτωση μη ικανοποίησής του, φιλτράρισμα του αρχικού αναλογικού σήματος έτσι ώστε να το ικανοποιεί), τα $c_{j+1}(n)$ είναι τα διακριτού χρόνου δείγματα του αρχικού σήματος.

Έτσι, μπορούμε άμεσα να εφαρμόσουμε τη διάταξη του σχ. 3.1 στην έξοδο ενός μετατροπέα αναλογικού-σε-ψηφιακό (ADC converter) και έτσι να λάβουμε τα σήματα προσέγγισης και λεπτομερειών του $1^{\text{ου}}$ επιπέδου. Στη συνέχεια, αν θέλουμε, μπορούμε να περάσουμε και το σήμα $c_j(2n)$ από μια ίδια διάταξη έτσι ώστε να λάβουμε τα σήματα προσέγγισης και λεπτομερειών για τα επόμενα επίπεδα (χαμηλότερες συχνότητες) και ούτω καθ' εξής.

Η διάταξη αυτή είναι γνωστή και ως *Αλγόριθμος Πυραμίδας* (Pyramid Algorithm) και ανακαλύφθηκε από τον Mallat στα τέλη της δεκαετίας του 1980. Αυτή η ανακάλυψη ήταν που έκανε εφικτό υπολογιστικά τον Διακριτό Μετασχηματισμό wavelet και άνοιξε το δρόμο για τις μετέπειτα εφαρμογές του. Ο αλγόριθμος

πυραμίδας ονομάζεται σε ορισμένα βιβλία και ως «Ταχύς Μετασχηματισμός Wavelet» ή «Fast Wavelet Transform». Στην πραγματικότητα, είναι ο μόνος εφαρμόσιμος αλγόριθμος για το μετασχηματισμό wavelet, συνεπώς στην εργασία αυτή όταν θα αναφερόμαστε στον Διακριτό Μετασχηματισμό Wavelet, θα εννοούμε τον αλγόριθμο πυραμίδας του Mallat. Έτσι, μπορούμε άμεσα να υλοποιήσουμε τον διακριτό μετασχηματισμό wavelet σε υλικό ως ένα ζευγάρι (ψηφιακών) φίλτρων με υποδειγματοληψία στην έξοδο.



Σχήμα 3.3 Μετασχηματισμός wavelet τριών σταδίων

Στις εφαρμογές της επεξεργασίας σήματος, το φίλτρο κλίμακας ονομάζεται *βαθυπερατό φίλτρο* και συμβολίζεται με $L(z)$ στο χώρο κατάστασης, ενώ το φίλτρο wavelet *υψιπερατό φίλτρο* και συμβολίζεται με $H(z)$. Επιπλέον, τα σήματα προσέγγισης και λεπτομερειών που προκύπτουν από ένα ζεύγος φίλτρων όπως συμβολίζονται με $x_m(n)$, $w_m(n)$ αντίστοιχα (αντί για $c_j(k)$, $d_j(k)$) όπου $m=0,1,2,\dots$ ένας ακέραιος που δείχνει πόσες συστοιχίες φίλτρων έχουν προηγηθεί από το αρχικό κβαντισμένο σήμα. Το m μπορεί να απουσιάζει όταν μελετούμε απλώς τη συστοιχία φίλτρων και δε μας ενδιαφέρει το ποιο στάδιο του μετασχηματισμού θα εκτελέσει. Οι συμβολισμοί αυτοί θα υιοθετηθούν και για το υπόλοιπο της εργασίας.

3.3 Επιλογή wavelet

Δείξαμε πριν πως ο (διακριτός) μετασχηματισμός wavelet ενός σήματος μπορεί αν υπολογιστεί εύκολα με τη χρήση ενός ζεύγους φίλτρων. Δεν αναφερθήκαμε όμως εν τέλει στο *ποιες* θα είναι οι συναρτήσεις κλίμακας και wavelet που ορίζουν το μετασχηματισμό.

Υπάρχει μια πληθώρα από μαθηματικά κριτήρια για να μπορεί μια συνάρτηση να αποτελεί συνάρτηση κλίμακας. Για παράδειγμα, θα πρέπει οπωσδήποτε να ικανοποιούνται οι σχέσεις που παραθέσαμε πριν. Ούτως ή άλλως, όπως αναφέραμε, ο αριθμός των υποψηφίων συναρτήσεων βάσης και wavelet είναι *άπειρος*. Ωστόσο, πάντα γίνεται κάποια επιλογή με *πρακτικά* κυρίως κριτήρια.

Για παράδειγμα, μας ενδιαφέρει άμεσα για μια οικογένεια σημάτων να βρούμε τέτοιες συναρτήσεις κλίμακας και wavelet έτσι ώστε τα σήματα αυτά να περιέχουν ένα μεγάλο ποσοστό της ενέργειάς τους σε λίγους μόνο συντελεστές του μετασχηματισμού, να μπορεί δηλαδή να γίνει μια *αραιή αναπαράσταση* του αρχικού σήματος με χρήση του Διακριτού Μετασχηματισμού wavelet. Αυτό είναι πολύ σημαντικό για τεχνικές συμπίεσης και αφαίρεσης θορύβου.

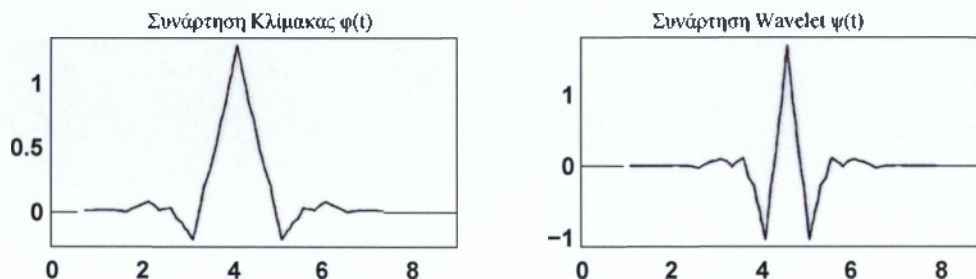
Ένα άλλο πολύ σημαντικό κριτήριο, που θα εξασφαλίζει και τον εύκολο υπολογισμό του μετασχηματισμού είναι τα φίλτρα κλίμακας και wavelet, $h(n)$ και $h_1(n)$ για τα ορθογώνια wavelet και $h'(n)$ και $h_1'(n)$ για τα διορθωγώνια, να είναι

πεπερασμένου μήκους (δηλαδή, FIR φίλτρα) και να έχουν λίγα σημεία. Επιπλέον ιδιότητες των φίλτρων που διευκολύνουν τους υπολογισμούς, π.χ. συμμετρικότητα των συντελεστών, είναι φυσικά επιθυμητές.

Εν γένει, η θεωρία της δημιουργίας μιας κατάλληλης συνάρτησης κλίμακας και wavelet με βάση κάποιες προδιαγραφές και απαιτήσεις είναι ένας ολόκληρος κλάδος των μαθηματικών από μόνος του και έχουν γραφεί πολλά ειδικευμένα βιβλία πάνω σε αυτό το ζήτημα (στη βιβλιογραφία παραθέτουμε μερικά). Αυτό είναι και ένα πολύ ισχυρό πλεονέκτημα του μετασχηματισμού wavelet, καθώς υπάρχει η ευελιξία της επιλογής των καταλληλότερων συναρτήσεων wavelet και βάσης για μια οικογένεια σημάτων ή για ένα συγκεκριμένο σκοπό επεξεργασίας. Έτσι, έχουν αναπτυχθεί wavelets (δηλαδή, συναρτήσεις wavelet και βάσης) για μια πληθώρα εφαρμογών, όπως γεωφυσική, συμπίεση εικόνας και ήχου, γραφικά, οικονομικές προβλέψεις κλπ. που σε κάθε περίπτωση δίνουν μετασχηματισμούς πολύ πιο «πληροφοριακούς» απ' ότι οι βάσεις τριγωνομετρικών συναρτήσεων.

Μια πολύ διαδεδομένη οικογένεια wavelet, με το μετασχηματισμό ενός μέλους της οποίας θα ασχοληθούμε και εμείς στη συνέχεια, είναι τα wavelets Cohen- Daubechies-Feauveau, η εν συντομία CDF. Πρόκειται διορθογώνια wavelets τα οποία έχουν συμμετρικά, FIR φίλτρα και των οποίων τα μήκη είναι κατά το δυνατόν παρόμοια. Τα CDF έχουν και άλλες επιθυμητές και σημαντικές ιδιότητες, βέβαια, αλλά αυτές είναι πέραν των σκοπών της εργασίας.

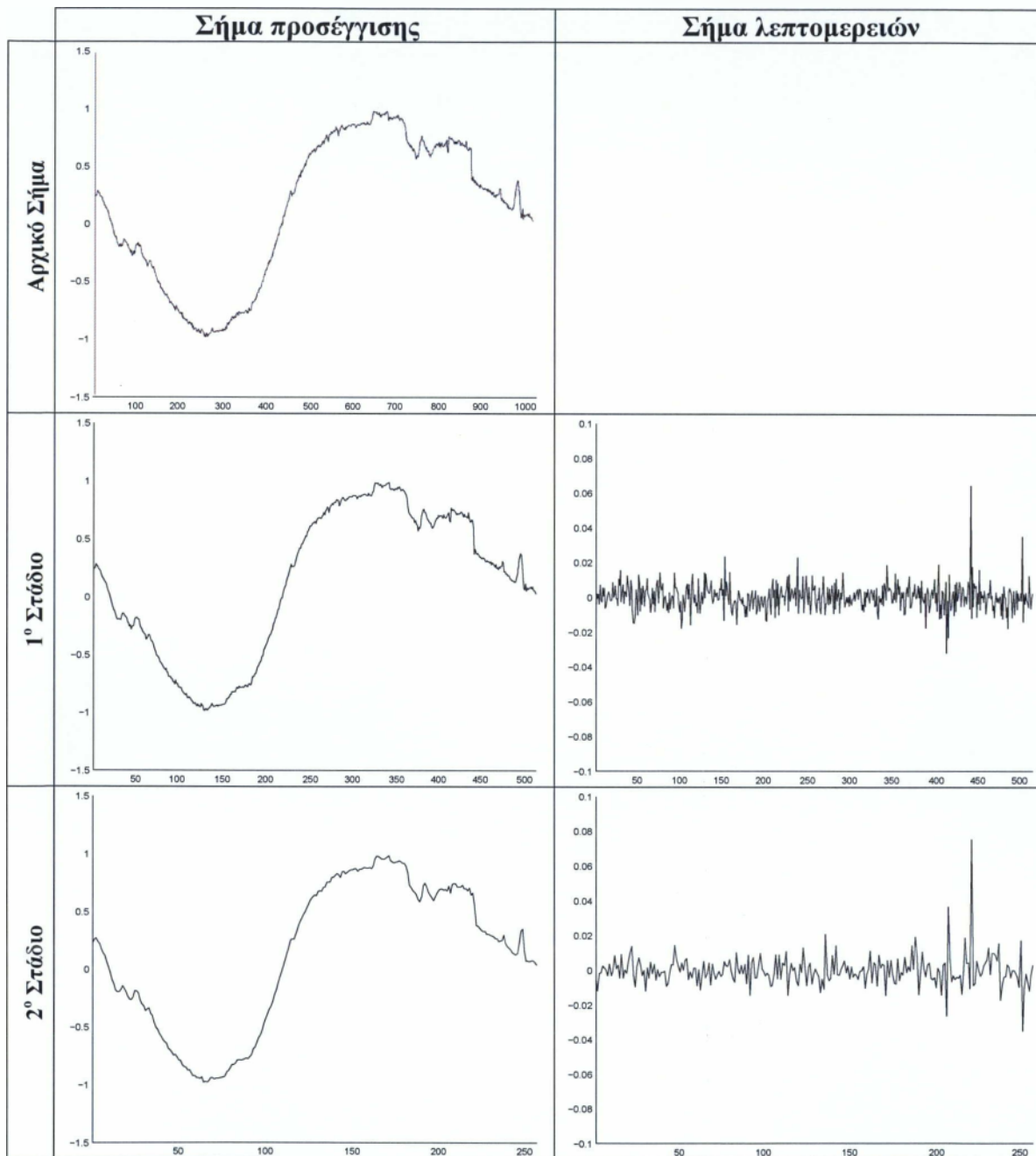
Ένα μέλος της οικογένειας CDF χαρακτηρίζεται από το μήκος (σε σημεία) του βαθυπερατού και του υψιπερατού φίλτρου που εκτελούν τον αντίστοιχο διακριτό μετασχηματισμό. Το πιο διαδεδομένο και γνωστό μέλος αυτής της οικογένειας είναι το CDF 9/7, στον ευθύ διακριτό μετασχηματισμό του οποίου το βαθυπερατό φίλτρο έχει μήκος 9 σημείων και το υψιπερατό 7.



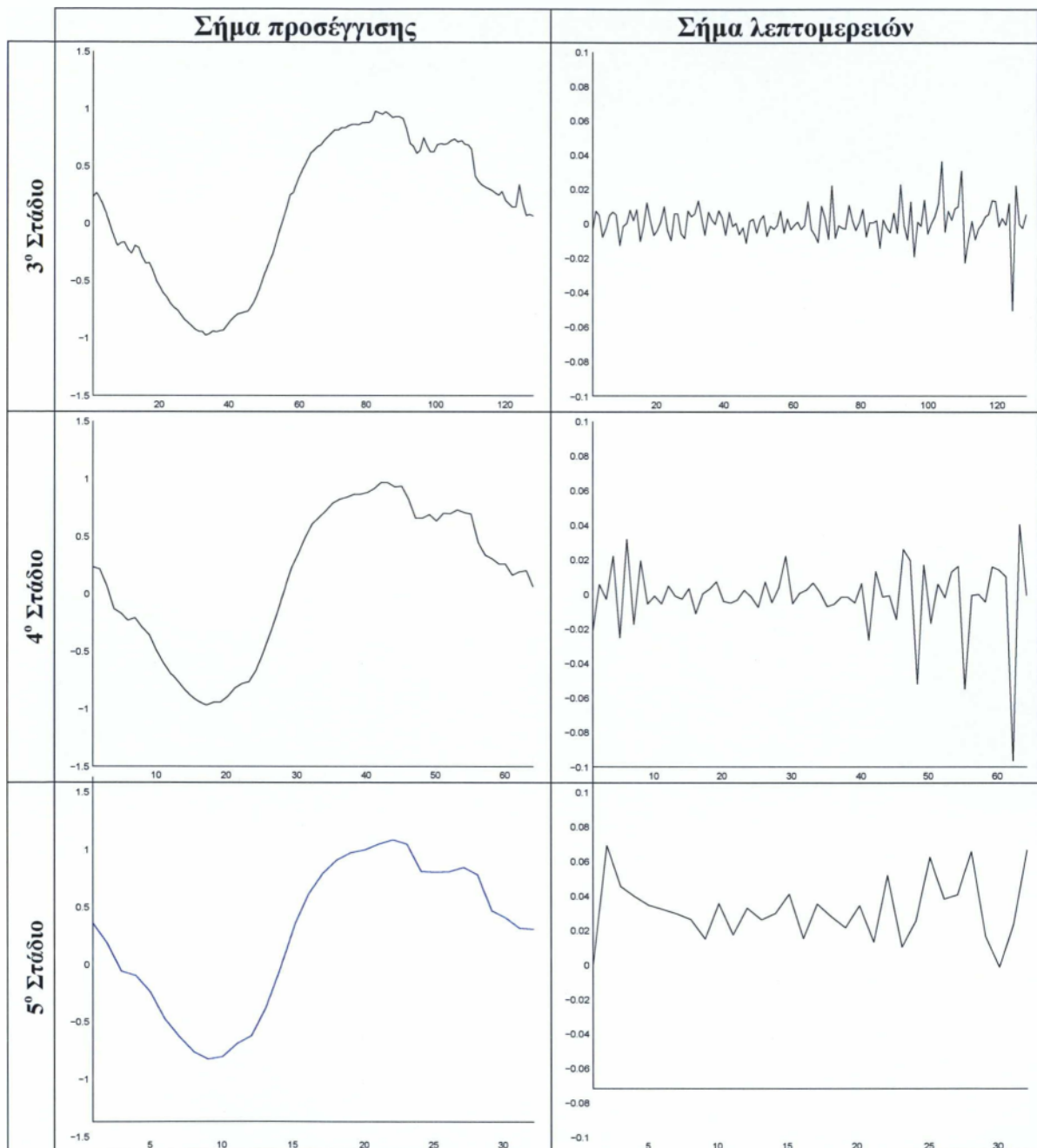
Σχήμα 3.4 Συνάρτηση κλίμακας και wavelet Cohen-Daubechies-Feauveau 9/7

Στην εργασία αυτή θα ασχοληθούμε με αρχιτεκτονικές αριθμητικών κυκλωμάτων που εκτελούν το διακριτό μετασχηματισμό wavelet CDF 9/7. Αν και όπως δείξαμε πριν, για την εκτέλεση του μετασχηματισμού δε χρειάζεται να ξέρουμε τη μορφή των συναρτήσεων κλίμακας και wavelet παρά μόνο τους συντελεστές των $H(z)$ και $L(z)$, παραθέτουμε εδώ τη μορφή τους για λόγους πληρότητας, καθώς και τα αποτελέσματα του διακριτού μετασχηματισμού τους (5 σταδίων) στα 1024 πρώτα δείγματα του σήματος *leleccum* από τις βιβλιοθήκες του Matlab.

Σχήμα 3.5 Σταδιακή αποδόμηση του σήματος leleccum με χρήση του μετασχηματισμού wavelet CDF9/7



Σχήμα 3.6 (συνέχεια 3.5) Σταδιακή αποδόμηση του σήματος *leleccum* με χρήση του μετασχηματισμού *wavelet CDF9/7*



Κεφάλαιο 4 Εφαρμογές

Από τα παραπάνω είναι εμφανές ότι οι μετασχηματισμοί wavelet έχουν μια πληθώρα εφαρμογών σε θέματα επεξεργασίας σημάτων. Είκοσι σχεδόν χρόνια μετά την ανακάλυψη των πρώτων σημαντικών wavelets από την Ingrid Daubechies, και παρά τα διάφορα προβλήματα νομικής φύσεως που έχουν προκύψει λόγω ευρεσιτεχνιών στα μαθηματικά, χρησιμοποιούνται συνεχώς σε τομείς όπου συναντάμε μη-εργοδικά, μη-περιοδικά σήματα και οι παραδοσιακοί μετασχηματισμοί Fourier αδυνατούν να δώσουν χρήσιμα αποτελέσματα.

Υπάρχει πλήθος από τέτοιες εφαρμογές. Οι κυριότερες είναι στην Ιατρική και τη Μηχανολογία, για εντοπισμό απότομων ασυνεχειών που υποδηλώνουν προβληματική λειτουργία και οι οποίες θα χάνονταν μέσα στο θόρυβο αν χρησιμοποιούσαμε τεχνικές Fourier, στην αστρονομία για τον διαχωρισμό σημάτων από το θόρυβο, ακόμα και στα μακρο-οικονομικά για μακροπρόθεσμες προβλέψεις της πορείας των μετοχών. Μια από τις πιο γνωστές εφαρμογές, στις οποίες χρησιμοποιείται το wavelet CDF 9/7 είναι ο αλγόριθμος συμπίεσης εικόνας JPEG2000, ο οποίος δημιουργήθηκε για να αντικαταστήσει τον πολύ γνωστό JPEG.

