Μέθοδοι Υπερανάλυσης Εικόνα

Μιχαήλ Βρίγκας

ΜΕΤΑΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ ΕΞΕΙΔΙΚΕΥΣΙ

— ♦ —

Ιωάννινα, Οκτώβριος 2010

ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ
ΠΑΝΕΠΙΣΤΗΜΙΟ ΙΩΑΝΝΙΝΩΝ

DEPARTMENT OF COMPUTER SCIENCE
UNIVERSITY OF IOANNINA

Μέθοδοι Υπερανάλυσης Εικόνας

Η ΜΕΤΑΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ ΕΞΕΙΔΙΚΕΥΣΗΣ

υποβάλλεται στην

ορισθείσα από την Γενική Συνέλευση Ειδικής Σύνθεσης

του Τμήματος Πληροφορικής Εξεταστική Επιτροπή

από τον

Μιχαήλ Βρίγκα

ως μέρος των Υποχρεώσεων για τη λήψη του

ΜΕΤΑΠΤΥΧΙΑΚΟΥ ΔΙΠΛΩΜΑΤΟΣ ΣΤΗΝ ΠΛΗΡΟΦΟΡΙΚΗ

ΜΕ ΕΞΕΙΔΙΚΕΥΣΗ

ΣΤΙΣ ΤΕΧΝΟΛΟΓΙΕΣ – ΕΦΑΡΜΟΓΕΣ

Οκτώβριος 2010

Dedication

Μερικοί αναγνώστες συνηθίζουν να βυθίζουν τη μύτη τους μέσα στις σελίδες ενός βιβλίου και ν' απολαμβάνουν το άρωμα της τυπωμένης μελάνης. Μολονότι αυτό γίνεται συνήθως στα κρυφά, πρέπει να σημειώσουμε ότι αποτελεί μία καθόλα νόμιμη πράξη. Η διατριβή απευθύνεται σε όλες τις αισθήσεις μας, κι αυτό είναι κάτι που το ξέρει καλά ο αναγνώστης.

Σωτήρης Τριβιζάς, *Η τέχνη της ανάγνωσης*.

ΕΥΧΑΡΙΣΤΙΕΣ

Θα ήθελα να ευχαριστήσω τον επιβλέποντα κ. Νίκου Χριστόφορο, Επίκουρο Καθηγητή του τμήματος Πληροφορικής του Πανεπιστημίου Ιωαννίνων, για την πολύ καλή συνεργασία που είχαμε κατά την εκπόνηση της διπλωματικής μου εργασίας. Ουσιαστική ήταν και η βοήθεια του κ. Κόντη Λυσίμαχου - Παύλου, Επίκουρου Καθηγητή του τμήματος Πληροφορικής του Πανεπιστημίου Ιωαννίνων, του οποίου οι συμβουλές ήταν πολύτιμες και τον οποίο ευχαριστώ βαθύτατα. Να ευχαριστήσω επίσης, τον φίλο μου Γερογιάννη Δημήτριο, διδακτορικό φοιτητή του τμήματος Πληροφορικής του Πανεπιστημίου Ιωαννίνων, για την βοήθεια που μου προσέφερε στην ολοκλήρωση της εργασίας μου. Θα ήθελα να ευχαριστήσω τις φίλες και συνεργάτιδες του εργαστηρίου Αγγελική, Κατερίνα και Ευαγγελία, για το ευχάριστο κλίμα συννεργασίας και τα χαρούμενα διαλείμματα. Τέλος, να ευχαριστήσω την αγαπημενη μου Ελένη για την κατανόηση, την υπομονή και την συμπαράσταση που μου προσέφερε κατά την περίοδο συγγραφής της διατριβής αύτης.

ΠΕΡΙΕΧΟΜΕΝΑ

ΕΥΡΕΤΗΡΙΟ ΣΧΗΜΑΤΩΝ

ΕΥΡΕΤΗΡΙΟ ΠΙΝΑΚΩΝ

Ευρετηριο Αλγοριθμων

ΠΕΡΙΛΗΨΗ

Μιχαήλ Βρίγκας του Στεφάνου και της Αγγελικής. MSc, Τμήμα Πληροφορικής, Πανεπιστημίου Ιωαννίνων, Οκτώβριος 2010. Τίτλος Διατριβής: Μέθοδοι Υπερανάλυσης Εικόνας. Επιβλέπων: Νίκου Χριστόφορος.

Το αντικείμενο της παρούσας εργασίας προέρχεται από τον τομέα της υπολογιστικής όρασης και της επεξεργασίας εικόνας και πραγματεύεται το πρόβλημα της υπερανάλυσης εικόνας. Στόχος της υπερανάλυσης εικόνας είναι να επιλύσει το ακόλουθο πρόβλημα: δοθέντος ενός συνόλου εικόνων χαμηλής ανάλυσης, να εκτιμήσει μια εικόνα υψηλότερης ανάλυσης. Αυτό μεταξύ των άλλων, περιλαμβάνει και τον υπολογισμό των άγνωστων παραμέτρων μετασχηματισμού υπέρθεσης μεταξύ των εικόνων. Οι άγνωστες παράμετροι υπέρθεσης ενημερώνονται επαναληπτικά μαζί με την υψηλής ανάλυσης εικόνα μέσω μιας διαδικασίας βελτιστοποίησης.

Αρχικά, μελετώνται διάφορες προσεγγίσεις που έχουν προταθεί μέχρι τώρα για την επίλυση του προβλήματος της υπερανάλυσης εικόνας. Παρουσιάζουμε μία μέθοδο εξαγωγής χαρακτηριστικών από τις χαμηλής ανάλυσης εικόνες, τα οποία μπορούν να χρησιμοποιηθούν για να εκτελέσουμε μία αξιόπιστη αντιστοιχία μεταξύ των εικόνων. Η τεχνική που ακολουθούμε βασίζεται στους περιγραφείς SIFT για να εξάγουμε τα χαρακτηριστικά αυτά σημεία. Τα χαρακτηριστικά είναι αμετάβλητα στο χώρο κλίμακας της εικόνας και την περιστροφή και παρέχουν μία πολύ καλή υπέρθεση των εικόνων χαμηλής ανάλυσης, ακόμα και όταν τα δεδομένα εισόδου περιέχουν προσθετικό θόρυβο.

Η ακρίβεια στην εύρεση των άγνωστων παραμέτρων μετασχηματισμού υπέρθεσης παίζει σημαντικό ρόλο στην διαδικασία της υπερανάλυσης εικόνας. Στο δεύτερο μέρος της διατριβής αυτής, προτείνουμε μια μέθοδο υπερανάλυσης εικόνας όπου ο υπολογισμός των παραμέτρων υπέρθεσης γίνεται αρχικά με χρήση των χαρακτηριστικών που παρουσιάστηκαν στο πρώτο μέρος της εργασίας. Στη συνέχεια, η εκτίμηση βελτιώνεται με την χρήση του κριτηρίου υπέρθεσης της μεγιστοποίησης της αμοιβαίας πληροφορίας των δύο εικόνων. Η αμοιβαία πληροφορία μεγιστοποιείται όταν οι δύο εικόνες έχουν υπερτεθεί σωστά μεταξύ τους. Από τα πειραματικά αποτελέσματα, συμπεραίνουμε ότι η προτεινόμενη μέθοδος επιτυγχάνει μεγαλύτερη ακρίβεια στην υπέρθεση, αλλά και στην ποιότητα της εικόνας υψηλής ανάλυσης σε σχέση με τη μέθοδο που βασίζεται αποκλειστικά στους περιγραφείς SIFT.

Τέλος, παρουσιάζεται μια νέα τεχνική υπερανάλυσης εικόνας με χρήση εύρωστων εκτιμητών. Οι παραδοσιακές μέθοδοι ανακατασκευής στηρίζονται στους L_1 και L_2 εκτιμητές και επομένως είναι αρκετά ευαίσθητες σε κρουστικό θόρυβο. Συνεπώς, είναι ανάγκη να χρησιμοποιήσουμε τεχνικές που εξαλείφουν την επίδραση μετρήσεων που δεν ακολουθούν

το κυρίαρχο μοντέλο. Στην προτεινομένη μέθοδο υπερανάλυσης εικόνας, χρησιμοποιούμε εύρωστους Μ-εκτιμητές για τον υπολογισμό του σφάλματος μεταξύ της εκτίμησης της υψηλής ανάλυσης εικόνας και κάθε χαμηλής ανάλυσης πλαισίου. Επιπλέον, η κύρια συνεισφορά της μεθόδου είναι ο εύρωστος υπολογισμός των παραμέτρων εξομάλυνσης και του βέλτιστου βήματος σύγκλισης του αλγορίθμου από τα δεδομένα. Τα πειραματικά αποτελέσματα επιβεβαιώνουν την αποτελεσματικότητα της μεθόδου, καταδεικνύοντας την ανωτερότητά της έναντι των κλασικών και εύρωστων μεθόδων υπερανάλυσης εικόνας.

Extended Abstract in English

Michail, Vrigkas, Stefanos. MSc, Computer Science Department, University of Ioannina, Greece. October, 2010. Thesis Title: Image Super-Resolution Methods. Thesis Supervisor: Nikou Christophoros.

The research topic of this dissertation comes from the field of computer vision and image processing and addresses the problem of image super-resolution. We use the term super-resolution to describe the process of obtaining a high-resolution image from a set of shifted, rotated, and degraded by noise low-resolution images. This procedure also involves the estimation of the registration parameters between the images. In the method presented here, the registration parameters between the low-resolution images are iteratively updated along with the high-resolution image in a iterative coordinate-descent optimization procedure. At first, we review several approaches that have been proposed to solve the super-resolution problem. We describe a method for extracting distinctive invariant features from low-resolution images that can be used to perform a reliable matching between the low-resolution images in the least squares sense. We follow a technique based on the SIFT descriptors to extract those features. These features are invariant to image scale and rotation, and provide a very good registration between the low-resolution frames, even when the input data suffers from change in illumination or additive noise. The accuracy of image registration plays a crucial role in super-resolution reconstruction process. In the second part of this dissertation, we propose a method of image super-resolution, where the computation of the registration parameters is initially performed as described in the first part of this work. Then, the estimation of the registration parameters is fine tuned by the maximization of the mutual information registration criterion. The basic idea is that the mutual information is maximized when the two images are correctly registered. The experimental results demonstrate that the proposed method yields sub-pixel registration accuracy and better quality of the reconstructed high-resolution image. Finally, we present our contribution to the robust super-resolution problem. The majority of image super-resolution algorithms in the literature are based on L_1 and L_2 error norm and therefore they are very sensitive to impulse noise or outliers. Our method uses robust M-estimators for computing the difference between the high-resolution estimate and each low-resolution frame. The main contribution with respect to other robust algorithms in super-resolution is that not only the high resolution image is computed by using a robust estimator but also the regularization parameters and the optimal step of the reconstruction method. The experimental results confirm the effectiveness of our method by suppressing the outliers and demonstrate the superior performance over other robust image super-resolution algorithms.

Κεφαλαιο 1

Εισαγωγη

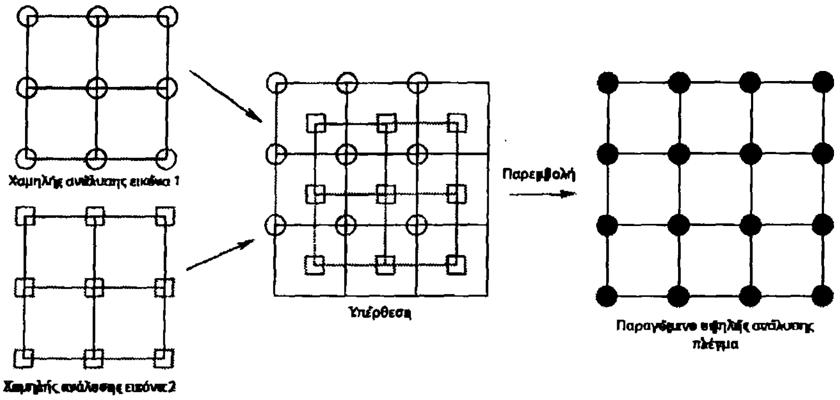
1.1 Εισαγωγή

Οι μέθοδοι υπερανάλυσης εικόνας χρησιμοποιούνται στην ανακατασκευή μιας εικόνας υψηλής ανάλυσης, από αρκετές κατάλληλα μετατοπισμένες εικόνες χαμηλής ανάλυσης. Ο σκοπός είναι να βελτιώσουμε την χωρική ανάλυση των εικόνων. Το πρόβλημα της ανακατασκευής της υψηλής ανάλυσης εικόνας προκαλεί αρκετό ενδιαφέρον, μιας και δεν είναι καλά ορισμένο λόγω της ύπαρξης πρόσθετου θορύβου.

Είναι δύσκολο να ορίσουμε αχριβώς το τι εννοούμε με τον όρο υψηλή ανάλυση στις μεθόδους υπερανάλυσης εικόνας, μιας και τα εικονοστοιχεία στην ανακατασκευασμένη εικόνα μπορούν να είναι αυθαίρετα μικρά σε σχέση με τις εικόνες χαμηλής ανάλυσης, αλλά ακόμα η εικόνα μας δεν περιέχει επιπλέον πληροφορία ή λεπτομέριες συγκρινόμενη με τις χαμηλής ανάλυσης εικόνες. Οι μέθοδοι υπερανάλυσης μπορούν να παράγουν αυτήν την επιπλέον πληροφορία, αλλά μόνο αν υπάρχει επικάλυψη στις εικόνες χαμηλής ανάλυσης. Διαφορετικά, δεν μπορούμε να αποκομίσουμε καμία χρήσιμη πληροφορία.

Το πεδίο της υπερανάλυσης έχει πρόσφατα μια εκρηκτική ανάπτυξη στην ερευνητική δραστηριότητα και κυρίως, στις εφαρμογές που έχουν να κάνουν με την ανακατασκευή μιας εικόνας υψηλής ανάλυσης. Τεχνικές σαν και αυτή χρησιμοποιούνται σε διάφορους τομείς, έτσι ώστε να λάβουμε καλύτερης ποιότητας εικόνες από βιντεοσειρές, τηλεοράσεις υψηλής ευχρίνειας ή ακόμη και ιατρικές εικόνες. Για παράδειγμα, οι μέθοδοι υπερανάλυσης εικόνας βρίσκουν εφαρμογή σε αστρονομικές και ιατρικές εικόνες, στρατιωτικές εφαρμογές, μετατροπή σήματος για τηλεοράσεις τυπικής ευκρίνειας (SDTV) σε τηλεοράσεις υψηλής ευκρίνειας (HDTV), αναγνώριση προσώπων και πολλές άλλες εφαρμογές. Η ανακατασκευή μιας εικόνας υψηλής ανάλυσης από μια και μόνο εικόνα χαμηλής ανάλυσης έχει μελετηθεί στην βιβλιογραφία [37], η μέθοδος αυτή ονομάζεται *quasi-super-resolution*. Μέθοδοι σαν και αυτή γενικά, δεν μας δίνουν καλής ποιότητας ανακατασκευασμένες εικόνες όπως αυτές που λαμβάνουμε όταν χρησιμοποιούμε αρκετές χαμηλής ανάλυσης εικόνες.

Οί μέθοδοι που χρησιμοποιούν πολλές χαμηλής ανάλυσης εικόνες υποθέτουν ότι υπάρχει μικρή (μη ακέραια) μετατόπιση μεταξύ των χαμηλής ανάλυσης εικόνων. Ο σκοπός είναι να ανακατασκευάσουμε μια υψηλής ανάλυσης εικόνα από αρκετές εικόνες χαμηλής ανάλυσης.



Σχήμα 1.1: Υπέρθεση δύο χαμηλής ανάλυσης εικόνων και παρεμβολή στην υψηλής ανάλυσης εικόνα. Οι κύκλοι και τα τετραγωνάκια αναπαριστούν τα εικονοστοιχεία. Στην φάση της υπέρθεσης, οι χαμηλής ανάλυσης εικόνες ευθυγραμμίζονται μεταξύ τους· στο υψηλής ανάλυσης πλέγμα τα νέα εικονοστοιχεία προσδιορίζονται από τους κοντινότερους γείτονες με κατάλληλη παρεμβολή.

Η ανακατασκευή εικόνας με χρήση μεθόδων υπερανάλυσης εικόνας περιλαμβάνει πολλά είδη αλγορίθμων επεξεργασίας εικόνας. Οι αλγόριθμοι αυτοί μπορούν χοντρικά να ταξινομηθούν σε τρεις κατηγορίες ανάλογα με το στάδιο που χρησιμοποιούνται στην διαδικασία της ανακατασκευής:

- Υπέρθεση
- Παρεμβολή
- Ανακατασκευή

Η υπέρθεση περιλαμβάνει τον καθορισμό άγνωστων παραμέτρων (π.χ. μετατόπιση, περιστροφή) οι οποίες σχετίζουν τις χαμηλής ανάλυσης εικόνες με την εικόνα υψηλής ανάλυσης (βλ. σχήμα 1.1). Η ανακατασκευή από την άλλη μεριά απομακρύνει το θόρυβο και τη θόλωση στις χαμηλής ανάλυσης εικόνες και τέλος η παρεμβολή παρεμβάλει τις χαμηλής ανάλυσης εικόνες στην εικόνα υψηλής ανάλυσης. Αυτά τα βήματα μπορούν επίσης να γίνουν μαζί, για παράδειγμα η ανακατασκευή και η παρεμβολή μπορούν να γίνουν μαζί.

Ένα από τα βασικότερα προβλήματα που προκύπτουν κατά την διαδικασία της υπερανάλυσης εικόνας είναι η εξαγωγή πληροφορίας από της χαμηλής ανάλυσης εικόνες. Στην πραγματικότητα σκοπός μας είναι να βρούμε τις άγνωστες παραμέτρους μετασχηματισμού υπέρθεσης. Για το λόγο αυτό, εφαρμόζονται μέθοδοι υπέρθεσης εικόνας πάνω στο πρόβλημα της υπερανάλυσης εικόνας.

Η ποιότητα της ανακατασκευασμένης εικόνας υψηλής ανάλυσης εξαρτάται από την ακρίβεια στην εκτίμηση των παραμέτρων μετασχηματισμού υπέρθεσης. Οι χαμηλής ανάλυσης εικόνες θα πρέπει να περιέχουν μη ακέραια μετατόπιση. Ο λόγος που συμβαίνει αυτό είναι ότι αν οι χαμηλής ανάλυσης εικόνες εχουν ακέραια μετατόπιση μεταξύ τους, τότε κάθε πλαίσιο θα περιέχει την ίδια πληροφορία. Στην περίπτωση αυτή, η ανακατασκευασμένη εικόνα υψηλής ανάλυσης είναι μία απλή εστιασμένη έκδοση μίας μόνο χαμηλής ανάλυσης εικόνας, η οποία δεν παράγει υψηλότερη ανάλυση.

Στην παρούσα εργασία χρησιμοποιούμε μεθόδους εξαγωγής χαρακτηριστικών σημείων από μία εικόνα [18] και μεγιστοποίηση της αμοιβαίας πληροφορίας [20], [36], οι οποίες σε συνδιασμό με την διαδικασία της υπερανάλυσης εικόνας, μας επιτρέπουν να ανακατασκευάσουμε την επιθυμητή εικόνα υψηλής ανάλυσης.

Στο κεφάλαιο 2 αναλύουμε το προβλημα της υπερανάλυσης εικόνας και παρέχουμε τα μαθηματικά εργαλεία για την επίλυσή του. Για την εκτίμηση της υψηλής ανάλυσης εικόνας και των παραμέτρων υπέρθεσης χρησιμοποιούμε τεχνικές μεγιστοποίησης εκ των υστέρων, δοθέντων των χαμηλής ανάλυσης εικόνων. Η επίλυση του προβήματος βασίζεται σε στοχαστικές μεθόδους.

Στο κεφάλαιο 3 παρουσιάζεται μία μεθοδος υπερανάλυσης εικόνας που βασίζεται στον εντοπισμό χαρακτηριστικών σημείων. Τα βήματα γι' αυτήν την διαδικασία συνοψίζονται στα εξής 4:

- Εντοπισμός χαρακτηριστικών σημείων: Στο βήμα αυτό εξάγονται βασικά χαρακτηριστικά σημεία για κάθε μία εικόνα χαμηλής ανάλυσης, τα οποία προκύπτουν από συγκεκριμένα κριτήρια και είναι αμετάβλητα στην κλίμακα της εικόνας και την περιστροφή.

- Ταίριασμα χαρακτηριστικών σημείων: Στην φάση αυτή, στόχος μας είναι να βρούμε μία αντιστοίχιση των χαρακτηριστικών σημείων που εντοπίστηκαν στο προηγούμενο βήμα.

- Εκτίμηση μετασχηματισμού υπέρθεσης: Με βάση την αντιστοίχιση που βρήκαμε στο προηγούμενο βήμα, θα πρέπει να εκτιμήσουμε τις άγνωστες παραμέτρους μετασχηματισμού υπέρθεσης.

- Υπερανάλυση εικόνας: Στην τελευταία φάση της διαδικασίας αυτής, μπορούμε να εφαρμόσουμε αλγορίθμους υπερανάλυσης εικόνας. Εχουμε ήδη εκτιμήσει τις άγνωστες παραμέτρους μετασχηματισμού υπέρθεσης, που μας επιτρέπουν να εξάγουμε πληροφορία από τις χαμηλής ανάλυσης εικόνες και να εκτιμήσουμε την άγνωστη υψηλής ανάλυσης εικόνα.

Στο κεφάλαιο 4 προτείνεται μία μέθοδος για υπερανάλυση εικόνας που βασίζεται στην χρήση αμοιβάιας πληροφορίας. Η μέθοδος αυτή παράγει πολύ καλα αποτελέσματα στην εκτίμηση των παραμέτρων υπέρθεσης. Όταν εφαρμόζουμε τις παραμέτρους αυτές στο πρόβλημα της υπερανάλυσης εικόνας, το PSNR της εκτίμησης της υψηλής ανάλυσης εικόνας είναι 1 έως 2 dB υψηλότερο από αυτό των εικόνων υψηλής ανάλυσης που ανακατασκευάστηκαν από άλλες μεθόδους.

Τέλος στο κεφάλαιο 5 μελετούμε το πρόβλημα της υπερανάλυσης εικόνας υπό την σκοπιά των εύρωστων εκτιμητών. Προτείνουμε μία νέα μέθοδο για υπερανάλυση εικόνας με χρήση εύρωστων εκτιμητών. Η τεχνική αυτή έχει αρκετά μεγάλη αξία και παρέχει πολύ καλά αποτελέσματα, υπό την παρουσία δεδομένων που δεν ακολουθούν το κυρίαρχο μοντέλο υπολογισμού. Τέτοιου είδους δεδομένα μπορεί να προέρχονται από σφάλματα στην εύρεση των παραμέτρων υπέρθεσης, θόρυβο, κινούμενα αντικείμενα που εισέρχονται στην σκηνή, "νεκρά" εικονοστοιχεία κ.α. Η τεχνική αυτή επιτυγχάνει να εξαλείψει την παρουσία των ανεπιθύμητων δεδομένων, εκει που οι κλασσικές μέθοδοι υπερανάλυσης εικόνας αποτυγχάνουν να δώσουν ικανοποιητικό αποτέλεσμα.

Κεφαλαιο 2

Εισαγωγη Στην Υπερανάλυση Εικονας

2.1 Εισαγωγή

Στο παρόν κεφάλαιο παρουσιάζεται η ιδέα της υπερανάλυσης εικόνας βρίσκοντας τις παραμέτρους υπέρθεσης με χρήση τεχνικών μεγιστοποίησης της εκ των υστέρων πιθανοφάνειας (*MAP*) [10], [13], [5], [28]. Στην παράγραφο 2.2 ορίζουμε το πρόβλημα της υπερανάλυσης εικόνας και περιγράφουμε το μοντέλο μας. Στην παράγραφο 2.3 ορίζουμε τους εκτιμητές μεγιστοποίησης εκ των υστέρων για τις παραμέτρους υπέρθεσης. Στην επόμενη παράγραφο 2.4 αναφερόμαστε σε μεθόδους εκτίμησης των παραμέτρων εξομάλυνσης της συνάρτησης κόστους. Έπειτα, κάνουμε μία αναφορά σε Μπεϋζιανές (*Bayesian*) μεθόδους υπερανάλυσης εικόνας και τέλος, παρουσιάζουμε κάποια πειραματικά αποτελέσματα του αλγορίθμου που υλοποιήσαμε.

2.2 Ορισμός του Προβλήματος

Η ανακατασκευή μιας εικόνας υψηλής ανάλυσης είναι μια πολύ ισχυρή μεθοδολογία για αύξηση της ανάλυσης της εικόνας από ένα σύνολο θολωμένων και με θόρυβο χαμηλής ανάλυσης εικόνων ή εικονοσειράς. Για να λυθεί αυτό το πρόβλημα, πρέπει να αναπτυχθεί ένα μοντέλο που να χαρακτηρίζει ολοκληρωτικά την ανάκτηση μιας εικόνας.

Θεωρούμε την επιθυμητή εικόνα υψηλής ανάλυσης μεγέθους $L_1N_1 \times L_2N_2$ γραμμένη σε λεξικογραφική μορφή σαν διάνυσμα $\mathbf{z} = [z_1, z_2, \ldots, z_N]^T$, όπου $N = L_1N_1L_2N_2$. Όπου το $\mathbf{z}$ αντιπροσωπεύει την ιδανική εικόνα ή τις υποκείμενες τιμές, οι οποίες έχουν ληφθεί σύμφωνα με το ρυθμό που ορίζει το κριτήριο Nyquist. Οι παράμετροι L_1 και L_2 αντιπροσωπεύουν τους συντελεστές υποδειγματοληψίας στο παρατηρούμενο μοντέλο κατά την οριζόντια και κάθετη κατεύθυνση, αντίστοιχα. Για το λόγο αυτό, κάθε παρατηρούμενη χαμηλής ανάλυσης εικόνα έχει μέγεθος $N_1 \times N_2$. Έστω η k-οστή χαμηλής ανάλυσης εικόνα η οποία μπορεί να γραφεί με λεξικογραφικό τρόπο ως εξής $\mathbf{y}_k = [y_{k,1}, y_{k,2}, \cdots, y_{k,M}]^T$ για $k = 1, 2, \cdots, p$ και όπου $M = N_1N_2$. Όλο το σύνολο των παρατηρούμενων εικόνων χαμηλής ανάλυσης το αναπαριστάμε με τον πίνακα:

$$\mathbf{y} = [\mathbf{y}_1^T, \mathbf{y}_2^T, \cdots, \mathbf{y}_p^T]^T = [y_1, y_2, \cdots, y_{p,M}]^T$$

Με αυτόν τον τρόπο, όλες οι παρατηρούμενες τιμές των εικονοστοιχείων περιέχονται στο $\mathbf{y}$.

Σκοπός μας είναι να ορίσουμε μια κατάλληλη σχέση μεταξύ της υποβαθμισμένης εικόνας υψηλής ανάλυσης και των παρατηρούμενων εικόνων χαμηλής ανάλυσης. Για το λόγο αυτό, χρησιμοποιείται ένα απλό αλλά γενικό μοντέλο παρατήρησης όπου τα χαμηλής ανάλυσης εικονοστοιχεία ορίζονται ως το σταθμισμένο άθροισμα των κατάλληλων υψηλής ανάλυσης εικονοστοιχείων με προσθετικό θόρυβο. Ένας σημαντικός παράγοντας στον καθορισμό των βαρών είναι η θέση του κάθε εικονοστοιχείου χαμηλής ανάλυσης σε σχέση με το σταθερό πλέγμα εικονοστοιχείων υψηλής ανάλυσης (δηλαδή, οι παράμετροι υπέρθεσης). Ειδικότερα, τα παρατηρούμενα χαμηλής ανάλυσης εικονοστοιχεία του πλαισίου k σχετίζονται με την εικόνα υψηλής ανάλυσης σύμφωνα με το παρακάτω μοντέλο:

$$y_{k,m} = \sum_{r=1}^{N} w_{k,m,r}(s_k) z_r + \eta_{k,m} \tag{2.1}$$

για $m = 1, 2, \cdots, M$ και $k = 1, 2, \cdots, p$. Το βάρος $w_{k,m,r}(s_k)$ αντιπροσωπεύει την συνεισφορά του r-οστού υψηλής ανάλυσης εικονοστοιχείου στο m-οστό χαμηλής ανάλυσης παρατηρούμενο εικονοστοιχείο του k-οστού πλαισίου. Το διάνυσμα $s_k = [s_{k,1}, s_{k,2}, \cdots, s_{k,K}]^T$, περιέχει τις K παραμέτρους υπέρθεσης για το πλαίσιο k. Ανάλογα με την εφαρμογή, αυτές οι παράμετροι μπορούν να αναπαραστούν καθολική (*global*) μετατόπιση κατά τις οριζόντιες και κατακόρυφες κατευθύνσεις, περιστροφή, μετασχηματισμούς ομοιότητας (*affine transformations*) ή οποιεσδήποτε άλλες παραμέτρους κίνησης. Ο όρος $\eta_{k,m}$ στην εξίσωση (2.1) αντιπροσωπεύει τον προσθετικό θόρυβο υπό την υπόθεση ότι είναι δείγματα ανεξάρτητα και ακολουθούν την ίδια κανονική κατανομή, με μέσο όρο μηδέν και διακύμανση σ_η^2. Γενικά, ένα μοντέλο που χρησιμοποιεί Gaussian θόρυβο θεωρούμε ότι είναι αρκετά χρήσιμο σε μια πληθώρα συστημάτων απεικόνισης.