Αν και οι πατέντες λογισμικού έχουν αποτρέψει την ευρείας κλίμακας χρησιμοποίησή του ως τώρα, πετυχαίνει πολύ ανώτερα αποτελέσματα από τον παραδοσιακό JPEG τόσο σε επίπεδο συμπίεσης όσο και επίπεδο ταχύτητας. Επιπλέον, η φύση των wavelets να μπορούν να δίνουν επιλεκτικά μέχρι ένα επίπεδο λεπτομερειών σε ένα σήμα, έχει βρει εφαρμογή σε περιπτώσεις που έχουμε μεγάλες εικόνες τις οποίες όταν τις βλέπουμε στο σύνολό τους δε

χρειαζόμαστε μεγάλη λεπτομέρεια, αλλά όταν εστιάζουμε σε μια συγκεκριμένη περιοχή περιμένουμε να δούμε και λεπτομέρειες. Τέτοιες είναι για παράδειγμα τοπογραφικές εικόνες που χρησιμοποιούνται από συστήματα GIS, για τις οποίες έχει αναπτυχθεί το πρότυπο συμπίεσης με wavelets ECW

Είναι φανερό ότι λόγω της ωρίμανσης των σχετικών αλγορίθμων αλλά και της επίλυσης των νομικών ζητημάτων στο προσεχές μέλλον, οι εφαρμογές των μετασχηματισμών wavelet θα πολλαπλασιαστούν και ενδεχομένως θα εκτοπίσουν από μερικούς τομείς τους παραδοσιακούς μετασχηματισμούς αναπαράστασης στη συχνότητα. Έτσι, κρίνεται επιτακτική η ανάπτυξη αποδοτικών και εξειδικευμένων αριθμητικών κυκλωμάτων που θα είναι σε θέση να εκτελούν γρήγορα και οικονομικά τους μετασχηματισμούς αυτούς σε ενσωματωμένα συστήματα, όπως φορητά ιατρικά όργανα, κινητά τηλέφωνα, «mp3» players και ψηφιακές φωτογραφικές μηχανές. Αυτό είναι και το ζήτημα με το οποίο πραγματεύεται η εργασία αυτή.

Κεφάλαιο 5 Εφαρμογή Matlab

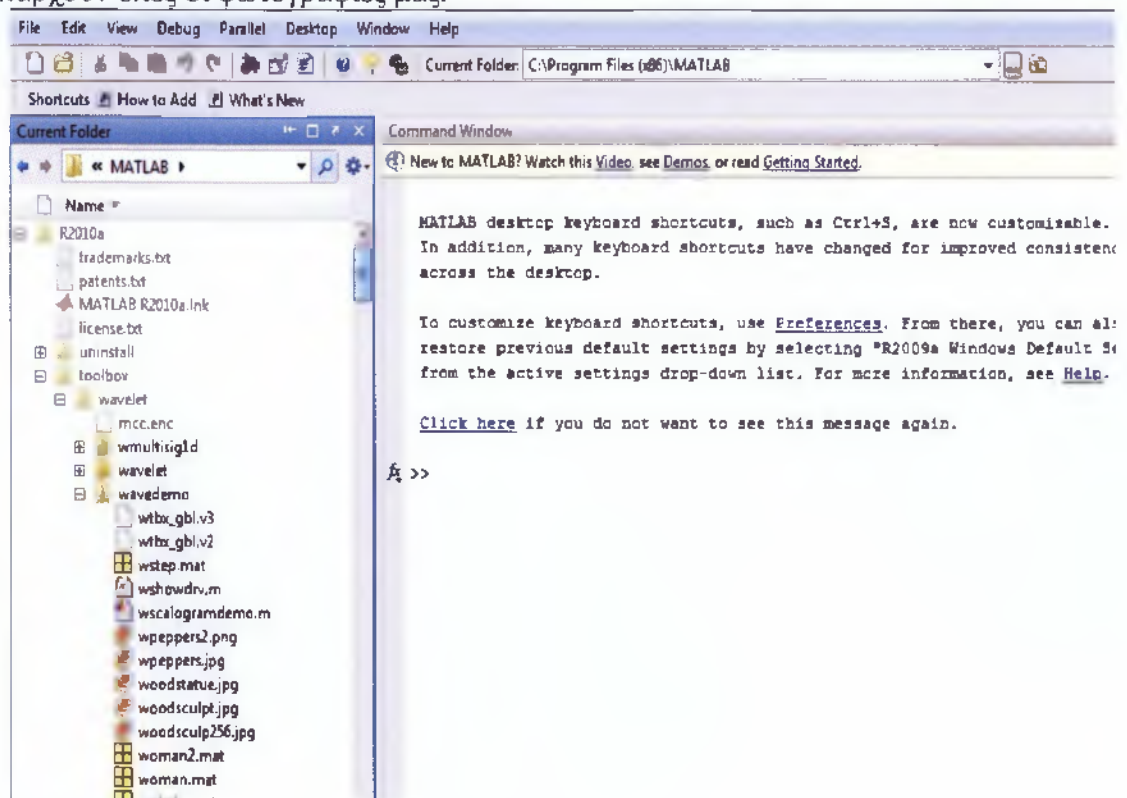
5.1 Εισαγωγή

Σε αυτό το κεφάλαιο με την βοήθεια του προγράμματος Matlab θα γράψουμε έναν κώδικα ο οποίος διαβάζει την εικόνα, την συμπιέζει και εκτυπώνει την αρχική και την ανακατασκευασμένη εικόνα σε διάφορα βήματα υπολογίζοντας παράλληλα το μέσο τετραγωνικό λάθος (MSE) και τον μέγιστο σηματοθορυβικό λόγο (PSNR).

5.2 Λεπτομέρειες Χρήσης

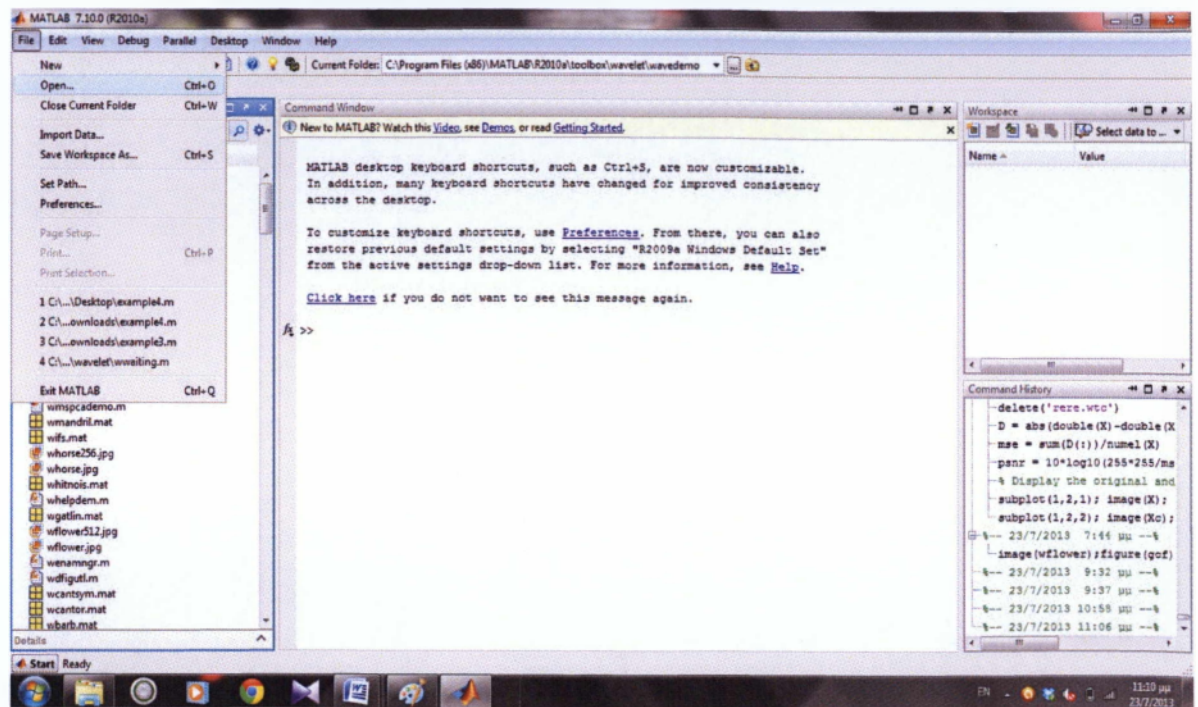
Για να μπορέσουμε να κάνουμε χρήση του κωδικοποιητή είναι απαραίτητο να έχουμε κάνει εγκατάσταση του προγράμματος Matlab στον υπολογιστή μας. Στη συνέχεια πρέπει να προσθέσουμε τις εικόνες που θέλουμε να επεξεργαστούμε οι οποίες πρέπει να είναι κατάλληλες για το Matlab. Έπειτα θα δείξουμε κάποια παραδείγματα και το πως φαίνεται η συμπεσμένη εικόνα.

Οπότε ανοίγουμε το εικονίδιο του υπολογιστή στην Επιφάνεια Εργασίας μετά συνεχίζουμε Program Files > MATLAB > R2010a > toolbox > wavelet > wavedemo (C:\Program Files (x86)\MATLAB\R2010a\toolbox\wavelet\wavedemo) και βάζουμε τις φωτογραφίες που θέλουμε εκεί μέσα. Στη συνέχεια αφού ανοίξουμε το περιβάλλον του MATLAB πάμε στην αναζήτηση και βρίσκουμε το wavedemo. Εκεί μέσα υπάρχουν όλες οι φωτογραφίες μας.



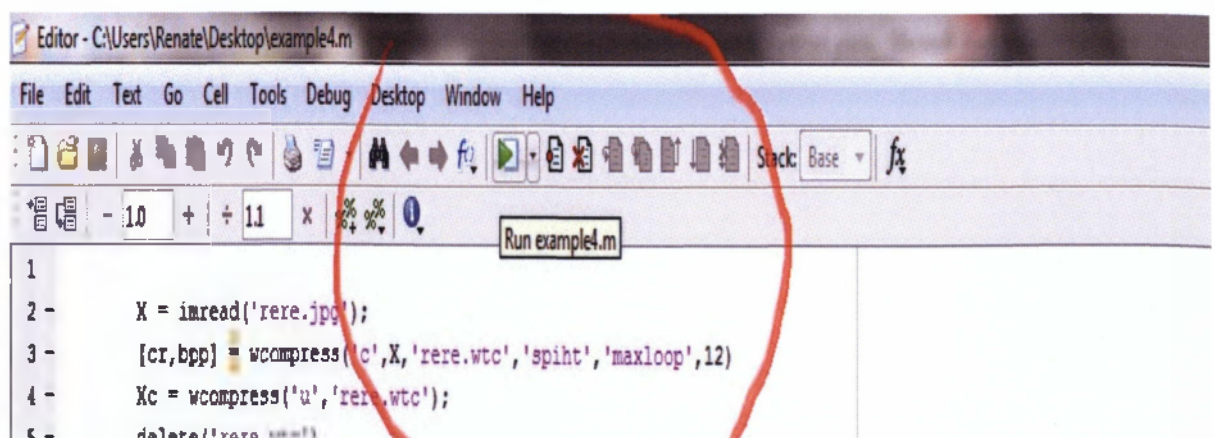
Εικόνα : 1 Matlab

Μετά πάμε file->Open->Desktop->example(mfile)



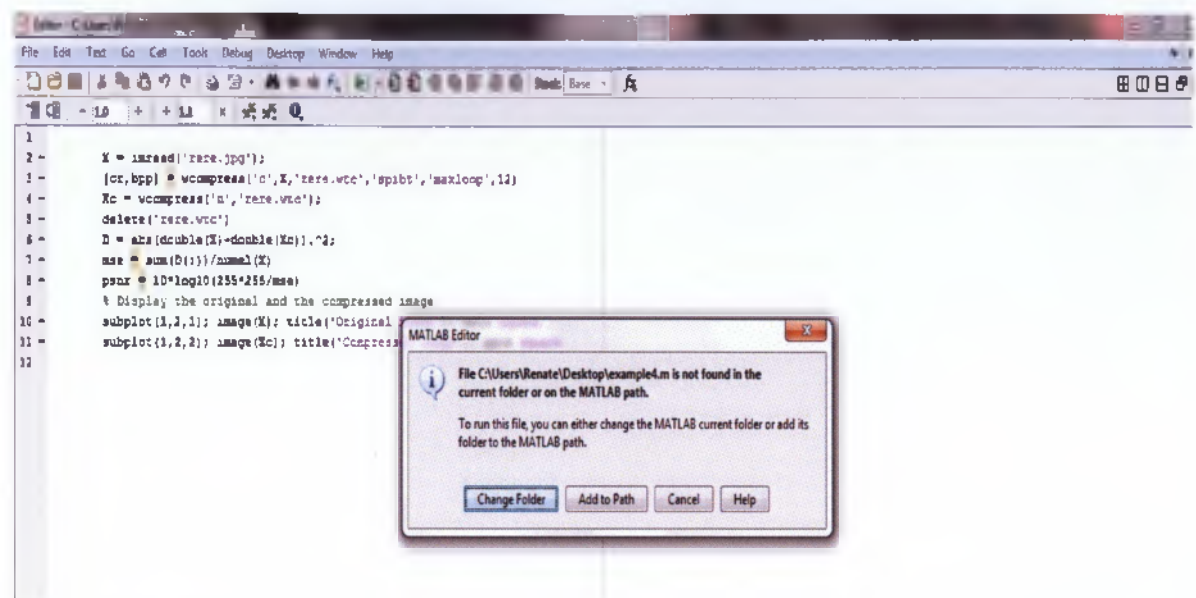
Εικόνα : 2 Matlab example

Σε κάθε παράδειγμα πιο κάτω χρησιμοποιούμε και άλλο m-file ωστόσο σε κάθε περίπτωση πατάμε το “run”.



Εικόνα : 3 Matlab run example

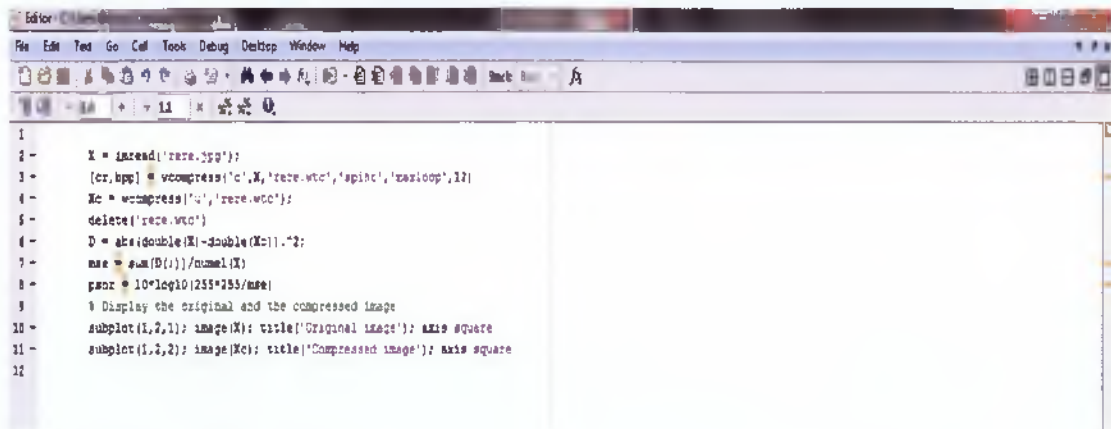
και συνεχίζουμε πατώντας “add to path” στο πινακάκι που μας βγάζει.



Εικόνα : 4 Matlab Editor

ΠΑΡΑΔΕΙΓΜΑ 1.

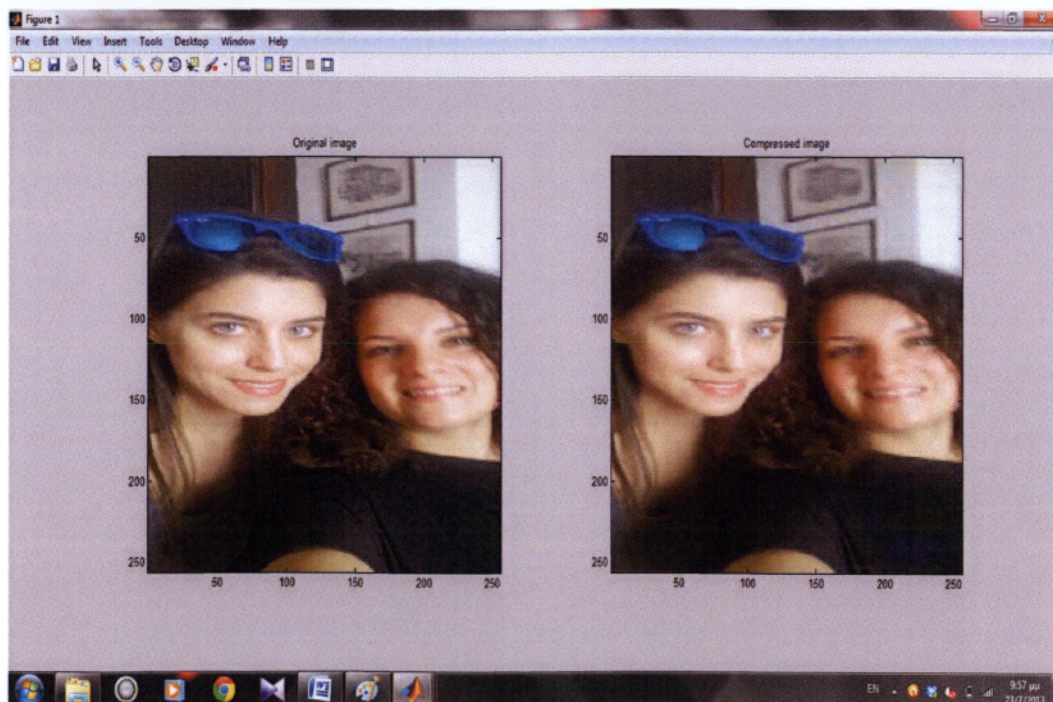
Σε αυτό το παράδειγμα συμπίεσης θα χρησιμοποιήσουμε τη μέθοδο «spiht_3d» διότι μας δίνει τα καλύτερα αποτελέσματα. Αρχικά θα χρησιμοποιήσουμε την έγχρωμη εικόνα `rere.jpg`, αυτή η εικόνα θα συμπιεστεί και τα δεδομένα της θα αποθηκευτούν στο αρχείο `rere.wtc`.



```
1
2 - X = imread('rere.jpg');
3 [cr, bpp] = wcompress('c', X, 'rere.wtc', 'spiht', 'maxloop', 12);
4 Xc = wcompress('u', 'rere.wtc');
5 delete('rere.wtc');
6 D = abs(double(X) - double(Xc)).^2;
7 mse = sum(D(:))/ numel(X);
8 pmse = 10*log10(255*255/mse);
9 % Display the original and the compressed image
10 subplot(1,2,1); image(X); title('Original image'); axis square
11 subplot(1,2,2); image(Xc); title('Compressed image'); axis square
12
```

Εικόνα : 5 Matlab Editor Rere

Στην συνέχεια το πρόγραμμα προβάλλει την αρχική εικόνα `X` και την ανακατασκευασμένη εικόνα `Xc`.



Εικόνα : 6 Matlab Renate-Demy

Μέσα στο “Command Window” του Matlab βλέπουμε κατευθείαν το μέσο τετραγωνικό λάθος (MSE) και τον μέγιστο σηματοθορυβικό λόγο (PSNR). Στη συγκεκριμένη φωτογραφία είναι $mse=25,837$ και $psnr=34,5928$.

```
To customize keyboard shortcuts, use Preferences. From there, you can also
restore previous default settings by selecting "R2009a Windows Default Set"
from the active settings drop-down list. For more information, see Help.

Click here if you do not want to see this message again.

>

mse =

    25.837

bpp =

    1.4324

Warning: File 'rere.wtc' not found.
> In examples at 5

mse =

    22.5837

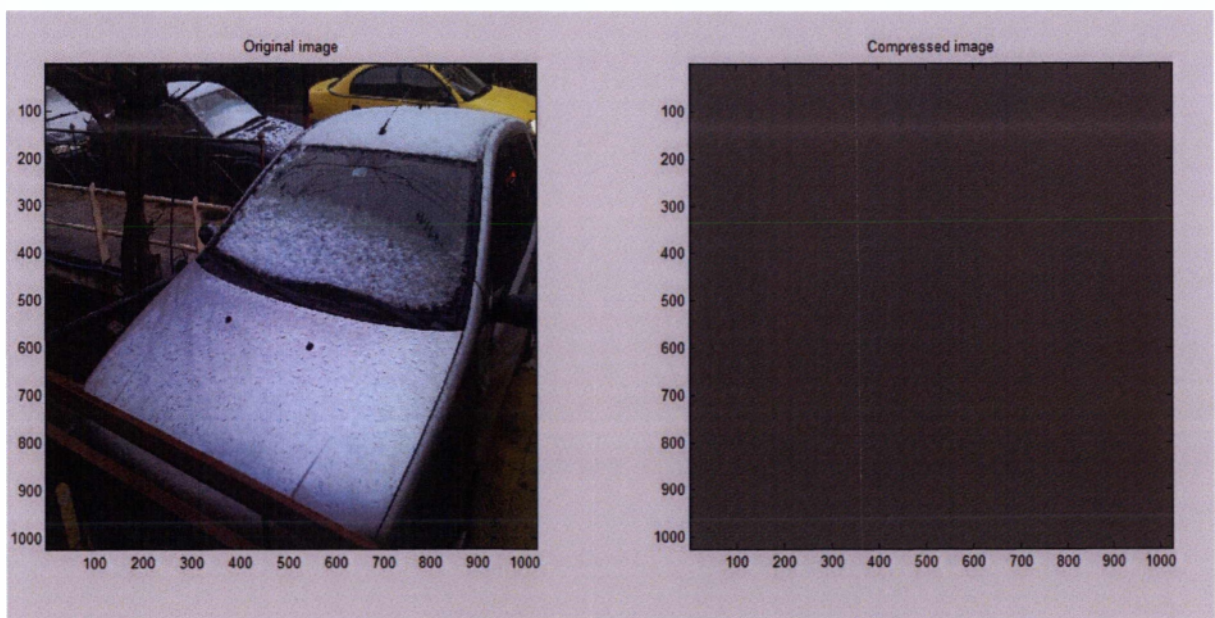
psnr =

    34.5928

fx \s
Εικόνα :7 mse , psnr
```

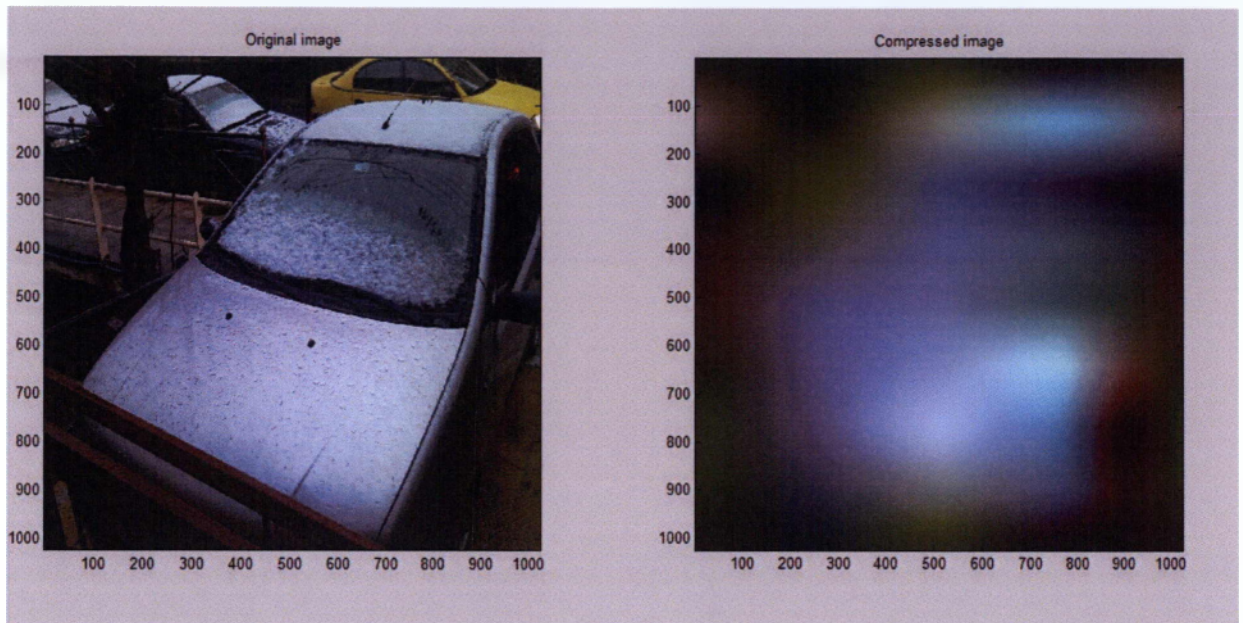
Σε λίγο θα παραθέσω και άλλες φωτογραφίες με αυτή τη μέθοδο σε διαφορετικό βήμα για να δείτε πως γίνεται η συμπίεση της εκάστοτε αρχικής εικόνας.

Με βήμα 2 το $mse=5.7229e+003$ και το $psnr=10.5546$.



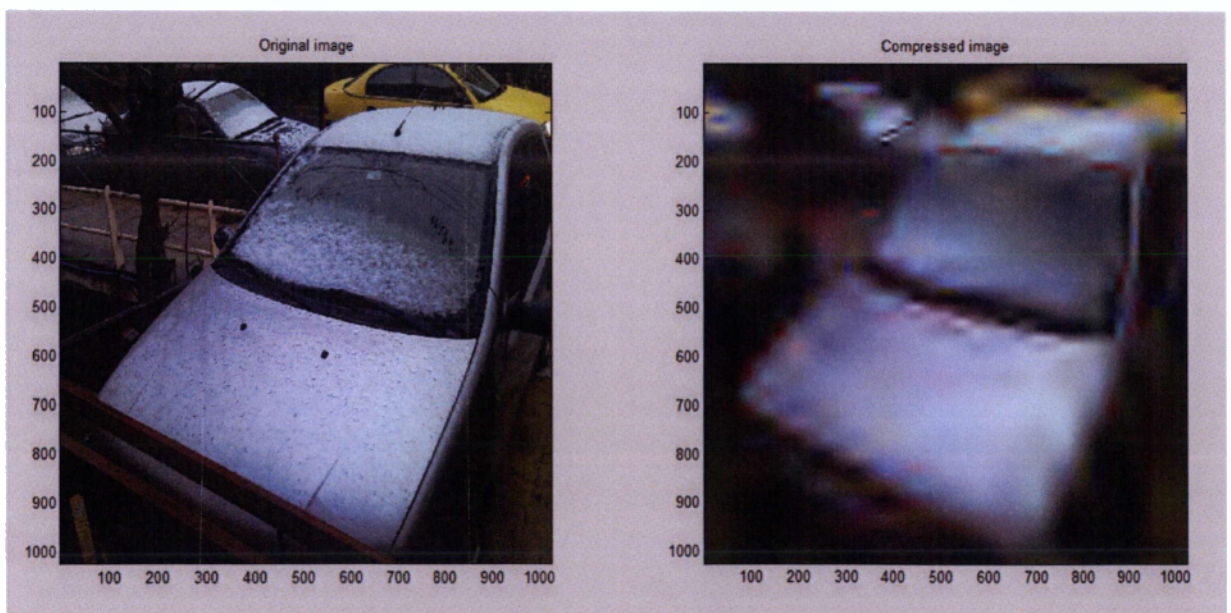
Εικόνα : 8 Αυτοκίνητο με βήμα 2 ($mse=5.7229e+003$ και $psnr=10.5546$)

Με βήμα 5 το $mse = 2.0796e+003$ και το $psnr = 14.9509$.



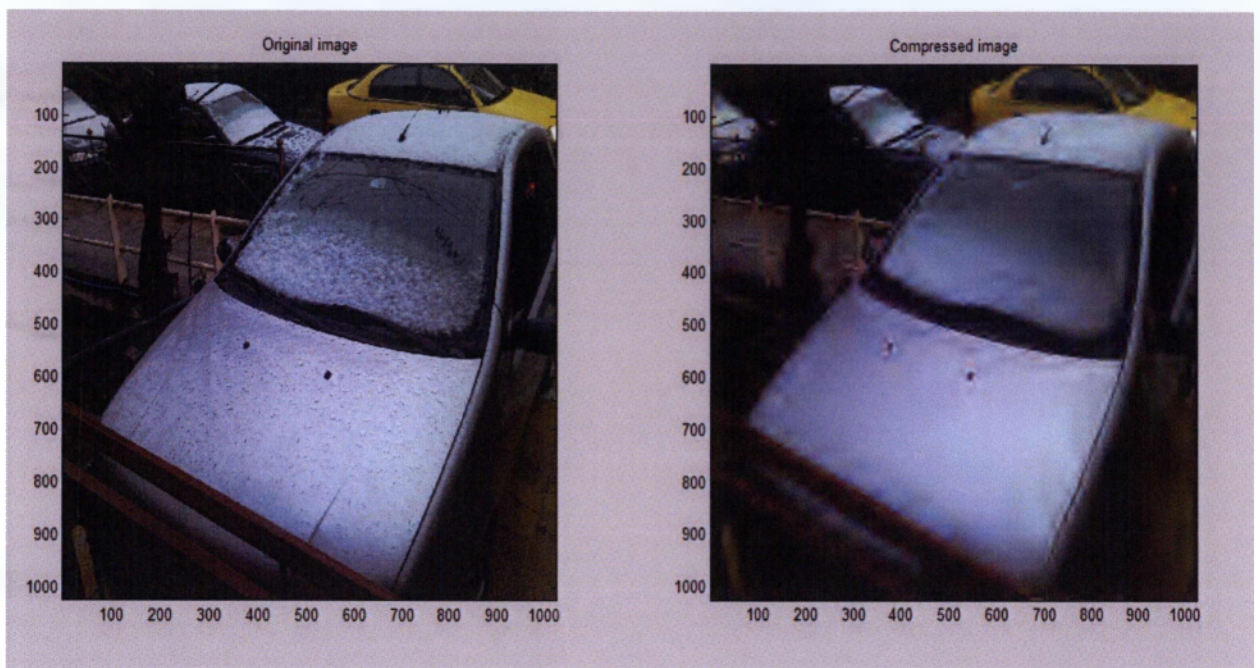
Εικόνα : 9 Αυτοκίνητο με βήμα 5 ($mse = 2.0796e+003$ και $psnr = 14.9509$)

Με βήμα 8 το $mse = 592.6064$ και το $psnr = 20.4031$.



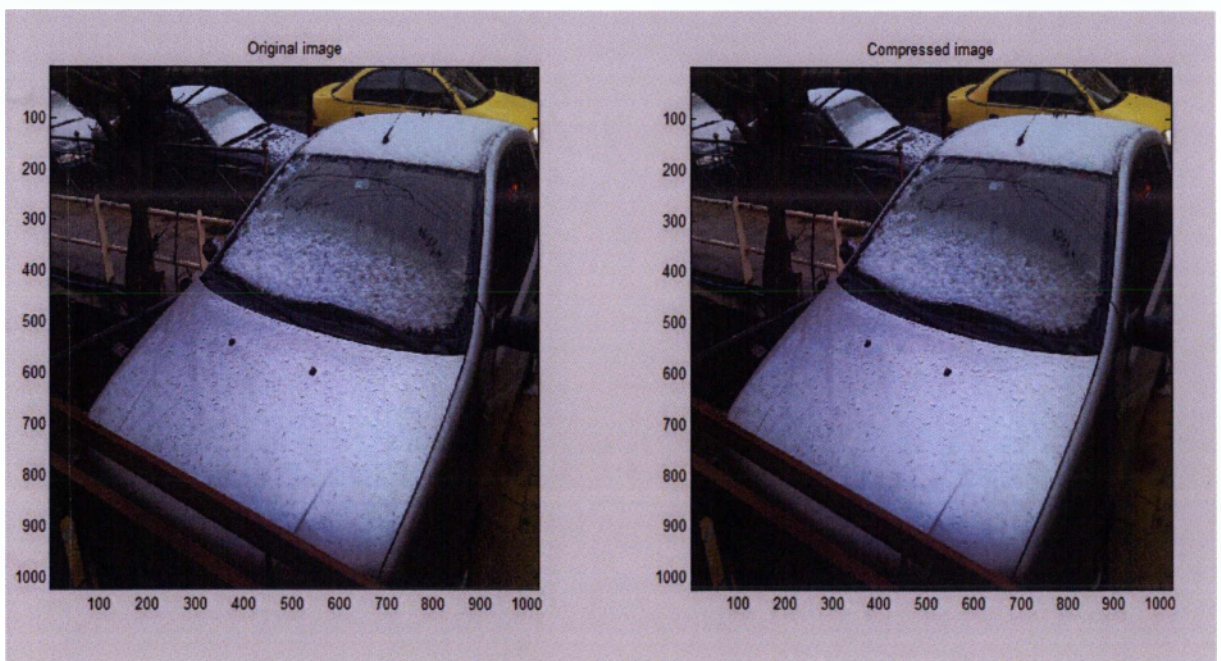
Εικόνα : 10 Αυτοκίνητο με βήμα 8 ($mse = 592.6064$ και $psnr = 20.4031$)

Με βήμα 10 το $mse = 249.4931$ και το $psnr = 24.1602$.



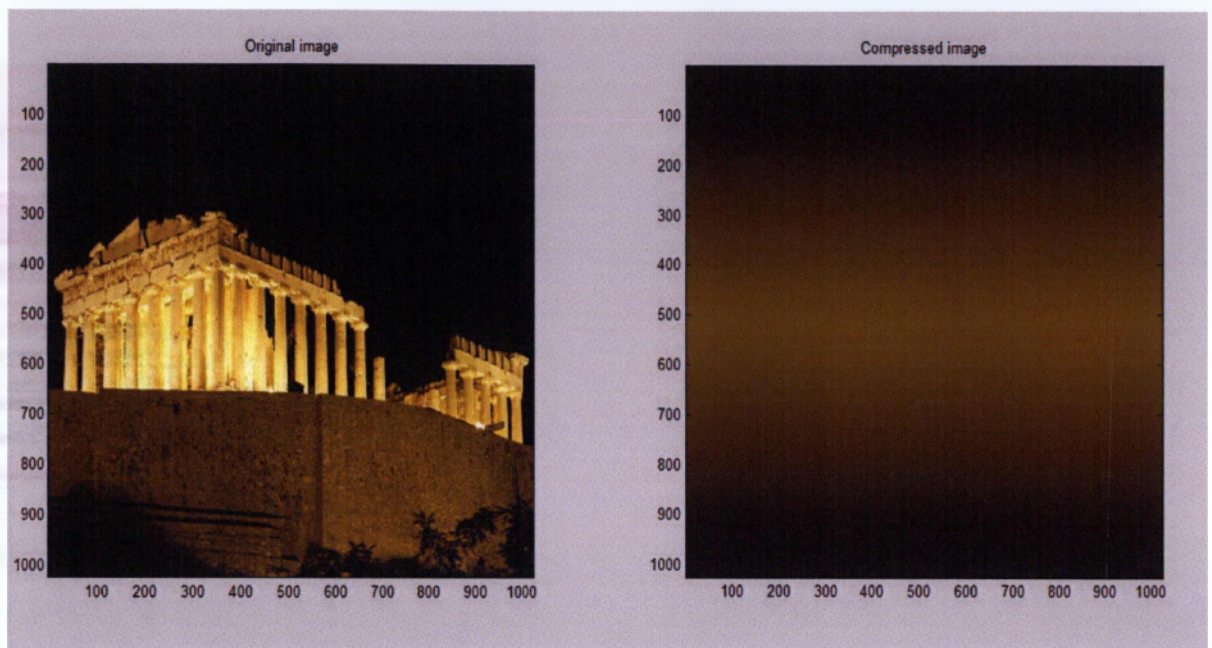
Εικόνα : 11 Αυτοκίνητο με βήμα 10 ($mse = 249.4931$ και $psnr = 24.1602$)

Με βήμα 15 το $mse = 11.1308$ και το $psnr = 37.6655$.