Το μοντέλο παρατήρησης στην εξίσωση (2.1) υποθέτει ότι τα δείγματα της εικόνας υψηλής ανάλυσης, $\mathbf{z}$, παραμένουν σταθερά κατά την διάρκεια της ανάκτησης από τα πλαίσια χαμηλής ανάλυσής. Έτσι οι διαφορές από το ένα πλαίσιο στο άλλο για τα βάρη του μοντέλου στην εξίσωση (2.1) προκύπτουν από την κίνηση κάθε εικονοστοιχείου χαμηλής ανάλυσης σε σχέση με το πλέγμα υψηλής ανάλυσης. Την πολύ μικρή κίνηση των εικονοστοιχείων είναι που προσπαθούμε να εκμεταλευτούμε, ώστε να κάνουμε μια εκτίμηση της υψηλής ανάλυσης εικόνας. Ένα απλό μοντέλο για να καθορίσουμε τα βάρη φάινεται στο σχήμα 2.1. Κάθε χαμηλής ανάλυσης εικονοστοιχείο (όπως φαίνεται στην δεξιά εικόνα (β) του σχήματος 2.1) λαμβάνεται αθροίζοντας όλα τα εικονοστοιχεία υψηλής ανάλυσης μέσα σε κάθε υποχώρο, τα οποία αντιστοιχούν σε αυτά της χαμηλής ανάλυσης (όπως φαίνεται στην αριστερή εικόνα (α)). Στο παράδειγμα αυτό, μόνο τα L_1L_2 υψηλής ανάλυσης εικονοστοιχεία συνεισφέρουν σε ένα συγκεκριμένο χαμηλής ανάλυσης εικονοστοιχείο. Για να αναπαραστήσουμε έναν ομοιόμορφο ανιχνευτή, τα βάρη που αντιστοιχούν σε εκείνα τα L_1L_2 υψηλής ανάλυσης εικονοστοιχεία, θα πρέπει να τεθούν σε $1/(L_1L_2)$. Τα υπόλοιπα βάρη θα πρέπει να τεθούν σε μηδέν. Αυτό το μοντέλο του διακριτού ανινχευτή προσομοιώνει την ένταση του φωτός καθώς πέφτει πάνω σε κάθε υποχώρο του χαμηλής ανάλυσης ανιχνευτή. Αν ολόκληρο το χαμηλής ανάλυσης πλέγμα κινείται σε σχέση με το σταθερό υψηλής ανάλυσης πλέγμα (δηλαδή, καθολική συμπαγής κίνηση), ένα διαφορετικό σύνολο από υψηλής ανάλυσης εικονοστοιχεία συνεισφέρουν σε κάθε χαμηλής ανάλυσης εικονοστοιχείο. Αυτό παράγει ένα νέο σύνολο γραμμικών ανεξάρτητων εξισώσεων από την εξίσωση (2.1).

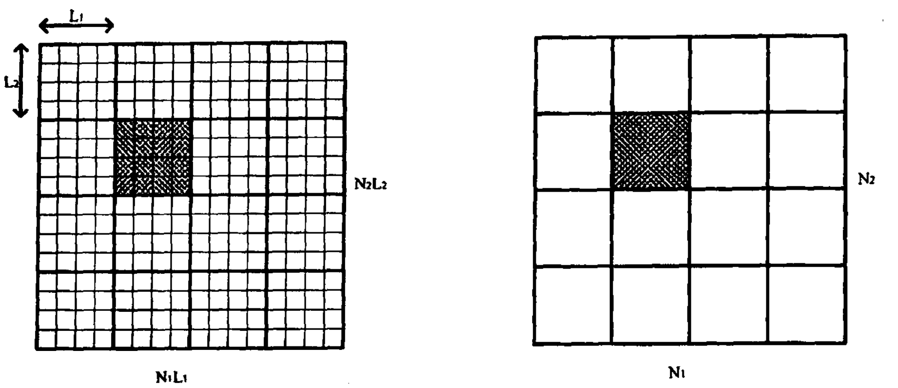
Εναλλακτικά, το μοντέλο στην εξίσωση (2.1) μπορεί να εκφραστεί με τους όρους ολόκληρου του συνόλου των χαμηλής ανάλυσης εικονοστοιχείων ως εξής:

$$y_m = \sum_{r=1}^{N} w_{m,r}(s) z_r + \eta_m \tag{2.2}$$

για $m = 1, 2, \cdots, pM$ και όπου $w_{m,r}(s)$ είναι η “συνεισφορά” του z_r στο y_m. Ολόκληρο το σύνολο των παραμέτρων κίνησης περιέχεται μέσα στο $\mathrm{s} = [s_1^T, s_2^T, \cdots, s_p^T]^T$. Σε αρκετές περιπτώσεις αυτές οι παράμετροι δεν είναι γνωστές εκ των προτέρων (*a priori*). Επομένως τις θεωρούμε ως τυχαίες παραμέτρους, τις οποίες θέλουμε να εκτιμήσουμε μαζί την υψηλής ανάλυσης εικόνα $\mathbf{z}$.

Είναι πιο βολικό να αναπαραστήσουμε το μοντέλο παρατήρηης σε συμβολισμό πινάκων. Έτσι, γράφοντας ξάνα την εξίσωση (2.2) προκύπτει η παρακάτω εξίσωση:

$$\mathbf{y} = \mathbf{W_s z} + \mathbf{n} \tag{2.3}$$



υψηλής ανάλυσης εικόνα $\mathbf{z}$ χαμηλής ανάλυσης εικόνα $\mathbf{y}_k$

(α)   (β)

Σχήμα 2.1: Μοντέλο διακριτού ανιχνευτή (α) που δείχνει τα εικονοστοιχεία υψηλής ανάλυσης τα οποία συνεισφέρουν (β) στα χαμηλής ανάλυσης εικονοστοιχεία. Η εικόνα $\mathbf{z}$ αντιπροσωπεύει την πραγματική εικόνα υψηλής ανάλυσης την οποία θέλουμε να εκτιμήσουμε και η $\mathbf{y}_k$ είναι η k-οστή χαμηλής ανάλυσης εικόνα. Να σημειώσουμε το διαφορετικό μέγεθος του πλέγματος για τις εικόνες $\mathbf{z}$ και $\mathbf{y}_k$.

όπου το στοιχείο (m, r) στο $\mathbf{W}_s$ είναι το $w_{m,r}(s)$ και το $\mathbf{n} = [\eta_1, \eta_2, \cdots, \eta_{pM}]^T$. Να σημειώσουμε ότι αφού τα στοιχεία του $\mathbf{n}$ είναι δείγματα ανεξάρτητα και ακολουθούν την ίδια κατανομή, με μέση τιμή μηδέν και διακύμανση σ_η^2, η συνάρτηση πυκνότητας πιθανότητας πολλών μεταβλητών του $\mathbf{n}$ δινέται από:

$$\begin{aligned} Pr(\mathbf{n}) &= \frac{1}{(2\pi)^{\frac{pM}{2}} \sigma_\eta^{pM}} \exp\left\{ -\frac{1}{2\sigma_\eta^2} \mathbf{n}^T \mathbf{n} \right\} \\ &= \frac{1}{(2\pi)^{\frac{pM}{2}} \sigma_\eta^{pM}} \exp\left\{ -\frac{1}{2\sigma_\eta^2} \sum_{m=1}^{pM} \eta_m^2 \right\}. \end{aligned} \tag{2.4}$$

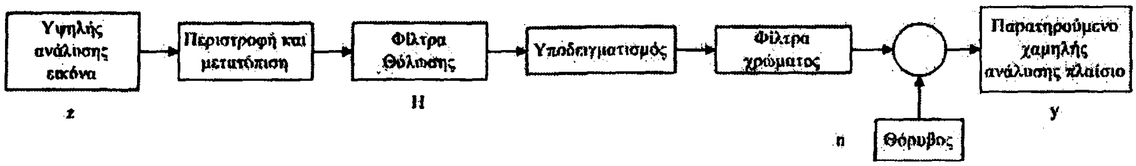
Η διαδικασία υποβαθμισμού της εικόνας μοντελοποιείται χρησιμοποιώντας φίλτρα για θόλωση, κίνηση και υποδειγματισμό παίρνοντας τον μέσο όρο των εικονοστοιχείων σε συνδιασμό με προσθετικό Gaussian θόρυβο. Ξαναγράφοντας την εξίσωση (2.3) προκύπτει η παρακάτω εξίσωση:

$$\mathbf{y} = \mathbf{W}\mathbf{z} + \mathbf{n} \tag{2.5}$$

όπου ο πίνακας υποδειγματισμού $\mathbf{W} = [\mathbf{W}_1, \mathbf{W}_2, \cdots, \mathbf{W}_k]$, για $k = 1, 2, \cdots, p$ αναπαριστά τη θόλωση, την κίνηση και την υποδειγματοληψία. Επομένως, για το πλαίσιο k, ο $\mathbf{W}_k$ μπορεί να γραφεί ως:

$$\mathbf{W_k} = \mathbf{S}\mathbf{B_k}\mathbf{M_k} \tag{2.6}$$

όπου $\mathbf{S}$ είναι ο $N_1 N_2 \times N$ πίνακας υποδειγματισμού, $\mathbf{B}_k$ είναι ο $N \times N$ πίνακας θόλωσης και $\mathbf{M}_k$ είναι ο $N \times N$ πίνακας κίνησης που περιέχει μηδενικά και άσσους και δίνει την θέση κάθε εικονοστοιχείου μετά την κίνηση. Το σχήμα 2.2 δείχνει το διάγραμμα ροής της εξίσωσης (2.5), το οποίο περιέχει περισσότερες παραμέτρους απ' ότι η εξίσωση αυτή. Το πρόβλημα αυτό μπορεί να λυθεί με διάφορους τρόπους. Στην παρούσα εργασία θα εξετάσουμε κάποιες από αυτές.



Σχήμα 2.2: Διάγραμμα ροής για το μοντέλο παρατήρησης. Ο πίνακας H μπορεί να περιέχει περιστροφή και μετατόπιση.

2.3 Εκτίμηση με Μεγιστοποίηση εκ των Υστέρων (MAP)

Στην γενική περίπτωση, επιθυμούμε να αναπτύξουμε μια εκτίμηση με μεγιστοποίηση εκ των υστέρων (*Maximimum a Posteriori, MAP*) της υψηλής ανάλυσης εικόνας $\mathbf{z}$ και των παραμέτρων υπέρθεσης $\mathbf{s}$ παράλληλα, δοθέντων των παρατηρήσεων $\mathbf{y}$. Οι εκτιμήσεις αυτές μπορούν να υπολογιστούν όπως φαίνεται παρακάτω:

$$\hat{\mathbf{z}}, \hat{\mathbf{s}} = \arg\max_{\mathbf{z},\mathbf{s}} Pr(\mathbf{z}, \mathbf{s}|\mathbf{y}). \tag{2.7}$$

Χρησιμοποιώντας τον κανόνα του Bayes, η εξίσωση (2.7) μπορεί εναλλακτικά να γραφεί ως εξης:

$$\hat{\mathbf{z}}, \hat{\mathbf{s}} = \arg\max_{\mathbf{z},\mathbf{s}} \frac{Pr(\mathbf{y}|\mathbf{z}, \mathbf{s})Pr(\mathbf{z}, \mathbf{s})}{Pr(\mathbf{y})}.$$

Παρατηρούμε ότι ο παρανομαστής δεν είναι συνάρτηση του $\mathbf{z}$ ή του $\mathbf{s}$, και θεωρώντας ότι τα $\mathbf{z}$ και $\mathbf{s}$ είναι στατιστικά ανεξάρτητα, συνεπώς οι εκτιμήσεις αυτών μπορούν να γραφούν στην παρακάτω μορφή:

$$\hat{\mathbf{z}}, \hat{\mathbf{s}} = \arg\max_{\mathbf{z},\mathbf{s}} Pr(\mathbf{y}|\mathbf{z}, \mathbf{s})Pr(\mathbf{z})Pr(\mathbf{s}). \tag{2.8}$$

Ισοδύναμα μπορούμε να ελαχιστοποιήσουμε τον αρνητικό λογάριθμο της εξίσωσης (2.8).

$$\begin{aligned}
\hat{\mathbf{z}}, \hat{\mathbf{s}} &= \arg\min_{\mathbf{z},\mathbf{s}} L(\mathbf{z}, \mathbf{s}) \\
&= \arg\min_{\mathbf{z},\mathbf{s}} -\log[Pr(\mathbf{y}|\mathbf{z}, \mathbf{s})] - \log[Pr(\mathbf{z})] - \log[Pr(\mathbf{s})].
\end{aligned} \tag{2.9}$$

Στό σημείο αυτό θα πρέπει να ορίσουμε τις εκ των προτέρων πυκνότητες πιθανότητας, της εικόνας $Pr(\mathbf{z})$, των παραμέτρων υπέρθεσης $Pr(\mathbf{s})$ και την υπό συνθήκη πυκνότητα πιθανότητα $Pr(\mathbf{y})|(\mathbf{z},\mathbf{s})$. Τέλος, θα ορίσουμε μια μέθοδο για βελτιστοποίηση της εξίσωσης (2.9) σε σχεση με τα $\mathbf{z}$ και $\mathbf{s}$.

2.3.1 Εκ των Προτέρων Μοντέλα

Το πρόβλημα της εκτίμησης της $\mathbf{z}$ από το $\mathbf{y}$ είναι γενικώς ένα κακώς ορισμένο πρόβλημα. Αυτό μπορεί να οδηγήσει σε εκτιμήσεις με υπερβολικά μεγάλο θόρυβο αν δεν το χειριστούμε με τον ανάλογο τρόπο. Παρ' όλα αυτά μια κατάλληλη επιλογή για το $Pr(\mathbf{z})$ μπορεί να ομαλοποιήσει το πρόβλημα αυτό. Η τυχαία μεταβλητή $\mathbf{z}$, ακολουθεί κανονική κατανομή μέσης τιμής μηδέν:

$$Pr(\mathbf{z}) = \frac{1}{(2\pi)^{\frac{N}{2}}|C_z|^{1/2}} \exp\left\{ -\frac{1}{2}\mathbf{z}^T C_z^{-1}\mathbf{z} \right\} \tag{2.10}$$

όπου C_z είναι ο $N \times N$ πίναχας συμμεταβλητότητας της $\mathbf{z}$. Ο εκθετικός όρος στην εξίσωση (2.10) ο οποίος περιέχει την συμμεταβλητότητα της εικόνας μπορεί να αναλυθεί σε ένα άθροισμα από γινόμενα παράγοντας την επόμενη εξίσωση:

$$Pr(\mathbf{z}) = \frac{1}{(2\pi)^{\frac{N}{2}}|C_z|^{1/2}} \exp\left\{ -\frac{1}{2\lambda}\sum_{i=1}^{N} \mathbf{z}^T \mathbf{d}_i \mathbf{d}_i^T \mathbf{z} \right\} \tag{2.11}$$

όπου το $\mathbf{d}_i = [d_{i,1}, d_{i,2}, \cdots, d_{i,N}]^T$ είναι ένα διάνυσμα συντελεστών και το λ μπορεί να θεωρηθεί ως μια "ριθμιστική" παράμετρος. Στην εργασία των Hu He and Lisimachos P. Kondi [12], αναφέρεται το πώς εκτιμώνται αυτές οι παράμετροι ομαλοποίησης. Επομένως, η εξίσωση μπορεί να ξαναγραφεί στην εξής μορφή:

$$Pr(\mathbf{z}) = \frac{1}{(2\pi)^{\frac{N}{2}}|C_z|^{1/2}} \exp\left\{ -\frac{1}{2\lambda}\sum_{i=1}^{N} \left(\sum_{j=1}^{N} d_{i,j} z_j \right)^2 \right\}. \tag{2.12}$$

Το διάνυσμα συντελεστών $\mathbf{d}_i$ για $i = 1, 2, \cdots, N$ εκφράζει την εκ των προτέρων γνώση σχετικά με την τοπική σχέση μεταξύ των εικονοστοιχείων της $\mathbf{z}$. Η παράμετρος λ ελέγχει τις ασυνέχειες των χαρακτηριστικών στην $\mathbf{z}$. Αν εξισώσουμε τα δύο μέλη των εξισώσεων (2.10) και (2.11), τα στοιχεία του αντίστροφου πίνακα συμμεταβλητότητας μπορούν να γραφούν όπως φαίνεται στην παρακάτω εξίσωση. Έστω το (i, j)-οστό στοιχείο στον πίνακα C_z^{-1} που συμβολίζουμε ως $C_{i,j}^{-1}$ το οποίο δίνεται από:

$$C_{i,j}^{-1} = \frac{1}{\lambda}\sum_{r=1}^{N} d_{r,i} d_{r,j}. \tag{2.13}$$

Ως συντελεστών έχουμε επιλέξει τις τιμές (ένα δισδιάστατο Laplacian πυρήνα):

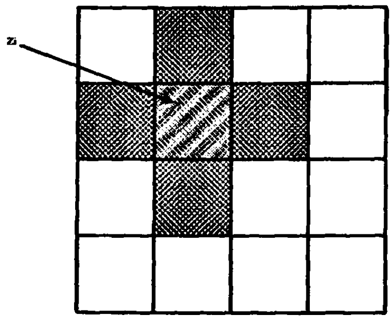
$$d_{i,j} = \begin{cases} 1, & \text{για } i = j \\ -1/4, & \text{για } j : z_j \text{ αν είναι άμμεσος γείτονας του } z_j \end{cases} \tag{2.14}$$

Το σχήμα 2.3 δείχνει τους τέσσερεις άμεσους γείτονες του εικονοστοιχείου z_i. Στο όριο του πλέγματος αυτού, οι συντελεστές τροποποιούνται με τέτοιο τρόπο έτσι ώστε λιγότεροι από τους τέσσερεις γείτονες να λαμβάνονται υπ' όψιν.

2.3.2 Επίλυση MAP με Στοχατικές Μεθόδους

Στην ενότητα 2.3 είχαμε αναφερθεί στην υπό συνθήκη πυκνότητα πιθανότητας της $Pr(\mathbf{y}|\mathbf{z}, \mathbf{s})$. Στην παρούσα φάση θα ορίσουμε την ποσότητα αυτή με την μορφή πινάκων. Δοθέντος του μοντέλου παρατήρησης στην εξίσωση (2.5) και της πυκνότητας πιθανότητας του θορύβου στην εξίσωση (2.4), η υπο συνθήκη πυκνότητα πιθανότητας μπορεί να γραφεί ως εξής:

$$Pr(\mathbf{y}|\mathbf{z}, \mathbf{s}) = \frac{1}{(2\pi)^{\frac{N}{2}}\sigma_\eta^N} \exp\left\{ -\frac{1}{2\sigma_\eta^2}(\mathbf{y} - \mathbf{W}_s\mathbf{z})^T(\mathbf{y} - \mathbf{W}_s\mathbf{z}) \right\}. \tag{2.15}$$



Σχήμα 2.3: Υψηλής ανάλυσης εικόνας που δείχνει τους άμμεσους γείτονες του εικονοστοιχείου z_i. Στην περίπτωση αυτή, το $d_{i,j}$ θα τεθεί στο μηδέν μόνο για εκείνα τα j έτσι ώστε το z_j είναι άμμεσος γείτονας του z_i (σκιασμένα εικονοστοιχεία).

Χρησιμοποιώντας της εξισώσεις (2.9), (2.10) και (2.15) και αγνοώντας τους όρους που προκύπτουν και οι οποίοι δεν είναι συναρτήσεις των $\mathbf{z}$ και $\mathbf{s}$, τότε οι MAP εκτιμήσεις μπορούν να εκφραστούν με τον παρακάτω τρόπο:

$$\hat{\mathbf{z}}, \hat{\mathbf{s}} = \arg\min_{\mathbf{z},\mathbf{s}} L(\mathbf{z}, \mathbf{s})$$

όπου στην παραπάνω έξίσωση το $L(\mathbf{z},\mathbf{s})$ δίνεται από:

$$L(\mathbf{z}, \mathbf{s}) = \frac{1}{2\sigma_\eta^2}(\mathbf{y} - \mathbf{W_s z})^T(\mathbf{y} - \mathbf{W_s z}) + \frac{1}{2}\mathbf{z}^T C_z^{-1}\mathbf{z}. \tag{2.16}$$

Η συνάρτηση (2.16) ονομάζεται συνάρτηση κόστους. Θα πρέπει να την ελαχιστοποιήσουμε σε σχέση με τα $\mathbf{z}$ και $\mathbf{s}$. Στην εξίσωση αυτή το $\mathbf{y}$ αναπαριστά τα δεδομένα μας. Αυτό παρέχει κατά ένα μεγάλο βαθμό αρκετό ενδιαφέρον οσον αφορά το πρόβλημα της βελτιστοποίησης. Παρατηρώντας την εξίσωση (2.16), βλέπουμε ότι δεν διαφοροποιείται σε σχέση με το $\mathbf{s}$ για διάφορα μοντέλα κίνησης. Ωστόσο, δοθέντος του ότι το $\mathbf{z}$ είναι σταθερό και δεν μεταβάλλεται, τότε είναι δυνατόν να εκτελέσουμε μία αναζήτηση σε ένα πεπερασμένο σύνολο διακριτών παραμέτρων κίνησης έτσι ώστε να ελαχιστοποιήσουμε την (2.16) σε σχέση με το $\mathbf{s}$. Επίσης, η ίδια εξίσωση σχηματίζει μία τετραγωνική συνάρτηση για το $\mathbf{z}$ και η οποία μπορεί να ελαχιστοποιηθεί εύκολα ως προς το $\mathbf{z}$ αν οι παράμετροι υπέρθεσης $\mathbf{s}$ δεν μεταβάλλονται. Κάθε φορά ελαχιστοποιούμε αυτή την συνάρτηση κόστους ως προς $\mathbf{z}$ και $\mathbf{s}$ ξεχωριστά και εναλλάξ και αυτή η διαδικασία θα συνεχίσει μέχρις ότου ο αλγόριθμος να συγκλίνει ή να φτάσουμε σε ένα συγκεκριμένο αριθμό από επαναλήψεις.

Κατά την διάρκεια της διαδικασίας της βελτιστοποίησης χρησιμοποιούμε μια αρχική εκτίμηση της εικόνας υψηλής ανάλυσης $\hat{z}^0$. Αυτή η εκτίμηση μπορεί να υπολογιστεί, για παράδειγμα, παρεμβάλοντας την πρώτη χαμηλής ανάλυσης εικόνα στην παρατηρούμενη ακολουθία. Σε κάθε επανάληψη n του αλγορίθμου, οι παράμετροι κίνησης μπορούν να υπολογιστούν ελαχιστοποιώντας την εξίσωση (2.16) ως προς το $\mathbf{s}$ δοθέντος ότι η τρέχουσα εκτίμηση της υψηλής ανάλυσης εικόνας $\hat{\mathbf{z}}^n = [\hat{z}_1^n, \hat{z}_2^n, \cdots, \hat{z}_N^n]^T$. Πιο αναλυτικά, η εκτίμηση της κίνησης για $n = 0, 1, 2, \ldots$ μπορεί να υπολογιστεί ως εξής:

$$\begin{aligned}\hat{s}^n &= \arg\min_{\mathbf{s}} L(\hat{\mathbf{z}}^{\mathbf{n}}, \mathbf{s}) \\ &= \arg\min_{\mathbf{s}} \{(\mathbf{y} - \mathbf{W_s}\hat{\mathbf{z}}^{\mathbf{n}})^T(\mathbf{y} - \mathbf{W_s}\hat{\mathbf{z}}^{\mathbf{n}})\}.\end{aligned}$$

Για να παράξουμε την διαδικασία της ενημέρωσης για την εικόνα υψηλής ανάλυσης, πρέπει να υπολογίσουμε την παράγωγο της συνάρτησης κόστους, όπως αυτή προκύπτει από την εξίσωση (2.16), ως προς την ποσότητα $\mathbf{z}$. Η παράγωγος αυτή δίνεται από τον παρακάτω τύπο:

$$\nabla_z L(\mathbf{z}, \mathbf{s}) = \frac{1}{\sigma_\eta^2}(\mathbf{W_s}^T\mathbf{W_s}\mathbf{z} - \mathbf{W}_{\mathbf{s}T}\mathbf{y}) + C_z^{-1}\mathbf{z} \qquad (2.17)$$

όπου

$$\nabla_z L(\mathbf{z}, \mathbf{s}) = \begin{bmatrix} \frac{\partial L(\mathbf{z},\mathbf{s})}{\partial z_1} \\ \frac{\partial L(\mathbf{z},\mathbf{s})}{\partial z_2} \\ \vdots \\ \frac{\partial L(\mathbf{z},\mathbf{s})}{\partial z_N} \end{bmatrix}$$

Στην επανάληψη n, του αλγορίθμου βάζουμε τις παραμέτρους κίνησης στην εξίσωση (2.17) έτσι ώστε $\mathbf{s} = \hat{s}^n$. Έπειτα, θέτουμε την ποσότητα $\nabla_z L(\mathbf{z},\mathbf{s})|_{\mathbf{s}} = \hat{s}^n$ ίση με το μηδέν και λύνουμε ως προς το $\mathbf{z}$ λαμβάνοντας την παρακάτω εκίμηση:

$$\hat{\mathbf{z}}^{\mathbf{n+1}} = [\mathbf{W}_{\hat{\mathbf{s}}^n}^T \mathbf{W}_{\hat{\mathbf{s}}^n} + \sigma_\eta^2 C_z^{-1}]^{-1} \mathbf{W}_{\hat{\mathbf{s}}^n}^T \mathbf{y}. \qquad (2.18)$$

Η διαδικασία της ενημέρωσης των παραμέτρων κίνησης και της εικόνας $\mathbf{z}$ συνεχίζεται για $n = 0, 1, 2, \ldots$ μέχρι η συνάρτηση κόστους $L(\hat{\mathbf{z}}^n, \hat{\mathbf{s}}^n)$ να σταθεροποιηθεί ή μέχρι να ισχύσει το επόμενο κριτήριο σύγκλισης $\|\hat{\mathbf{z}}^{n+1} - \hat{\mathbf{z}}^n\| / \|\hat{\mathbf{z}}^n\| < \epsilon$, όπου το ϵ είναι ένα προκαθορισμένο κατώφλι. Είναι αρκετά πιθανό η συνάρτηση (2.16) να έχει τοπικά ελάχιστα και η διαδικασία της βελτιστοποίησης να παγιδευτεί σε κάποιο από αυτά. Επομένως, είναι σημαντικό να ξεκινήσει η διαδικασία που περιγράψαμε με την καλύτερη (όσο το δυνατόν) εκτίμηση της $\mathbf{z}$.

Ένα πρακτικό πρόβλημα που μπορεί να προκύψει κατά την υλοποίηση της εξίσωσης (2.18), είναι αυτό της μεγάλης κατανάλωσης χώρου και μνήμης, λόγω των μεγάλων διαστάσεων που έχουν οι πίνακες που χρησιμοποιούμε. Αν και οι πίνακες $\mathbf{W_s}$ και C_z^{-1} είναι στην γενική περίπτωση αραιοί, είναι αρκετά συχνά στην πράξη πιο καλό να χρησιμοποιούμε μια διαδικασία ελαχιστοποίησης με χρήση της παραγώγου πάρα να εκτελούμε αντιστροφή πινάκων στην εξίσωση (2.18). Η προσέγγιση αυτή περιγράφεται στην αμέσως επόμενη ενότητα.

2.3.3 Βελτιστοποίηση με Χρήση της Παραγώγου

Στην ενότητα αυτή, περιγράφουμε μια επαναληπτική διαδικασία ελαχιστοποίησης της παραγώγου ως προς την υψηλής ανάλυσης εικόνα $\mathbf{z}$. Αντικαθιστώντας όλες τις πυκνότητες πιθανότητας στην εξίσωση (2.9) και λαμβάνοντας υπ' όψιν τις εξισώσεις (2.2) και (2.4) προκύπτει:

$$Pr(\mathbf{y} \mid \mathbf{z}, \mathbf{s}) = \frac{1}{(2\pi)^{\frac{pM}{2}} \sigma_\eta^{pM}} \exp\left\{ -\frac{1}{2\sigma_\eta^2} \sum_{m=1}^{pM} \left(y_m - \sum_{r=1}^{N} w_{m,r}(\mathbf{s}) z_r \right)^2 \right\}. \tag{2.19}$$

Χρησιμοποιώντας την εξίσωση (2.9) σε συνδιασμό με την υπό συνθήκη πυκνότητα πιθανότητας της εξίσωσης (2.19) και την εκ των προτέρων πυκνότητα πιθανότητας της εξίσωσης (2.12), τότε οι MAP εκτιμήσεις των $\mathbf{z}$ και $\mathbf{s}$ διατυπώνονται:

$$\hat{\mathbf{z}}, \hat{\mathbf{s}} = \arg\min_{\mathbf{z},\mathbf{s}} L(\mathbf{z}, \mathbf{s})$$

όπου τώρα η MAP συνάρτηση κόστους μπορεί να εκφραστεί ως εξής:

$$L(\mathbf{z}, \mathbf{s}) = \frac{1}{2\sigma_\eta^2} \sum_{m=1}^{pM} \left(y_m - \sum_{r=1}^{N} w_{m,r}(\mathbf{s}) z_r \right)^2 + \frac{1}{2\lambda} \sum_{i=1}^{N} \left(\sum_{j=1}^{N} d_{i,j} z_j \right)^2. \tag{2.20}$$

Όπως φαίνεται ξεκάθαρα από την εξίσωση (2.20), η συνάρτηση κόστους κυμαίνεται μεταξύ δύο τύπων λάθους. Ο πρώτος όρος αναφέρεται ως μια γραμμική εξίσωση σφάλματος. Το σφάλμα αυτό ελαχιστοποιείται όταν το $\mathbf{z}$, προβάλλεται μέσω του μοντέλου παρατήρησης, έτσι ώστε να ταιριάζει πάνω στα δεδομένα. Η ελαχιστοποίηση του όρου αυτού σε μερικές περιπτώσεις μπορεί να οδηγήσει σε υπερβολικά μεγάλο θόρυβο για κάποιες εφαρμογές, λόγω της κακής φύσης του αντίστροφου προβλήματος. Ο δεύτερος όρος αναφέρεται ως το εκ των προτέρων σφάλμα της εικόνας και λειτουργεί σαν ομαλοποίηση για την εικόνα. Αυτός ο όρος στην γενική περίπτωση ελαχιστοποιείται όταν η εικόνα $\mathbf{z}$ είναι ομαλή. Τα βάρη κάθε συνιστώσας στην συνάρτηση κόστους ελέγχονται από της ποσότητες σ_η^2 και λ. Για παράδειγμα, αν η ακρίβεια των δεδομένων είναι υψηλή (αυτό σημαίνει ότι το σ_η^2 είναι μικρό), τότε υπερισχύει η γραμμική εξίσωση σφάλματος καθώς ο δεύτερος όρος τείνει στο μηδέν όταν το λ τείνει στο άπειρο. Αν τα δεδομένα περιέχουν πολύ θόρυβο, τότε η συνάρτηση κόστους θα δώσει μεγαλύτερη έμφαση στο εκ των προτέρων σφάλμα για την εικόνα. Αυτό γενικά, οδηγεί σε πιο ομαλές εκτιμήσεις για την εικόνα. Επομένως, ο όρος αυτός θεωρείται σαν ένας όρος ποινής που ελέγχεται από την παράμετρο λ, και αυτό που κάνει είναι να μας απομακρύνει όσο είναι δυνατόν από θορυβώδεις λύσεις.