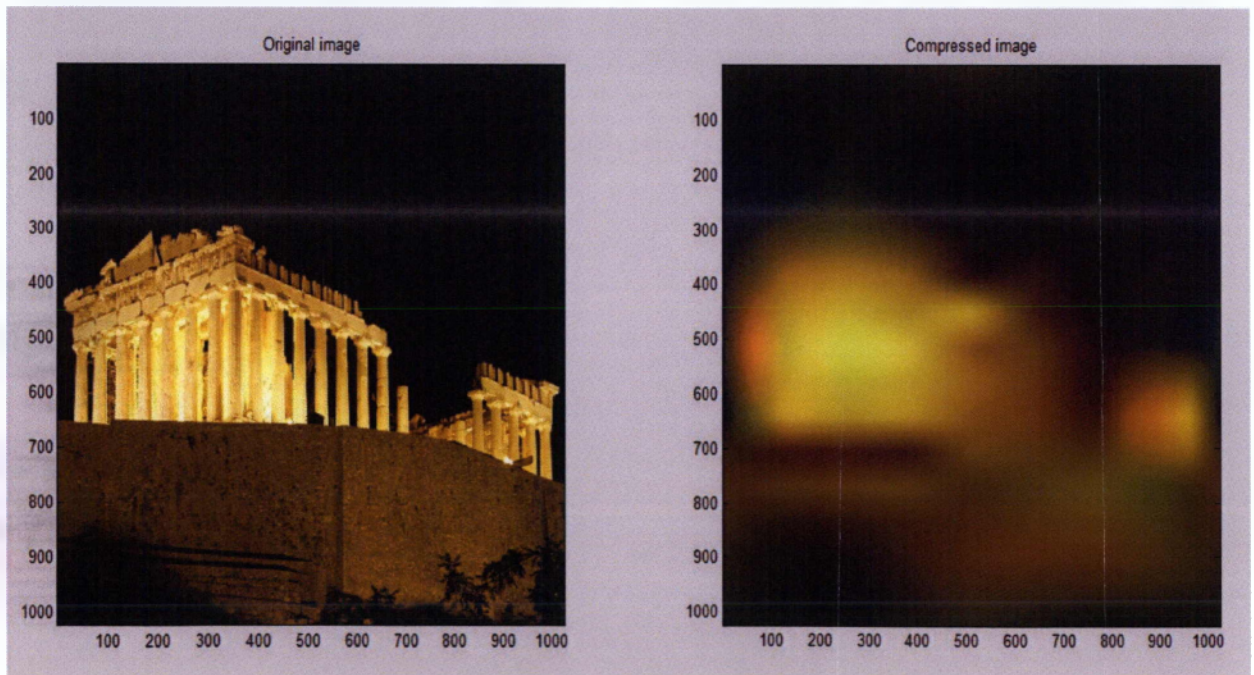
Εικόνα: 12 Αυτοκίνητο με βήμα 15 ($mse = 11.1308$ και $psnr = 37.6655$)

Με βήμα 2 το $mse = 3.2450e+003$ και το $psnr = 13.0186$.



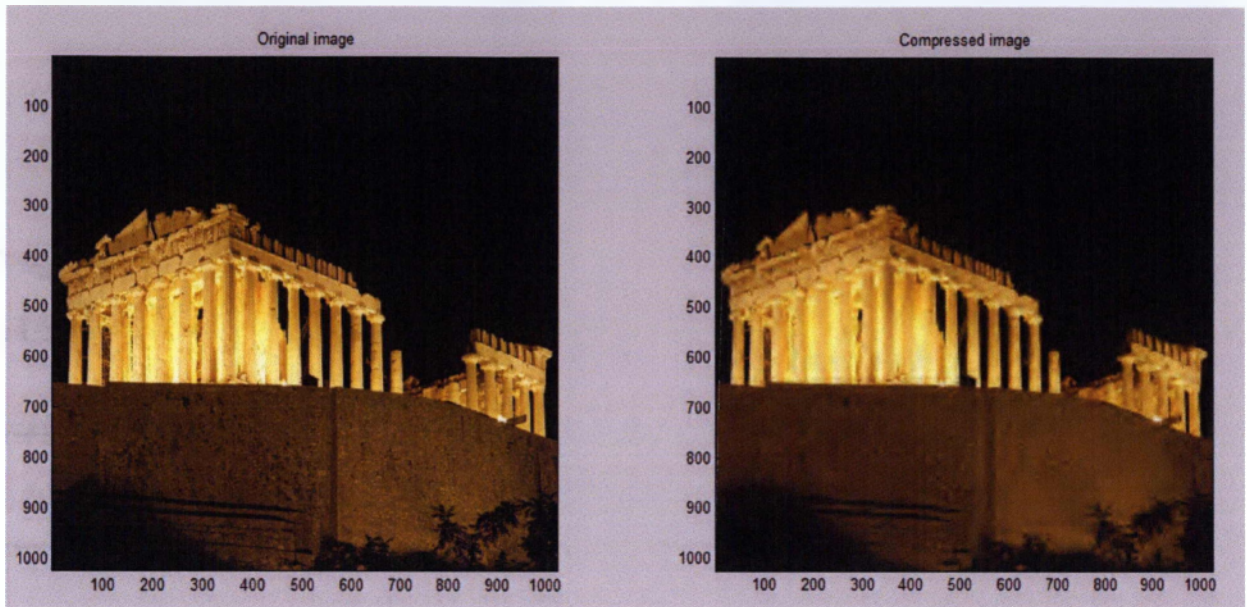
Εικόνα :13 Παρθενώνας με βήμα 2 ($mse = 3.2450 \cdot 10^3$ και $psnr = 24.1602$)

Με βήμα 5 το $mse = 1.0584e+003$ και το $psnr = 17.8841$.



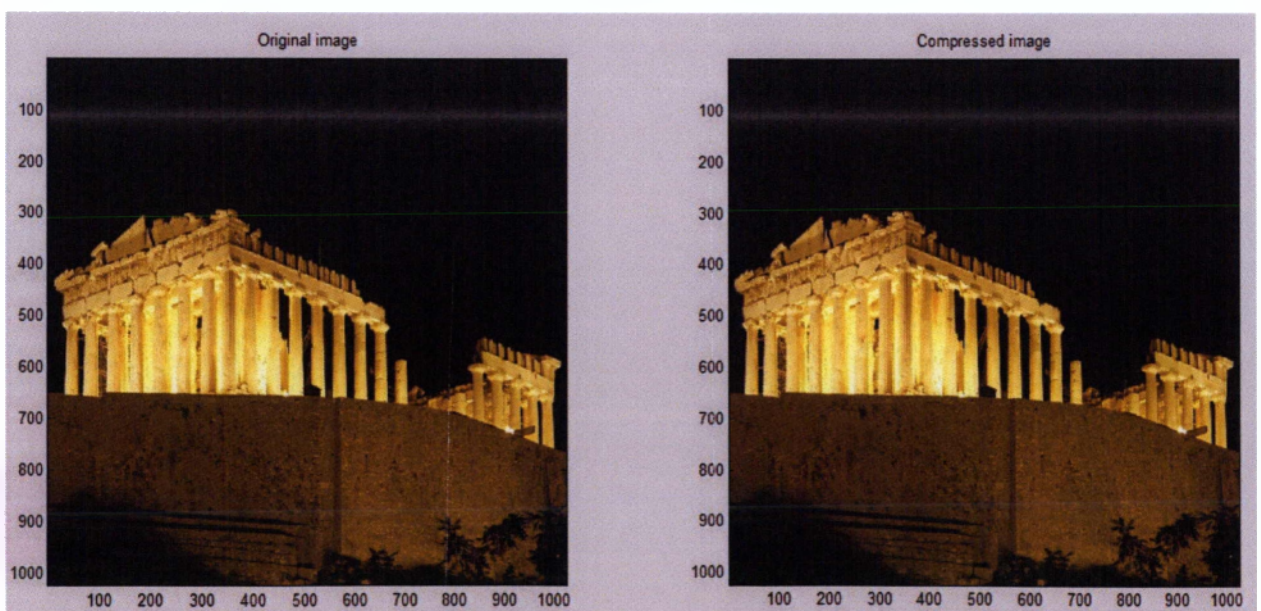
Εικόνα : 14 Παρθενώνας με βήμα 5 ($mse = 1.0584 \cdot 10^3$ και $psnr = 17.8841$)

Με βήμα 10 το $mse = 115.2553$ και το $psnr = 27.5142$.



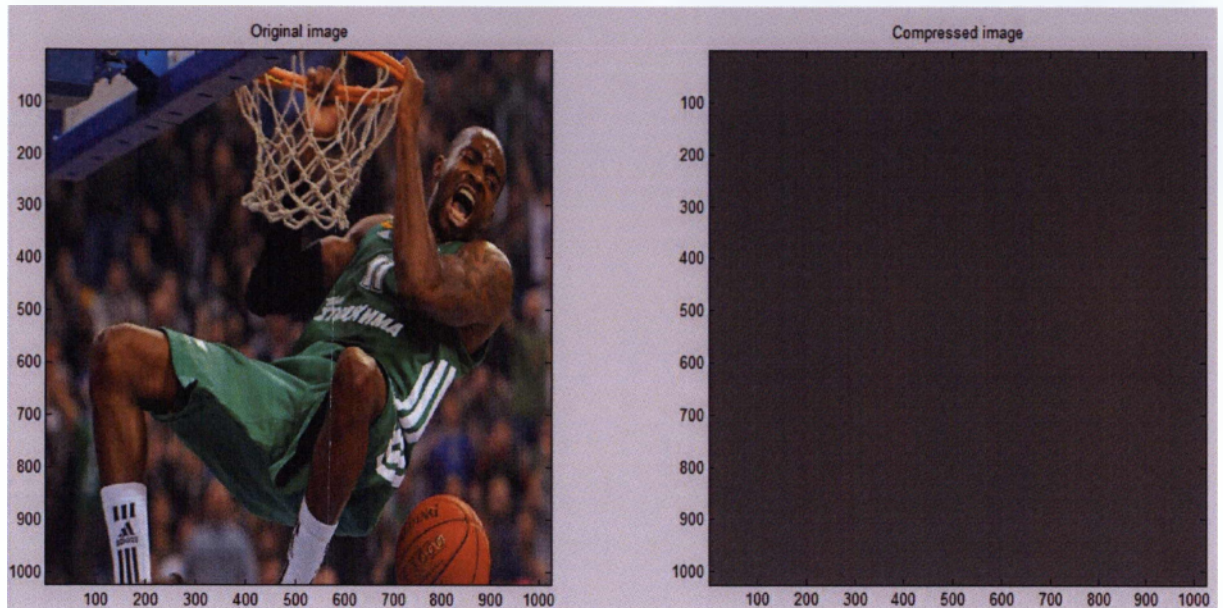
Εικόνα : 15 Παρθενώνας με βήμα 10 ($mse = 115.2553 \text{ e}+003$ και $psnr = 27.5142$)

Με βήμα 15 το $mse = 2.7866$ και το $psnr = 43.6801$.



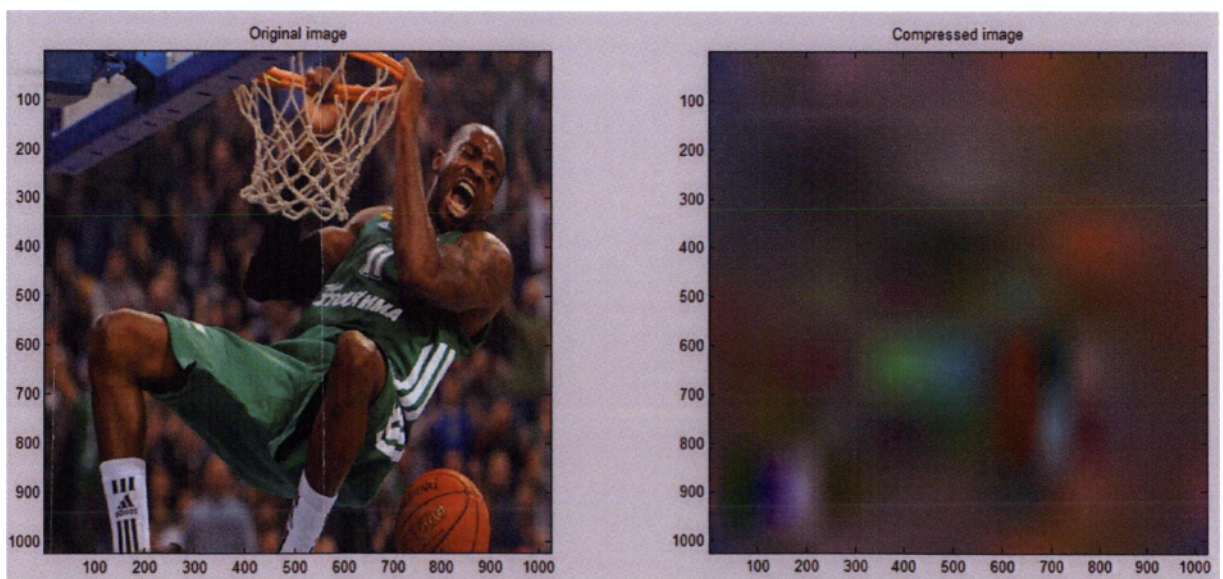
Εικόνα : 16 Παρθενώνας με βήμα 15 ($mse = 2.7866 \text{ e}+003$ και $psnr = 43.6801$)

Με βήμα 2 το $mse \approx 1.9421e+003$ και το $psnr = 15.2482$.



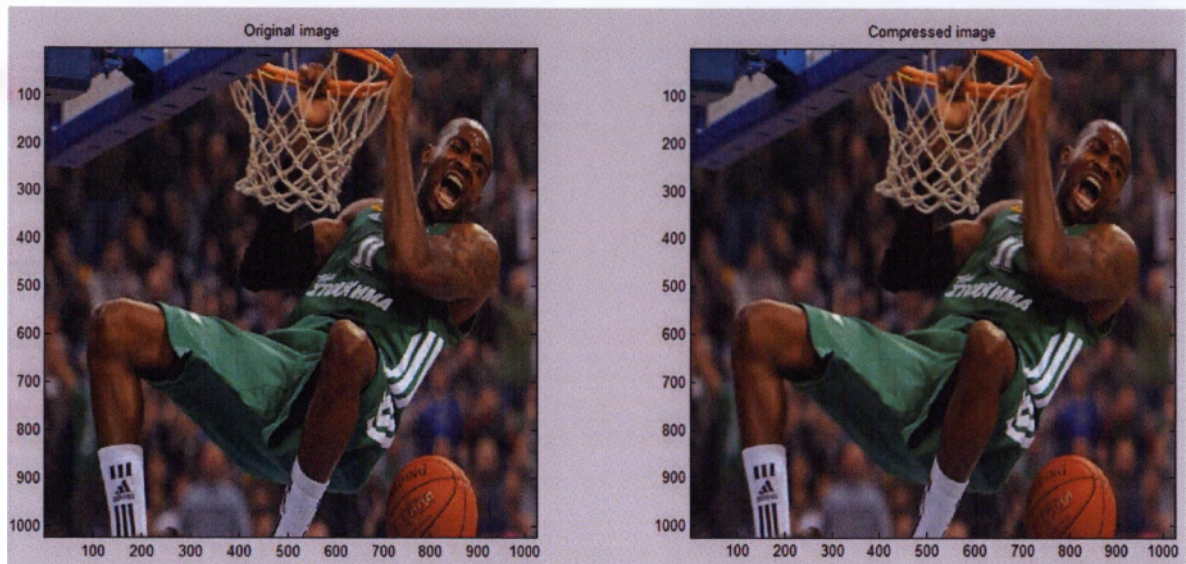
Εικόνα : 17 Λασμε με βήμα 2 ($mse = 1.9421 e+003$ και $psnr = 15.2482$)

Με βήμα 5 το $mse = 1.4849e+003$ και το $psnr = 16.4139$.



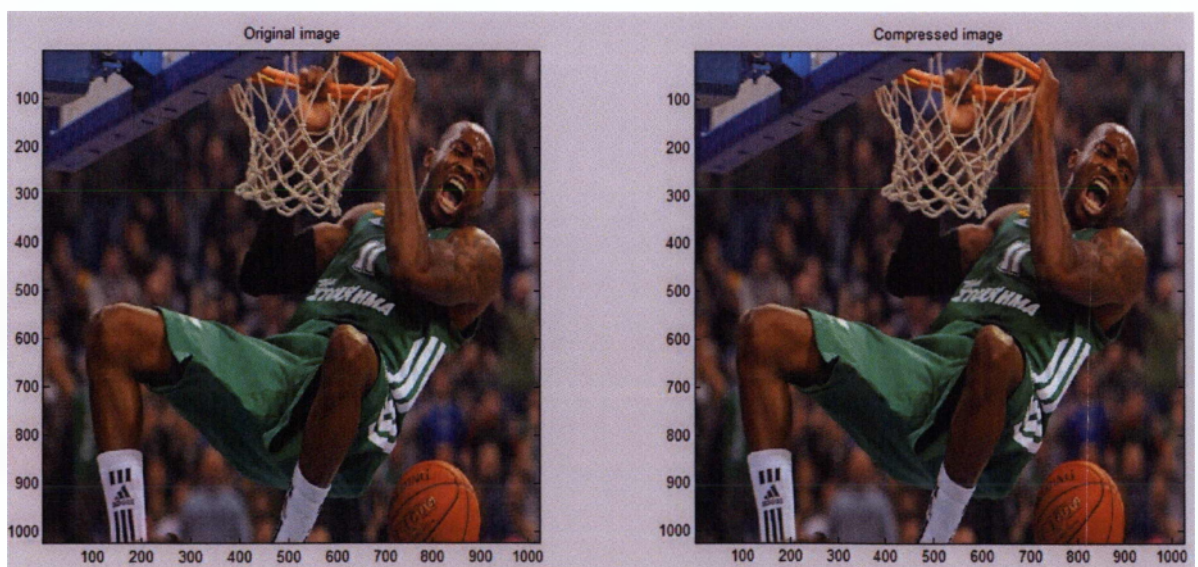
Εικόνα : 18 Λασμε με βήμα 5 ($mse = 1.4849 e+003$ και $psnr = 16.4139$)

Με βήμα 12 το mse =32.2998 και το psnr =33.0388.



Εικόνα : 19 Lasme με βήμα 12 (mse =32.2998 e+003 και psnr = 33.0388)

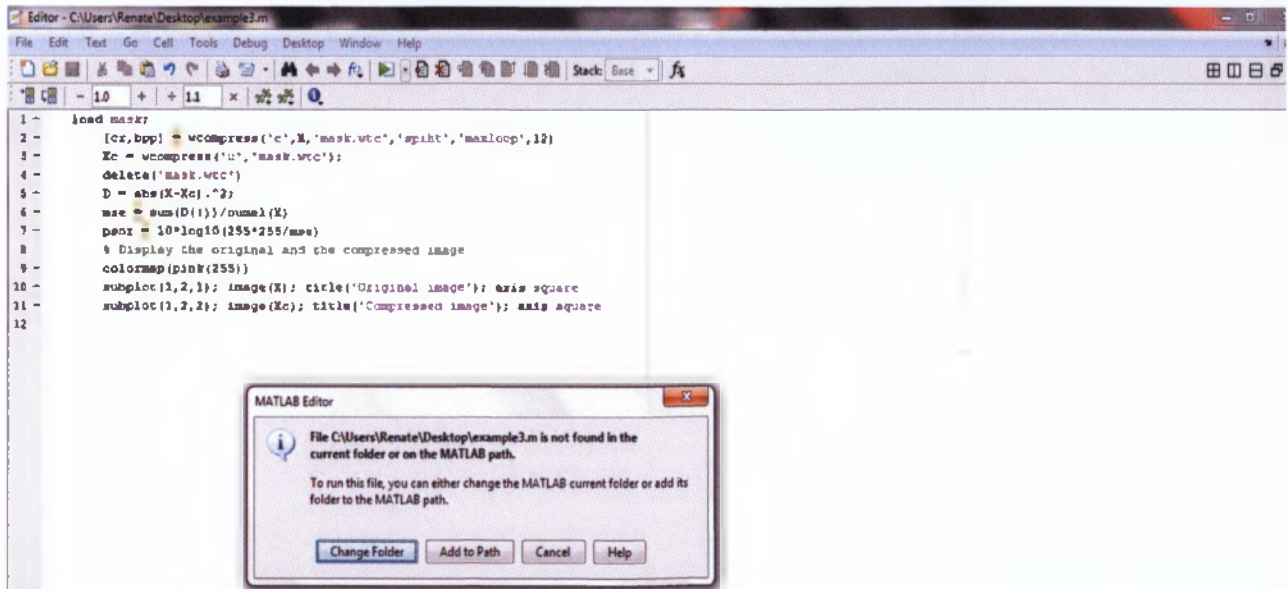
Με βήμα 15 το mse =3.8268 και το psnr = 42.3025.



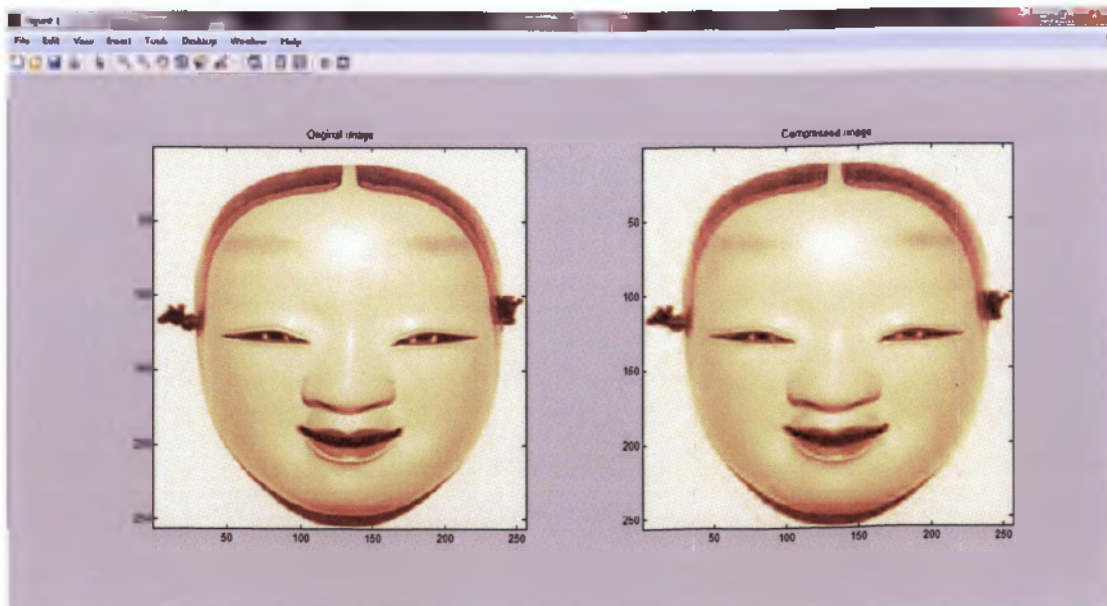
Εικόνα : 20 Lasme με βήμα 15 (mse =3.8268 e+003 και psnr = 42.3025)

ΠΑΡΑΔΕΙΓΜΑ 2.

Μια άλλη μέθοδος που μπορούμε να χρησιμοποιήσουμε είναι η 'spih1'. Εδώ βλέπουμε πως συμπιέζουμε την αρχική μας εικόνα και την αποθηκεύουμε στο αρχείο 'mask.wtc'. Στη συνέχεια παίρνουμε πάλι το αρχείο το αποσυμπιέζουμε και διαγράφουμε το 'mask.wtc'.

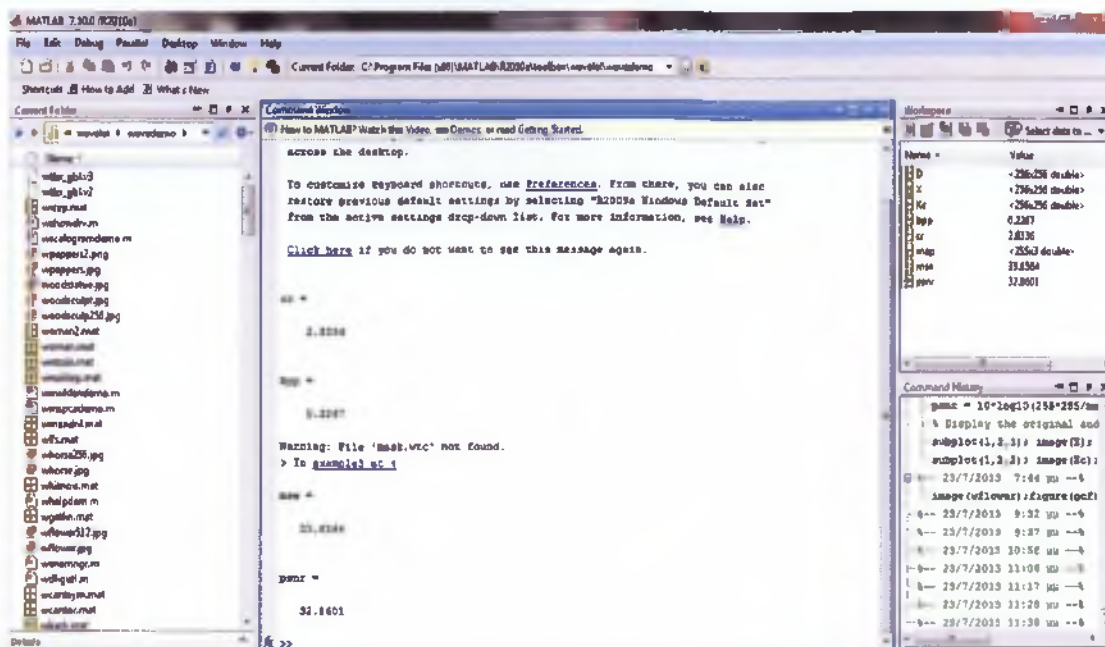


Εικόνα : 21 Matlab Example 2



Εικόνα : 22 Mask Original image - Compressed image

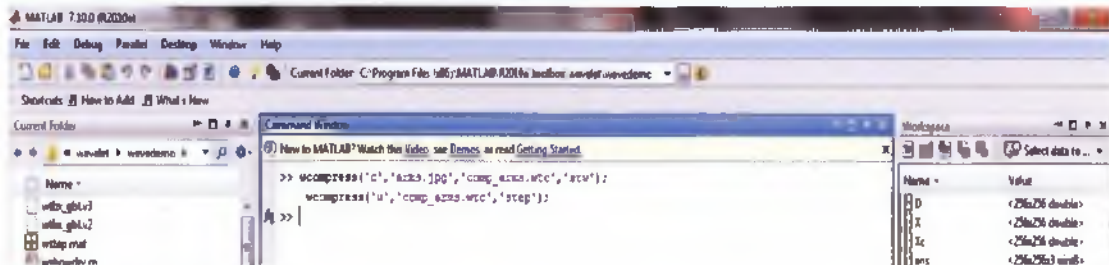
Βλέπουμε έντονη διαφορά στην αρχική εικόνα «mask» με τη συμπιεσμένη. Στη συγκεκριμένη φωτογραφία είναι mse= 33.6564 και psnr= 32.8601.



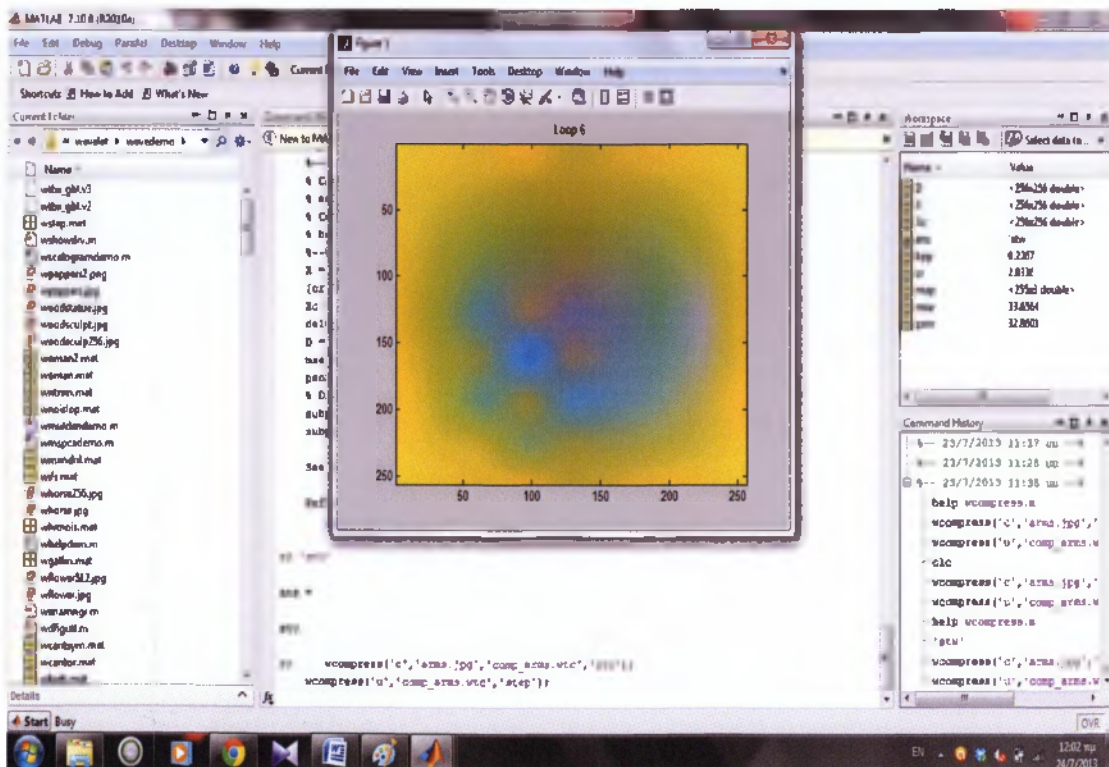
Εικόνα : 23 Matlab mse,psnr

ΠΑΡΑΔΕΙΓΜΑ 3.

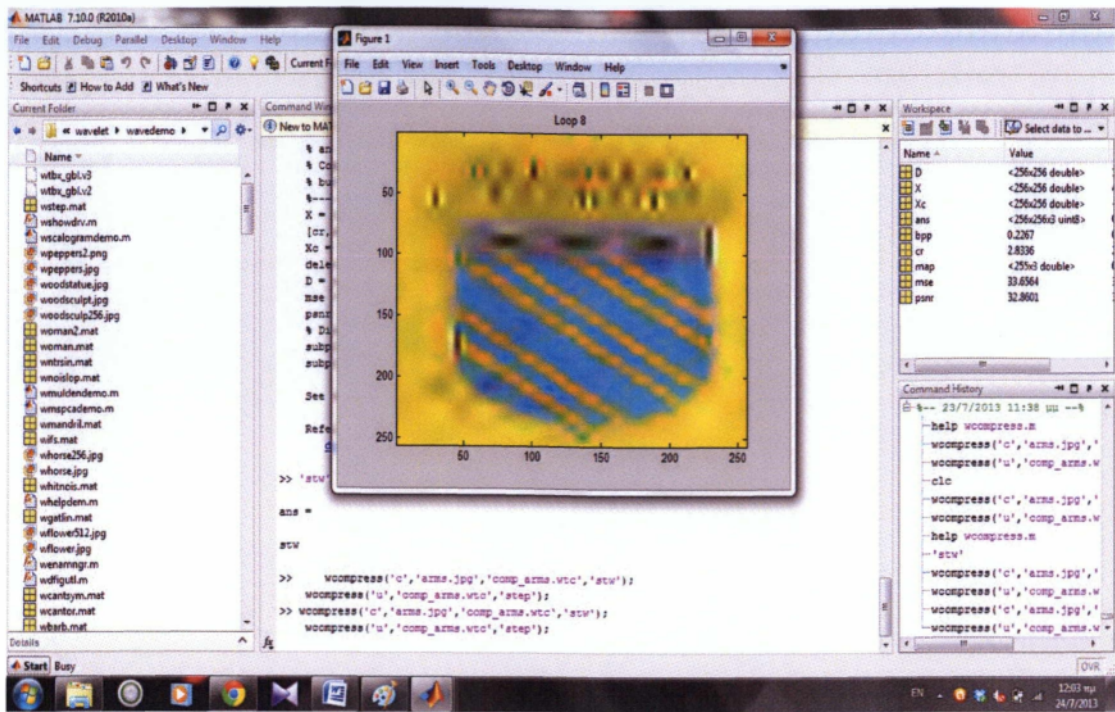
Σε αυτό το παράδειγμα βλέπουμε πώς να συμπίεζουμε μια εικόνα χρησιμοποιώντας τη μέθοδο συμπίεσης 'stw' να το αποθηκεύσουμε σε ένα αρχείο. Μετά βλέπουμε πως γίνεται βήμα βήμα η αποσυμπίεση της εικόνας.