Οι παράμετροι κίνησης ενημερώνονται μέσω μιας διαδικασίας αναζήτησης. Σε κάθε επανάληψη n του αλγορίθμου, θέλουμε να ελαχιστοποιήσουμε την εξίσωση (2.20) ως προς $\mathbf{s}$, δοθέντος του γεγονότος ότι $\mathbf{z} = \hat{\mathbf{z}}^n$. Έτσι, χρησιμοποιώντας μόνο τους όρους που περιλαμβάνουν το $\mathbf{s}$ και ελαχιστοποιώντας την εξίσωση για κάθε έναν από τους p όρους ανεξάρτητα, εκτιμώντας παράλληλα τις παραμέτρους υπέρθεσης για το πλαίσιο k στην επανάληψη n, προκύπτει η παρακάτω εξίσωση:

$$\hat{s}_k^n = \arg\min_{\mathbf{s}_k} \left\{ \sum_{m=1}^{M} \left(y_{k,m} - \sum_{r=1}^{N} w_{k,m,r}(\mathbf{s}_k) \hat{z}_r^n \right)^2 \right\} \tag{2.21}$$

για $k = 1, 2, \cdots, p$.

Για την ελαχιστοποίηση της εξίσωσης (2.21) απαιτείται ένα είδος αναζήτησης για τα s_k. Η αναζήτηση αυτή μπορεί να είναι ένας απλός αλγόριθμος ταιριάσματος κατά μπλόκ ή ένας πιο σύνθετος αλγόριθμος υπέρθεσης όπως θα δούμε σε επόμενα κεφάλαια. Στην παρούσα φάση έχουμε υλοποιήσει ένα παραδοσιακό αλγόριθμο ταιριάσματος κατά μπλόκ, στον οποίο τα χαμηλής ανάλυσης πλαίσια συγκρίνονται μεταξύ τους και κάθε ένα από αυτά με την εκτίμηση της υψηλής ανάλυσης εικόνα, έτσι ώστε να καθορίσουμε τις παραμέτρους κίνησης.

Για να πάρουμε την ενημέρωση της παραγώγου για την εκτίμηση της εικόνας, παραγωγίζουμε την συνάρτηση κόστους (2.20) ως προς το εικονοστοιχείο z_k για $k = 1, 2, \cdots, N$. Η μερική παράγωγος δίνεται από:

$$\begin{aligned} g_k(\mathbf{z}, \mathbf{s}) &= \frac{\partial L(\mathbf{z}, \mathbf{s})}{\partial z_k} \\ &= \frac{1}{\sigma_\eta^2} \sum_{m=1}^{pM} w_{m,k}(\mathbf{s}) \left(\sum_{r=1}^{N} w_{m,r}(\mathbf{s}) z_r - y_m \right) + \frac{1}{\lambda} \sum_{i=1}^{N} d_{i,k} \left(\sum_{j=1}^{N} d_{i,j} z_j \right). \end{aligned} \tag{2.22}$$

Να σημειώσουμε ότι ο πρώτος όρος στην εξίσωση (2.21) είναι το άθροισμα των διαφορών μεταξύ των δεδομένων που έχουμε προβλέψει ως τώρα μείον τα πραγματικά χαμηλής ανάλυσης δεδομένα. Κάθε όρος στο άθροισμα αυτό περιέχει ένα βάρος από την συνεισφορά του z_k στο χαμηλής ανάλυσης εικονοστοιχείο, το $w_{m,k}(\mathbf{s})$. Ο δεύτερος όρος είναι ένας γραμμικός συνδιασμός των υψηλής ανάλυσης εικονοστοιχείων για κάθε k. Αυτό το τμήμα της παραγώγου μπορεί να υπολογιστεί για όλα τα υψηλής ανάλυσης εικονοστοιχεία μέσω της διαδικασίας της συνέλιξης. Για τους συντελεστές της εξίσωσης (2.14) οι κατάλληλοι συντελεστες συνέλιξης φαίνονται στο σχήμα 2.4. Τέλος, η ενημέρωση της λύσης με χρήση της παραγώγου για κάθε εκτίμηση του εικονοστοιχείου, δοθέντος ότι οι παράμετροι κίνησης δεν μεταβάλονται έτσι ώστε $\mathbf{s} = \hat{s}^n$, είναι:

$$\hat{z}_k^{n+1} = \hat{z}_k^n - \varepsilon^n g_k(\hat{\mathbf{z}}^n, \hat{\mathbf{s}}^n) \tag{2.23}$$

για $n = 0, 1, 2, \ldots$ και $k = 1, 2, \cdots, N$. Εναλλακτικά, η ενημέρωση αυτή μπορεί να γραφεί και με την μορφή:

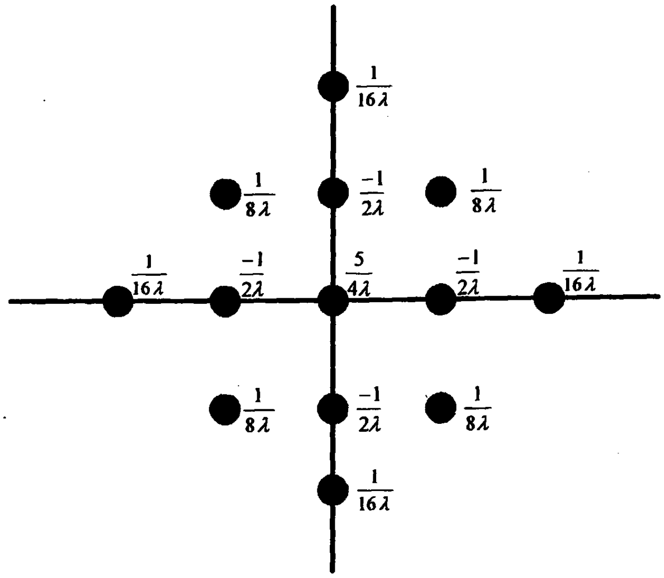
$$\hat{\mathbf{z}}^{n+1} = \hat{\mathbf{z}}^n - \varepsilon^n \nabla_{\mathbf{z}} L(\mathbf{z}, \mathbf{s})|_{\mathbf{z}=\hat{\mathbf{z}}^n, \mathbf{s}=\hat{\mathbf{s}}^n}. \tag{2.24}$$

Η παράμετρος ε^n στις εξισώσεις (2.23) και (2.24) αντιπροσωπεύει το μέγεθος του βήματος στην n-οστή επανάληψη. Εν γένει, η παράμετρος αυτή θα πρέπει να είναι αρκούντως μικρή έτσι ώστε να αποφεύγουμε την απόκλιση και αρκετά μεγάλη για να εξασφαλίζεται η σύγκλιση σε λογικό αριθμό επαναλήψεων. Το βέλτιστο βήμα εξασφαλίζεται ελαχιστοποιώντας την εξίσωση που ακολουθεί ως προς το ε^n.

$$L(\hat{\mathbf{z}}^{n+1}, \hat{\mathbf{s}}^n) = L(\hat{\mathbf{z}}^n - \varepsilon^n \nabla_{\mathbf{z}} L(\mathbf{z}, \mathbf{s})|_{\mathbf{z}=\hat{\mathbf{z}}^n, \mathbf{s}=\hat{\mathbf{s}}^n}, \hat{\mathbf{s}}^n). \tag{2.25}$$

Μετά από την ελαχιστοποίηση προκύπτει η παρακάτω εξίσωση για το μέγεθος του βήματος:

$$\varepsilon^n = \frac{\frac{1}{\sigma_\eta^2} \sum_{m=1}^{pM} \gamma_m \left(\sum_{r=1}^{N} w_{m,r}(\hat{\mathbf{s}}^n) \hat{z}_r^n - y_m \right) + \frac{1}{\lambda} \sum_{i=1}^{N} \bar{g}_i \left(\sum_{j=1}^{N} d_{i,j} \hat{z}_j^n \right)}{\frac{1}{\sigma_\eta^2} \sum_{m=1}^{pM} \gamma_m^2 + \frac{1}{\lambda} \sum_{i=1}^{N} \bar{g}_i^2} \tag{2.26}$$



Σχήμα 2.4: Συντελεστές συνέλιξης που χρησιμοποιούνται για να λάβουμε την εκ των προτέρων κατανομή της Λαπλασιανής (δεύτερη παραγώγος) της εικόνας.

όπου γ_m είναι η παράγωγος που προβάλεται πάνω στο μοντέλο των χαμηλής ανάλυσης εικονοστοιχείων και δίδεται από:

$$\gamma_m = \sum_{r=1}^{N} w_{m,r}(\hat{\mathbf{s}}^n) g_r(\hat{\mathbf{z}}^n, \hat{\mathbf{s}}^n) \tag{2.27}$$

και η ποσότητα $\bar{g}_i$, η οποία είναι το σταθμισμένο άθροισμα των γειτονικών τιμών της παραγώγου, που φαίνεται στη σχέση που ακολουθεί:

$$\bar{g}_i = \sum_{j=1}^{N} d_{i,j} g_j(\hat{\mathbf{z}}^n, \hat{\mathbf{s}}^n) \tag{2.28}$$

Μία συνολική εκτίμηση της διαδικασίας αυτής παρέχεται στον αλγόριθμο 1:

2.4 Εκτίμηση των Παραμέτρων Εξομάλυνσης

Μπορούμε να παρατηρήσουμε ότι η συνάρτηση κόστους που αναφέραμε στην υποενότητα 2.3.3 είναι μια συνάρτηση εξομάλυνσης Tikhonov [16]. Ξαναγράφουμε την συνάρτηση αυτή με τη μορφή πινάκων

$$L(\mathbf{z}, \mathbf{s}) = \|\mathbf{y} - \mathbf{W}_s \mathbf{z}\|^2 + \alpha \|\mathbf{D}\mathbf{z}\|^2$$

όπου $\alpha = \frac{\sigma_\eta^2}{\lambda}$. Η συνάρτηση κόστους που γράψαμε πριν έχει δύο όρους: ο πρώτος που εκφράζει την ακρίβεια της λύσης σε σχέση με τα δεδομένα ($\|\mathbf{y} - \mathbf{W}_s\mathbf{z}\|^2$) και ο όρος που εκφράζει την εκ των προτέρων πληροφορία της υψηλής ανάλυσης εικόνας ($\|\mathbf{D}\mathbf{z}\|^2$). Ο δεύτερος όρος αυτής της εξίσωσης περιέχει ένα υψιπερατό φίλτρο και με αυτόν τον τρόπο "αναγκάζει" την λύση να γίνει πιο ομαλή βάζοντας έναν όρο "ποινής" στις ασυνέχειες. Το σχετικό βάρος

Αλγόριθμος 1 Επαναληπτικός αλγόριθμος MAP εκτίμησης για υπερανάλυση εικόνας [10].

βήμα 1: Ξεκινάμε για $n = 0$ με την αρχική εκτίμηση της εικόνας να είναι η $\hat{\mathbf{z}}^0$ που παράγεται από παρεμβολή με το πρώτο πλαίσιο χαμηλής ανάλυσης.

βήμα 2: Για $k = 1, 2, \ldots, p$, βρίσκουμε τα $\hat{\mathbf{s}}_k$ σύμφωνα με την εξίσωση (2.21) ώστε να παραχθούν τα $\hat{\mathbf{s}}^n$

βήμα 3: Υπολογιζουμε την παράγωγο $g_k(\hat{\mathbf{z}}^n, \hat{\mathbf{s}}^n)$ από την εξίσωση (2.22) για $k = 1, 2, \ldots, N$.

βήμα 4: Υπολογίζουμε το βέλτιστο βήμα ε^n από την (2.26)

βήμα 5: Θέτουμε $\hat{z}_k^{n+1} = \hat{z}_k^n - \varepsilon^n g_k(\hat{\mathbf{z}}^n, \hat{\mathbf{s}}^n)$ για $k = 1, 2, \ldots, N$ παράγοντας την $\hat{\mathbf{z}}^{n+1}$.

βήμα 6: Αν $\|\hat{\mathbf{z}}^{n+1} - \hat{\mathbf{z}}^n\| / \|\hat{\mathbf{z}}^n\| < \epsilon$ ή έχουμε φτάσει σε ένα σύνολο από επαναλήψεις, σταματάμε.

βήμα 7: Θέτουμε $n = n + 1$ και επιστρέφουμε στο βήμα 2.

μεταξύ αυτών των δύο όρων καθορίζεται από μία παράμετρο ομαλοποίησης α, που όπως δείξαμε και πιο πάνω είναι ο λόγος της δύναμης του θορύβου σ_η^2 προς την παράμετρο λ. Στην πιο γενική περίπτωση, δεν έχουμε καμία εκ των προτέρων γνώση για κανένα από τα σ_η^2 και λ. Στην περίπτωση αυτή η παράμετρος εξομάλυνσης μπορεί να ρητά εκφραστεί ως συνάρτηση της υψηλής ανάλυσης εικόνας [14]. Συνεπώς, ξαναγράφουμε την ομαλοποιημένη συνάρτηση κόστους ως το άθροισμα κάθε ξεχωριστής συνιστώσας εξομάλυνσης για κάθε μία από τις p χαμηλής ανάλυσης εικόνες:

$$L(\mathbf{z}, \mathbf{s}) = \sum_{k=1}^{p} \left\{ \|\mathbf{y}_k - \mathbf{W}_{\mathbf{s},k}\mathbf{z}\|^2 + \alpha_k(\mathbf{z}) \|\mathbf{D}_k \mathbf{z}\|^2 \right\}. \tag{2.29}$$

Απαλείφουμε τον δείκτη k από την ποσότητα $\mathbf{D}_k$ της παραπάνω εξίσωσης καθώς $\mathbf{D} = \mathbf{D}_k$, δηλαδή το ίδιο υψηπερατό φίλτρο χρησιμοποιείται για όλα τα χαμηλής ανάλυσης πλαίσια $k = 1, 2, \cdots, p$. Σε αυτήν την περίπτωση μπορούμε να ξαναορίσουμε τη συνάρτηση για κάθε εικόνα χαμηλής ανάλυσης

$$L(\alpha_k(\mathbf{z}), \mathbf{z}, \mathbf{s}) = \|\mathbf{y}_k - \mathbf{W}_{\mathbf{s},k}\mathbf{z}\|^2 + \alpha_k(\mathbf{z}) \|\mathbf{D}\mathbf{z}\|^2. \tag{2.30}$$

για $k = 1, 2, \cdots, p$. Παράλληλα θέτουμε τις ακόλουθες απαιτήσεις για το $\alpha_k(\mathbf{z})$: θα πρέπει να είναι συνάρτηση του θορύβου και η επιλογή του να παράγει μία κυρτή συνάρτηση, την οποία όταν ελαχιστοποιούμε να μας δίνει την υψηλής ανάλυσης εικόνα. Στις εξισώσεις που έχουμε αναφερθεί μέχρι τώρα ισχύει ότι $\alpha(\mathbf{z}) = \sum_{k=1}^{p} \alpha_k(\mathbf{z})$.

Με βάση τις προϋποθέσεις που θέσαμε για την παράμετρο εξομάλυνσης προκύπτει μια γραμμική συνάρτηση μεταξύ της $\alpha_x(\mathbf{z})$ και κάθε όρου της συνάρτησης κόστους. Τελικά, λύνοντας ως προς την παράμετρο αυτή προκύπτει η παρακάτω σχέση:

$$\alpha_k(\mathbf{z}) = \frac{\|\mathbf{y}_k - \mathbf{W}_{s,k}\mathbf{z}\|^2}{\frac{1}{\gamma_k} - \|\mathbf{D}\mathbf{z}\|^2} \tag{2.31}$$

Επίσης, ακολουθώντας την την διαδικασία της σύγκλισης των κριτηρίων όπως στην εργασία [14] παίρνουμε:

$$\frac{1}{\gamma_k} > \frac{\varepsilon p\|\mathbf{y}_k - \mathbf{W}_{s,k}\mathbf{z}\|^2\phi_{\max}(\mathbf{D}^T\mathbf{D})}{2 - \varepsilon p\ \phi_{\max}(\mathbf{W}_{s,k}^T\mathbf{W}_{s,k})} + \|\mathbf{D}\mathbf{z}\|^2 \tag{2.32}$$

όπου $\phi_{\max}(\cdot)$ είναι η μέγιστη ιδιοτιμή του πίνακα (που εσωκλείεται στις παρενθέσεις). Επομένως λαμβάνοντας υπ' όψιν και την εξίσωση (2.6) προκύπτει:

$$\phi_{\max}(\mathbf{W}_{s,k}^T\mathbf{W}_{s,k}) = \phi_{\max}(\mathbf{B}_k^T\mathbf{M}_k^T\mathbf{S}^T\mathbf{B}_k\mathbf{M}_k\mathbf{S}). \tag{2.33}$$

Αν κάνουμε υποδειγματοληψία παίρνοντας τόν μέσο όρο των εικονοστοιχείων μπορούμε πολύ εύκολα να δείξουμε ότι ισχύει

$$\mathbf{S}^T\mathbf{S} = \frac{1}{(L_1L_2)^2}\mathbf{I}$$

όπου $\mathbf{I}$ είναι ο $N \times N$ μοναδιαίος πίνακας. Μιας και δεν χάνεται αλλά ούτε και προστίθεται πληροφορία κατά την διάρκεια της κίνησης $\mathbf{M}_k$, τα στοιχεία του πίνακα αυτού είναι άσσοι και μηδενικά, με κάθε στήλη και κάθε γραμμη να περιέχει μόνο έναν άσσο. Στην περίπτωση αυτή, ισχύει

$$\mathbf{M}_k^T\mathbf{M}_k = \mathbf{I}$$

Χρησιμοποιώντας Gaussian φίλτρα για θόλωση, τότε υποθέτουμε ότι οι συντελεστές απόκρισης κανονικοποιούνται στην μονάδα, το οποίο είναι ισοδύναμο με

$$\phi_{\max}(\mathbf{B}_k^T\mathbf{B}_k) = 1.$$

Αντικαθιστώντας τις τρεις αυτές τελευταίες εξισώσεις στην παρακάτω εξίσωση πλέον έχουμε:

$$\phi_{\max}(\mathbf{W}_{s,k}^T\mathbf{W}_{s,k}) = \frac{1}{(L_1L2)^2}\phi_{\max}(\mathbf{B}_k^T\mathbf{B}_k) = \frac{1}{(L_1L2)^2} \tag{2.34}$$

Συνεπώς η ανισότητα (2.32) μεταβάλλεται και γίνεται:

$$\frac{1}{\gamma_k} > \frac{\varepsilon p\ \phi_{\max}(\mathbf{D}^T\mathbf{D})}{2 - \left(\varepsilon p/(L_1L_2)^2\right)}\|\mathbf{y}_k - \mathbf{W}_{s,k}\mathbf{z}\|^2 + \|\mathbf{D}\mathbf{z}\|^2 \tag{2.35}$$

Τώρα μπορούμε να ορίσουμε το μέγεθος του βήματος ε έτσι ώστε:

$$\frac{\varepsilon p\ \phi_{\max}(\mathbf{D}^T\mathbf{D})}{2 - \left(\varepsilon p/(L_1L_2)^2\right)} = 1$$

Η σχέση για την οποία ισχύει το παραπάνω είναι:

$$\varepsilon = \frac{2}{p}\left(\frac{(L_1L_2)^2}{(L_1L_2)^2\phi_{\max}(\mathbf{D}^T\mathbf{D}) + 1}\right) \tag{2.36}$$

Η ανισότητα (2.35) απλοποιείται και γίνεται:

$$\frac{1}{\gamma_k} > \|\mathbf{y}_k - \mathbf{W}_{s,k}\mathbf{z}\|^2 + \|\mathbf{D}\mathbf{z}\|^2 \tag{2.37}$$

Τώρα έχουμε ότι $\|\mathbf{y}_k\|^2 \geq \|\mathbf{y}_k - \mathbf{W}_{s,k}\mathbf{z}\|^2$, μιας και έχουμε υποθέσει ότι οι χαμηλής ανάλυσης εικόνες έχουν μεγαλύτερη ενέργεια από τον πρόσθετο θόρυβο και επίσης ισχύει $\|\mathbf{y}_k\|^2 \approx \frac{\|\mathbf{z}\|^2}{L_1 L_2} > \|\mathbf{D}\mathbf{z}\|^2$ για μικρούς λόγους υποδειγματοληψίας $L_1 = L_2 = 2$, αφού έχουμε υποθέσει ότι η εικόνα $\mathbf{z}$ έχει πολύ λίγη ενέργεια στις υψηλές συχνότητες απ' ότι στις χαμηλές συχνότητες και κάθε χαμηλής ανάλυσης εικόνα $\mathbf{y}_k$ έχει $1/L_1L_2$ από την ενέργεια της $\mathbf{z}$ για περιπτώσεις χωρίς θόρυβο. Στην παρούσα εργασία χρησιμοποιούμε λόγο υποδειγματοληψίας $L_1 = L_2 = 2$. Επομένως για αυτήν την επιλογή μπορούμε να ορίσουμε την ποσότητα:

$$\frac{1}{\gamma_k} = 2\|\mathbf{y}_k\|^2 \tag{2.38}$$

Η παραπάνω ποσότητα ικανοποιεί την συνθήκη της σύγκλισης και επίσης δίνει ένα συντελεστή α

$$\alpha_k(\mathbf{z}) = \frac{\|\mathbf{y}_k - \mathbf{W}_{s,k}\mathbf{z}\|^2}{2\|\mathbf{y}_k\|^2 - \|\mathbf{D}\mathbf{z}\|^2} \tag{2.39}$$

2.5 Μπεϋζιανές Μέθοδοι Υπερανάλυσης Εικόνας

Έχουν προταθεί διάφορες Μπεϋζιανές τεχνικές για υπερανάλυση εικόνας. Θα παρουσιάσουμε μερικές από αυτές οι οποίες ανάλογα με το κριτήριο που χρησιμοποιούν για την εύρεση των παραμέτρων υπέρθεσης ανήκουν σε διαφορετική κατηγορία.

2.5.1 Περιθωριοποίηση Εικόνας Υψηλής Ανάλυσης

Η μέθοδος που έχει προταθεί στην βιβλιογραφία [33] υιοθετεί την Μπεϋζιανή προσέγγιση κάνοντας περιθωριοποίηση της άγνωστης εικόνας υψηλής ανάλυσης.

Με την μέθοδο αυτή, βελτιστοποιήουμε την περιθώρια πιθανοφάνεια ως προς τις παραμέτρους υπέρθεσης και αυτό γίνεται μπορεί να γίνει με δύο τρόπους. Η πρώτη προσέγγιση είναι να χρησιμοποιήσουμε τον αλγόριθμο αναμενόμενης μεγιστοποίησης (*expectation maximization EM*) [2]. Στο E-step εκτιμάμε την εκ των υστέρων κατανομή της υψηλής ανάλυσης εικόνας. Στο M-step μεγιστοποιούμε την πρόβλεψη της $\mathbf{z}$. Η μεγιστοποίηση αυτή μπορεί να γίνει με τον αλγόριθμο τον συζιγών κλίσεων (scaled conjugate gradients SCG). Η δεύτερη προσέγγιση είναι να μεγιστοποιήσουμε την περιθώρια πιθανοφάνεια κατευθείαν χρησιμοποιώντας τον αλγόριθμο SCG. Εν γένει, θεωρείται ότι η απευθείας μεγιστοποίηση είναι ταχύτερη από τον αλγόριθμο EM.

2.5.2 Περιθωριοποίηση Παραμέτρων Υπέρθεσης

Στην υποενότητα αυτή παρουσιάζουμε μια τεχνική για υπερανάλυση εικόνας και η οποία χρησιμοποιεί Bayesian μεθόδους περιθωριοποιώτας τις άγνωστες παραμέτρους υπέρθεσης, οι οποίες σχετίζονται με το σύνολο των χαμηλής ανάλυσης δεδομένων. Η διαφορά με την εργασία των Tipping και Bishop [33] που παρουσιάστηκε προηγουμένως, είναι ότι σε εκείνη η περιθωριοποίηση γίνεται πάνω στην υψηλής ανάλυσης εικόνα, "αναγκάζοντάς" μας σε μία δυσμενή χρήση μιας εκ των προτέρων κατανομής για την εικόνα. Ολοκληρώνοντας ως προς τις παραμέτρους υπέρθεσης και οχι ως προς την εικόνα, η μέθοδος αυτή δίνει μια πιο ρεαλιστική κατανομή και επίσης μειώνει την διάσταση τοτ ολοκληρλωματος αισθητά, απαλείφοντας το κύριο υπολογιστικό κόστος του προηγούμενου αλγορίθμου. Σε αντίθεση με το μοντέλο κίνησης που χρησιμοποιείται στην [33], εδώ παρουσιάζεται ένα πιο γενικό μοντέλο που επιτρέπει αλλαγές στον φωτισμό αλλά και την κίνηση. Ένα πλεονέκτημα της προσέγγισης αυτής είναι ότι οι απαιτήσεις σε μνήμη μειώνονται αισθητά σε σχέση με την προηγούμενη μέθοδο, παρ' όλα αυτά είναι πιο χρονοβόρα σε σχέση με τεχνικές υπερανάλυσης εικόνας που χρησιμοποιούν μέθοδο MAP. Για περισσότερες πληροφορίες σχετικά με την μέθοδο αυτή παραπέμπουμε τον αναγνώστη στην αναφορά [27], καθώς δεν θα αναφερθούμε εκτενέστερα σε μεθόδους αυτής της κατηγορίας.

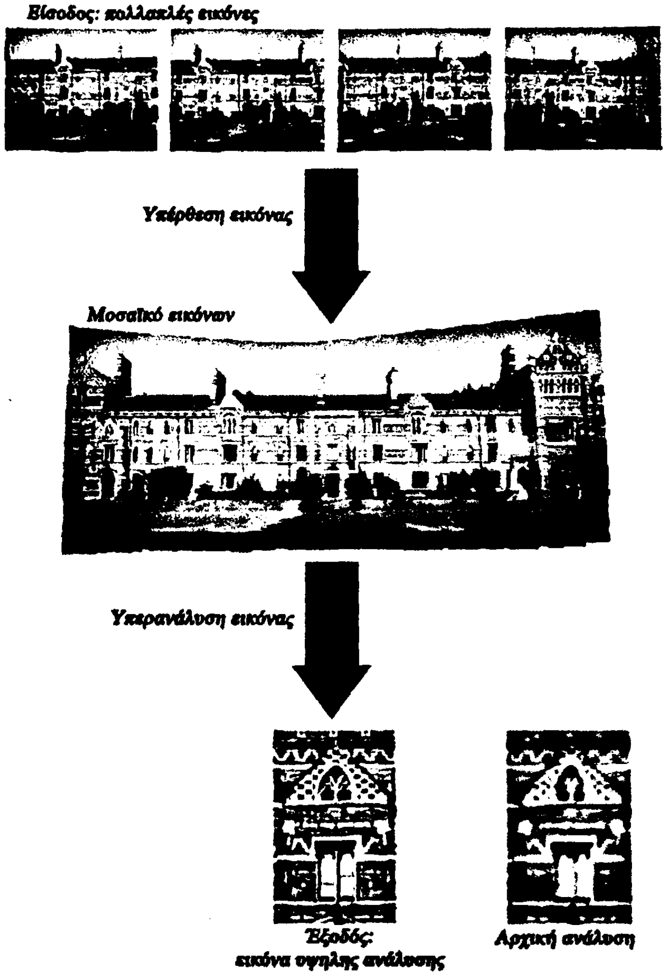
2.6 Υπέρθεση Βασισμένη σε Χαρακτηριστικά Σημεία της Εικόνας

Ένα πρόβλημα που προκύπτει και ανήκει στο πεδίο της Υπολογιστικής Όρασης, είναι το πρόβλημα της υπέρθεσης εικόνας με βαση κάποια χαρακτηριστικά σημεία πάνω στην εικόνα αυτή. Απαραίτητο για την επιτυχεία κάθε αλγορίθμου υπερανάλυσης εικόνας είναι η εύρεση αντίστοιχων σημείων μεταξύ των εικόνων στην ακολουθεία εισόδου. Το πρόβλημα της "αντιστοίχισης" μπορεί να δοθεί ως εξής: *δοθέντων δύο διαφορετικών όψεων της ίδιας σκηνής, για κάθε σημείο στην μία όψη να βρούμε το αντιστοιχό του στην άλλη όψη.*

Μια ευρέως διαδεδομένη τεχνική για υπερανάλυσης εικόνας είναι που σχετίζεται με το προβλημα της ομογραφίας δύο ή περισσοτέρων εικόνων, όπων φαίνεται στο σχήμα 2.5. Αρχικά οι εικόνες εισόδου ευθηγραμμίζονται μεταξύ τους πάνω σε μία κοινή εικόνα αναφοράς. Η φάση της ευθηγράμμισης των εικόνων, περιλαμβάνει τόσο γεωμετρικές συνιστώσες όσο και φωτομετρικές συνιστώσες που έχουν να κάνουμε με την φωτεινότητα της εικόνας. Μετά από το βήμα αυτό ακολουθεί το βήμα της ομογραφίας των εικόνων, συνδιάζοντας όλες τις εικόνες εισόδου σε μία ώστε να προκύψει μία σκήνη και στην συνέχεια εφαρμόζουμε τον αλγόριθμο της υπερανάλυσης εικόνας, σε οποιαδήποτε περιοχή ενδιαφέροντος.

Όσον αφορά την υπέρθεση αναφερόμαστε σε επίπεδους προβολικούς μετασχηματισμούς, που επίσης ονομάζονται ως επίπεδη ομογραφία (***planar homography***), ένας γεωμετρικός μετασχηματισμός έχει οχτώ βαθμούς ελευθερίας. Στις εξίσώσεις που ακουλουθούν αναφέρουμε τους στοιχειώδεις συμβολισμούς. Θεωρούμε ότι τα σημεία μας αναπαριστώνται σε ομογενείς συντεταγμένες, επομένως το σημείο (x, y) αναπαριστάται ως $(x, y, 1)$. Ο προβολικός μετασχηματισμός σημείων είναι:

$$\begin{pmatrix} x_1' \\ x_2' \\ x_3' \end{pmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$



Σχήμα 2.5: Φάσεις υπερανάλυσης εικόνας με χρήση της τεχνικής της ομογραφίας.

ή ισοδύναμα $\mathbf{x}' = \mathbf{H}\mathbf{x}$.

Στο πεδίο της υπολογιστικής όρασης είναι αρκετά συνηθησμένο να προσπαθούμε να εκτιμήσουμε τις παραμέτρους ενός γεωμετρικού μετασχηματισμού, όπως η ομογραφία $\mathbf{H}$ με αυτόματο εντοπισμό αντίστοιχων χαρακτηριστικών σημείων στις εικόνες εισόδου. Εν γένει, μέσα σε μια εικόνα υπάρχουν εκατοντάδες σημεία ενδιαφέροντος τα οποία μπορούν με διάφορους τρόπους να υπολογιστούν [18], χρησιμοποιώντας για παράδειγμα τους περιγραφείς SIFT, τους οποίους θα παρουσιάσουμε αναλυτικά σε επόμενο κεφάλαιο. Επειδή τα εξαγώμενα χαρακτηριστικά σημεία είναι πολλά, λόγω της μετόπισης και της περιστροφής των χαμηλής ανάλυσης πλαισίων, ενδέχεται να απορρίψουμε κάποια από αυτά καθώς δεν υπάρχουν υπάρχουν τα αντιστοιχά τους. Για το λόγω αυτό, ακολουθεί ένα βήμα κατά το οποίο βρίσκουμε τα σημεία εκείνα που συμφωνούν με την ομογραφία χρησιμοποιώντας για παράδειγμα τον αλγόριθμο RANSAC [11]. Η συνολική διαδικασία που ακολουθείται φαίνεται στον αλγόριθμο που ακολουθεί:

Για περισσότερες λεπτομέριες σχετικά με την τεχνική της ομογραφίας και της εύρεσης χαρακτηριστικών σημείων σε μία εικόνα παραπέμπουμε στις εργασίες των D. Capel και A. Zisserman [11] και David G. Lowe [18].