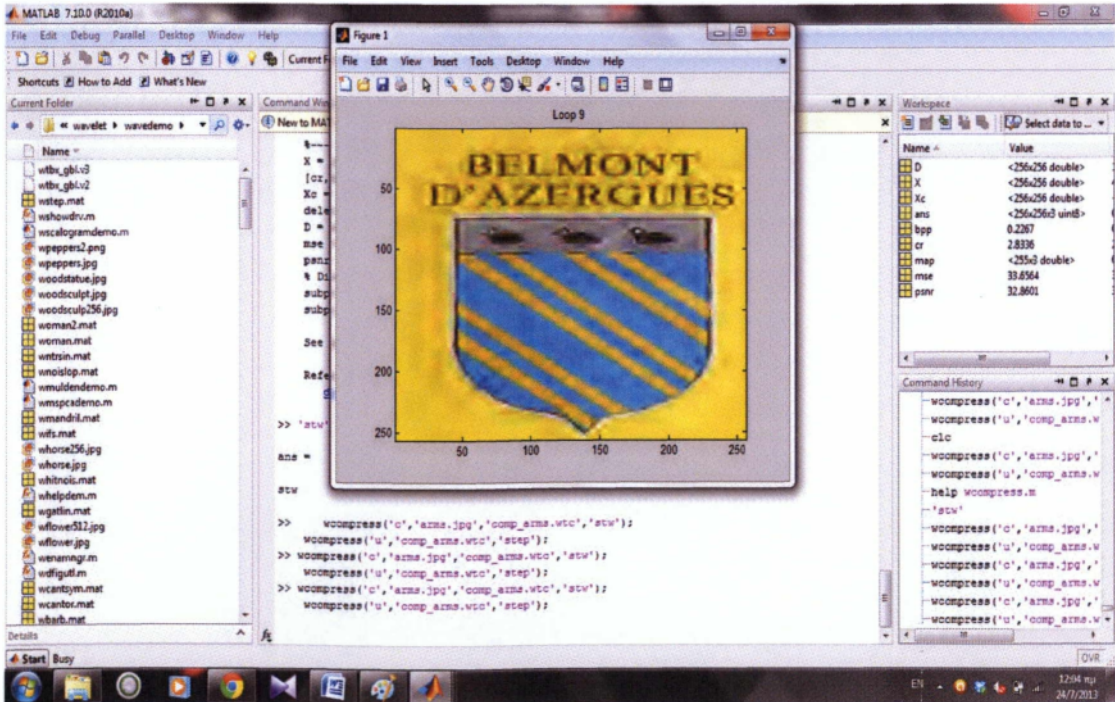
Εικόνα : 24 Matlab example 3



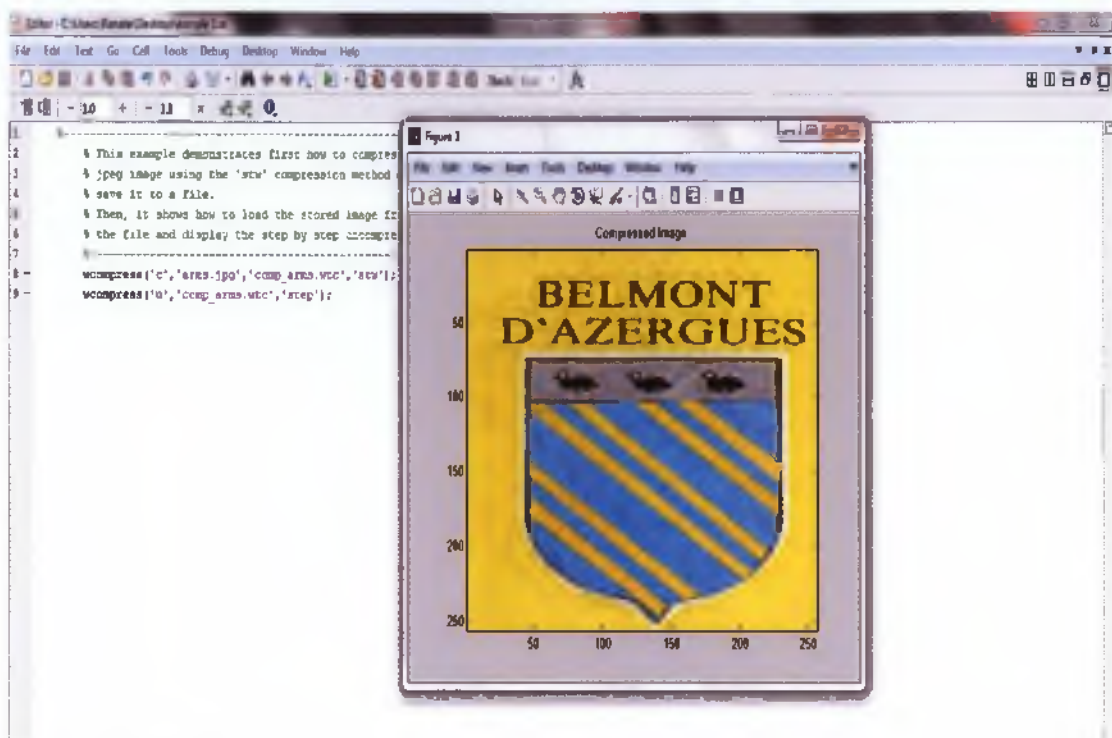
Εικόνα : 25 Matlab arms 1



Εικόνα : 26 Matlab arms 2



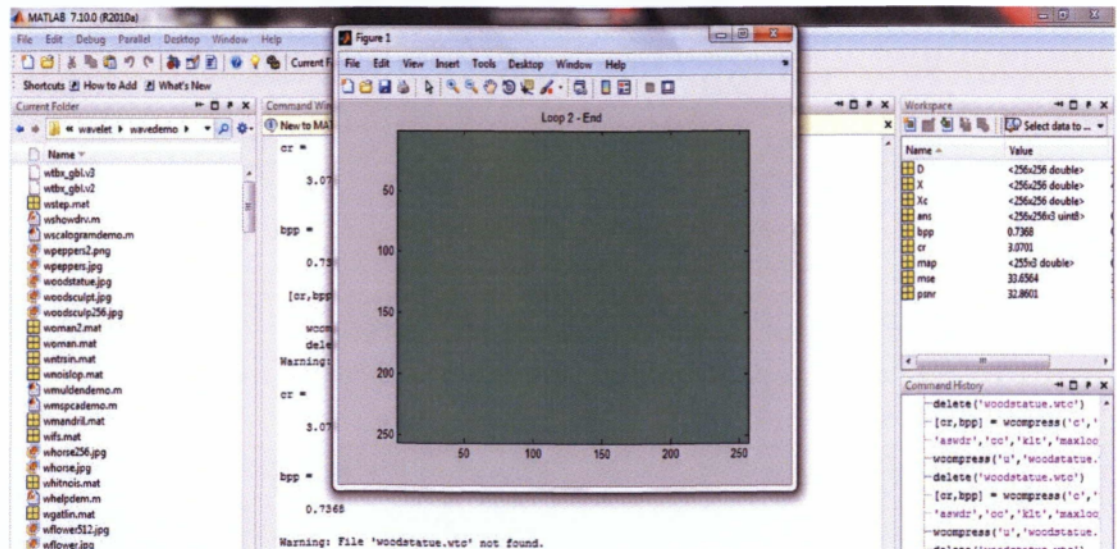
Εικόνα : 27 Matlab arms 3



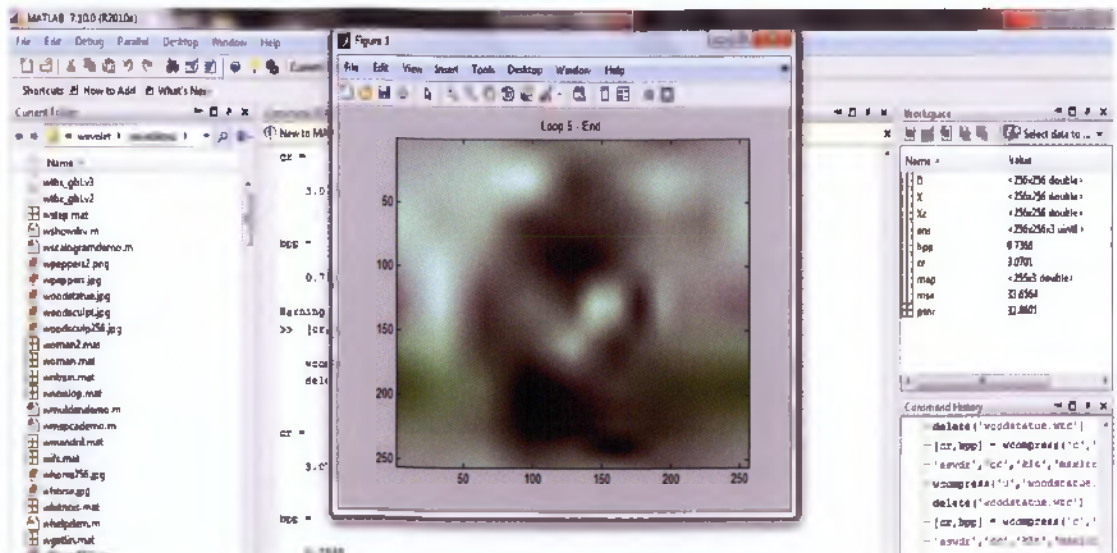
Εικόνα : 28 Matlab arms 4

ΠΑΡΑΔΕΙΓΜΑ 4.

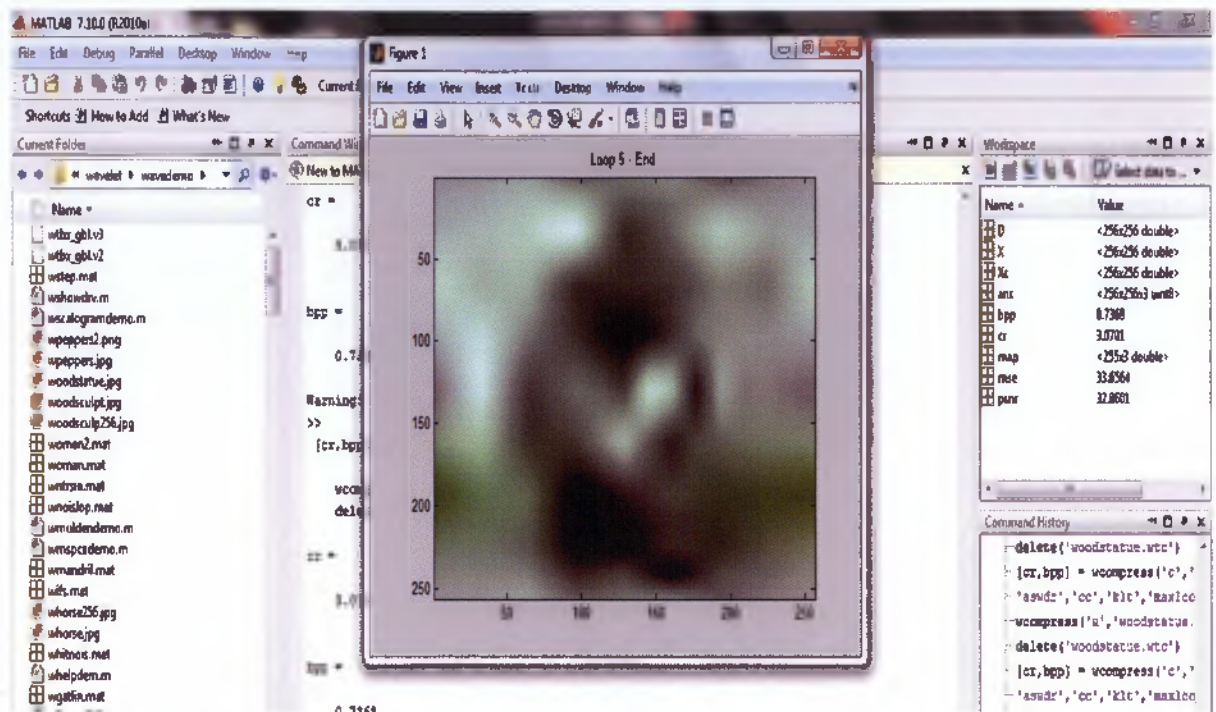
Σε αυτό το παράδειγμα βλέπουμε πώς να συμπίεζουμε μια εικόνα χρησιμοποιώντας τη μέθοδο συμπίεσης 'aswdr' και να το αποθηκεύσουμε σε ένα αρχείο. Κατά τη διάρκεια της συμπίεσης χρησιμοποιούμε τρεις παρεμέτρους - Conversion color (cc) - Maximum number of loops for computing (maxloop)- Plot type (plotpar)