Αλγόριθμος 2 Αυτόματη υπέρθεση δύο εικόνων.

βήμα **1:** Υπολογισμός των σημείων ενδιαφέροντος για κάθε εικόνα (περιγραφείς SIFT).

βήμα **2:** Υπολογίζουμε ένα σύνολο σημείων ενδιαφέροντος τα οποία ταιριάζουν στην γειτονιά με αυτά της γειτονιάς των υπολοίπων εικόνων.

βήμα **3:** Χρήση αλγορίθου RANSAC. Επαναλαμβάνουμε τα παρακάτω για N δείγματα.

(α) Επιλέγουμε 4 τυχαία αντίστοιχα σημεία και υπολογίζουμε την ομογραφία $\mathbf{H}$.

(β) Υπολογίζουμε ένα σφάλμα απόστασης για κάθε αντιστοιχία.

(γ) Υπολογίζουμε τον αριθμό σημείων εκείνων που σημφωνούν με την $\mathbf{H}$ από τα αντίστοιχα σημεία για τα οποία το σφάλμα απόστασης είναι μικρότερό από ένα κατώφλι.

βήμα **4:** Επαναϋπολογίζουμε τον $\mathbf{H}$ από όλα τα αντίστοια σημεία που θεωρούνται ως αποδεκτά.

βήμα **5:** Επιπλέον στοιχεία ενδιαφέροντος καθορίζονται χρησιμοποιώντας την εκτίμηση του $\mathbf{H}$ για να ορίσουμε μια περιοχή αναζήτησης.
Τα δυο τελευταία βήματα επαναλαμβάνονται μέχρι ο αριθμός των αντιστοιχιών να είναι σταθερός.

2.7 Πειραματικά Αποτελέσματα

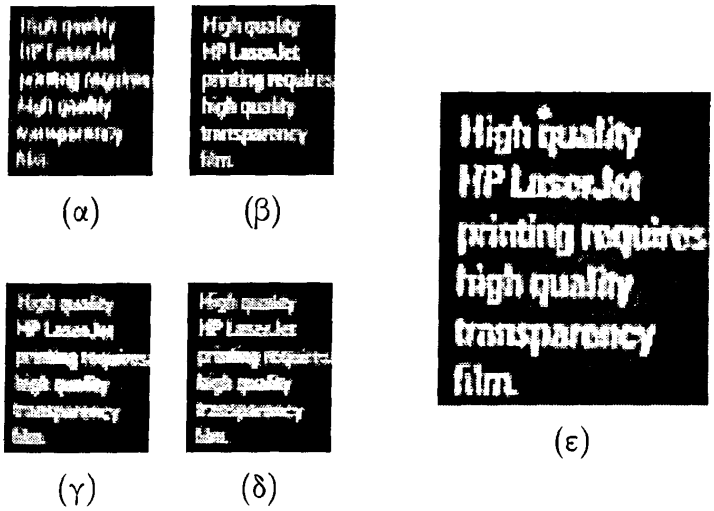
Κλείνοντας το κεφάλαιο αυτό θα παρουσιάσουμε κάποια πειραματικά αποτελέσματα για την συμπεριφορά της μεθόδου της υπερανάλυσης εικόνας που αναλύσαμε παραπάνω στον αλγόριθμο 1.

Στα πειράματά μας έχουμε χρησιμοποιήσει εικόνες που τις έχουμε θολώσει με Gaussian χαμηλοπερατό φίλτρο τυπικής απόκλισης 1 και μέγεθος παραθύρου 5×5 ενώ παράλληλα έχουν υποβαθμιστεί με ένα παράγοντα $L_1 = L_2 = 2$. Το μέγιστο πλήθος επαναλήψεων που χρησιμοποιήσαμε είναι 100. Οι χαμηλής ανάλυσης εικόνες έχουν τυχαίες μετατοπίσεις που κυμαίνονται στο διάστημα $[-3, 3]$. Για την εύρεση των διανυσμάτων κίνησης χρημοποιήσαμε έναν απλό αλγόριθμο ταιριάσματος κατά μπλοκ με μέγεθος παραθύρου αναζήτησης $[-4, 4]$ και βήμα αναζήτησης 0.15. Οι μετρικές που χρησιμοποιούμε για την ποσοτικοποίηση της ποιότητας των υποβαθμισμένων εικόνων, τα επίπεδα του θορύβου στις υποβαθμισμένες εικόνες χαμηλής ανάλυσης και την ποιότητα των ανακατασκευασμένων αποτελεσμάτων είναι η μέγιστη τιμή του λόγου σήματος προς θόρυβο (*PSNR*). Οι μετρικές αυτές ορίζονται ως:

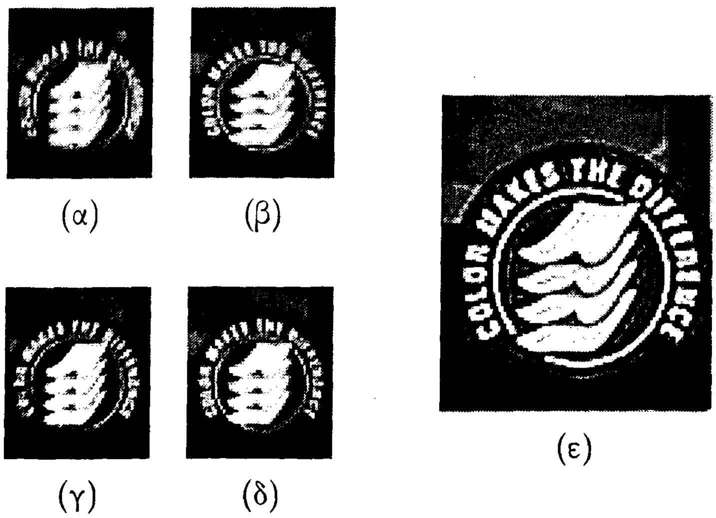
$$PSNR = 10 \log_{10} \frac{(255)^2}{\|\mathbf{f} - \mathbf{g}\|^2}$$

όπου $\mathbf{f}$ και $\mathbf{g}$ είναι η πραγματική και η ανακατασκευασμένη εικόνα, αντιστοίχως. Για την δημιουργία των τεχνητών παραδειγμάτων χρησιμοποιήσαμε προσθετικό Gaussian λευκό θόρυβο.

Η δομή με την οποία θα παρουσιάσουμε τα αποτελέσματά μας είναι η ακόλουθη: δίνονται οι χαμηλής ανάλυσης εικόνες από τις οποίες θα προκύψει το αποτέλεσμα, παρουσιάζεται η ανακατασκευασμένη εικόνα υψηλής ανάλυσης, δύο πίνακες που περιέχουν τις τιμές των σφαλμάτων για τα διανύσματα κίνησης (μέση τιμή, τυπική απόκλιση και ενδιάμεσση τιμή) και δύο πίνακες που περιέχουν αριθμητικά αποτελέσματα (μέση τιμή, τυπική απόκλιση και ενδιάμεση τιμή) για το PSNR, όπως προέκυψαν από τον αλγόριθμο 1. Στους πίνακες αυτούς τα πειράματα έχουν γίνει με 30 dB και 20 dB θόρυβο, αντίστοιχα.

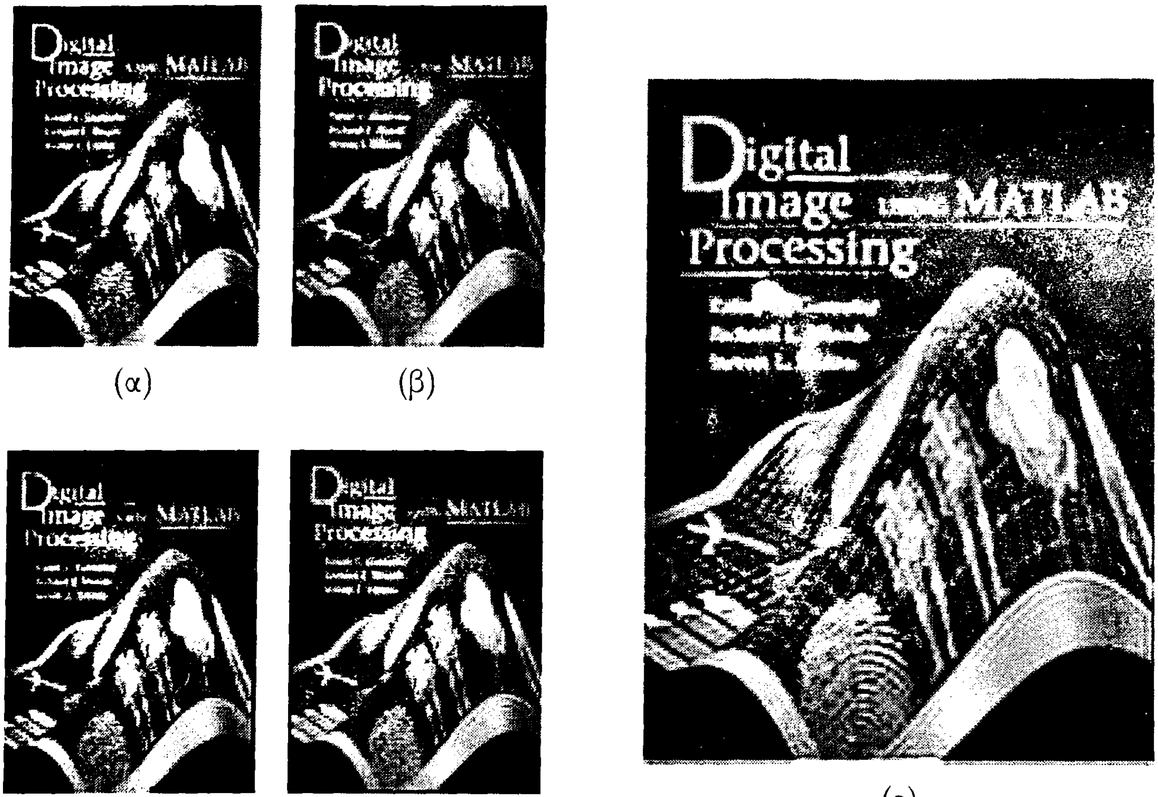


Σχήμα 2.6: *Κείμενο.* (α) – (δ) Εικόνες χαμηλής ανάλυσης. (ε) Ανακατασκευασμένη εικόνα υψηλής ανάλυσης. Ο αριθμός των χαμηλής ανάλυσης εικόνων που χρησιμοποιήθηκαν για την ανακασκευή της (ε) είναι 30.



Σχήμα 2.7: *Δίσκος.* (α) – (δ) Εικόνες χαμηλής ανάλυσης. (ε) Ανακατασκευασμένη εικόνα υψηλής ανάλυσης. Ο αριθμός των χαμηλής ανάλυσης εικόνων που χρησιμοποιήθηκαν για την ανακασκευή της (ε) είναι 20.

Οι χαμηλής ανάλυσης εικόνες των σχημάτων 2.6 και 2.7 πάρθηκαν από την ιστοσελ του Peyman Milanfar [1]. Γι' αυτό το σύνολο δεδομένων δεν γνωρίζουμε τις πραγματι εικόνες υψηλής ανάλυσης (*ground truth*).

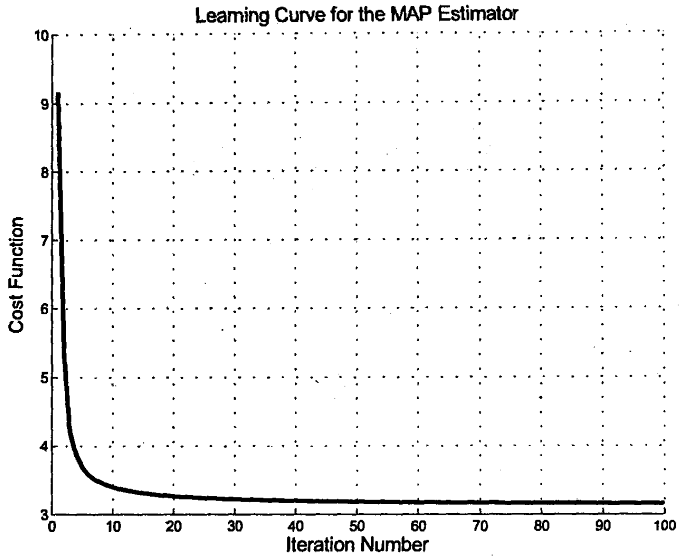


(α) (β) (γ) (δ) (ε)

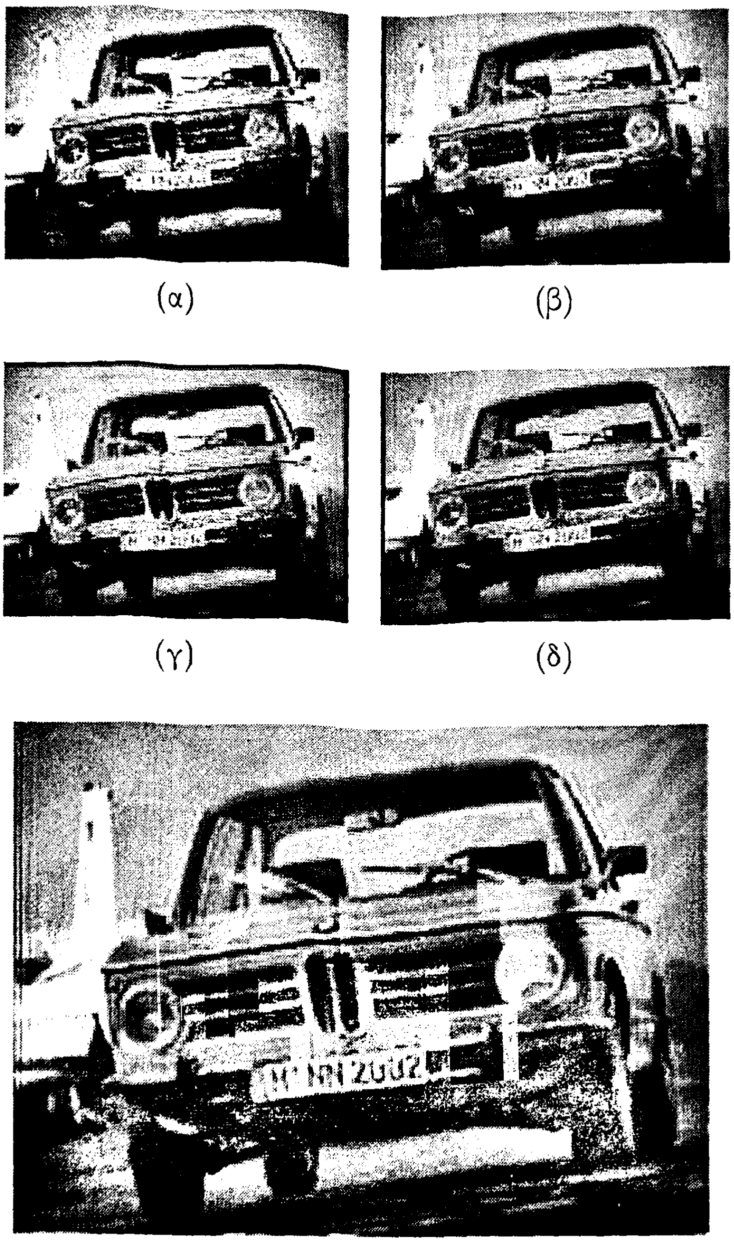
Σχήμα 2.8: *Εξώφυλλο.* (α) – (δ) Εικόνες χαμηλής ανάλυσης. (ε) Ανακατασκευασμένη εικό υψηλής ανάλυσης. Ο αριθμός των χαμηλής ανάλυσης εικόνων που χρησιμοποιήθηκαν για τ ανακασκευή της (ε) είναι 4. Η ανακατασκευασμένη εικόνα υψηλής ανάλυσης έχει PSNR 22. dB.

Οι χαμηλής ανάλυσης εικόνες που παρουσιάζονται στα σχήματα 2.8, 2.10, 2.12, 2.1 και 2.16 παράχθηκαν με θόρυβο 30 dB, ενώ οι χαμηλής ανάλυσης εικόνες του σχήματ 2.18 παράχθηκαν με μεγαλύτερο θόρυβο της τάξεως των 20 dB. Το σύνολο πειραμάτων γι κάθε εικόνα είναι 10.

[1] `http://users.soe.ucsc.edu/~milanfar/software/sr-datasets.html`



Σχήμα 2.9: Η συνάρτηση κόστους $L(\mathrm{z}, \mathrm{s})$ σε σχέση με τον αριθμό των επαναλήψεων για την εικόνα *Εξώφυλλο* του σχήματος 2.8.

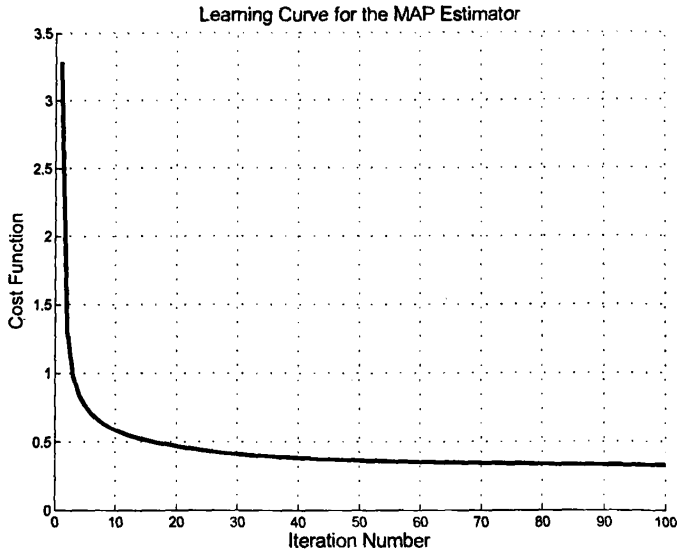


(α) (β)

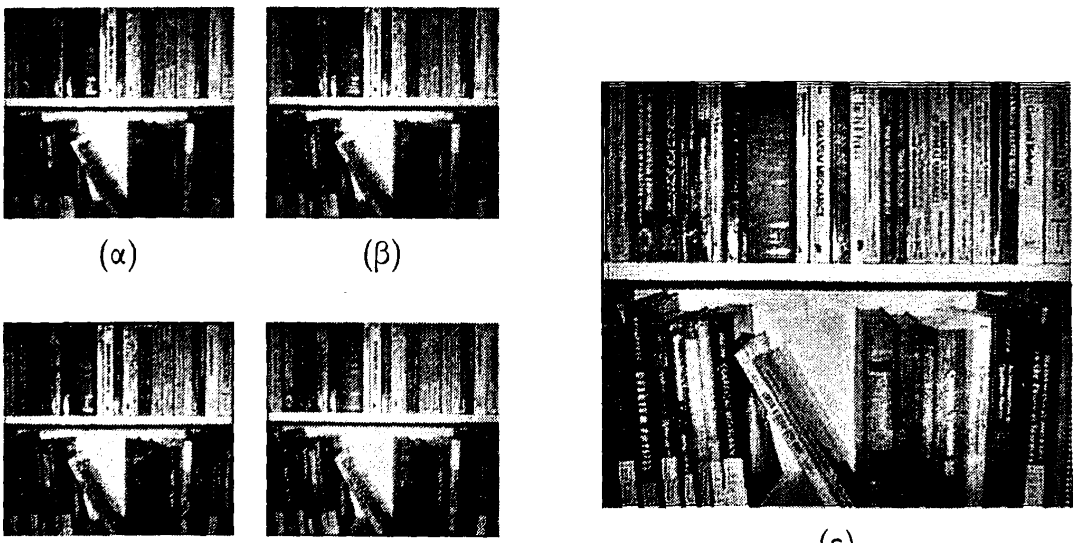
(γ) (δ)

(ε)

Σχήμα 2.10: *Αυτοκίνητο.* (α) – (δ) Εικόνες χαμηλής ανάλυσης. (ε) Ανακατασκε
υψηλής ανάλυσης. Ο αριθμός των χαμηλής ανάλυσης εικόνων που χρησιμοποι
ανακασκευή της (ε) είναι 4. Η ανακατασκευασμένη εικόνα υψηλής ανάλυσης έ;
dB.



Σχήμα 2.11: Η συνάρτηση κόστους $L(\mathrm{z}, \mathrm{s})$ σε σχέση με τον αριθμό των επαναλήψεων για την εικόνα *Αυτοκίνητο* του σχήματος 2.10.

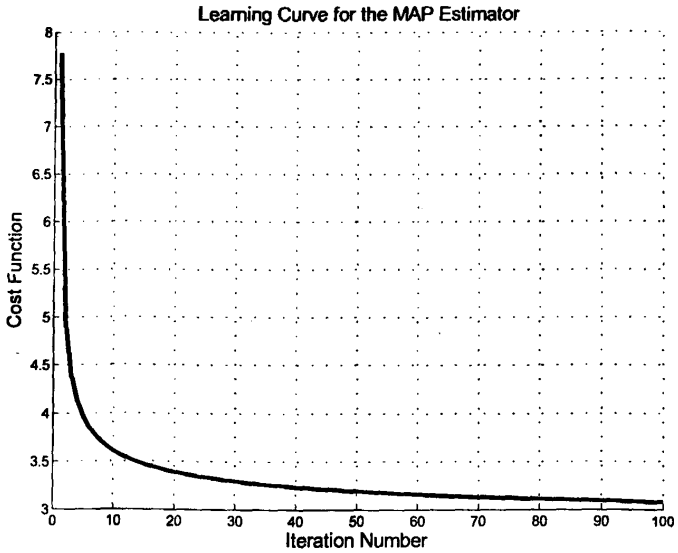


(α) (β)

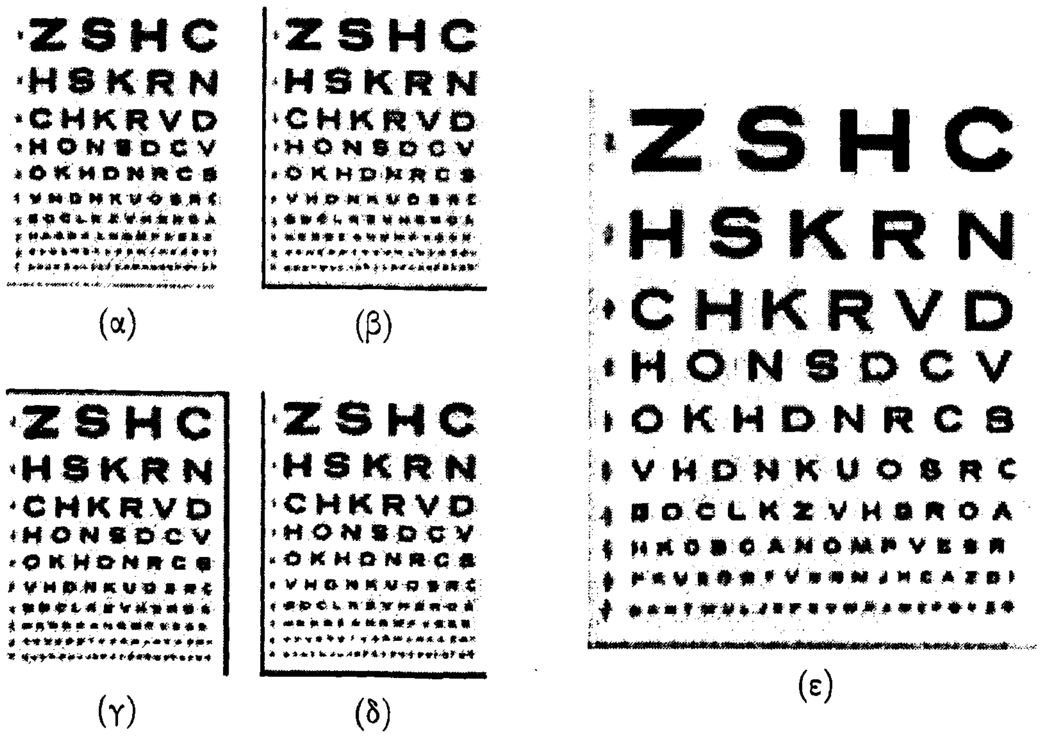
(γ) (δ)

(ε)

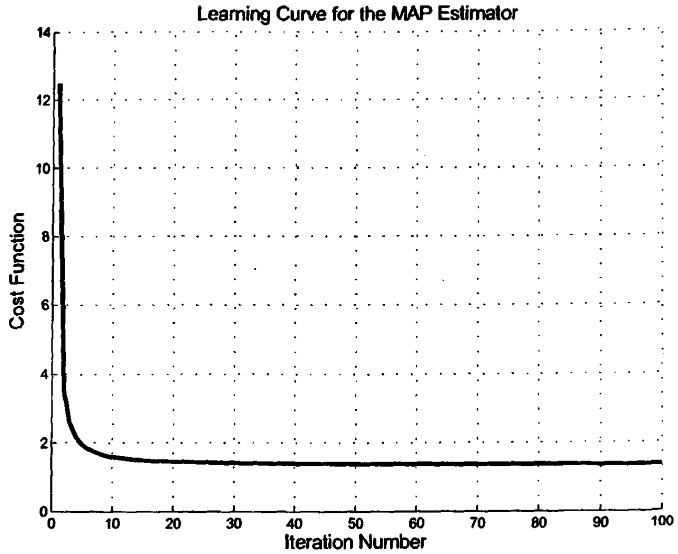
Σχήμα 2.12: *Βιβλία.* (α) – (δ) Εικόνες χαμηλής ανάλυσης. (ε) Ανακατασκευασμένη εικόνα υψηλής ανάλυσης. Ο αριθμός των χαμηλής ανάλυσης εικόνων που χρησιμοποιήθηκαν για την ανακασχευή της (ε) είναι 4. Η ανακατασκευασμένη εικόνα υψηλής ανάλυσης έχει PSNR 21.14 dB.



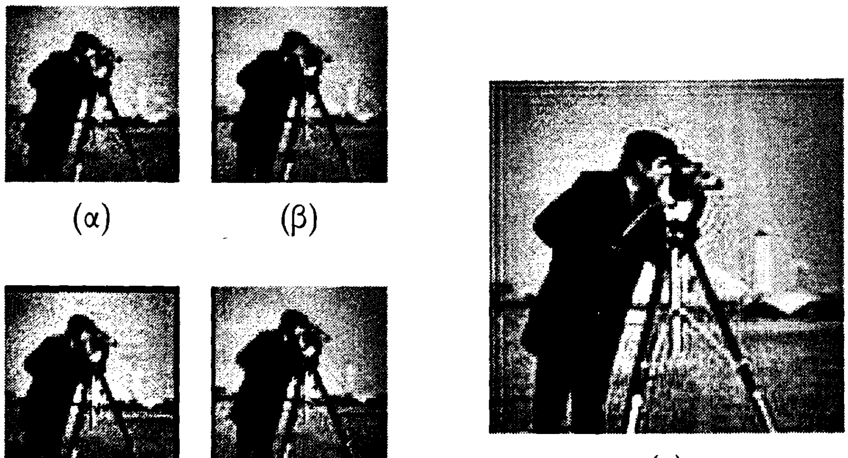
Σχήμα 2.13: Η συνάρτηση κόστους $L(\mathrm{z}, \mathrm{s})$ σε σχέση με τον αριθμό των επαναλήψεων για την εικόνα *Βιβλία* του σχήματος 2.12.



Σχήμα 2.14: *Πίνακας οφθαλμίατρου.* (α) – (δ) Εικόνες χαμηλής ανάλυσης. (ε) Ανακατασκευασμένη εικόνα υψηλής ανάλυσης. Ο αριθμός των χαμηλής ανάλυσης εικόνων που χρησιμοποιήθηκαν για την ανακασκευή της (ε) είναι 4. Η ανακατασκευασμένη εικόνα υψηλής ανάλυσης έχει PSNR 23.90 dB.

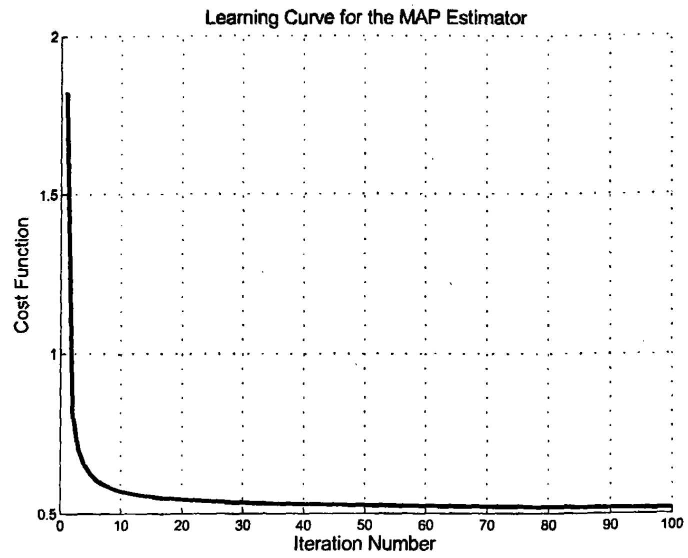


Σχήμα 2.15: Η συνάρτηση κόστους $L(\mathbf{z}, \mathbf{s})$ σε σχέση με τον αριθμό των επαναλήψεων για την εικόνα *Πίνακας οφθαλμίατρου* του σχήματος 2.14.

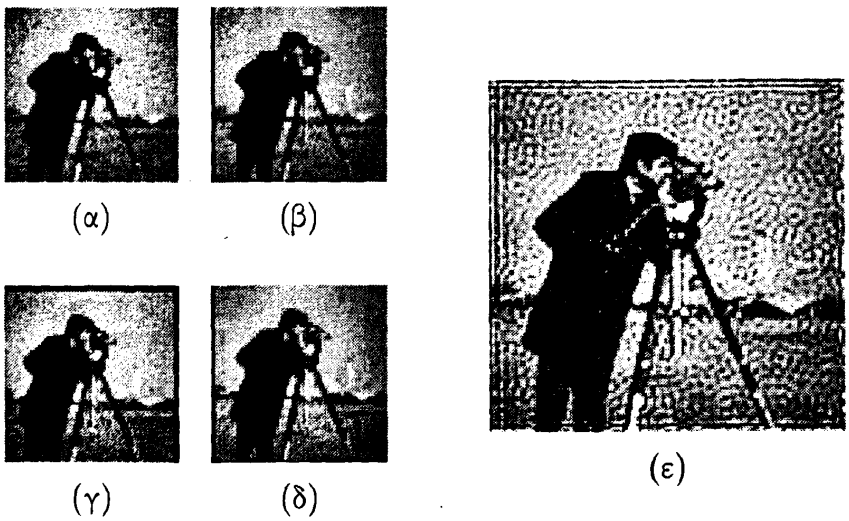


(α) (β) (γ) (δ) (ε)

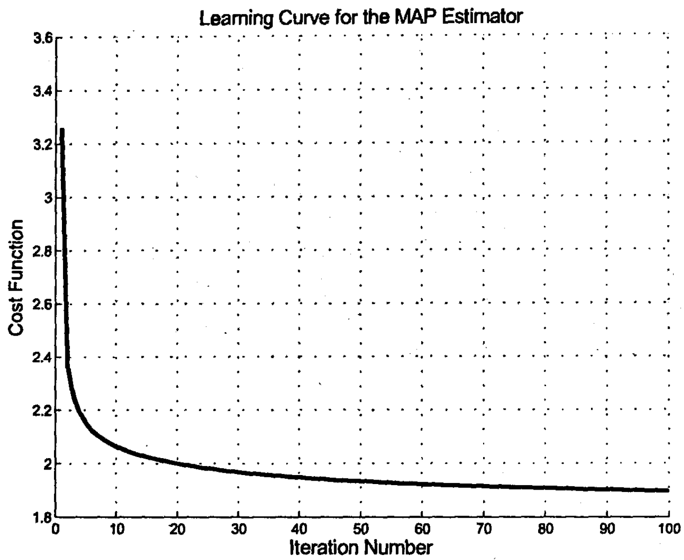
Σχήμα 2.16: *Cameraman 1.* (α) – (δ) Εικόνες χαμηλής ανάλυσης. (ε) Ανακατασκευασμένη εικόνα υψηλής ανάλυσης. Ο αριθμός των χαμηλής ανάλυσης εικόνων που χρησιμοποιήθηκαν για την ανακασκευή της (ε) είναι 4. Η ανακατασκευασμένη εικόνα υψηλής ανάλυσης έχει PSNR 26.39 dB.



Σχήμα 2.17: Η συνάρτηση κόστους $L(\mathbf{z}, \mathbf{s})$ σε σχέση με τον αριθμό των επαναλήψεων για την εικόνα *Cameraman 1* του σχήματος 2.16.



Σχήμα 2.18: *Cameraman 2.* (α) – (δ) Εικόνες χαμηλής ανάλυσης. (ε) Ανακατασκευασμένη εικόνα υψηλής ανάλυσης. Ο αριθμός των χαμηλής ανάλυσης εικόνων που χρησιμοποιήθηκαν για την αναχασχευή της (ε) είναι 4. Η ανακατασκευασμένη εικόνα υψηλής ανάλυσης έχει PSNR 23.83 dB.



Σχήμα 2.19: Η συνάρτηση κόστους $L(\mathrm{z}, \mathrm{s})$ σε σχέση με τον αριθμό των επαναλήψεων για την εικόνα *Cameraman 2* του σχήματος 2.18.

Πίνακας 2.1: Αριθμητικά αποτελέσματα για τις παραμέτρους υπέρθεσης που παρουσιάζονται στα σχήματα 2.8, 2.10, 2.12, 2.14 και 2.16 με μέγεθος θορύβου 30 dB.

Εικόνα	Σφάλμα s^x			Σφάλμα s^y		
	mean	**std**	**median**	**mean**	**std**	**median**
Εξώφυλλο (σχ. 2.8)	0.39	0.14	0.30	0.11	0.06	0.22
Αυτοκίνητο (σχ. 2.10)	0.18	0.19	0	0.21	0.33	0
Βιβλία (σχ. 2.12)	-0.15	0.02	0.11	0.43	0.13	0.08
Πίνακας Οφθαλμίατρου (σχ. 2.14)	-0.10	0.26	-0.02	-0.40	0.11	-0.45
Cameraman 1 (σχ. 2.16)	-0.43	0.39	-0.37	0.06	0.02	0.1

Πίνακας 2.2: Αριθμητικά αποτελέσματα για τα PSNR που παρουσιάζονται στα σχήματα 2.8, 2.10, 2.12, 2.14 και 2.16 με μέγεθος θορύβου 30 dB.

Εικόνα	PSNR		
	mean	**std**	**median**
Εξώφυλλο (σχ. 2.8)	22.35	0.88	22.06
Αυτοκίνητο (σχ. 2.10)	22.08	0.86	21.82
Βιβλία (σχ. 2.12)	20.88	0.59	20.67
Πίνακας Οφθαλμίατρου (σχ. 2.14)	21.41	0.67	20.78
Cameraman 1 (σχ. 2.16)	25.88	0.39	25.81

Πίνακας 2.3: Αριθμητικά αποτελέσματα για τις παραμέτρους υπέρθεσης που παρουσιάζονται στα σχήματα 2.8, 2.10, 2.12, 2.14 και 2.18 με μέγεθος θορύβου 20 dB.

Εικόνα	Σφάλμα s^x			Σφάλμα s^y		
	mean	**std**	**median**	**mean**	**std**	**median**
Εξώφυλλο (σχ. 2.8)	0.09	0.20	0.01	0.03	0.04	0
Αυτοκίνητο (σχ. 2.10)	0.21	0.01	0.12	-0.27	0.10	-0.24
Βιβλία (σχ. 2.12)	-0.06	0.23	0	-0.11	0.02	-0.03
Πίνακας Οφθαλμίατρου (σχ. 2.14)	-0.25	0.05	-0.21	-0.33	0.09	-0.29
Cameraman 2 (σχ. 2.18)	-0.54	0.15	-0.62	0.15	0.22	0.37

Πίνακας 2.4: Αριθμητικά αποτελέσματα για τα PSNR που παρουσιάζονται στα σχήματα 2.8, 2.10, 2.12, 2.14 και 2.18 με μέγεθος θορύβου 20 dB.

Εικόνα	**PSNR**		
	mean	**std**	**median**
Εξώφυλλο (σχ. 2.8)	22.43	0.41	22.56
Αυτοκίνητο (σχ. 2.10)	19.31	0.27	19.37
Βιβλία (σχ. 2.12)	17.37	0.42	17.21
Πίνακας Οφθαλμίατρου (σχ. 2.14)	18.69	0.32	18.68
Cameraman 2 (σχ. 2.18)	22.65	0.27	22.51

ΚΕΦΑΛΑΙΟ 3

ΥΠΕΡΑΝΑΛΥΣΗ ΕΙΚΟΝΑΣ ΜΕ ΑΝΤΙΣΤΟΙΧΙΣΗ ΣΗΜΕΙΩΝ ΕΝΔΙΑΦΕΡΟΝΤΟΣ

3.1 Εισαγωγή

Στο κεφάλαιο αυτό θα εξετάσουμε την υπέρθεση εικόνων με βάση την αντιστοίχιση χαρακτηριστικών σημείων και πως αυτή ενσωματώνεται στη διαδικασία της υπερανάλυσης εικόνας.

Η μέθοδος της εξαγωγής διακριτών σταθερών χαρακτηριστικών σημείων από μία εικόνα [18] μπορεί να χρησιμοποιηθεί για να εκτελέσουμε ένα αξιόπιστο ταίριασμα μεταξύ διαφορετικών όψεων ενός αντικειμένου ή μίας σκηνής. Τα χαρακτηριστικά αυτά είναι αμετάβλητα στο χώρο της κλίμακας της εικόνας και την περιστροφή, και φαίνεται ότι παρέχουν μία πολύ καλή υπέρθεση των εικόνων ακόμα και αν τα δεδομένα εισόδου είναι παραμορφωμένα, περιέχουν προσθετικό θόρυβο ή υπάρχει διαφορά στην φωτεινότητα. Στην ενότητα αυτή θα περιγράψουμε μια μέθοδο εξαγωγής χαρακτηριστικών σημείων από εικόνα, η οποία έχει πολλές ιδιότητες ώστε την καθιστά κατάλληλη για υπέρθεση διαφορετικών εικόνων μεταξύ τους. Τα χαρακτηριστικά αυτά, όπως αναφέραμε είναι αμετάβλητα στον χώρο κλίμακας της εικόνας και στην περιστροφή και μερικώς αμετάβλητα στις αλλαγές της φωτεινότητας και στην τρισδιάστατη αντίληψη της κάμερας. Στην πράξη ο αριθμός των χαρακτηριστικών σημείων που εξάγονται από τυπικές εικόνες, είναι αρκετά μεγάλος. Επιπλέον το γεγονός ότι τα σημεία αυτά είναι διακριτά, μας επιτρέπει να αντιστοιχίσουμε ένα σημείο με κάποιο άλλο με αρκετά μεγάλη πιθανότητα. Στην συνέχεια παρουσιάζουμε τις κύριες φάσεις υπολογισμού χαρακτηριστικών σημείων:

- Ανίχνευση ακρότατων στον χώρο κλίμακας: Το πρώτο στάδιο υπολογισμού εκτελεί μία αναζήτηση σε όλες τις κλίμακες και τις τοποθεσίες της εικόνας. Είναι υπολογισμένο αποδοτικά χρησιμοποιώντας μία συνάρτηση διαφορών Gaussian συναρτήσεων, για να αναγνωρίσει πιθανά σημεία ενδιαφέροντος, τα οποία είναι αμετάβλητα στην κλίμακα και τον προσανατολισμό.

- Εντοπισμός σημείων ενδιαφέροντος: Σε κάθε υποψήφια θέση, ένα λεπτομερές μοντέλο ταιριάζει κάθε φορά για να καθορίσει την τοποθεσία και την κλίμακα. Τα σημεία ενδιαφέροντος επιλέγονται κάθε φορά με βάση το κριτήριο της σταθερότητάς τους.

- Ανάθεση προσανατολισμού: Ένας ή περισσότεροι προσανατολισμοί ανατίθενται σε κάθε σημείο ενδιαφέροντος βασιζόμενοι στις τοπικές κατευθύνσεις της κλίσης της παραγώγου της εικόνας. Όλες οι λειτουργίες που θα ακολουθήσουν, εκτελούνται πάνω στα δεδομένα της εικόνας τα οποία έχουν μετασχηματιστεί σε σχέση με τον προσανατολισμό, την κλίμακα και την τοποθεσία κάθε χαρακτηριστικού, επομένως παρέχουν σταθερότητα σε αυτούς τους μετασχηματισμούς.

- Περιγραφέας σημείων ενδιαφέροντος: Οι τοπικές κλίσεις της παραγώγου της εικόνας μετρώνται στην επιλεγμένη κλίμακα γύρω από την περιοχή των σημείων ενδιαφέροντος. Αυτές μετασχηματίζονται σε μία αναπαράσταση που επιτρέπει τοπικές μεταβολές στο σχήμα και αλλαγές στην φωτεινότητα.

Η προσέγγιση αυτή ονομάζεται Μετασχηματισμός Χαρακτηριστικών Αμετάβλητα στην Κλίμακα (*Scale Invariant Feature Transform, SIFT*), αφού μετασχηματίζει τα δεδομένα της εικόνας σε σταθερής κλίμακας συντεταγμένες τα οποία σχετίζονται με τα τοπικά χαρακτιριστικά.

Μια σημαντική ιδιότητα της προσέγγισης αυτής, είναι ότι παράγει ένα μεγάλο αριθμό χαρακτηριστικών σημείων τα οποία καλύπτουν την εικόνα. Για την υπέρθεση των εικόνων, τα σημεία SIFT αρχικά εξάγονται από ένα σύνολο εικόνων αναφοράς. Η νέα εικόνα προκύπτει συγκρίνοντας κάθε χαρακτηριστικό της νέας εικόνας με αυτά που έχουμε εξάγει από την εικόνα αναφοράς, βρίσκοντας μία αντιστοίχιση χαρακτηριστικών σημείων, που βασίζεται στην Ευκλείδια απόσταση μεταξύ των χαρακτηριστικών αυτών. 'Οπως έχουμε αναφέρει και πιο πριν, οι περιγραφείς SIFT είναι διακριτοί και αυτό μας επιτρέπει να βρούμε για ένα χαρακτηριστικό τη σωστή αντιστοίχισή του με αρκετά καλή πιθανότητα. Ωστόσο, σε μια μη ομαλή εικόνα, πολλά χαρακτηριστικά από το φόντο αυτής ενδέχεται να μην έχουν σωστή αντιστοίχιση, προκαλώντας έτσι σφάλμα στις τιμές των παραμέτρων υπέρθεσης.

3.2 Ανίχνευση Ακρότατων στον Χώρο Κλίμακας

'Οπως περιγράψαμε και στην εισαγωγή, κάνουμε ανίχνευση των σημείων ενδιαφέροντος χρησιμοποώντας μία σειριακή προσέγγιση που μας επιτρέπει να χρησιμοποιήσουμε αποτελεσματικούς αλγορίθμους για να αναγνωρίσουμε υποψήφια χαρακτηριστικά σημεία, το οποία θα εξεταστούν στη συνέχεια για το αν είναι αποδεκτά ή οχι. Το πρώτο στάδιο της ανίχνευσης των σημείων ενδιαφέροντος είναι να αναγνωρίσουμε θέσεις και αναλογίες οι οποίες μπορούν επαναληπτικά να προσδιοριστούν κάτω από διαφορετικές όψεις του ίδιου αντικειμένου. Η ανίχνευση θέσεων που είναι αμετάβλητες στις αλλαγές κλιμακας της εικόνας μπορεί να επιτευχθεί εκτελώντας μία αναζήτηση για όλες τις πιθανές κλίμακες, χρησιμοποιώντας μία συνεχή συνάρτηση κλίμακας γνωστή ως συνάρτηση χώρου κλίμακας.

Μία αρκετά καλή υπόθεση για την συνάρτηση χώρου κλίμακας είναι η Gaussian συνάρτηση. Επομένως, η συνάρτηση χώρου κλίμακας μιας εικόνας ορίζεται ως η συνάρτηση, $L(x, y, \sigma)$, η οποία παράγεται από την συνέλιξη μιας μεταβλητής κλίμακας Gaussian συνάρτησής $G(x, y, \sigma)$, με την εικόνα εισόδου $I(x, y)$:

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \tag{3.1}$$

όπου με το σύμβολο $*$ εννοούμε την πράξη της συνέλιξης και

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2}.$$

Για να ανιχνεύσουμε αποδοτικά τις θέσεις σημείων ενδιαφέροντος στον χώρο κλίμακας, μπορούμε να υπολογίσουμε τα ακρότατα της διαφοράς των συνελίξεων της εικόνας με δύο γειτονικές Gaussian συναρτήσεις οι οποίες διαφέρουν κατά ένα ένα σταθερό παράγωντα k:

$$\begin{aligned} D(x, y, \sigma) &= \big(G(x, y, k\sigma) - G(x, y, \sigma)\big) * I(x, y) \\ &= L(x, y, k\sigma) - L(x, y, \sigma). \end{aligned} \tag{3.2}$$

Ο λόγος που χρησιμοποιούμε αυτή την συνάρτηση είναι ότι υπολογίζεται πολύ εύκολα, καθώς προκύπτει από μια απλή αφαίρεση. Επιπλέον, η διαφορά της Gaussian συνάρτησης παρέχει μια αρκετά καλή προσέγγιση για την κανονικοποιημένη Laplacian της Gaussian συνάρτησης, $\sigma^2\nabla^2 G$. 'Εχει βρεθεί πειραματικά ότι η μέγιστη και η ελάχιστη τιμή της ποσότητας $\sigma^2\nabla^2 G$ δίνει τα πιο σταθερά χαρακτηριστικά σημεία σε μια εικόνα εν συγκρίσει με άλλες πιθανές συναρτήσεις όπως η παράγωγος, η Hessian, ή τα Harris corners [11].

Η σχέση μεταξύ της D και της ποσότητας $\sigma^2\nabla^2 G$ μπορεί να κατανοηθεί από τη παρακάτω εξίσωση:

$$\frac{\partial G}{\partial \sigma} = \sigma\nabla^2 G$$

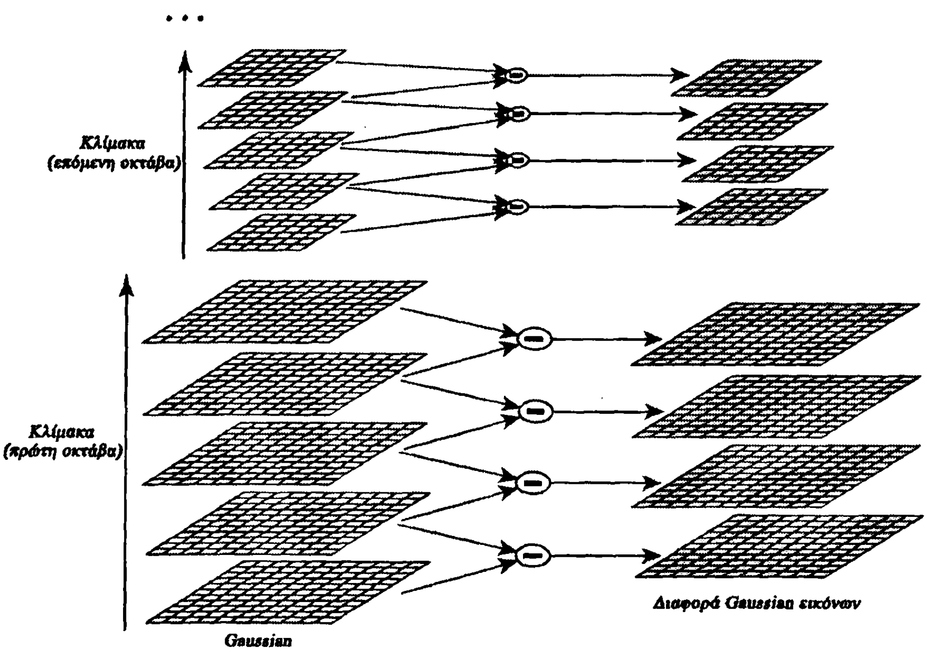
Από την εξίσωση αυτή μπορούμε να δούμε ότι η ποσότητα $\nabla^2 G$ από την προσέγγιση πεπερασμένων διαφορών $\partial G/\partial\sigma$, χρησιμοποιώντας διαφορές γειτονικών κλιμάκων στο $k\sigma$ και σ:

$$\sigma\nabla^2 G = \frac{\partial G}{\partial \sigma} \approx \frac{G(x,y,k\sigma) - G(x,y,\sigma)}{k\sigma - \sigma}$$

επομένως προκύπτει:

$$G(x,y,k\sigma) - G(x,y,\sigma) \approx (k-1)\sigma^2\nabla^2 G.$$

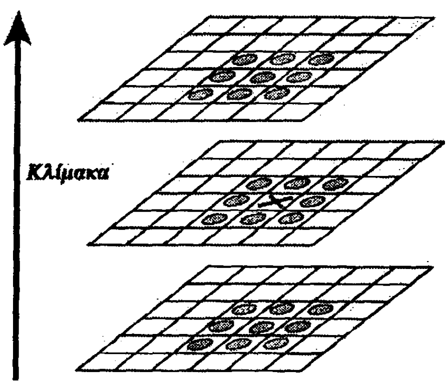
Από την παραπάνω εξίσωση μπορούμε να συμπεράνουμε ότι όταν η διαφορά των Gaussian συναρτήσεων έχει κλίμακες που διαφέρουν κατά έναν σταθερό παράγοντα τότε αυτό ενσωματώνεται στην κανονικοποιημένη κλίμακα σ^2 που απαιτείται για την αμετάβλητη Laplacian κλίμακα. Ο παράγοντας $(k-1)$ είναι μία σταθερά για όλες τις κλίμακες και επομένως δεν επηρεάζει την θέση των ακρότατων. Το σφάλμα προσέγγισης τείνει στο μηδέν καθώς το k πηγαίνει στο 1.



Σχήμα 3.1: Για κάθε οκτάβα του χώρου κλίμακας η αρχική εικόνα επαναληπτικά συνελίσσεται με Gaussian συναρτήσεις για να παράγουν ένα σύνολο από εικόνες χώρου κλίμακας, οπως φαίνεται στα αριστερά του σχήματος. Γειτονικές Gaussian εικόνες αφαιρούνται για να παράγουν την διαφορά των Gaussian εικόνων, στα δεξιά του σχήματος. Μετά από κάθε οκτάβα, η Gaussian εικόνα υποδειγματοληπτείται στο μισό και η διαδικασία επαναλαμβάνεται. Το σχήμα αντιγράφηκε από το [18].

Μία αποδοτική προσέγγιση για την κατασκευή της $D(x,y,\sigma)$ φαίνεται στο σχήμα 3.1. Η αρχική εικόνα σταδιακά συνελίσσεται με Gaussian συναρτήσεις για να παράγει εικόνες που διαφέρουν κατά μία σταθερά k στο χώρο κλίμακας, όπως βλέπουμε στην αριστερή πλευρά του σχήματος. Διαλέγουμε να διαιρέσουμε κάθε οκτάβα του χώρου κλίμακας σε έναν ακέραιο αριθμό διαστημάτων s, έτσι ώστε $k = 2^{1/s}$. Πρέπει να παράγουμε $s+3$ εικόνες στην στοίβα των θολωμένων εικόνων για κάθε οκτάβα, έτσι ώστε η ανίχνευση του τελικού ακρότατου να καλύπτει όλη την οκτάβα των εικόνων. Γειτονικές κλίμακες των εικόνων αφαιρούνται για να παράγουν τη διαφορά των Gaussian εικόνων, όπως φαίνεται στα δεξιά του σχήματος. Μόλις υπολογιστεί μια οκτάβα επαναδειγματοληπτούμε την Gaussian εικόνα στο μισό του αρχικού μεγέθους της κάθε φορά, παίρνοντας κάθε δεύτερο εικονοστοιχείο κάθε γραμμής και κάθε στήλης.

3.2.1 Ανίχνευση Τοπικών Ακρότατων



Σχήμα 3.2: Η μέγιστη και ελάχιστη τιμή της διαφοράς των εικόνων που προκύπτουν από τις συνελίξεις, ανιχνεύεται συγκρίνοντας ένα εικονοστοιχείο με τους 26 γείτονές του σε μία 3×3 περιοχή γειτονικών κλιμάκων. Το σχήμα αναπαράχθηκε από το [18].

Για να βρούμε το τοπικό μέγιστο και ελάχιστο της $D(x, y, \sigma)$, κάθε σημείο ελέγχεται με τους οχτώ γείτονές του στην τρέχουσα εικόνα και με τους εννιά γείτονές του στις ακριβώς από πάνω και από κάτω κλίμακες, όπως βλέπουμε και στο σχήμα 3.2. Το σημείο αυτό επιλέγεται μόνο αν είναι μεγαλύτερο από όλους τους γείτονές του ή μικρότερο από αυτούς.

Ένα αρκετά σημαντικό θέμα, είναι ο καθορισμός της συχνότητας δειγματοληψίας στην εικόνα και στα πεδία κλίμακας, η οποία χρειάζεται για την αξιόπιστη εύρεση των ακρότατων. Δυστυχώς, όπως αποδεικνύεται δεν υπάρχει κάποια ελάχιστη απόσταση μεταξύ των δειγμάτων στην οποία εντοπίζονται όλα τα ακρότατα, μιας και αυτά μπορούν να βρίσκονται αυθαίρετα κοντα μεταξύ τους. Αυτό μπορούμε να το καταλάβουμε από το εξής παράδειγμα: θεωρούμε μια εικόνα μαύρου φόντου η οποία περιέχει έναν άσπρο κύκλο. Αυτή θα έχει ένα μόνο μέγιστο χώρου κλίμακας εκεί όπου η κυκλική περιοχή της συνάρτησης διαφορών Gaussian συνάρτησεων ταιριάζει με την θέση που βρίσκεται ο κύκλος στην αρχική εικόνα. Για μία αρκετά επιμηκυμένη έλλειψη, υπάρχουν δύο μέγιστα κοντά στα άκρα της έλλειψης. Καθώς οι θέσεις των μεγίστων της έλλειψης είναι συνεχής συνάρτηση της εικόνας, για μία έλλειψη με μεσαία επιμήκυνση θα υπάρχει μία μετάβαση από ένα μέγιστο σε δύο, με τα μέγιστα να βρίσκονται αυθαίρετα κοντά μεταξύ τους. Εν γένει, ακρότατα που βρίσκονται αρκετά κοντά μεταξύ τους, είναι ασταθή για μικρές διαταραχές της εικόνας.

3.3 Ακριβής Εντοπισμός Σημείων Ενδιαφέροντος

Μόλις βρούμε ένα υποψήφιο σημείο ενδιαφέροντος, συγκρίνοντας ένα εικονοστοιχείο με τους γείτονές του, το επόμενο βήμα είναι να εκτελέσουμε μία λεπτομερή αντιστοίχιση στα γειτονικά δεδομένα για την θέση, την κλίμακα και τον λόγο της κύριας καμπυλότητας. Οι πληροφορίες που εξάγουμε από το βήμα αυτό, μας επιτρέπουν να απορρίψουμε κάποια σημεία που έχουν μικρή αντίθεση και επομένως είναι ευαίσθητα σε θόρυβο ή εντοπίζονται ελλιπώς κατά μήκος μίας ακμής.

Μία μέθοδος που έχει αναπτυχθεί [18] για να καθορίσουμε την παρεμβάλλουσα θέση του μεγίστου, χρησιμοποιεί σειρές Taylor (μέχρι δεύτερου βαθμού) της συνάρτησης κλίμακας χώρου $D(x, y, \sigma)$, η οποία μετατοπίζεται έτσι ώστε η αρχή των αξόνων της να βρίσκεται στο σημείο που την υπολογίζουμε

$$D(\mathbf{x}) = D + \frac{\partial D^T}{\partial \mathbf{x}}\mathbf{x} + \frac{1}{2}\mathbf{x}^T\frac{\partial^2 D}{\partial \mathbf{x}^2}\mathbf{x} \tag{3.3}$$

όπου το D και οι παράγωγοι του αποτιμώνται στο σημείο ενδιφέροντος και $\mathbf{x} = (x, y, \sigma)^T$ είναι η απόσταση από αυτό το σημείο. Η θέση του ακρότατου, $\hat{\mathbf{x}}$, καθορίζεται παίρνοντας την παράγωγο της συνάρτησης αυτής ως προς x και θέτοντάς την ίση με μηδέν:

$$\hat{\mathbf{x}} = -\frac{\partial^2 D^{-1}}{\partial \mathbf{x}^2}\frac{\partial D}{\partial \mathbf{x}}. \tag{3.4}$$

Ο Hessian πίνακας και η παράγωγος της D προσεγγίζονται χρησιμοποιώντας τις διαφορές των γειτονικών σημείων. Το προκύπτον 3×3 γραμμικό σύστημα μπορεί να λυθεί με ελάχιστο κόστος. Αν η απόσταση $\hat{\mathbf{x}}$ είναι μεγαλύτερη από 0.5 σε κάθε διάσταση, αυτό σημαίνει ότι το ακρότατο βρίσκεται κοντά σε κάποιο άλλο εικονοστοιχείο. Στην περίπτωση αυτή το σημείο αλλάζει και η παρεμβολή γίνεται γύρω από αυτό το νέο σημείο. Η τελική απόσταση $\hat{\mathbf{x}}$ προστίθεται στην τρέχουσα θέση για να πάρουμε την εκτίμηση που προκύπτει από την παρεμβολή για το τοπικό ακρότατο.

Η τιμή που βρίσκουμε για το ακρότατο $D(\hat{\mathbf{x}})$ είναι χρήσιμη για να απορρίψουμε ασταθή ακρότατα με χαμηλή αντίθεση. Αυτό γίνεται αντικαθστώντας την εξίσωση (3.4) στην (3.3):

$$D(\hat{\mathbf{x}}) = D + \frac{1}{2}\frac{\partial D^T}{\partial \mathbf{x}}\hat{\mathbf{x}}.$$

3.3.1 Εξάλειψη Απόκρισης Ακμών

Όσον αφορά στην σταθερότητα, δεν είναι αρκετό να απορρίπτουμε σημεία ενδιαφέροντος με χαμηλή αντίθεση. Η συνάρτηση διαφορών Gaussian συναρτήσεων έχει μεγάλη απόκριση στις ακμές, ακόμα και αν τα σημεία αυτά γύρω από τις ακμές είναι λίγα και επομένως είναι ασταθή σε μικρές ποσότητες θορύβου.

Ένα κακώς ορισμένο άκροτατο της συνάρτησης διαφοράς Gaussian συναρτήσεων έχει μεγάλη κύρια καμπυλότητα γύρω από την ακμή αλλά μικρή κατά την κάθετη κατέυθυνση. Οι κύριες καμπηλότητες μπορούν να υπολογιστούν από έναν 2×2 Hessian πίνακα, $\mathbf{H}$, που υπολογίζεται στη θέση και την κλίμακα των σημείων ενδιαφέροντος:

$$\mathbf{H} = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix} \tag{3.5}$$

Οι παράγωγοι εκτιμώνται παίρνοντας τις διαφορές των γειτονικών δειγματικών σημείων.