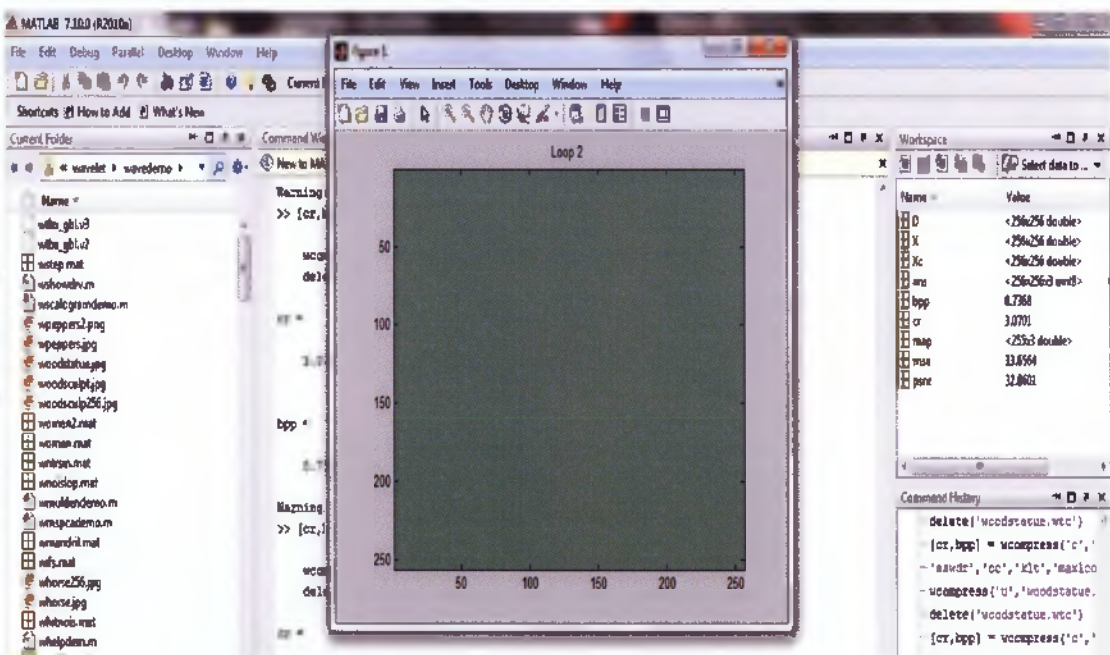
Εικόνα: 29 woodstatue loop 2



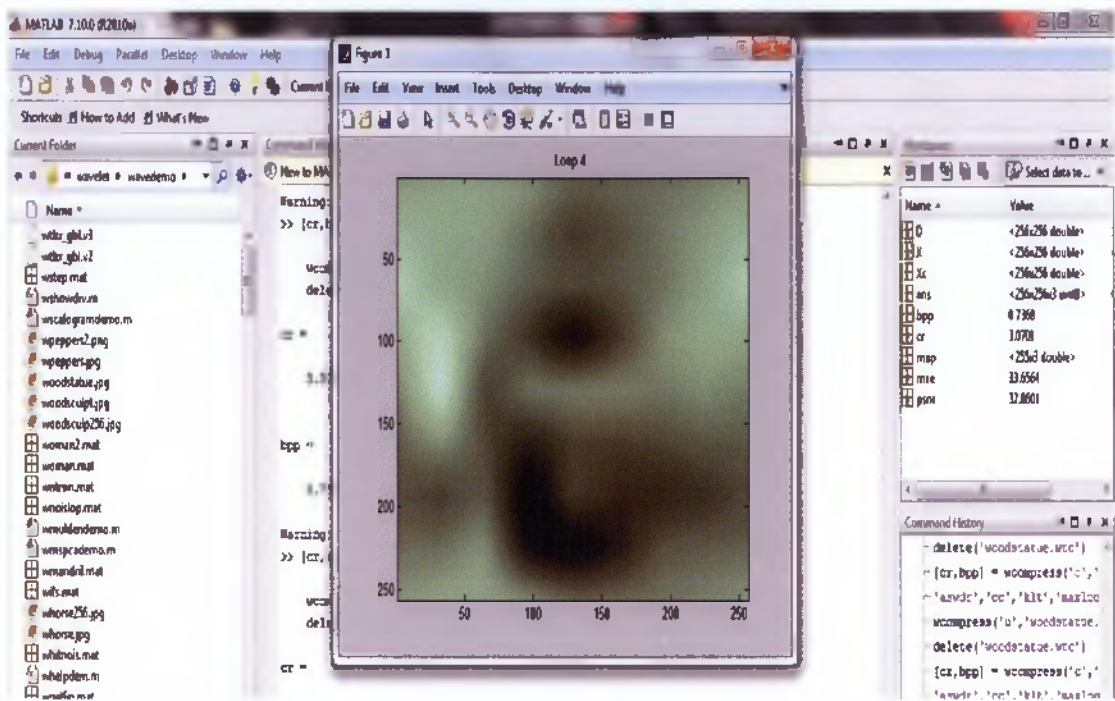
Εικόνα: 30 woodstatue loop 2



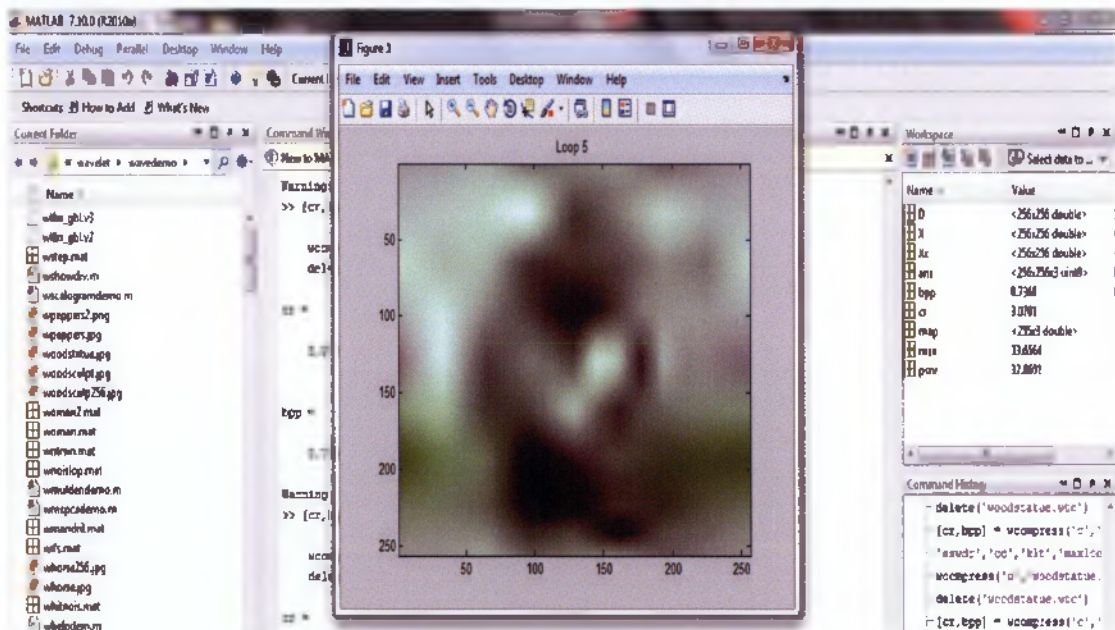
Εικόνα: 31 woodstatue loop 5



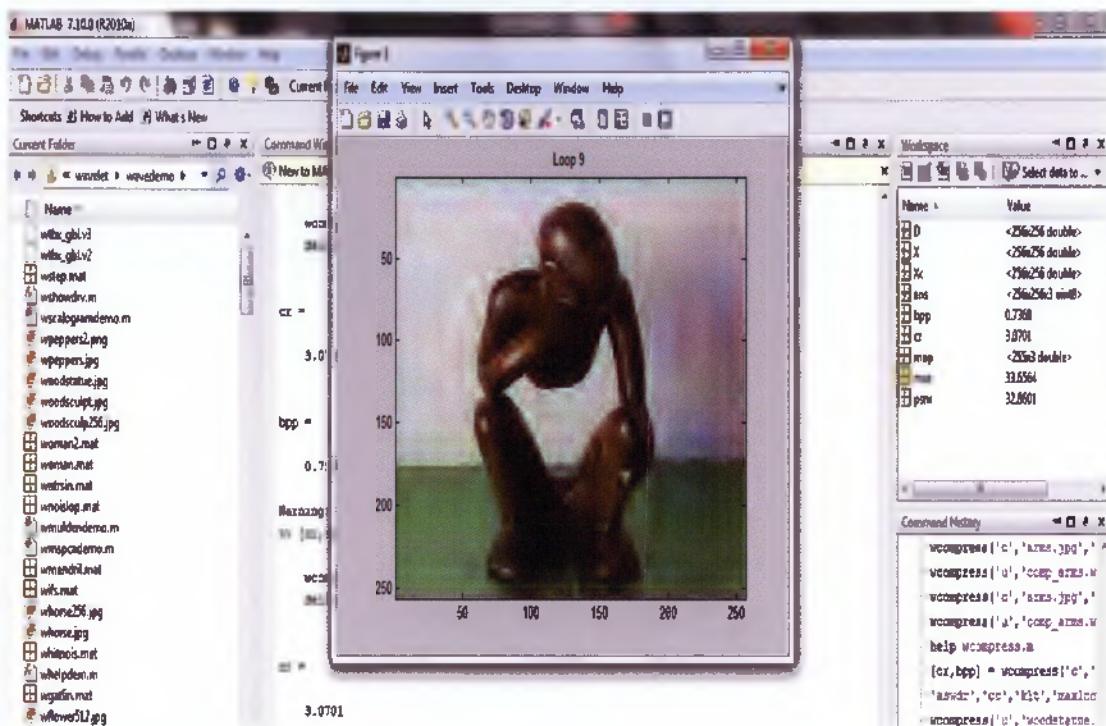
Εικόνα: 32 woodstatue loop 2 new



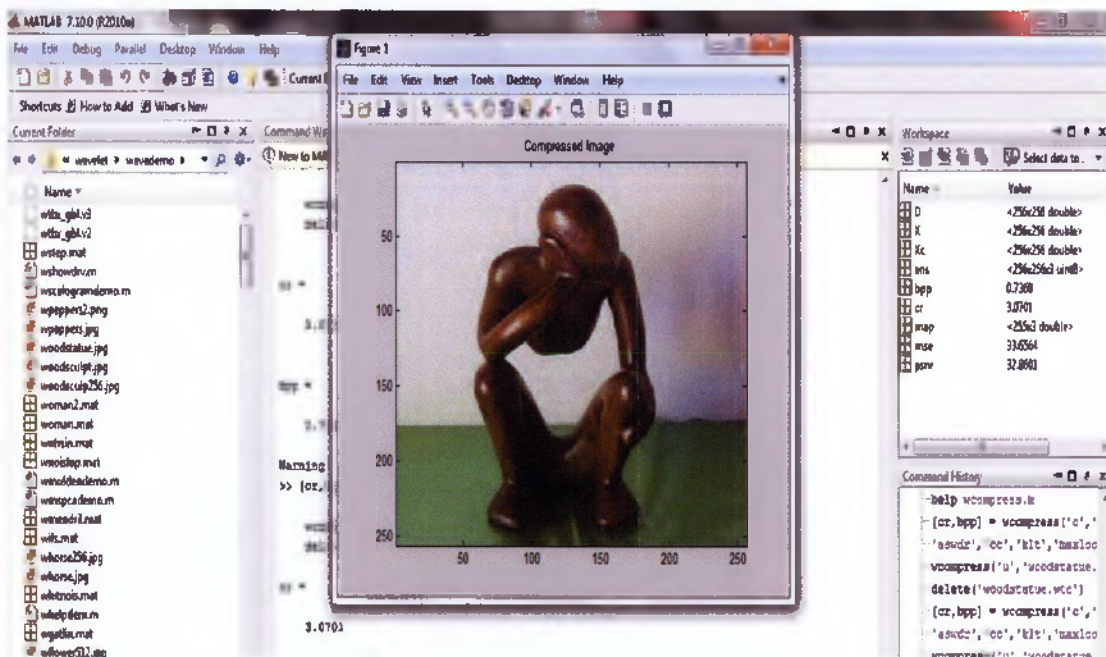
Εικόνα: 33 woodstatue loop 4



Εικόνα: 34 woodstatue loop 5



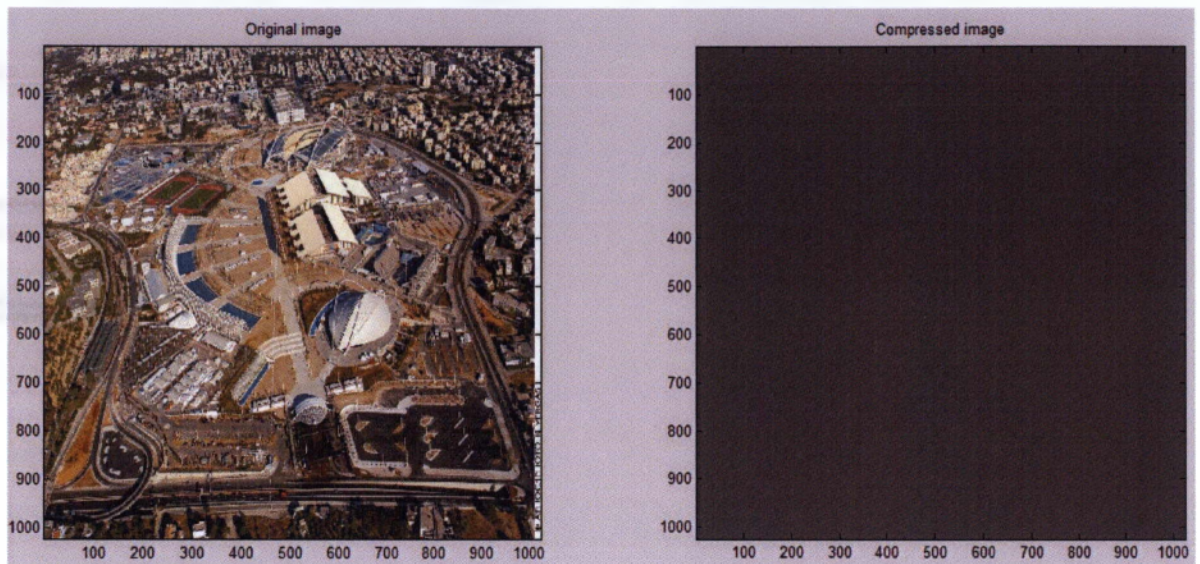
Εικόνα: 35 woodstatue loop 9



Εικόνα: 36 woodstatue compressed image

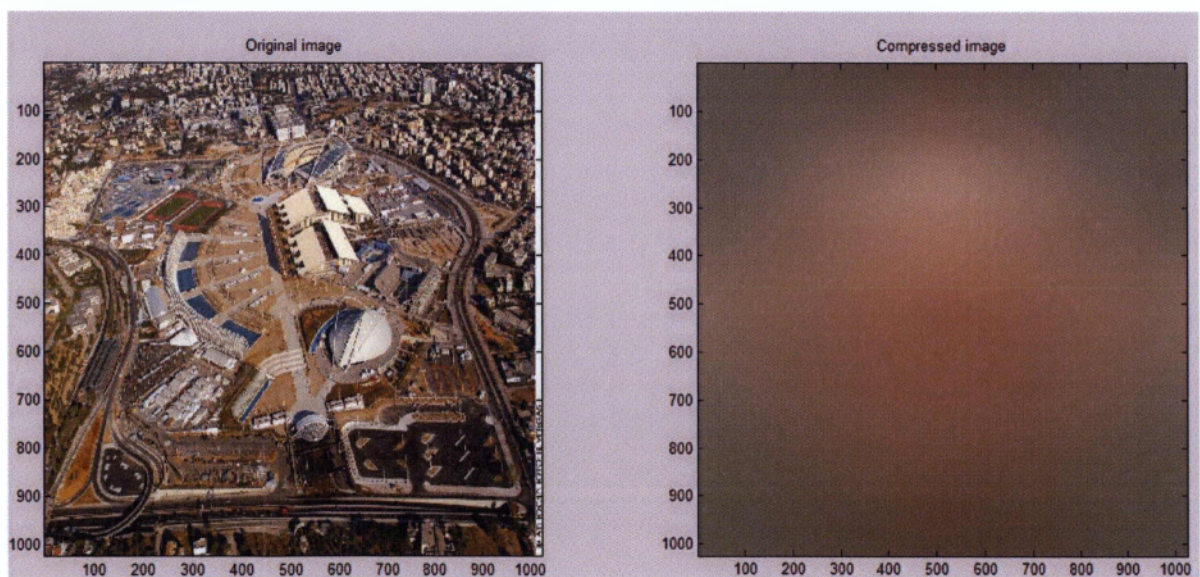
Στη συγκεκριμένη φωτογραφία είναι $mse = 33.6564$ και $psnr = 32.8601$.

Με βήμα 2 το $mse = 6.4849e+003$ και το $psnr = 10.0118$.



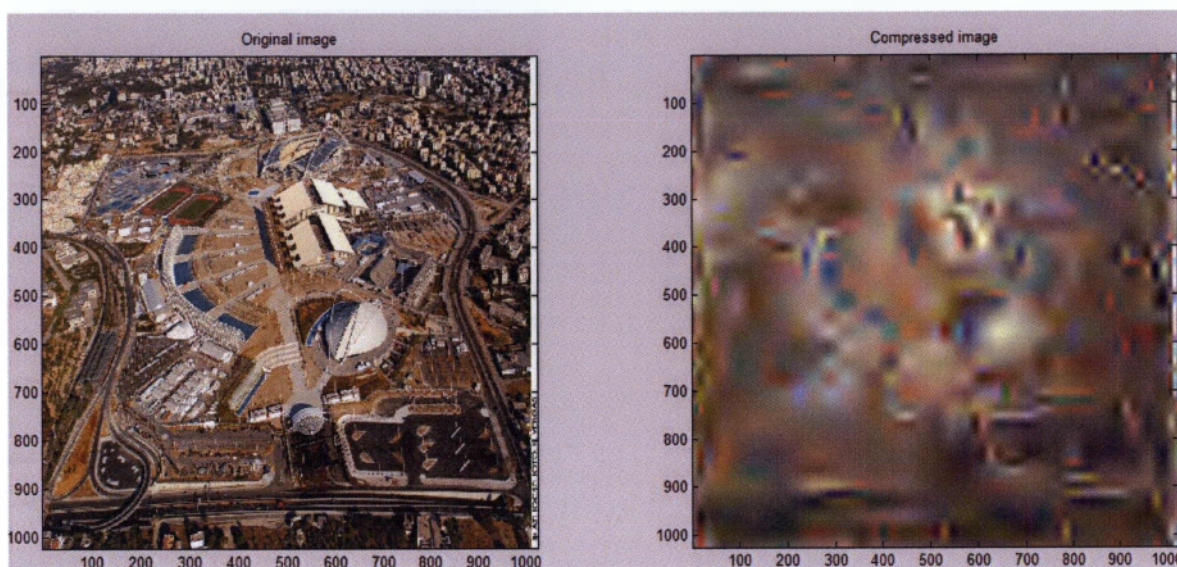
Εικόνα: 37 Ολόκληρο το Ο.Α.Κ.Α. με βήμα 2

Με βήμα 5 το $mse = 3.6742e+003$ και το $psnr = 12.4791$.



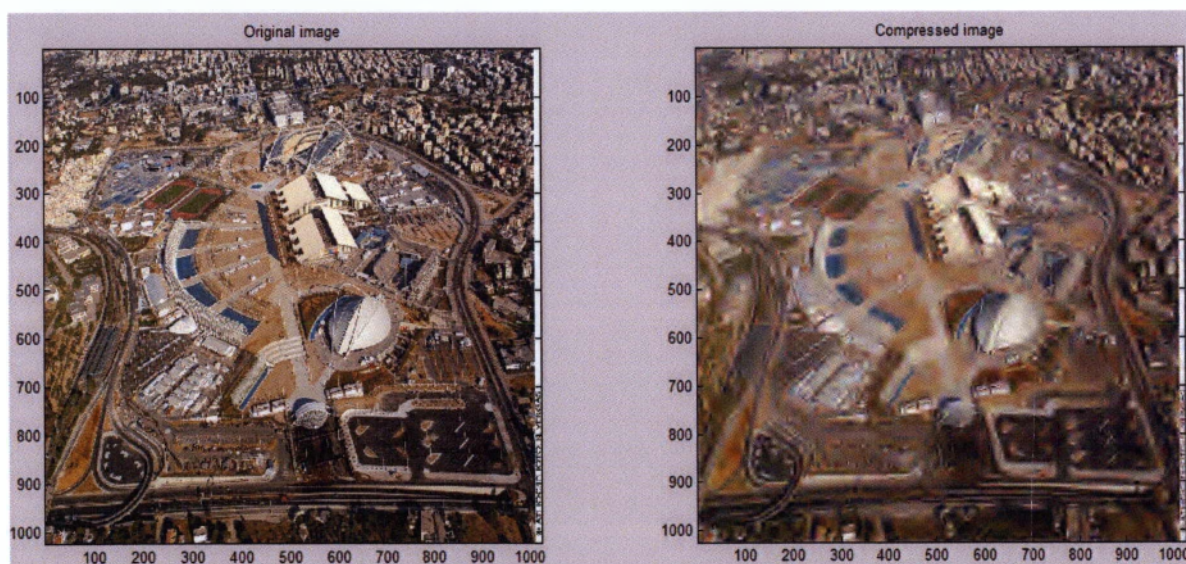
Εικόνα: 38 Ολόκληρο το Ο.Α.Κ.Α. με βήμα 5

Με βήμα 8 το $mse = 2.4157e+003$ και το $psnr = 14.3003$.



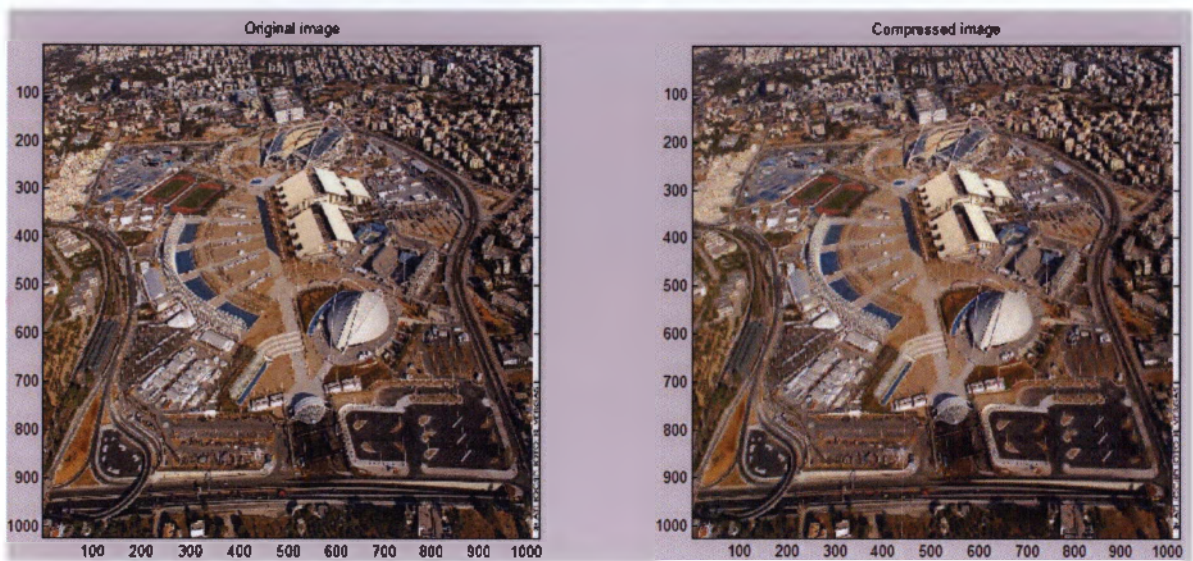
Εικόνα: 39 Ολόκληρο το Ο.Α.Κ.Α. με βήμα 8

Με βήμα 10 το $mse = 1.3419e+003$ και το $psnr = 16.8536$.



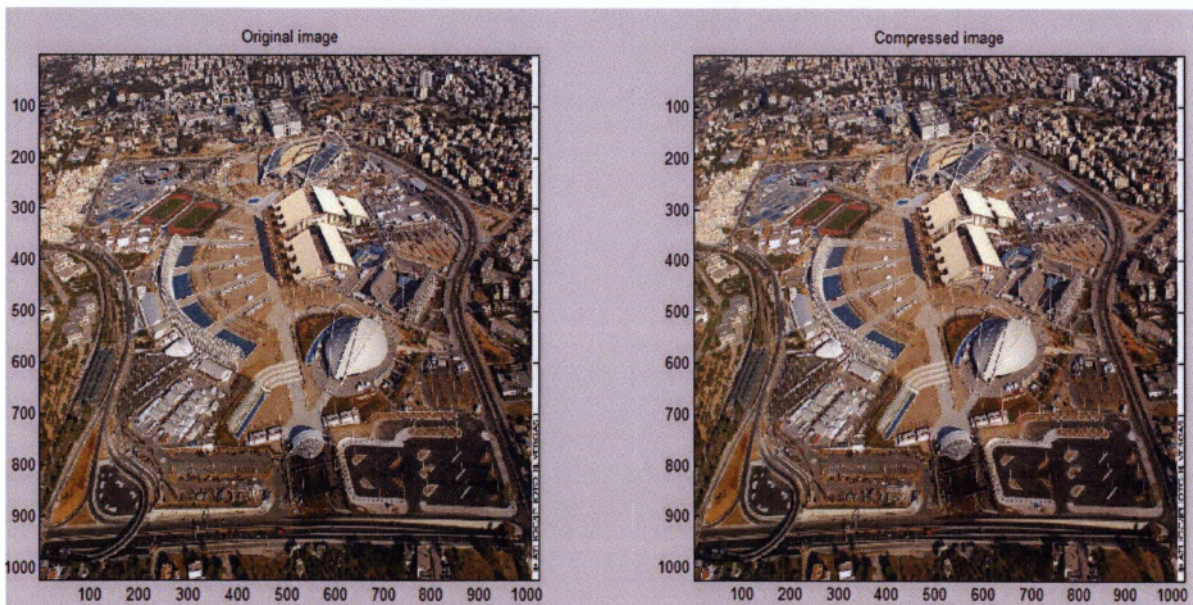
Εικόνα: 40 Ολόκληρο το Ο.Α.Κ.Α. με βήμα 10

Με βήμα 12 το mse =399.4215 και το psnr =22.1165.



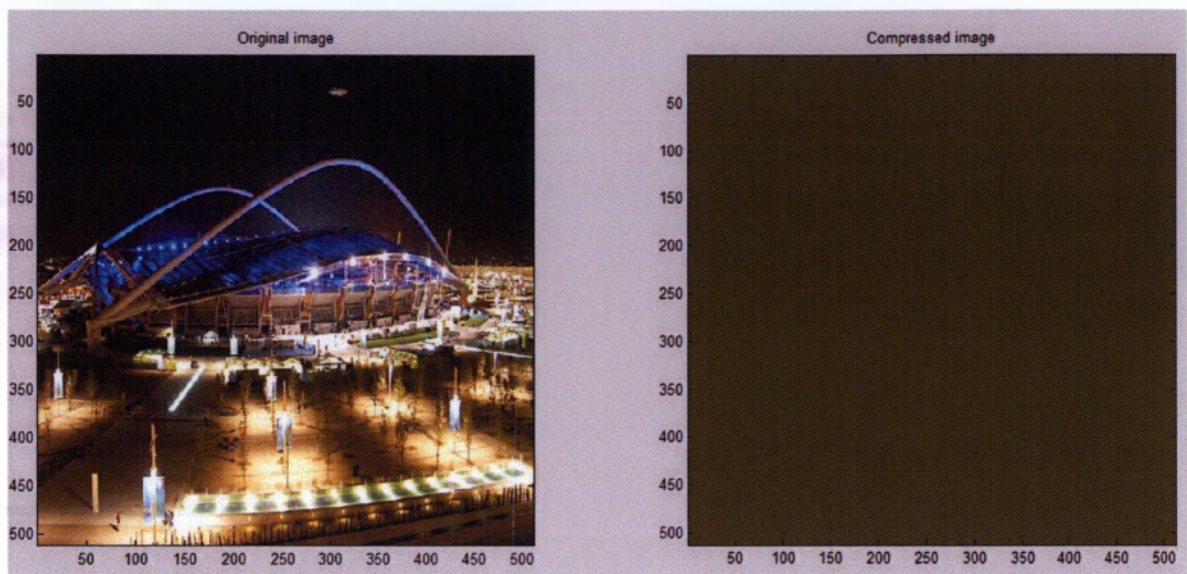
Εικόνα: 41 Ολόκληρο το Ο.Α.Κ.Α. με βήμα 12

Με βήμα 15 το mse =60.1104 και το psnr =30.3413.



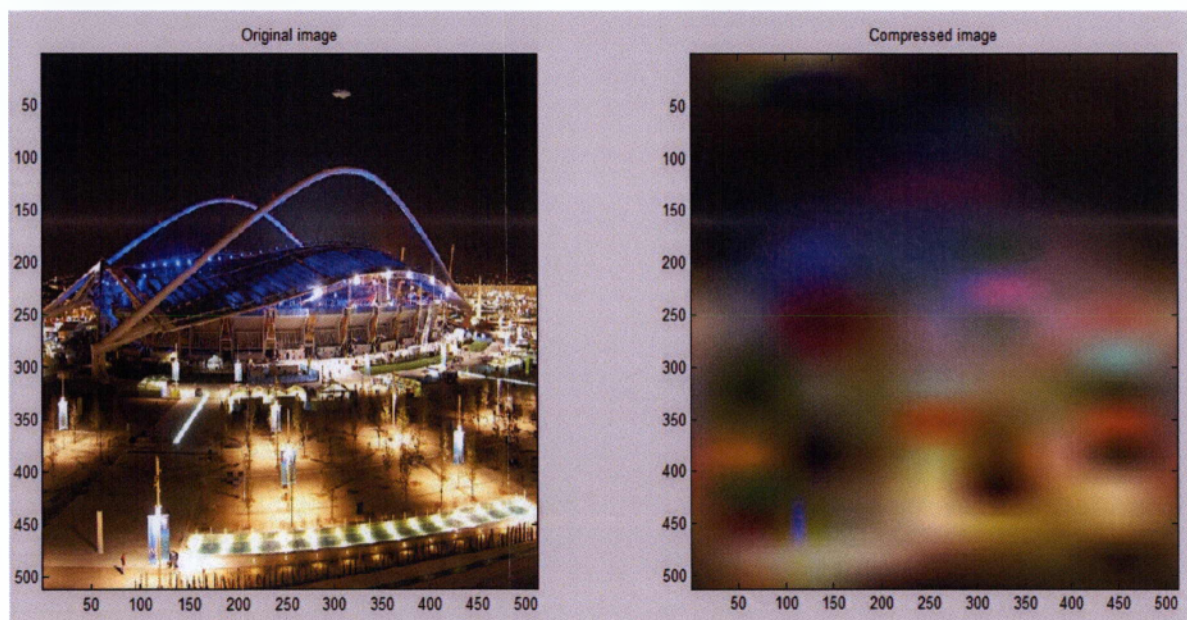
Εικόνα: 42 Ολόκληρο το Ο.Α.Κ.Α. με βήμα 15

Με βήμα 2 το $mse = 6.0430e+003$ και το $psnr = 10.3183$.