Οι ιδιοτιμές του πίνακα $\mathbf{H}$ είναι ανάλογες των κύριων καμπυλοτήτων της D. Μπορούμε να αποφύγουμε τον υπολογισμό των ιδιοτιμών, η οποία είναι μία αρκετά χρονοβόρα διαδικασία, μιας και ενδιαφερόμαστε μόνο για το λόγο τους. Έστω λοιπόν, α η μεγαλύτερη ιδιοτιμή και β η μικρότερη ιδιοτιμή. Τότε, μπορούμε να υπολογίσουμε το άθροισμα των ιδιοτιμών από το ίχνος του Hessian πίνακα $\mathbf{H}$ και το γινόμενο από την ορίζουσα:

$$Tr(\mathbf{H}) = D_{xx} + D_{yy} = \alpha + \beta,$$

$$Det(\mathbf{H}) = D_{xx}D_{yy} - (D_{xy})^2 = \alpha\beta$$

Στην περίπτωση που η ορίζουζα είναι αρνητική, οι καμπυλότητες έχουν διαφορετικά πρόσιμα, επομένως το σημείο απορρίπτεται αφού δεν είναι ακρότατο. Ορίζουμε ως r το λόγο μεταξύ της μεγαλύτερης και της μικρότερης ιδιοτιμής, συνεπώς ισχύει ότι $\alpha = r\beta$. Τότε προκύπτει:

$$\frac{Tr(\mathbf{H})^2}{Det(\mathbf{H})} = \frac{(\alpha+\beta)^2}{\alpha\beta} = \frac{(r\beta+\beta)^2}{r\beta^2} = \frac{(r+1)^2}{r}.$$

Η ποσότητα αυτή εξαρτάται μόνο από τον λόγο των ιδιοτιμών. Η ποσότητα $\frac{(r+1)^2}{r}$ είναι ελάχιστη όταν οι δύο ιδιοτιμές είναι ίσες και αυξάνει με ρυθμό r. Επομένως, για να ελέγξουμε αν ο λόγος των κύριων καμπυλοτήτων είναι μικρότερος από ένα κατώφλι r, αρκεί να ελέγξουμε αν ισχύει η ανισότητα:

$$\frac{Tr(\mathbf{H})^2}{Det(\mathbf{H})} < \frac{(r+1)^2}{r}.$$

Στο σχήμα 3.3 φαίνονται τα αποτελέσματα της επιλογής των σημείων ενδιαφέροντος σε μια εικόνα. Τα σημεία ενδιαφέροντος απεικονίζονται ως διανύσματα δίνοντας για κάθε ένα από αυτα τη θέση, την κλίμακα και τον προσανατολισμό τους. Η εικόνα (α) δείχνει την αρχική εικόνα, η εικόνα (β) δείχνει τα 832 σημεία ενδιαφέροντος για όλα τα ελάχιστα και μέγιστα που ανιχνεύτηκαν από την συνάρτηση διαφοράς Gaussian συναρτήσεων. Στην εικόνα (γ) βλέπουμε τα 729 σημεία ενδιαφέροντος που απομένουν αν αφαιρέσουμε αυτά που η τιμή της $D(\hat{\mathbf{x}})$ είναι μικρότερη από 0.03. Τέλος η (δ) δείχνει τα αποτελέσματα της διαδικασίας της εξάλειψης των σημείων ενδιαφέροντος που έχουν λόγο μεταξύ των κύριων καμπυλοτήτων μεγαλύτερο από 10.

Σχήμα 3.3: Στάδια επιλογής σημείων ενδιαφέροντος. (α) Είναι η αρχική εικόνα. (β) Οι αρχικές θέσεις των 832 σημείων ενδιαφέροντος στα μέγιστα και ελάχιστα της συνάρτησης διαφορών Gaussian συναρτήσεων. Τα σημεία ενδιαφέροντος απεικονίζονται ως διανύσματα τα οποία δηλώνουν την κλίμακα, τον προσανατολισμό και τη θέση. (γ) Αφού εφαρμόσουμε ένα κατώφλι ελάχιστης φωτεινότητας, απομένουν 729 σημεία ενδιαφέροντος. (δ) Τα τελικά 536 σημεία ενδιαφέροντος που απομένουν αφού εφαρμόσουμε ένα κατώφλι στο λόγο της κύριας καμπυλότητας. Το σχήμα έχει ληφθεί από το [18].

3.4 Ανάθεση Προσανατολισμού

Αναθέτοντας ένα προσανατολισμό σε κάθε χαρακτηριστικό σημείο βασιζόμενοι στις τοπικές ιδιότητες της εικόνας, ο περιγραφέας των σημείων ενδιαφέροντος μπορεί να αναπαρασταθεί σε σχέση με τον προσανατολισμό που έχουμε αναθέσει και επομένως με τον τρόπο αυτό, να πετύχουμε σταθερότητα ως προς την περιστροφή της εικόνας.

Για την επιλογή της Gaussian εικόνας, L, χρησιμοποιούμε την κλίμακα των σημείων ενδιαφέροντος, έτσι ώστε όλοι οι υπολογισμοί να εκτελούνται σε σταθερής κλίμακας κατάσταση. Για κάθε εικόνα $L(x, y)$, το μέτρο της παραγώγου συμβολίζεται ως $m(x, y)$, και ο προσανατολιμός ως $\theta(x, y)$, οι ποσότητες αυτές υπολογίζονται χρησιμοποιώντας τις διαφορές των εικονοστοιχείων:

$$m(x,y) = \sqrt{\left(L(x+1,y) - L(x-1,y)\right)^2 + \left(L(x,y+1) - L(x,y-1)\right)^2}$$

$$\theta(x,y) = \tan^{-1}\left(\frac{L(x,y+1) - L(x,y-1)}{L(x+1,y) - L(x-1,y)}\right)$$

Σχηματίζεται ένα ιστόγραμμα προσανατολισμών από τους προσανατολισμούς της παραγώγου γύρω από την περιοχή των σημείων ενδιαφέροντος. Το ιστόγραμμα αυτό περιέχει 36 κάδους καλύπτοντας όλη την περιοχή προσανατολισμού των 360 μοιρών. Κάθε δείγμα που προστίθεται στο ιστόγραμμα έχει ένα βάρος, που προέρχεται από το μέτρο της παραγώγου αυτού και ένα κυκλικό Gaussian παράθυρο με βάρη, του οποίου το σ είναι 1.5 φορές μεγαλύτερο από την κλίμακα του σημείου ενδιαφέροντος.

Οι κορυφές του ιστογράμματος αντιστοιχούν στις κυρίαρχες κατευθύνσεις των τοπικών παραγώγων. Βρίσκουμε την μεγαλύτερη κορυφή, και έπειτα όλες οι υπόλοιπες που βρίσκονται στο 80% αυτής, χρησιμοποιούνται για να δημιουργήσουμε ένα σημείο ενδιαφέροντος με αυτόν τον προσανατολισμό. Συνεπώς, για θέσεις που έχουν πολλές κορυφές ίδιου μεγέθους, θα δημιουργηθούν πολλά χαρακτηριστικά σημεία για την ίδια θέση και κλίμακα αλλά διαφορετικού προσανατολισμού. Μόνο το 15% των σημείων περίπου, έχει περισσότερους από έναν προσανατολισμούς, ωστόσο αυτοί συνεισφέρουν σημαντικά στην σταθερότητα της αντιστοίχισης. Τέλος, μία παραβολή προσαρμόζεται πάνω σε τρεις τιμές του ιστογράμματος, που είναι κοντινότερες σε κάθε κορυφή ώστε να κάνουμε παρεμβολή στις θέσεις των κορυφών για μεγαλύτερη ακρίβεια.

3.5 Τοπικός Περιγραφέας Εικόνας

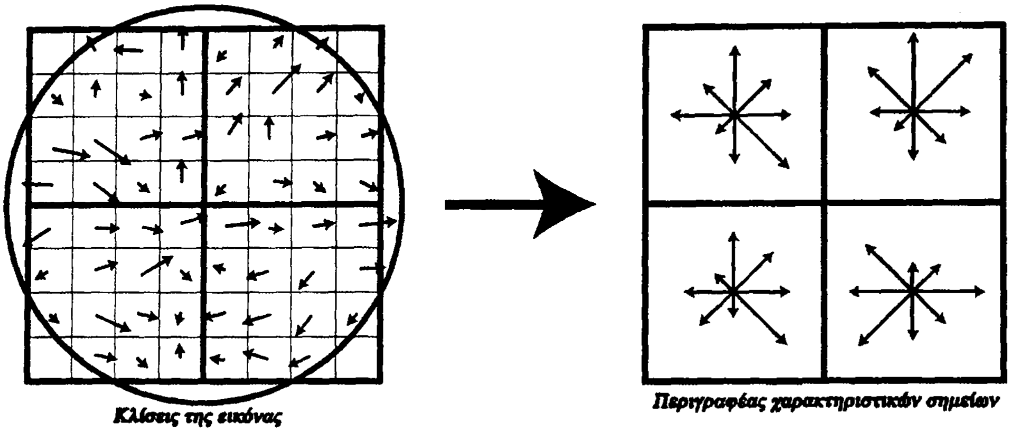
Στό βήμα αυτό θέλουμε να υπολογίσουμε ένα περιγραφέα για την τοπική περιοχή της εικόνας, ο οποίος να είναι διακριτός και επίσης να είναι όσο το δυνατόν πιο αμετάβλητος στις μεταβολές, όπως η αλλαγή της φωτεινότητας για παράδειγμα.

Μία προφανής προσέγγιση θα ήταν να δειγματοληπτήσουμε τις τοπικές εντάσεις της εικόνας γύρω από τα σημεία ενδιαφέροντος με κατάλληλη κλίμακα και να ταιριάξουμε αυτά χρησιμοποιώντας ένα κανονικοποιημένο συντελεστή συσχέτισης. Ωστόσο, μία απλή συσχέτιση πάνω στην εικόνα είναι αρκετά ευαίσθητη σε αλλαγές, οι οποίες μπορούν να προκαλέσουν λάθος αντιστοίχιση των σημείων, όπως για παράδειγμα μη συμπαγείς παραμορφώσεις.

3.5.1 Αναπαράσταση Περιγραφέα SIFT

Το σχήμα 3.4 απεικονίζει τον υπολογισμό ένος περιγραφεά σημείου κλειδιού. Αρχικά, τα μέτρα των κλίσεων της εικόνας και οι προσανατολισμοί δειγματολητούνται γύρω από την περιοχή του σημείου ενδιαφέροντος, χρησιμοποιώντας την κλίμακα του σημείου ενδιαφέροντος για να επιλέξουμε τα επίπεδα της Gaussian θόλωσης για την εικόνα. Για να πετύχουμε την αμεταβλητότητα του προσανατολισμού, οι συντεταγμένες του περιγραφέα και οι προσανατολισμοί των κλίσεων περιστρέφονται ως προς τον προσανατολισμό των σημείων ενδιαφέροντος. Οι κλίσεις απεικονίζονται με μικρά βελάκια, στο αριστερό μέρος του σχήματος.

Μία Gaussian συνάρτηση βάρους με σ ίσο με το μισό του πλάτους του παραθύρου του περιγραφέα χρησιμοποιείται για να αναθέσουμε ένα βάρος σε κάθε δειγματικό σημείο. Αυτό απεικονίζεται μέ ένα κυκλικό παράθυρο στο σχήμα 3.4. Σκοπός αυτού του Gaussian παραθύρου είναι να αποφύγουμε ξαφνικές αλλαγές του περιγραφέα σε μικρές αλλαγές της θέσης του παραθύρου και να δώσουμε λιγότερη έμφαση στις κλίσεις εκείνες που είναι μακριά



Σχήμα 3.4: Ο περιγραφέας ενός σημείου ενδιαφέροντος δημιουργείται υπολογίζοντας τα μέτρα των παραγώγων και τους προσανατολισμούς για κάθε δειγματικό σημείο της εικόνας σε μία περιοχή γύρω από τη θέση του σημείου ενδιαφέροντος. Ο κύκλος απεικονίζει την προβολή ενός Gaussian παραθύρου (αριστερή εικόνα). Τα δείγματα αυτά στη συνέχεια συγκεντρώνονται σε ιστογράμματα αθροίζοντας τα περιεχόμενά τους σε 4×4 περιοχές, με το μήκος του κάθε βέλους να αντιστοιχεί στο άθροισμα των κλίσεων γύρω από αυτή την κατεύθυνση μέσα σε αυτήν την περιοχή (δεξιά εικόνα). Το σχήμα αναπαράχθηκε από το [18].

από το κέντρο του περιγραφέα, μιας και αυτές επηρεάζονται περισσότερο από τα λάθη της υπέρθεσης.

Ο περιγραφέας του σημείου κλειδιού απεικονίζεται στο δεξί μέρος του σχήματος 3.4. Αυτό επιτρέπει μια σημαντική μετατόπιση στις θέσεις των κλίσεων δημιουργώντας ιστογράμματα σε μία 4×4 περιοχή. Η εικόνα δείχνει οχτώ κατευθύνσεις για κάθε ιστόγραμμα, με το μήκος του κάθε βέλους να αντιστοιχεί στο εύρος αυτού του ιστογράμματος. Μία κλίση στα αριστερά του σχήματος μπορεί να μετατοπιστεί μέχρι 4 θέσεις και παρ' όλα αυτά να συνεισφέρει στο ίδιο ιστόγραμμα στα δεξία του σχήματος, επομένως πετυχαίνουμε τον αντικειμενικό στόχο του να επιτρέπονται μεγάλες τοπικές μετατοπίσεις θέσεων.

Σημαντικό είναι να αποφύγουμε όλες τις οριακές επιρροές, όπου ο περιγραφέας απότομα αλλάζει όταν ένα δείγμα μετατοπίζεται αλλάζοντας ιστόγραμμα ή προσανατολισμό. Γι' αυτό χρησιμοιείται τριπλά γραμμική παρεμβολή για να κατανείμουμε την τιμή κάθε κλίσης σε γειτονικούς κάδους στο ιστόγραμμα. Δηλαδή, κάθε είσοδος σε ένα κάδο πολλαπλασιάζεται με ένα βάρος αντιστρόφως ανάλογο της απόστασης του δείγματος από την κεντρική τιμή του κάδου.

Ο περιγραφέας αποτελείται από ένα διάνυσμα που περιέχει τις τιμές όλων των ιστογραμμάτων. Η δεξία εικόνα του σχήματος 3.4 απεικονίζει ένα 2×2 πίνακα ιστογραμμάτων. Τα καλύτερα αποτελέσματα επιτυγχάνονται χρησιμοποιώντας ένα 4×4 πίνακα ιστογραμμάτων 8 με κάδους το κάθε ένα. Επομένως, για κάθε χαρακτηριστικό σημείο δημιουργείται ένα $4 \times 4 \times 8 = 128$ διάνυσμα χαρακτηριστικών.

3.6 Περιγραφή της Μεθόδου

Στην ενότητα αυτή, θα περιγράψουμε την εφαρμογή των περιγραφέων SIFT στο πρόβλημα της υπερανάλυσης εικόνας.

Σκοπός μας είναι να βρούμε τις παραμέτρους μετασχηματισμού υπέρθεσης (διάνυσμα μετατόπισης $\mathbf{T}$ και πίνακας γωνιών περιστροφής $\mathbf{R}$). Για το λόγο αυτό χρησιμοποιήσαμε τους περιγραφείς SIFT. Για την εκτίμηση των παραμέτρων μετασχηματισμού υπέρθεσης εφαρμόσαμε εκτίμησης ελάχιστων τετραγώνων, για περισσότερρες πληροφορίες με την τεχνική αυτή παραπέμπουμε τον αναγνώστη στην αναφορά [34].

Αρχικά, βρίσκουμε τις παραμέτρους μετασχηματισμού $\mathbf{R}$ και $\mathbf{T}$ και κάνουμε υπέρθεση στις εικόνες με τη χρήση των περιγραφέων SIFT. Στην συνέχεια εφαρμόζουμε τον αλγόριθμο υπερανάλυσης εικόνας (αλγόριθμος 1) που παρουσιάσαμε στο κεφάλαιο 2 και σε κάθε επανάληψη του αλγορίθμου βρίσκουμε τις παραμέτρους μετασχηματισμού υπέρθεσης με τη χρήση των περιγραφέων SIFT. Σε γενικές γραμμές ο αλγόριθμος που προτείνουμε ειναι:

Αλγόριθμος 3 Μέθοδος υπερανάλυσης εικόνας βασισμένη στην αντιστοίχιση σημείων.

βήμα 1: Εύρεση των παραμέτρων μετασχηματισμού υπέρθεσης ($\mathbf{R}$, $\mathbf{T}$).

βήμα 2: Αρχική υπέρθεση εικόνων με χρήση περιγραφέων SIFT.

while Δεν έχει συγκλίνει ακόμα.
- **βήμα 3:** Εύρεση των παραμέτρων μετασχηματισμού υπέρθεσης ($\mathbf{R}$, $\mathbf{T}$).
- **βήμα 4:** Υπέρθεση εικόνων με χρήση περιγραφέων SIFT.
- **βήμα 5:** Εκτίμηση της εικόνας υψηλής ανάλυσης.

3.7 Πειραματικά Αποτελέσματα

Στην ενότητα αυτή παρουσιάζουμε κάποια πειραματικά αποτελέσματα από την χρήση των περιγραφέων SIFT στην υπερανάλυση εικόνας. Η υπέρθεση των εικόνων χαμηλής ανάλυσης γίνεται μία φορά πριν την εκτέλεση του αλγορίθμου υπερανάλυσης εικόνας και έπειτα σε κάθε κάθε επανάληψη του αλγορίθμου βρίσκοντας μία αντιστοίχιση μεταξύ των σημείων ενδιαφέροντος που προκύπτουν σε κάθε βήμα από την εκτίμηση της υψηλής ανάλυσης εικόνας σε σχέση με τις χαμηλής ανάλυσης εικόνες. Η μέθοδος που ακολουθήσαμε παρουσιάστηκε επιγραμματικά στον αλγόριθμο 3.

Για τα πειράματα μας χρησιμοποιήσαμε τόσο πραγματικά όσο και τεχνητά σύνολα δεδομένων. Τα τεχντητά παραδείγματα δημιουργήθηκαν με τυχαία περιστροφή στο διάστημα $[-5, 5]$ και τυχαία μη ακέραια μετατόπιση στο διάστημα $[-1.5, 1.5]$. Έχουμε θολώσει τις εικόνες με Gaussian χαμηλοπερατό φίλτρο τυπικής απόκλισης 1 και μέγεθος παραθύρου 5×5 ενώ παράλληλα, οι εικόνες έχουν υποβαθμιστεί με ένα παράγοντα $L_1 = L_2 = 2$. Το μέγιστο πλήθος επαναλήψεων που χρησιμοποιήσαμε για την σύγκλιση του αλγορίθμου 3 είναι 50. Τέλος, έχουμε προσθέσει Gaussian λευκό θόρυβο στις τεχνητά παραχθείσες εικόνες χαμηλής ανάλυσης, που κυμαίνεται από 20 dB έως 30 dB. Για τα πραγματικά συνολα δεδομένων (σχήμα 3.6) δεν γνωρίζουμε τις σωστές λύσεις (*ground truth*).

Για την ποσοτικοποίηση της ποιότητας των υποβαθμισμένων εικόνων, τα επίπεδα του θορύβου στις υποβαθμισμένες εικόνες χαμηλής ανάλυσης και την ποιότητα των ανακατασκευασμένων αποτελεσμάτων χρησιμοποιούμε τη μέγιστη τιμή του λόγου σήματος προς θόρυβο (*PSNR*). Οι μετρικές αυτές ορίζονται ως:

$$PSNR = 10\log_{10}\frac{(255)^2}{\|\mathbf{f}-\mathbf{g}\|^2}$$

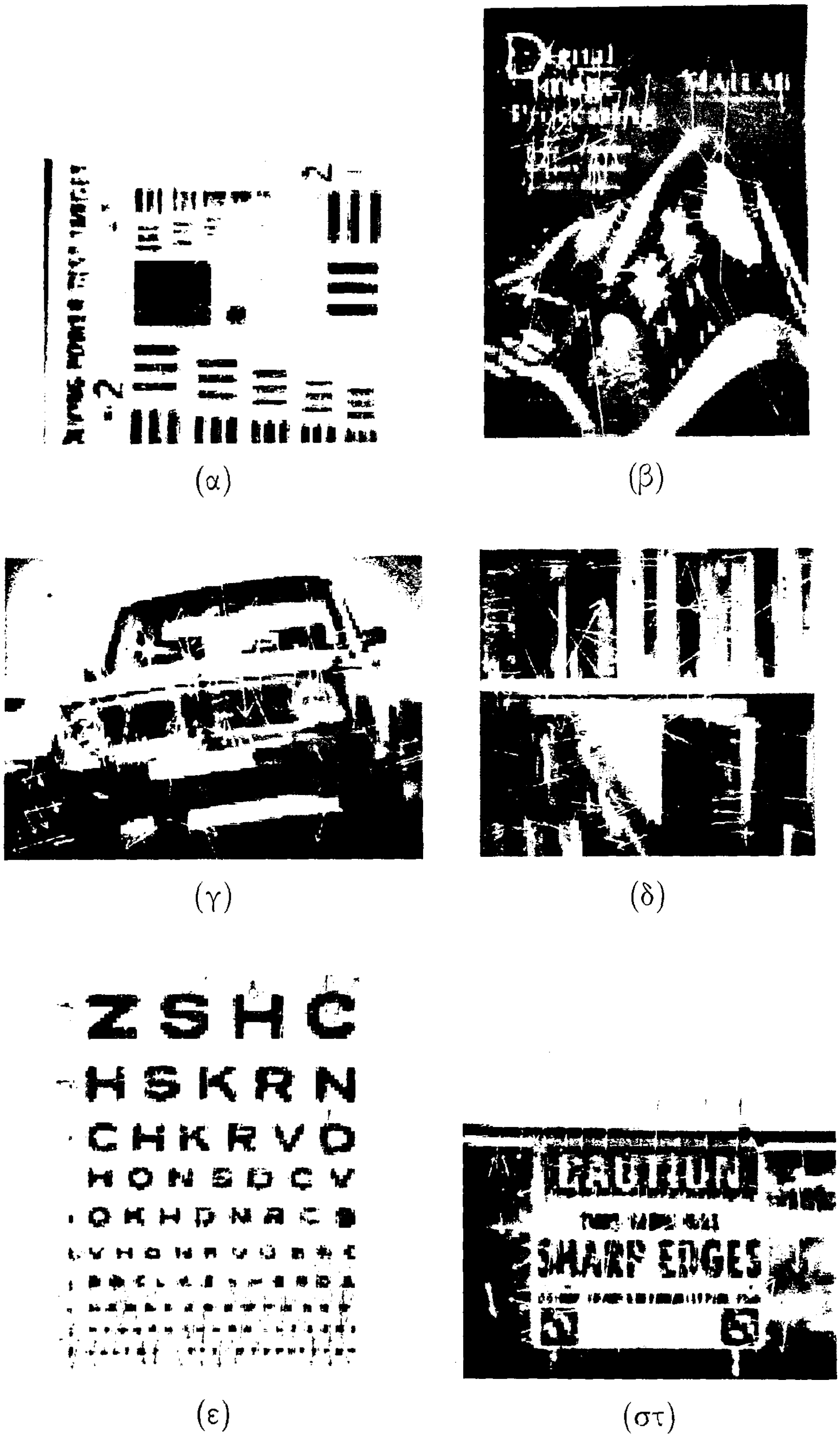
όπου $\mathbf{f}$ και $\mathbf{g}$ είναι η πραγματική και η ανακατασκευασμένη εικόνα, αντιστοίχως.

Η μέθοδος αυτή δίνει αρκετά ικανοποιητικά αποτελέσματα. Η ποιότητα των ανακατασκευασμένων εικόνων υψηλής ανάλυσης είναι αρκετά καλή και προσεγγίζουν την πραγματική λύση. Παρ' ολη την καλή ποιότητα των αποτελεσμάτων δεν μπορούμε να αποφύγουμε το φαινόμενο του *ringing*. Αυτό οφείλεται κατά κύριο λόγο, στα υψηλά επίπεδα θορύβου που παρουσιάζουν οι χαμηλής ανάλυσης εικόνες και την μεγάλη απόκλιση σε περιστροφή και μετατόπιση μεταξύ των χαμηλής ανάλυσης εικόνων. Για την εύρεση των χαρακτηριστικών σημείων χρησιμοποιήσαμε τον κώδικα που παρέχεται από τους συγγραφείς[1] [18]. Η εύρεση των σημείων ενδιαφέροντος είναι μία αρκετά γρήγορη διαδικασία και αυτό επιταχύνει τον αλγόριθμό μας.

Η δομή με την οποία θα παρουσιάσουμε τα αποτελέσματά μας είναι η ακόλουθη: δίνονται οι χαμηλής ανάλυσης εικόνες από τις οποίες θα προκύψει το αποτέλεσμα, για κάθε πείραμα παρουσιάζεται μία ενδεικτική εικόνα όπου απεικονίζονται τα χαρακτηριστικά σημεία αυτής. Παρουσιάζεται η ανακατασκευασμένη εικόνα υψηλής ανάλυσης και δύο πίνακες που περιέχουν τις τιμές των σφαλμάτων (μέση τιμή και τυπική απόκλιση) για τις παραμέτρους υπέρθεσης (μετατόπιση και περιστροφή) που προέκυψαν από τον αλγόριθμο 3, για μεγέθη θορύβου 30 dB και 20 dB, αντοίστιχα. Τέλος, παρουσιάζονται δύο ακόμα πίνακες που περιέχουν αριθμητικά αποτελέσματα (μέση τιμή, τυπική απόκλιση και ενδιάμεση τιμή) για το PSNR, ο κάθε ένας παρουσιάζει τα αποτελέσματα για μεγέθη θορύβου 30 dB και 20 dB, αντοίστιχα.

Στο σχήμα 3.5 ενδεικτικά απεικονίζονται τα χαρακτηριστικά σημεία για μία από τις χαμηλής ανάλυσης εικόνα των πειραμάτων που παρουσιάζουμε. Πληροφοριακά να αναφέρουμε ότι για την εικόνα (α) ανιχνεύσαμε 138 χαρακτηριστικά σημεία, για την (β) 156, για την (γ) 136, για την (δ) 105, για την (3) 364 και για την (στ) 170 χαρακτηριστικά σημεία.

[1] `http://people.cs.ubc.ca/~lowe/keypoints/`

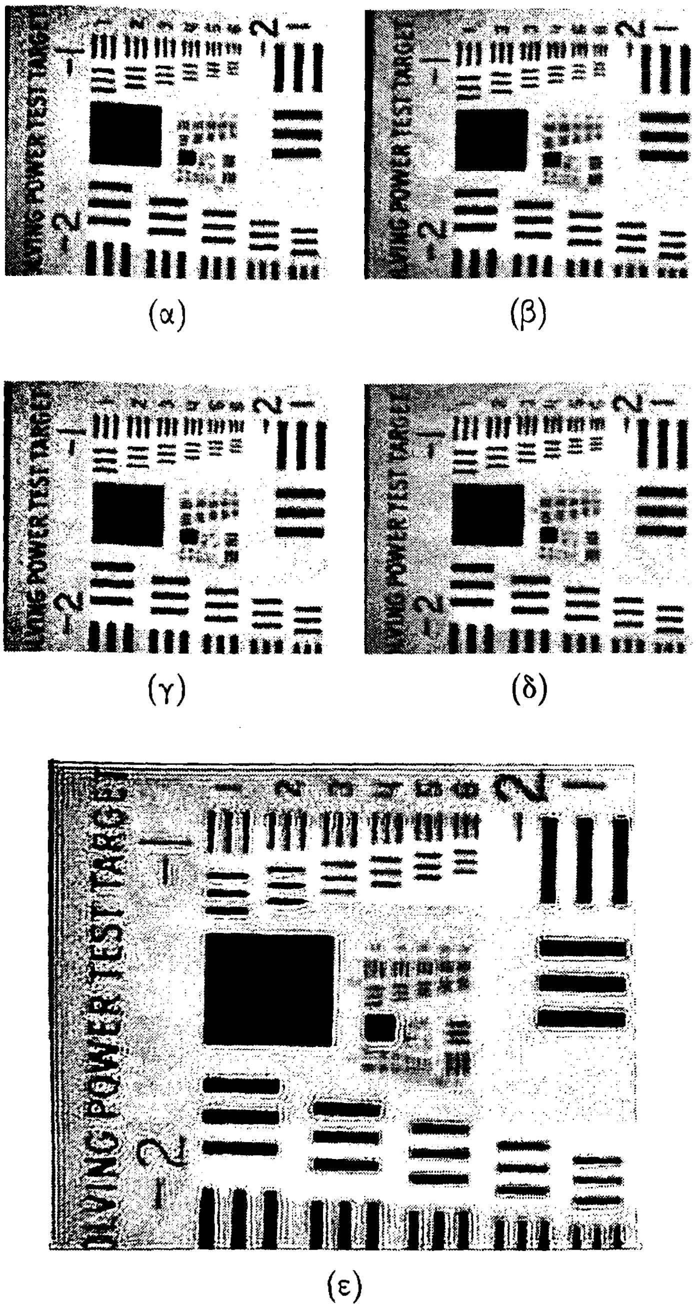


(α) (β)

(γ) (δ)

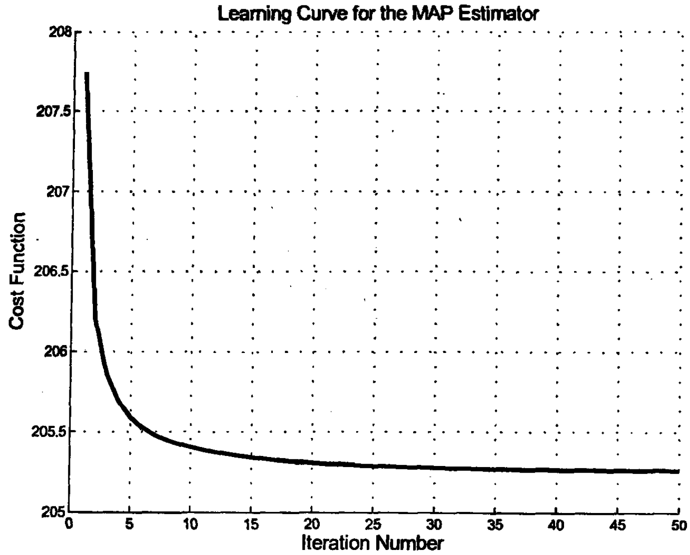
(ε) (στ)

Σχήμα 3.5: (α) - (στ) Χαρακτηριστικά σημεία για τις χαμηλής ανάλυσης εικόνες τ
μας. Τα σημεία ενδιαφέρντος αναπαριστώνται με βέλη που δηλώνουν το μέτρο, τ
προσανατολισμό τους.

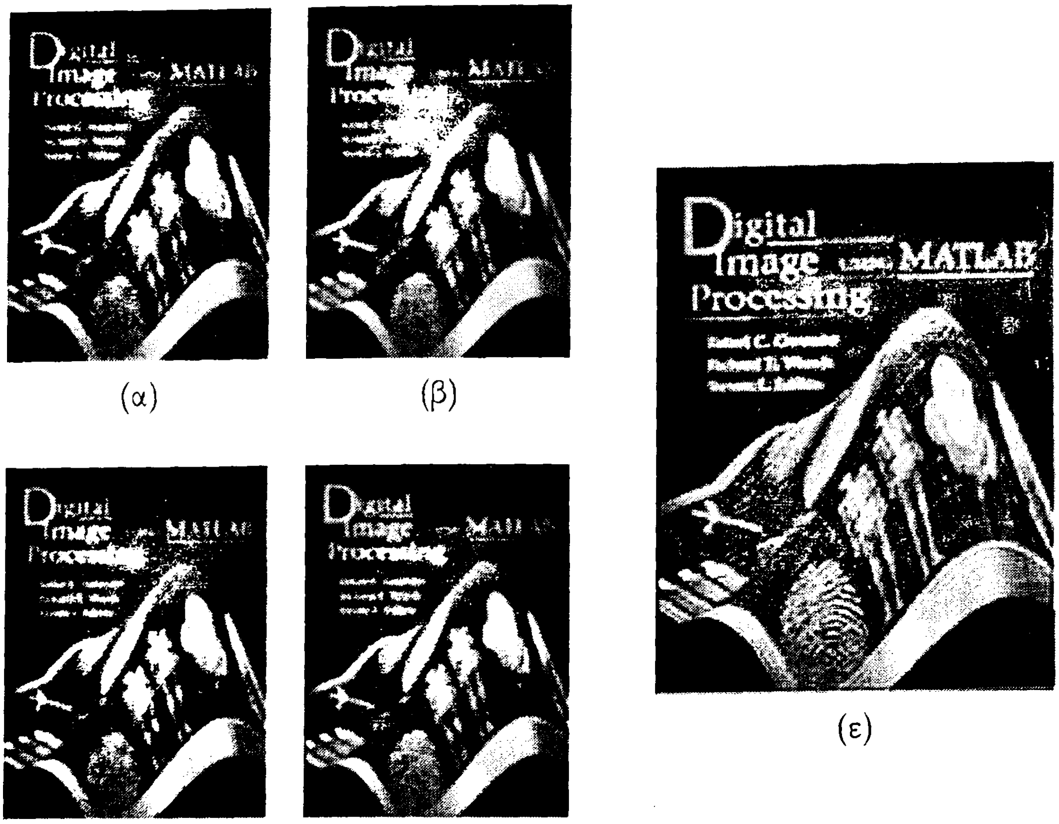


Σχήμα 3.6: *Resolution Chart.* (α) – (δ) Εικόνες χαμηλής ανάλυσης. (ε) Ανακατασκευασμένη εικόνα υψηλής ανάλυσης. Ο αριθμός των χαμηλής ανάλυσης εικόνων που χρησιμοποιήθηκαν για την ανακασκευή της (ε) είναι 30.

Στο σχήμα 3.6 οι χαμηλής ανάλυσης εικόνες προέκυψαν από πραγματικά δεδομένα. Προέρχεται από εικονοσειρά 742 πλαισίων, η οποία πάρθηκε κουνώντας με τυχαίο τρόπο την κάμερα.

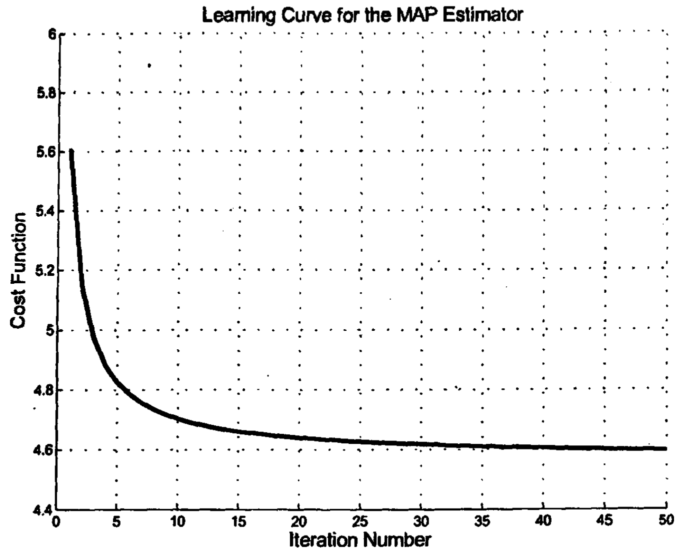


Σχήμα 3.7: Καμπύλη μάθησης που δείχνει την συνάρτηση κόστους $L(\mathbf{z}, \mathbf{s})$ σε σχέση με τον αριθμό των επαναλήψεων για την εικόνα *Resolution Chart* του σχήματος 3.6.

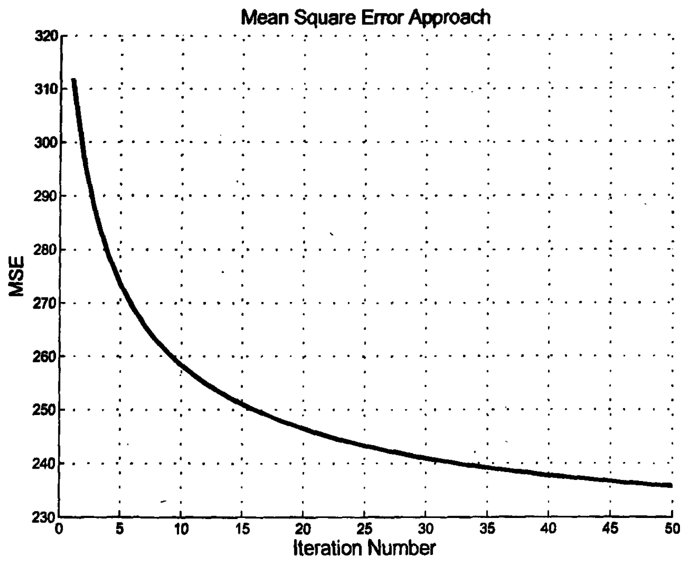


(α) (β) (γ) (δ) (ε)

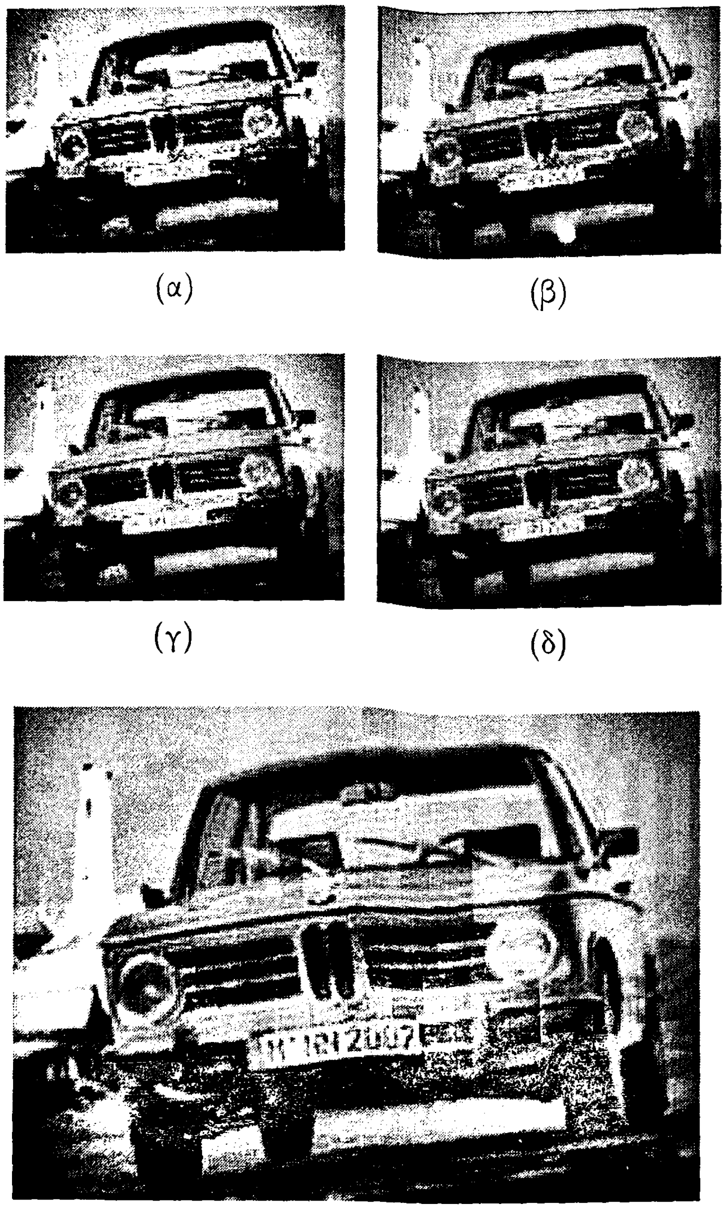
Σχήμα 3.8: *Εξώφυλλο.* (α) – (δ) Εικόνες χαμηλής ανάλυσης. (ε) Ανακατασκευασμένη ε υψηλής ανάλυσης. Ο αριθμός των χαμηλής ανάλυσης εικόνων που χρησιμοποιήθηκαν γι αναχασχευή της (ε) είναι 4.



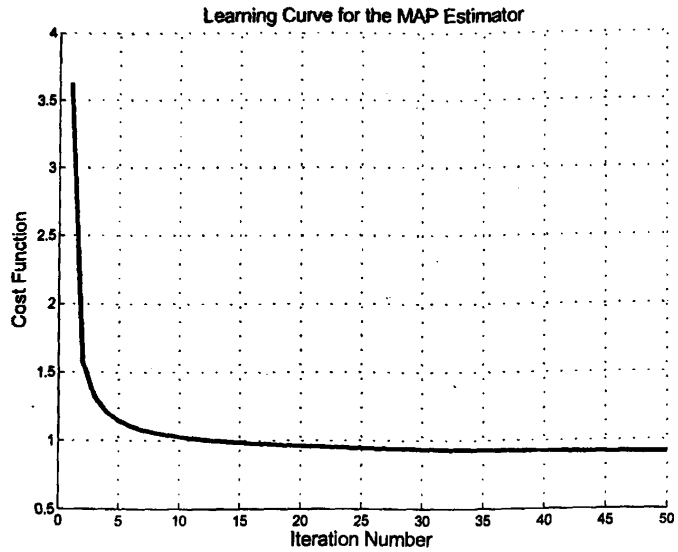
Σχήμα 3.9: Η συνάρτηση κόστους $L(\mathbf{z}, \mathbf{s})$ σε σχέση με τον αριθμό των επαναλήψεων για την εικόνα *Εξώφυλλο* του σχήματος 3.8.



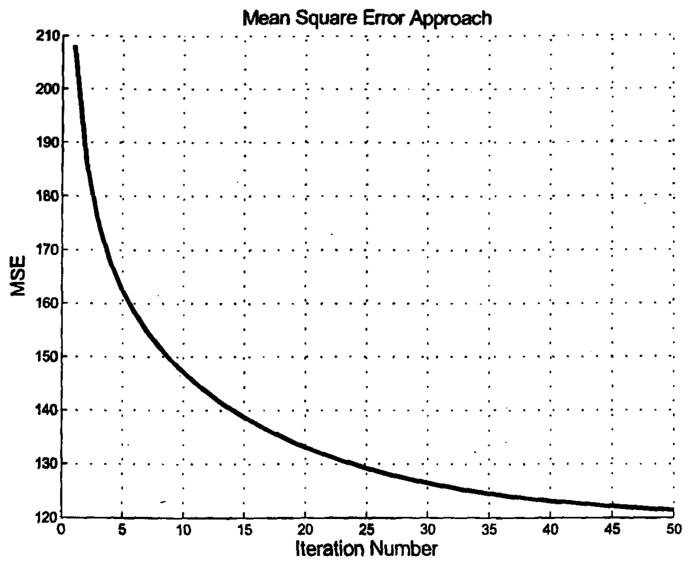
Σχήμα 3.10: Καμπύλη μέσου τετραγωνικού σφάλματος μεταξύ της πραγματικής εικόνας υψηλής ανάλυσης και της εκτίμησης της εικόνας υψηλής ανάλυσης για την εικόνα *Εξώφυλλο* του σχήματος 3.8. Η ανακατασκευασμένη εικόνα υψηλής ανάλυσης έχει PSNR 24.42 dB.



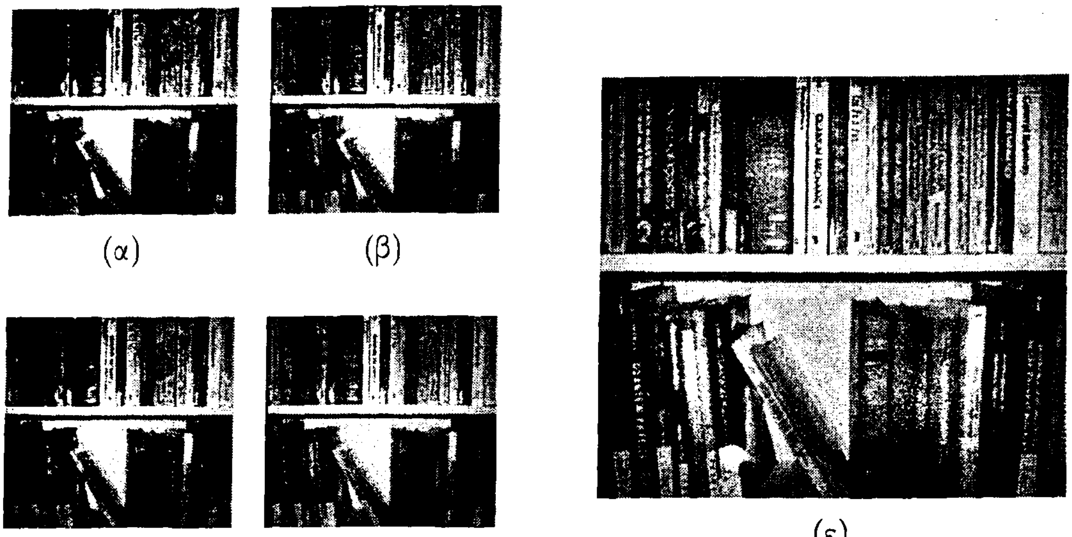
Σχήμα 3.11: *Αυτοκίνητο.* (α) – (δ) Εικόνες χαμηλής ανάλυσης. (ε) Ανακατασκευ υψηλής ανάλυσης. Ο αριθμός των χαμηλής ανάλυσης εικόνων που χρησιμοποιή ανακασκευή της (ε) είναι 4.



Σχήμα 3.12: Η συνάρτηση κόστους $L(\mathbf{z}, \mathbf{s})$ σε σχέση με τον αριθμό των επαναλήψεων για την εικόνα *Αυτοκίνητο* του σχήματος 3.11.

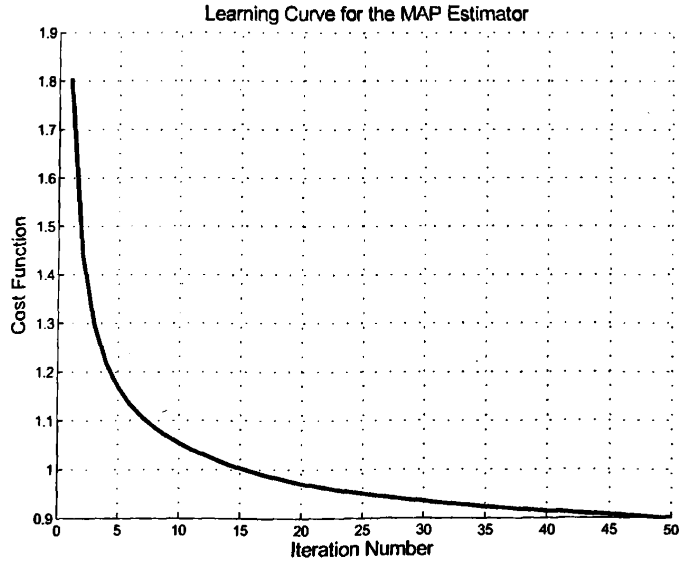


Σχήμα 3.13: Καμπύλη μέσου τεραγωνικού σφάλματος μεταξύ της πραγματικής εικόνας υψηλής ανάλυσης και της εκτίμησης της εικόνας υψηλής ανάλυσης για την εικόνα *Αυτοκίνητο* του σχήματος 3.11. Η ανακατασκευασμένη εικόνα υψηλής ανάλυσης έχει PSNR 27.30 dB.

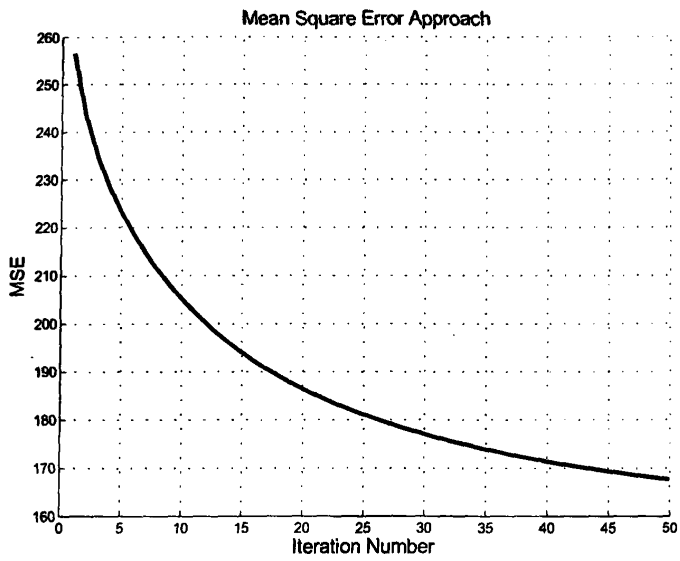


(α) (β) (γ) (δ) (ε)

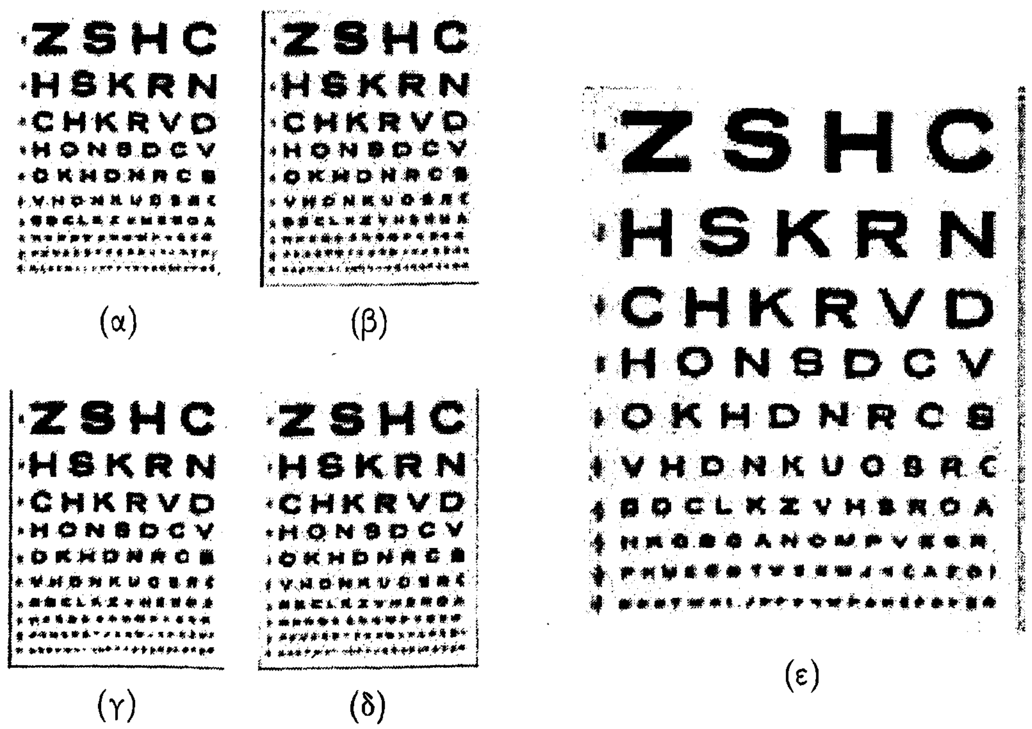
Σχήμα 3.14: *Βιβλία.* (α) – (δ) Εικόνες χαμηλής ανάλυσης. (ε) Ανακατασκευασμένη εικόνα υψηλής ανάλυσης. Ο αριθμός των χαμηλής ανάλυσης εικόνων που χρησιμοποιήθηκαν για την ανακασκευή της (ε) είναι 4.



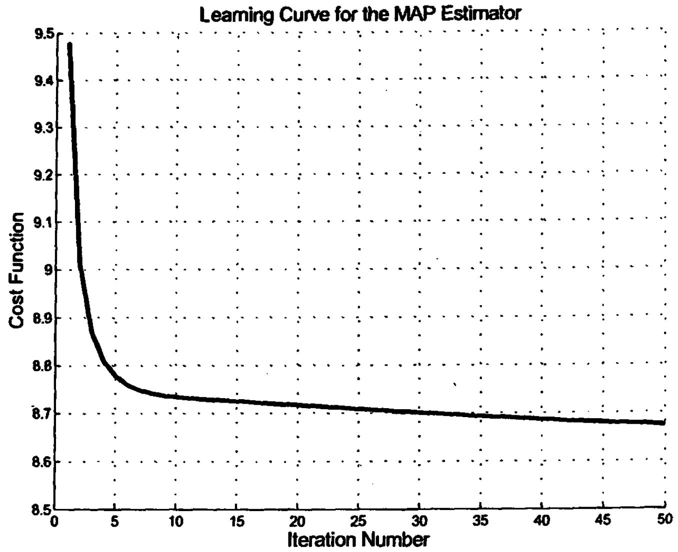
Σχήμα 3.15: Η την συνάρτηση κόστους $L(\mathbf{z}, \mathbf{s})$ σε σχέση με τον αριθμό των επαναλήψεων για την εικόνα *Βιβλία* του σχήματος 3.14.



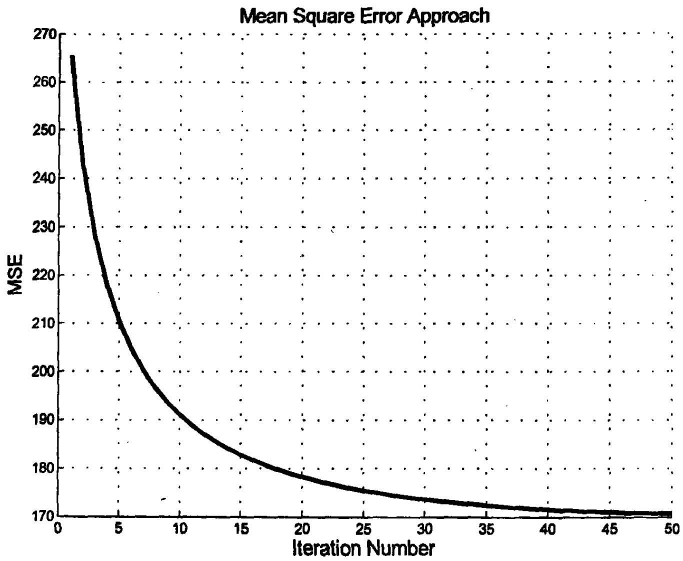
Σχήμα 3.16: Καμπύλη μέσου τεραγωνικού σφάλματος μεταξύ της πραγματικής εικόνας υψηλής ανάλυσης και της εκτίμησης της εικόνας υψηλής ανάλυσης για την εικόνα *Βιβλία* του σχήματος 3.14. Η ανακατασκευασμένη εικόνα υψηλής ανάλυσης έχει PSNR 25.87 dB.



Σχήμα 3.17: *Πίνακας Οφθαλμίατρου.* (α) – (δ) Εικόνες χαμηλής ανάλυσης. (ε) Ανακατασκευασμένη εικόνα υψηλής ανάλυσης. Ο αριθμός των χαμηλής ανάλυσης εικόνων που χρησιμοποιήθηκαν για την ανακασκευή της (ε) είναι 4.



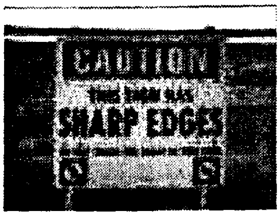
Σχήμα 3.18: Η την συνάρτηση κόστους $L(\mathbf{z}, \mathbf{s})$ σε σχέση με τον αριθμό των επαναλήψεων για την εικόνα *Πίνακας Οφθαλμίατρου* του σχήματος 3.17.



Σχήμα 3.19: Καμπύλη μέσου τεραγωνικού σφάλματος μεταξύ της πραγματικής εικόνας υψηλής ανάλυσης και της εκτίμησης της εικόνας υψηλής ανάλυσης για την εικόνα *Πίνακας Οφθαλμίατρου* του σχήματος 3.17. Η ανακατασκευασμένη εικόνα υψηλής ανάλυσης έχει PSNR 25.82 dB.

(α)

(β)

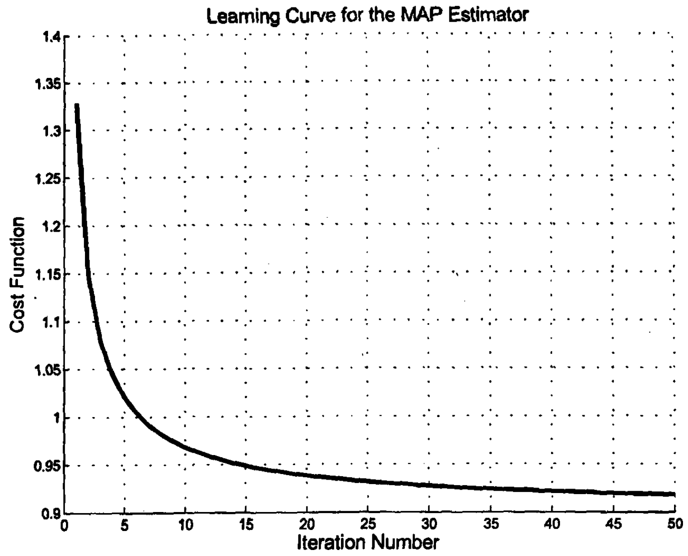


(γ)

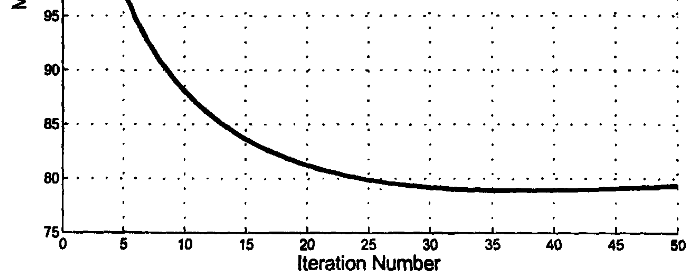
(δ)

(ε)

Σχήμα 3.20: *Πινακίδα.* (α) – (δ) Εικόνες χαμηλής ανάλυσης. (ε) Ανακατασκευασμένη εικόνα υψηλής ανάλυσης. Ο αριθμός των χαμηλής ανάλυσης εικόνων που χρησιμοποιήθηκαν για την ανακασκευή της (ε) είναι 4.



Σχήμα 3.21: Η συνάρτηση κόστους $L(\mathrm{z}, \mathrm{s})$ σε σχέση με τον αριθμό των επαναλήψεων για την εικόνα *Πιναχίδα* του σχήματος 3.20.



Σχήμα 3.22: Καμπύλη μέσου τετραγωνικού σφάλματος μεταξύ της πραγματικής εικόνας υψηλής ανάλυσης και της εκτίμησης της εικόνας υψηλής ανάλυσης για την εικόνα *Πιναχίδα* του σχήματος 3.20. Η ανακατασκευασμένη εικόνα υψηλής ανάλυσης έχει PSNR 29.09 dB.

Πίνακας 3.1: Αριθμητικά αποτελέσματα για τις παραμέτρους υπέρθεσης που παρουσιάζονται στα σχήματα 3.8, 3.11, 3.14, 3.17 και 3.20 με μέγεθος θορύβου 30 dB.

Εικόνα	Σφάλμα δ^x		Σφάλμα δ^y		Σφάλμα θ°	
	mean	**std**	**mean**	**std**	**mean**	**std**
Εξώφυλλο (σχ. 3.8)	0.20	0.08	0.19	0.08	-0.01	0.07
Αυτοκίνητο (σχ. 3.11)	0.61	0.03	0.02	0.01	0.16	0.07
Βιβλία (σχ. 3.14)	-0.09	0.13	-0.26	0.03	-0.08	0.01
Πίνακας Οφθαλμίατρου (σχ. 3.17)	0.32	0.14	-0.07	0.07	0.68	0.28
Πινακίδα (σχ. 3.20)	0.47	0.02	-0.01	0.06	0.96	0.55

Πίνακας 3.2: Αριθμητικά αποτελέσματα για το PSNR που παρουσιάζονται στα σχήματα 3.8, 3.11, 3.14, 3.17 και 3.20 με μέγεθος θορύβου 30 dB.

Εικόνα	**PSNR**		
	mean	**std**	**median**
Εξώφυλλο (σχ. 3.8)	24.34	0.34	24.31
Αυτοκίνητο (σχ. 3.11)	26.67	0.41	26.45
Βιβλία (σχ. 3.14)	25.22	0.28	25.09
Πίνακας Οφθαλμίατρου (σχ. 3.17)	25.47	0.42	25.35
Πιναχίδα (σχ. 3.20)	28.29	0.27	28.19

Πίνακας 3.3: Αριθμητικά αποτελέσματα για τις παραμέτρους υπέρθεσης που παρουσιάζονται στα σχήματα 3.8, 3.11, 3.14, 3.17 και 3.20 με μέγεθος θορύβου 20 dB.

Εικόνα	Σφάλμα δ^x		Σφάλμα δ^y		Σφάλμα θ^o	
	mean	**std**	**mean**	**std**	**mean**	**std**
Εξώφυλλο (σχ. 3.8)	0.08	0.08	0.04	0.03	-0.03	0.14
Αυτοκίνητο (σχ. 3.11)	0.58	0.11	0.05	0.09	-0.11	0.07
Βιβλία (σχ. 3.14)	0.61	0.12	0.02	0.10	0.36	0.03
Πίνακας Οφθαλμίατρου (σχ. 3.17)	0.86	0.16	-0.21	0.09	0.09	0.02
Πιναχίδα (σχ. 3.20)	0.76	0.14	-0.25	0.07	0.80	0.08

Πίνακας 3.4: Αριθμητικά αποτελέσματα για το PSNR που παρουσιάζονται στα σχήματα 3.8, 3.11, 3.14, 3.17 και 3.20 με μέγεθος θορύβου 20 dB.

Εικόνα	**PSNR**		
	mean	**std**	**median**
Εξώφυλλο (σχ. 3.8)	22.39	0.30	22.34
Αυτοκίνητο (σχ. 3.11)	21.23	0.37	21.22
Βιβλία (σχ. 3.14)	23.45	0.19	23.44
Πίνακας Οφθαλμίατρου (σχ. 3.17)	23.30	0.34	23.24
Πιναχίδα (σχ. 3.20)	23.18	0.53	23.24

Κεφαλαιο 4

Υπερανάλυση Εικονάς με Μεγιστοποιήση της Αμοιβαίας Πληροφορίας

4.1 Εισαγωγή

Η υπέρθεση εικόνας απαιτεί τον μετασχηματισμό μίας εικόνας σε μια άλλη με τέτοιο τρόπο ώστε τα αντίστοιχα εικονοστοιχεία να βρίσκονται στις ίδιες χωρικές συντεταγμένες. Στην ενότητα αυτή, θα περιγράψουμε τη χρήση της αμοιβαίας πληροφορίας ως μέτρο για την υπέρθεση δύο εικόνων και προτείνουμε έναν αλγόριθμο για υπερανάλυση εικόνας, που στηρίζεται στην υπέρθεση με χρήση αμοιβαίας πληροφορίας. Η αναφορά [38] κάνει λόγο για την χρήση της αμοιβαίας πληροφορίας στο πρόβλημα της υπερανάλυσης εικόνας. Από τις βασικότερες εργασίες στον τομέα της υπέρθεσης με χρήση αμοιβαίας πληροφορίας είναι αυτές των P. Viola και W. Wells [36] και Maes et al. [20]. Οι μέθοδοι που στηρίζονται την αμοιβαία πληροφορία είναι ιδιαίτερα διαδεδομένες. Μπορούν να εφαρμοστούν πάνω σε δεδομένα που έχουν αλλοιωθεί από θόρυβο, ή να εφαρμοστούν σε προβλήματα υπέρθεσης πολυτροπικών (*multimodal*) εικόνων, δηλαδή εικόνων που έχουν ληφθεί με διαφορετικό μηχανισμό λήψης. Μάλιστα, η μέθοδος αυτή θεωρείται ως η καλύτερη μέθοδος για υπέρθεση εικόνων.