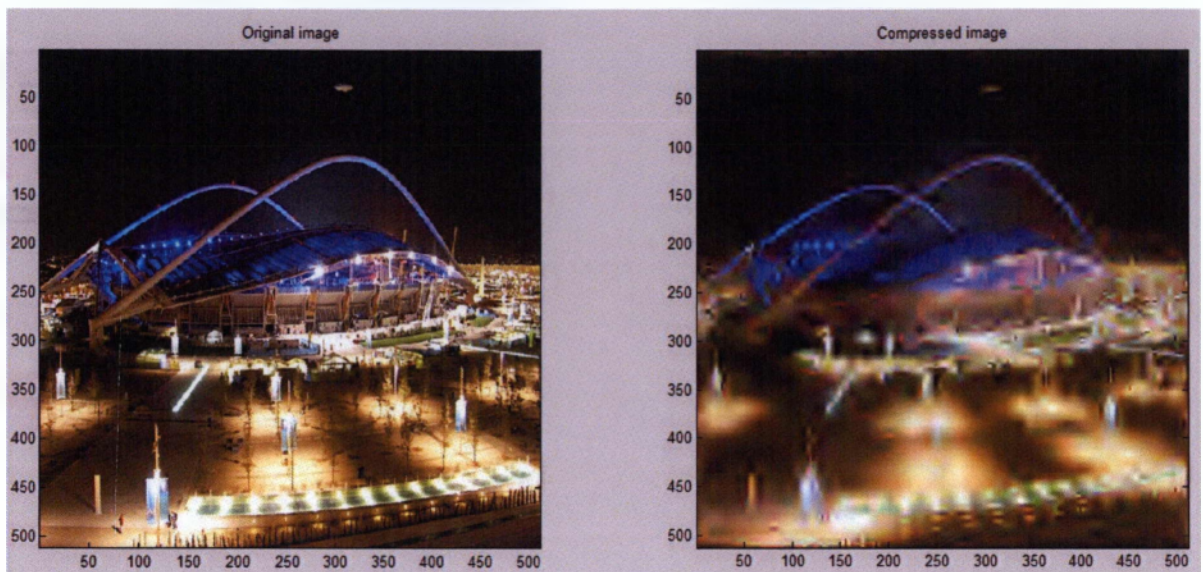
Εικόνα: 43 Ο.Α.Κ.Α. με βήμα 2

Με βήμα 5 το $mse = 2.4269e+003$ και το $psnr = 14.2803$.



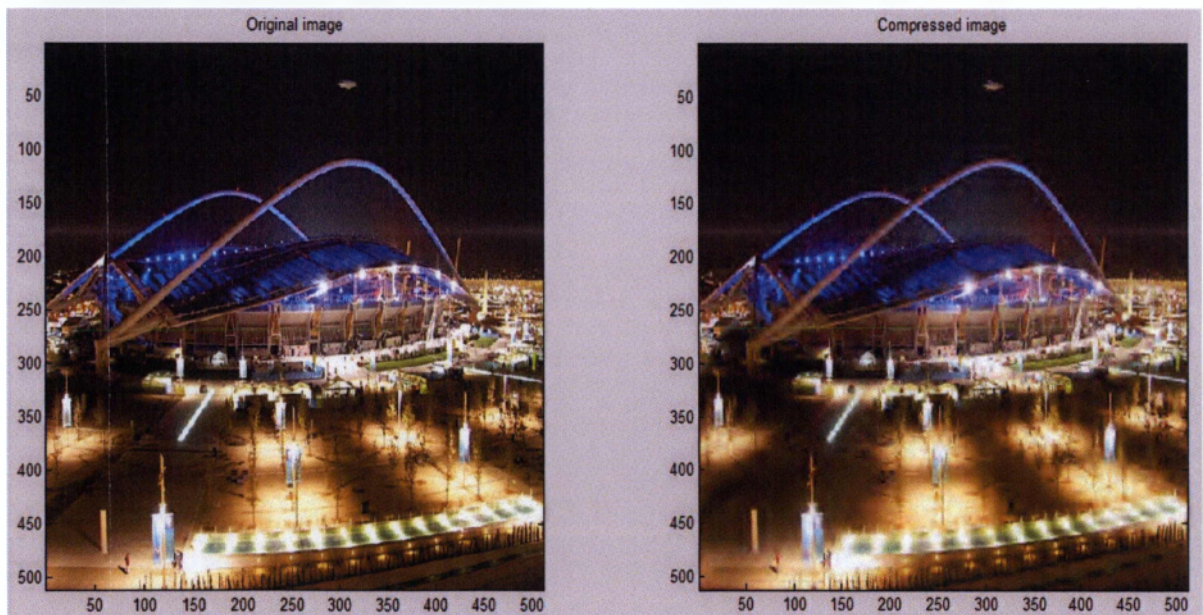
Εικόνα: 44 Ο.Α.Κ.Α. με βήμα 5

Με βήμα 8 το mse =721.1377 και το psnr =19.5506.



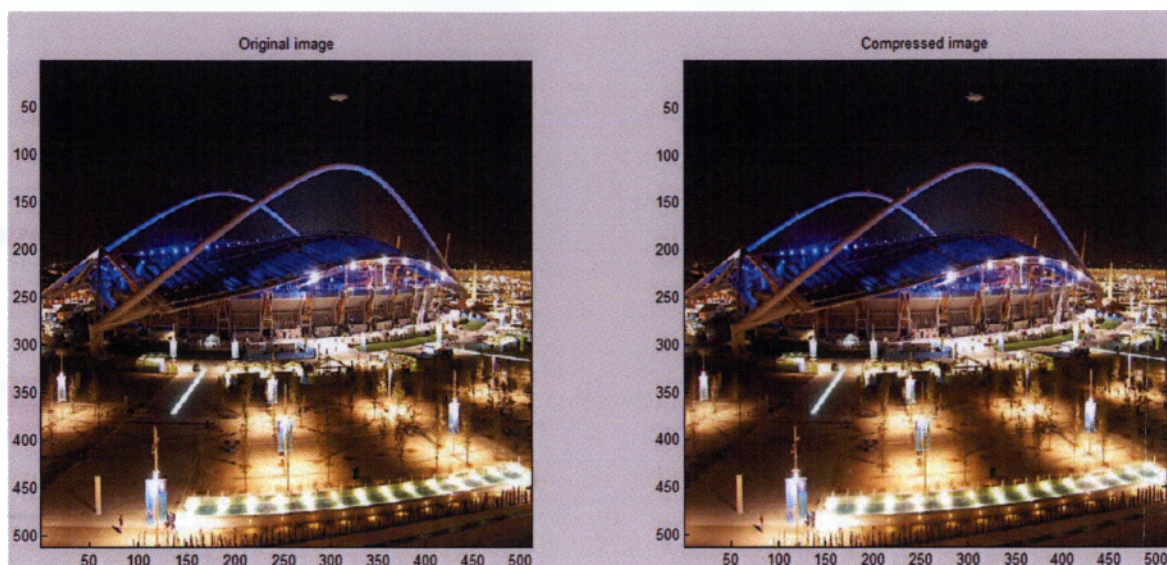
Εικόνα: 45 Ο.Α.Κ.Α. με βήμα 8

Με βήμα 10 το mse =206.2716 και το psnr =24.9864.



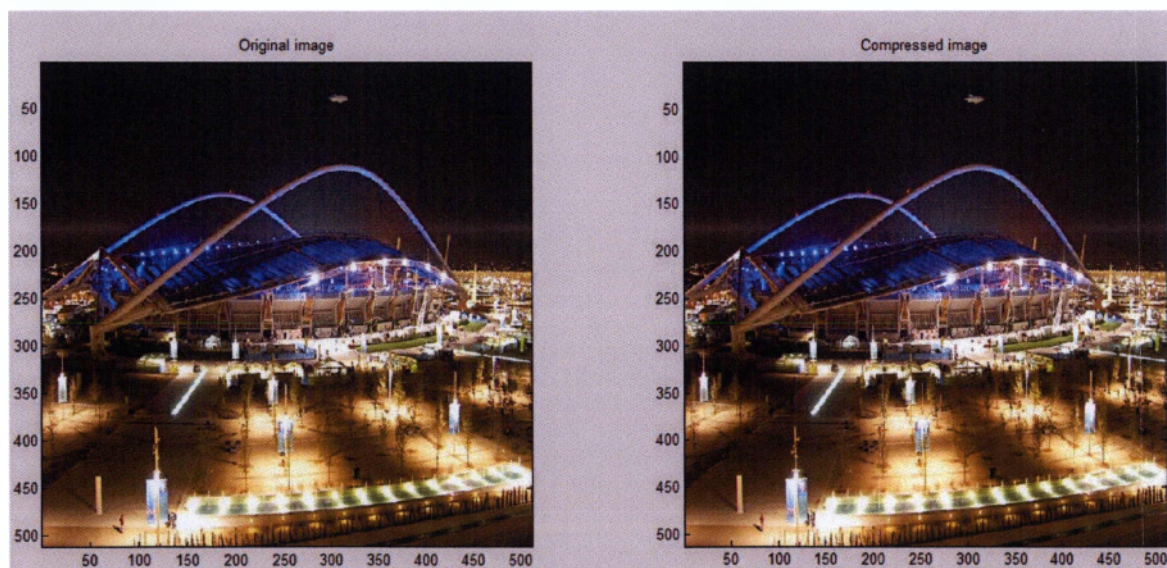
Εικόνα: 46 Ο.Α.Κ.Α. με βήμα 10

Με βήμα 12 το $mse = 33.3542$ και το $psnr = 32.8993$.



Εικόνα: 47 Ο.Α.Κ.Α. με βήμα 12

Με βήμα 15 το $mse = 2.6759$ και το $psnr = 43.8562$.



Εικόνα: 48 Ο.Α.Κ.Α. με βήμα 15

ΠΑΡΑΡΤΗΜΑ

Example: 1.

```
%-----  
  
% Compression and uncompression of a truecolor image  
% and computed MSE and PSNR error values.  
% Compression parameters are the same as those used  
for example 3,  
% but using the 'spiht_3d' method give better  
performance yet.  
  
%-----  
  
X = imread('wpeppers.jpg');  
[cr,bpp] =  
wcompress('c',X,'wpeppers.wtc','spiht','maxloop',12)  
Xc = wcompress('u','wpeppers.wtc');  
delete('wpeppers.wtc')  
D = abs(double(X)-double(Xc)).^2;  
mse = sum(D(:))/numel(X)  
psnr = 10*log10(255*255/mse)  
% Display the original and the compressed image  
subplot(1,2,1); image(X); title('Original image');  
axis square  
subplot(1,2,2); image(Xc); title('Compressed image');  
axis square
```


Example 2.

```
% Compression and uncompression of a grayscale image
% and computed MSE and PSNR error values.
%
% Two measures are commonly used to quantify the error between
% two images: the Mean Square Error(MSE) and the Peak Signal
% to Noise Ratio (PSNR) which is expressed in decibels.
%
% This example demonstrates first how to compress the mask image
% using the 'spiht' compression method and save it to the file
% 'mask.wtc'.
% Then, it shows how to load the stored image from the file
% 'mask.wtc', uncompress it and delete the file 'mask.wtc'.
```

```
load mask;
[cr,bpp] = wcompress('c',X,'mask.wtc','spiht','maxloop',12)
Xc = wcompress('u','mask.wtc');
delete('mask.wtc')
D = abs(X-Xc).^2;
mse = sum(D(:))/numel(X)
psnr = 10*log10(255*255/mse)
% Display the original and the compressed image
colormap(pink(255))
subplot(1,2,1); image(X); title('Original image'); axis square
subplot(1,2,2); image(Xc); title('Compressed image'); axis square
```

Example 3.

```
% This example demonstrates first how to compress a
% jpeg image using the 'stw' compression method and
% save it to a file.
% Then, it shows how to load the stored image from
% the file and display the step by step
uncompression.
```

```
wcompress('c','arms.jpg','comp_arms.wtc','stw');
wcompress('u','comp_arms.wtc','step');
```

Example 4.

```
%-----  
  
% This example demonstrates first how to compress a  
% jpeg image using the 'aswdr' compression method and  
% save it to a file.  
% During the compression process 3 parameters are  
used:  
% - Conversion color (cc) set to Karhunen-Loeve  
transform 'klt'  
% - Maximum number of loops for computing (maxloop)  
set to 11  
% - Plot type (plotpar) set to step by step display  
% Then, it shows how to load the stored image from  
% the file and display the step by step  
uncompression.
```

```
%-----  
  
[cr,bpp] =  
wcompress('c','woodstatue.jpg','woodstatue.wtc', ...  
  
'aswdr','cc','klt','maxloop',11,'plotpar','step')  
wcompress('u','woodstatue.wtc','step');  
delete('woodstatue.wtc')
```


Τρόπος Υπολογισμού:

Στα παραδείγματα υπολογίζεται το μέσο τετραγωνικό λάθος (MSE) και ο μέγιστος σηματοθορυβικός λόγος (PSNR).

$$1) D = \text{abs}(\text{double}(X) - \text{double}(X_c)).^2$$

Οι εικόνες X και X_c αναπαρίστανται ως δυσδιάστατοι πίνακες. Η συνάρτηση $\text{double}(X)$ μετατρέπει όλα τα στοιχεία του δυσδιάστατου πίνακα X σε πραγματικούς διπλής ακρίβειας. Ο δυσδιάστατος πίνακας $\text{abs}(\text{double}(X) - \text{double}(X_c))$ είναι το απόλυτο της διαφοράς των δύο πινάκων/ εικόνων. Η πράξη $\text{abs}(\text{double}(X) - \text{double}(X_c))^2$ σημαίνει ύψωση στο τετράγωνο όλων των στοιχείων του πίνακα $\text{abs}(\text{double}(X) - \text{double}(X_c))$. Ο δυσδιάστατος πίνακας D επομένως αναπαριστά το τετραγωνικό λάθος μεταξύ των δύο εικόνων για κάθε pixel των εικόνων.

$$2) \text{mse} = \text{sum}(D(:))/\text{numel}(X)$$

Με την εντολή $\text{sum}(D(:))$ αθροίζονται όλα τα στοιχεία του δυσδιάστατου πίνακα D , επομένως το $\text{sum}(D(:))$ είναι ένας αριθμός. Η εντολή $\text{numel}(X)$ επιστρέφει τον αριθμό των στοιχείων του X , επομένως το $\text{sum}(D(:))/\text{numel}(X)$ είναι το άθροισμα των τετραγωνικών λαθών μεταξύ των δύο εικόνων για όλα τα στοιχεία τους (pixels) δια τον αριθμό των στοιχείων (pixels), δηλαδή είναι το μέσο τετραγωνικό λάθος της συμπίεσης- αποσυμπίεσης.

$$3) \text{psnr} = 10 * \log_{10}(255 * 255 / \text{mse})$$

Το psnr είναι ο μέγιστος σηματοθορυβικός λόγος της συμπίεσης/αποσυμπίεσης. Εδώ ως θόρυβο εννοούμε το μέσο τετραγωνικό λάθος.

ΒΙΒΛΙΟΓΡΑΦΙΑ

1. Mathworks, <http://www.mathworks.com/help/wavelet/ug/two-dimensional-true-compression.html>
2. Math.bgsu <http://www-math.bgsu.edu/z/wavelets/>
3. J. Proakis and D. Manolakis, *Digital Signal Processing*, 3rd Edition. Prentice Hall, Upper Saddle River NJ, 1996
4. S. Burrus, A. Gopinath and H. Guo, *Introduction to Wavelets and Wavelet Transforms: A Primer*. Prentice Hall, Upper Saddle River NJ, 1998
5. G. Strang and T. Nguyen, *Wavelets and Filter Banks*. Wellesley-Cambridge, Wellesley MA, 1996
6. S. Mallat, *A Wavelet Tour of Signal Processing*. Academic, San Diego CA. 1998
7. M. Vitterli and J. Kovacevic, *Wavelets and Subband Coding*. Prentice Hall, Upper Saddle River NJ, 1995
8. M. Frazier, *An Introduction to Wavelets Through Linear Algebra*. Springer-Verlag, New York NY, 1999
9. I. Daubechies, *Ten Lectures on Wavelets*. SIAM, Philadelphia PA, 1992
10. H.G. Stark, *Wavelets and Signal Processing*. Springer Verlag, Berling Heidelberg, 2005
11. M. Misiti, Y. Misiti, G. Oppenheim and J-M Poggi *Wavelet Toolbox for use with MATLAB[®]*
12. E. Stollnitz, T. DeRose and D. Salesin, *Wavelet Computer Graphics A Primer*

13. Stan McClellan and Jerry D. Gibson, "Speech Signal Processing: Coding, Transmission and Storage"
14. Γουμενίδης Θεόδωρος, "Κωδικοποίηση Φωνής – Κωδικοποιητές Κυματομορφής",
Φεβρουάριος 2004.
15. Jeremy Bradbury, "Linear Predictive Coding", December 2000.
16. NTT DoCoMo, "Proposed Contents of Operational Handbook", Asia – Pacific Telecommunity (The 3rd Meeting of the APT IMT-2000 Forum), 2-3 September 2002, Busan, Republic of Korea.
17. Jason Woodard, "Speech Coding", <http://www.ecs.soton.ac.uk/~jpw/index.html>
18. Niranjana Dhanakoti, "Speech Signal Processing", project report 2002.
19. Bryan Douglas, "Voice Encoding Methods for Digital Wireless Communications Systems", Fall 1997.
20. Eddie L. T. Choy, "Waveform Interpolation Speech Coder at 4 kb/s", August 1998
21. Susanna Varho, "New Linear Predictive Methods for Digital Speech Processing", 2001
22. Nadim Batri, "Robust Spectral Parameter Coding in Speech Processing", May 1998
23. Alan McCree and Jan Carlos De Martin, "A 1.7 Kb/s Melp Coder with Improved Analysis and Quantization", DSPS R&D. Texas Instruments, Dallas, Texas
24. Wesley Pereira, "Modifying LPC Parameter Dynamics to Improve Speech Coder Efficiency", September 2001

25. Barnewell Thomas P III., Kambiz NAYebi, and Craig H. Richardon, "Speech Coding: A Computer Laboratory Textbook", *John Wiley & Sons, Inc.*, 1996.
26. Hwai-Tsu Hu and Hsi-Tsung Wu, "A Glottal-Excited Linear Prediction (GELP) Model for Low-Bit-Rate Speech Coding", may 1999
27. <http://www.owl.net.rice.edu/~elec532/PROJECTS00/speechBIG/> Notorious LPC
28. Andreas Spanias, "Speech Coding: A Tutorial Review".
29. Andreas Spanias, "Multimedia Signal Processing Lecture Notes"
30. Schussler Mare, "Design and Simulation of a Speech Coder for Mobile Communication Systems", Master's Thesis, 1994
31. Antti Kiviluoto, "Speech Coding Standards"
32. B. S. Atal and Suzanne L. Hanauer, "Speech Analysis and Synthesis by Linear Prediction of the Speech Wave", Bell Telephone Laboratories