Ορισμός 4.1. Έστω δύο τυχαίες μεταβλητές A και B με περιθώριες κατανομές $p_A(a)$ και $p_B(b)$ και από κοινού κατανομή $p_{AB}(a, b)$. Η αμοιβαία πληροφορία $I(A, B)$, μετρά τον βαθμό εξάρτησης μεταξύ των τυχαίων μεταβλητών A και B και ορίζεται ως η ποσότητα

$$I(A,B) = \sum_a \sum_b p_{AB}(a,b) \log \left(\frac{p_{AB}(a,b)}{p_A(a) \cdot p_B(b)} \right)$$

Οι μέθοδοι που βασίζονται στην αμοιβαία πληροφορία έχουν να κάνουν με την αναζήτηση των παραμέτρων εκείνων που μεγιστοποιούν την αμοιβαία πληροφορία. Η ιδέα έχει παρθεί από την Θεωρία Πληροφορίας και είναι ένα μέτρο σύγκρισης μεταξύ δύο τυχαίων μεταβλητών. Με άλλα λόγια, αυτό μπορεί να εκφραστεί ως η διαφορά μεταξύ της από κοινού εντροπίας $H(A,B)$ των τυχαίων μεταβλητών A και B και του αθροίσματος των περιθώριων εντροπιών $H(A)$, $H(B)$

$$I(A,B) = H(A) + H(B) - H(A,B) \tag{4.1}$$

όπου

- H(A) είναι η εντροπία της τυχαίας μεταβλητής A

$$H(A) = -\sum_a p_A(a) \log \left(p_A(a)\right)$$

- H(A,B) είναι η από κοινού εντροπία των τυχαίων μεταβλητών A και B

$$H(A,B) = -\sum_a \sum_b p_{AB}(a,b) \log \left(p_{AB}(a,b)\right)$$

Πλήθος εργασιών έχουν δημοσιευθεί στον τομέα αυτό. Οι σημαντικότερες ανήκουν στους P. Viola και W. Wells [36] και Maes et al. [20] ενώ ενδεικτικά αναφέρουμε τις εργασίες των J. Pluim et al. [30], Y. Chen et al. [6] και J. Pluim et al. [29]. Σημαντική είναι και η εργασία του C. Studholme [32], στην οποία εισάγεται η έννοια της κανονικοποιημένης αμοιβαίας πληροφορίας, $NMI(A,B)$ που δίνεται από την παρακάτω εξίσωση:

$$NMI(A,B) = \frac{H(A) + H(B)}{H(A,B)} \tag{4.2}$$

Η κανονικοποιημένη αμοιβαία πληροφορία αποτελεί καλύτερο μέτρο σύγκρισης οταν υπολογίζεται σε περιοχές που επικαλύπτονται. Στην παρούσα εργασία, έχει γίνει χρήση αυτής της ποσότητας.

Για να πετύχουμε αμεταβλητότητα στην επικάλυψη των χαμηλής ανάλυσης εικόνων, χρειαζόμαστε ένα μέτρο ανεξάρτητο από τις μεταβολές στις περιθώριες εντροπίες $H(A)$ και $H(B)$, στην περιοχή επικάλυψης των δύο εικόνων. Για το λόγο αυτό, χρησιμοποιήσαμε ένα κανονικοποιημένο μέτρο, που είναι ο λόγος των περιθώριων εντροπιών προς την από κοινού εντροπία.

Κάθε αλλαγή στην αβεβαιότητα της εικόνα και επομένως στις περιθώριες εντροπίες, δεν θα έχει αποτέλεσμα στην ευθυγράμμιση των εικόνων. Η μεγιστοποίηση της ποσότητας $NMI(A,B)$, σκοπό έχει την εύρεση του μετασχηματισμού εκείνου όπου η από κοινού εντροπία ελαχιστοποιείται σε σχέση με τις περιθώριες.

Στον πίνακα 4.1 παρουσιάζονται μερικές ιδιότητες της αμοιβαίας πληροφορίας. Για την απόδειξη αυτών παραπέμπουμε τον αναγνώστη στην αναφορά [35].

Πίνακας 4.1: Ιδιότητες αμοιβαίας πληροφορίας.

Μη-αρνητικότητα:	$I(A,B) \geq 0$
Ανεξαρτησία:	$I(A,B) = 0 \Leftrightarrow p_{AB}(a,b) = p_A(a) \cdot p_B(b)$
Συμμετρία:	$I(A,B) = I(B,A)$
Αυτο-πληροφορία:	$I(A,A) = H(A)$
Κάτω όριο:	$I(A,B) \geq 0$
Άνω όριο:	$I(A,B) \leq \min\big(H(A),H(B)\big)$
	$\leq \big(H(A)+H(B)\big)/2$
	$\leq \max\big(H(A),H(B)\big)$
	$\leq H(A,B)$
	$\leq \big(H(A)+H(B)\big)$

4.2 Κριτήριο Υπέρθεσης Αμοιβαίας Πληροφορίας

Οι πολυτροπικές εικόνες που παρουσιάζουν το ίδιο αντικείμενο, αναπαριστούν μετρήσεις διαφορετικών ιδιοτήτων για το αντικείμενο αυτό. Αν και οι εντάσεις της εικόνας αντιστοιχούν στο ίδιο αντικείμενο, μπορεί να είναι πολύ διαφορετικές μεταξύ των εικόνων που έχουν ληφθεί με διαφορετικό τρόπο. Γενικά, δεν είναι ανεξάρτητες παρατηρήσεις όπως συμβαίνει στην πραγματικότητα για το ίδιο το αντικείμενο. Οι τιμές των εντάσεων μεταξύ διαφορετικών εικόνων που αναπαριστούν το ίδιο αντικείμενο, δεν είναι ανεξάρτητες ποσότητες, αλλά στατιστικά συσχετισμένες μετρήσεις.

Έστω $\mathcal{A}$ και $\mathcal{B}$ είναι δύο εικόνες που σχετίζονται γεωμετρικά με το μετασχηματισμό υπέρθεσης $\mathbf{T}_\alpha$ με παραμέτρους α, έτσι ώστε τα εικονοστοιχεία $\mathbf{p}$ της $\mathcal{A}$ με ένταση a αντιστοιχούν στα εικονοστοιχεία $\mathbf{T}_\alpha(\mathbf{p})$ της $\mathcal{B}$ με ένταση b. Η εξάρτηση μεταξύ των a και b ή αλλιώς η πληροφορία που περιέχει η μία τιμή για την άλλη, μετράται με την αμοιβαία πληροφορία $I(A,B)$ των μεταβλητών $A = \{a\}$ και $B = \{b\}$

$$a = \mathcal{A}(\mathbf{p})$$
$$b = \mathcal{B}\big(\mathbf{T}_\alpha(\mathbf{p})\big)$$
$$I(A,B) = \sum_a \sum_b p_{AB}(a,b) \log_2 \left(\frac{p_{AB}(a,b)}{p_A(a) \cdot p_B(b)} \right) \tag{4.3}$$

όπου $p_{AB}(a,b)$, $p_A(a)$ και $p_B(b)$ είναι οι από κοινού και οι περιθώριες κατανομές των εντάσεων a και b, αντίστοιχα. Οι εκτιμήσεις γι' αυτές τις κατανομές, μπορούν να ληφθούν από τα από κοινού και περιθώρια ιστογράμματα των τμημάτων που επικαλύπτονται και για τις δύο εικόνες. Η σχέση $p_{AB}(a,b)$ μεταξύ των εντάσεων a και b και επομένως η αμοιβαία πληροφορία $I(A,B)$ εξαρτάται από τον μετασχηματισμό υπέρθεσης $\mathbf{T}_\alpha$, δηλαδή την υπέρθεση των εικόνων. Το κριτήριο υπέρθεσης που βασίζεται στην αμοιβαία πληροφορία υποθέτει πως οι εικόνες είναι ευθυγραμμισμένες από τον μετασχηματισμό υπέρθεσης $\mathbf{T}_{\alpha^*}$ για τον οποίο η αμοιβαία πληροφορία $I(A,B)$ είναι μέγιστη

$$\alpha^* = \arg\max_{\alpha} I(A,B).$$

Αν θεωρήσουμε ότι και οι δύο περιθώριες κατανομές $p_A(a)$ και $p_B(b)$, είναι ανεξάρτητες από τις παραμέτρους υπέρθεσης α, τότε το κριτήριο υπέρθεσης της αμοιβαίας πληροφορίας, ελάχιστοποιεί την από κοινού εντροπία $H_{AB}(A,B)$. Αν καμία από τις δύο κατανομές $p_A(a)$ ή $p_B(b)$, δεν είναι ανεξάρτητη των παραμέτρων α, όπου είναι η περίπτωση που η μία εικόνα περιέχεται ακριβώς στην άλλη, τότε το κριτήριο υπέρθεσης της αμοιβαίας πληροφορίας, ελάχιστοποιεί την υπό συνθήκη εντροπία $H(A|B)$ ή $H(B|A)$. Ωστόσο, αν και οι δύο εικόνες επικαλύπτονται μερικώς, το ποσοστό της επικάλυψης θα αλλάζει όταν οι παράμετροι υπέρθεσης α μεταβάλλονται και τότε οι περιθώριες κατανομές $p_A(a)$ και $p_B(b)$ και συνεπώς και οι εντροπίες αυτών, θα εξαρτώνται από το α. Το κριτήριο υπέρθεσης της αμοιβαίας πληροφορίας λαμβάνει υπ' όψιν του το παραπάνω, επομένως η εξίσωση (4.1) ερμηνεύεται ως: "Η μεγιστοποίηση της αμοιβαίας πληροφορίας τείνει στην μεγιστοποίηση των δύο πρώτων όρων και την ελαχιστοποίηση του τελευταίου όρου."

Υπάρχουν περιπτώσεις που το κριτήριο υπέρθεσης της αμοιβαίας πληροφορίας αποτυγχάνει. Τέτοιες αποτυχίες συμβαίνουν λόγω της ανεπαρκούς αμοιβαίας πληροφορίας που έχουμε για τις εικόνες, λόγω της ασάφειας ως προς τη σχέση της έντασης μεταξύ των δύο εικόνων ή λόγω της ανικανότητας να εκτιμήσουμε σωστά την αμοιβαία πληροφορία εξ' αιτίας του μικρού αριθμού των εικονοστοιχείων της εικόνας.

Για χαμηλής ανάλυσης εικόνες ή στην περίπτωση που η περιοχή επικάλυψης των δύο εικόνων είναι μικρή, η στατιστική σχέση μεταξύ των δύο εικόνων πρεπει να εξαχθεί από ένα μικρό αριθμό εικονοστοιχείων. Στις περιπώσεις αυτές η αμοιβαία πληροφορία μπορεί να παρουσιάσει πολλά τοπικά μέγιστα γύρω από τη σωστή λύση υπέρθεσης.

Επιλέγουμε την μία από τις εικόνες που θα υπερτεθούν να είναι η κινούμενη εικόνα $\mathcal{A}$, από την οποία θα πάρουμε τα δείγματα $s \in S$ και θα μετασχηματιστεί με ένα γεωμετρικό μετασχηματισμό $\mathbf{T}_\alpha$ με παραμέτρους α, στην εικόνα αναφοράς $\mathcal{B}$. Το σύνολο S μπορεί να περιέχει όλα τα εικονοστοιχεία της $\mathcal{A}$ ή ένα υποσύνολο αυτών.

Το από κοινού ιστόγραμμα των εντάσεων της εικόνας $H_\alpha(a,b)$ για το ποσοστό της επικάλυψης $s \in S_\alpha \subset S$ της $\mathcal{A}$ και $\mathcal{B}$ κατασκευάζεται τοποθετώντας σε κάδους ζευγάρια εντάσεων $\big(a(s), b(\mathbf{T}_\alpha(s))\big)$ για κάθε $s \in S_\alpha$. Στην γενική περίπτωση, ο μετασχηματισμός $\mathbf{T}_\alpha(s)$ για το εικονοστοιχείο s, δεν θα συμπίπτει με ένα σημείο του πλέγματος στην εικόνα $\mathcal{B}$ και επομένως, είναι αναγκαία η χρήση παρεμβολής στην εικόνα αναφοράς για να βρούμε την τιμή της έντασης $b(\mathbf{T}_\alpha(s))$. Έχουν προταθεί διάφοροι τρόποι παρεμβολής. Για παράδειγμα, η παρεμβολή του κοντινότερου γείτονα για την εικόνα $\mathcal{B}$, δεν εγγυάται καλή ακρίβεια, μιας και είναι ευαίσθητη σε μετατοπίσεις μεγαλύτερες του ενός εικονοστοιχείου. Άλλες μέθοδοι παρεμβολής, όπως η τριπλά-γραμμική (*trilinear*) ή ανώτερης τάξης, ενδέχεται να εισάγουν νέες εντάσεις που δεν υπαρχουν στην εικόνα αναφοράς $\mathcal{B}$. Με αυτόν το τρόπο, οδηγούμαστε σε απρόβλεπτες αλλαγές της περιθώριας κατανομής $p_{\mathcal{B},\alpha}(b)$ της εικονας αναφοράς, για μικρές αλλαγές της παραμέτρου α. Στη βιβλιογραφία, έχει προταθεί μια καινούρια μέθοδος παρεμβολής που επιλύει το παραπάνω πρόβλημα και ονομάζεται μερικού όγκου τριπλά-γραμμική παρεμβολή [19], για περισσότερες λεπτομέριες παραπέμπουμε τον αναγνώστη στην αναφορά αυτή.

Οι βέλτιστες παράμετροι υπέρθεσης α^*, υπολογίζονται από την μεγιστοποίηση της ποσότητας $I(\alpha)$. Για την μεγιστοποίηση αυτής της ποσότητας (ή ελαχιστοποίηση της $-I(\alpha)$), χρησημοποιήσαμε την μέθοδο ελαχιστοποίησης SIMPLEX [17], που είναι ενσωματωμένη στο πακέτο βελτιστοποίησης της Matlab (συνάρτηση fminsearch).

4.3 Περιγραφή της Μεθόδου

Στην ενότητα αυτή, θα περιγράψουμε την εφαρμογή της υπέρθεσης εικόνας με χρήση αμοιβαίας πληροφορίας στο πρόβλημα της υπερανάλυσης εικόνας.

Σκοπός μας είναι να εκτιμήσουμε τις παραμέτρους μετασχηματισμού υπέρθεσης (διάνυσμα μετατόπισης $\mathbf{T}$ και πίνακας γωνιών περιστροφής $\mathbf{R}$). Δεν έχουμε λοιπόν, παρά να εφαρμόσουμε την ιδέα που αναπτύχθηκε στην προηγούμενη ενότητα μεγιστοποιώντας την αμοιβαία πληροφορία $I(A, B)$, για τις δύο χαμηλής ανάλυσης εικόνες A και B. Στην συνέχεια, εφαρμόζουμε τον αλγόριθμο υπερανάλυσης εικόνας (αλγόριθμος 1), όπως αυτός αναπτύχθηκε στο κεφάλαιο 2. Σε γενικές γραμμές ο αλγόριθμος που προτείνουμε είναι:

Αλγόριθμος 4 Μέθοδος υπερανάλυσης εικόνας βασισμένη στην αμοιβαία πληροφορία.

βήμα 1: Αρχικοποίηση των παραμέτρων μετασχηματισμού υπέρθεσης ($\mathbf{R}$, $\mathbf{T}$) με χρήση περιγραφέων SIFT και πρώτη εκτίμηση της υψηλής ανάλυσης εικόνας.

βήμα 2: Μεγιστοποίηση της αμοιβαίας πληροφορίας (MI) και εύρεση των παραμέτρων μετασχηματισμού υπέρθεσης ($\mathbf{R}$, $\mathbf{T}$).

βήμα 3: Εκτίμηση της εικόνας υψηλής ανάλυσης.

βήμα 4: Επιστροφή στο βήμα 2 μέχρι να συγκλίνει.

4.4 Πειραματικά Αποτελέσματα

Στην ενότητα αυτή θα παρουσιάσουμε πειραματικά αποτελέσματα από την χρήση της αμοιβαίας πληροφορίας στο πρόβλημα της υπερανάλυσης εικόνας, ώστε να δείξουμε τις δυνατότητες της μεθόδου που προτείνουμε. Τα πειράματα που εκτελέσαμε, στόχο έχουν να συγκρίνουν τη μέθοδο αυτή με την αντίστοιχη μέθοδο που χρησιμοποιεί τους περιγραφείς SIFT (βλ. κεφάλαιο 3).

Οι μετρικές που χρησιμοποιούμε για την ποσοτικοποίηση της ποιότητας των υποβαθμισμένων εικόνων, τα επίπεδα του θορύβου στις υποβαθμισμένες εικόνες χαμηλής ανάλυσης και την ποιότητα των ανακατασκευασμένων αποτελεσμάτων χρησιμοποιούμε τη μέγιστη τιμή του λόγου σήματος προς θόρυβο (*PSNR*).

$$PSNR = 10 \log_{10} \frac{(255)^2}{\|\mathbf{f} - \mathbf{g}\|^2}$$

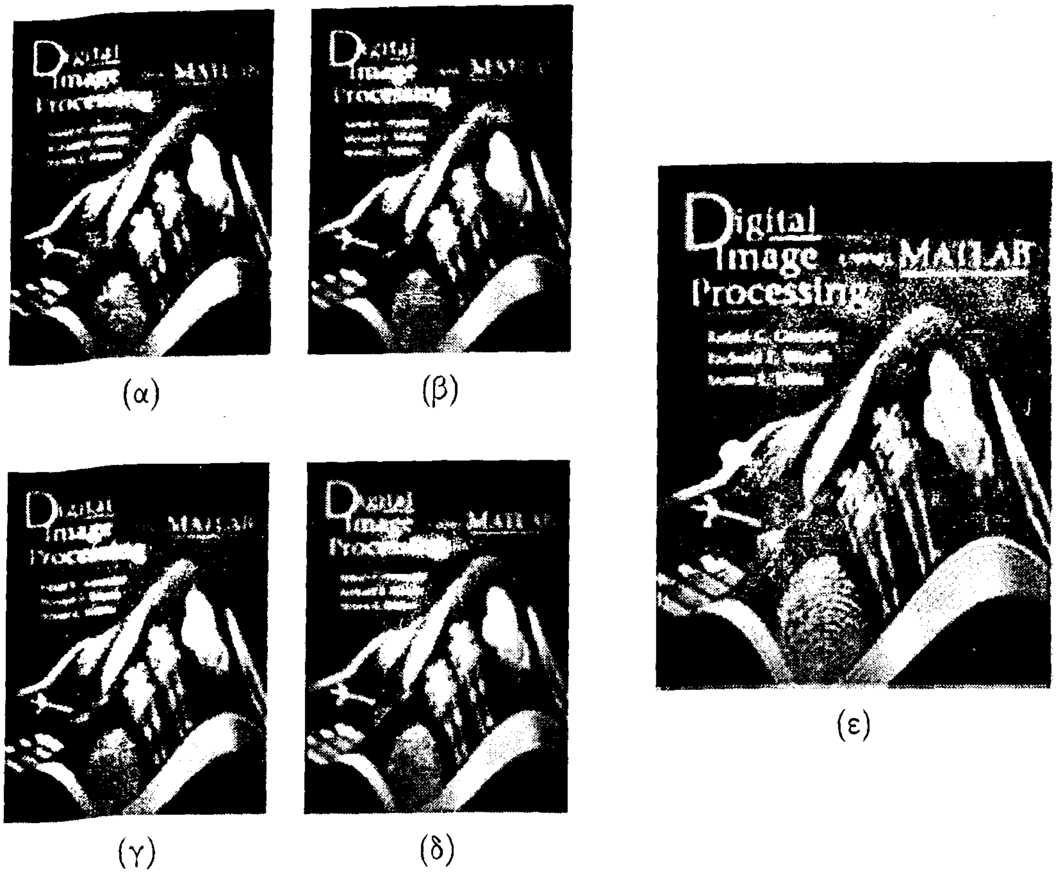
όπου $\mathbf{f}$ και $\mathbf{g}$ είναι η πραγματική και η ανακατασκευασμένη εικόνα, αντιστοίχως.

Τα πειράματά μας έχουν διεξαχθεί τόσο σε πραγματικά, όσο και σε τεχνητά σύνολα δεδομένων. Να αναφέρουμε ότι, για την εκτέλεση των πειραμάτων μας θεωρήσαμε συμπαγείς μετασχηματισμούς για περιστροφή και μετατόπιση. Με βάση τις παρατηρήσεις μας, σύμπεράναμε ότι η μέθοδος δίνει ικανοποιητικά αποτελέσματα για εικόνες που είναι περιστραμένες μέσα στο διάστημα $[-5, 5]$ και μετατοπισμένες στο διάστημα $[-10, 10]$. Για μεγαλύτερες τιμές παρατηρήσαμε ότι η μέθοδος παρουσιάζει σημαντικά επίπεδα θορύβου *ringing*. Αυτο οφείλεται στον αλγόριθμο ελαχιστοποίησης SIMPLEX που χρησιμοποιήσαμε, ο οποίος είναι ένας αλγόριθμος τοπικής ελαχιστοποίησης και η αρχικοποίηση καθορίζει σημαντικά το αποτέλεσμα.

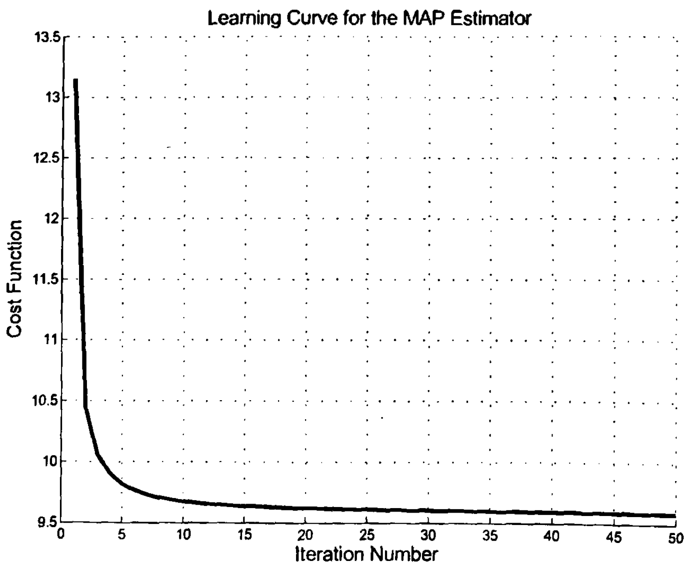
Έχουμε θολώσει τις εικόνες μας με ένα Gaussian χαμηλοπερατό φίλτρο τυπικής απόκλισης 1 και μεγέθους παραθύρου 5×5 ενώ παράλληλα οι εικόνες έχουν υποβαθμιστεί με έναν παράγοντα $L_1 = L_2 = 2$. Μετά από δοκιμές, παρατηρήσαμε ότι το μέγιστο πλήθος επαναλήψεων που απαιτείται για να συγκλίνει ο αλγόριθμος σε μία ικανοποιητική λύση είναι 50. Τέλος, έχει προστεθεί Gaussian λευκός θόρυβος στις τεχντητά παραχθείσες εικόνες, ενώ τα πειράματα έχουν διεξαχθεί με διάφορα μεγέθη θορύβου.

Η δομή με την οποία θα παρουσιάσουμε τα αποτελέσματά μας είναι η ακόλουθη: δίνονται οι χαμηλής ανάλυσης εικόνες, παρουσιάζεται η ανακατασκευασμένη εικόνα υψηλής ανάλυσης και δύο πίνακες που περιέχουν τις τιμές των σφαλμάτων (μέση τιμή και τυπική απόκλιση) για τις παραμέτρους υπέρθεσης (μετατόπιση και περιστροφή) που προέκυψαν από τον αλγόριθμο 4, για μεγέθη θορύβου 30 dB και 20 dB, αντοίστιχα. Τέλος, παρουσιάζονται δύο ακόμα πίνακες που περιέχουν αριθμητικά αποτελέσματα (μέση τιμή, τυπική απόκλιση και ενδιάμεση τιμή) για το PSNR, ο κάθε ένας παρουσιάζει τα αποτελέσματα για μεγέθη θορύβου 30 dB και 20 dB, αντοίστιχα. Το σύνολο πειραμάτων για κάθε εικόνα είναι 10.

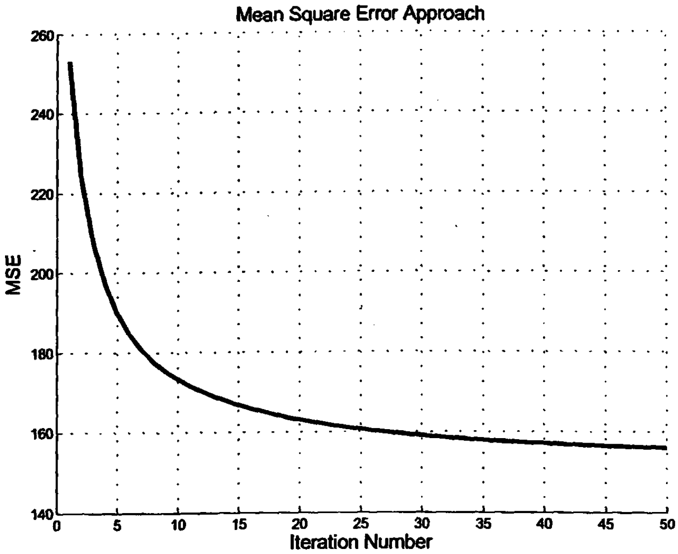
Όπως παρατηρούμε από τις καμπύλες του μέσου τετραγωνικού σφάλματος, εύκολα μπορούμε να δουμε την σημαντική βελτίωση που παρουσιάζει η μέδοθος αυτή σε σχέση με την προηγούμενη μέδοδο, που χρησιμοποιεί τους περιγραφείς SIFT (βλ. κεφάλαιο 3). Η διαφορά των PSNR στις ανακατασχευασμένες εικόνες υψηλής ανάλυσης φτάνει κατα μέσο όρο τα 2 dB, που είναι σημαντική βελτίωση.



Σχήμα 4.1: *Εξώφυλλο.* (α) – (δ) Εικόνες χαμηλής ανάλυσης. (ε) Ανακατασκευασμένη ε υψηλής ανάλυσης. Ο αριθμός των χαμηλής ανάλυσης εικόνων που χρησιμοποιήθηκαν γι ανακασκευή της (ε) είναι 4.



Σχήμα 4.2: Η συνάρτηση κόστους $L(\mathbf{z}, \mathbf{s})$ σε σχέση με τον αριθμό των επαναλήψεων γιο εικόνα *Εξώφυλλο* του σχήματος 4.1.



Σχήμα 4.3: Καμπύλη μέσου τεραγωνικού σφάλματος μεταξύ της πραγματικής εικόνας υψηλής ανάλυσης και της εκτίμησης της εικόνας υψηλής ανάλυσης για την εικόνα *Εξώφυλλο* του σχήματος 4.1. Η ανακατασκευασμένη εικόνα υψηλής ανάλυσης έχει PSNR 26.14 dB.

Πίνακας 4.2: Αριθμητικά αποτελέσματα για τις παραμέτρους υπέρθεσης που παρουσιάζονται στα σχήματα 4.1, 4.4, 4.7, 4.10 και 4.13 με μέγεθος θορύβου 30 dB.

Εικόνα	Σφάλμα δ^x		Σφάλμα δ^y		Σφάλμα θ^o	
	mean	**std**	**mean**	**std**	**mean**	**std**
Εξώφυλλο (σχ. 4.1)	-0.12	0.02	0.41	0.07	0.59	0.23
Αυτοκίνητο (σχ. 4.4)	0.21	0.12	0.31	0.21	0.95	0.07
Βιβλία (σχ. 4.7)	0.34	0.20	0.31	0.09	0.57	0.18
Πίνακας Οφθαλμίατρου (σχ. 4.10)	0.25	0.13	0.08	0.11	0.01	0.09
Πινακίδα (σχ. 4.13)	0.03	0.01	-0.04	0.02	0.41	0.02

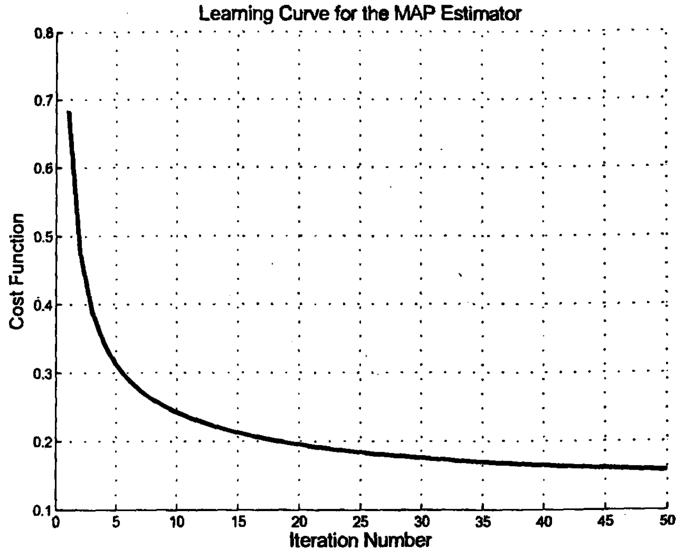
(α) (β)

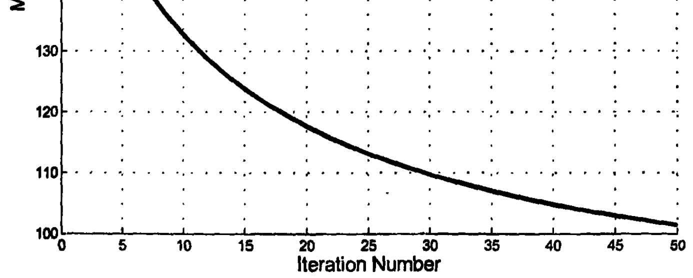
(γ) (δ)

(ε)

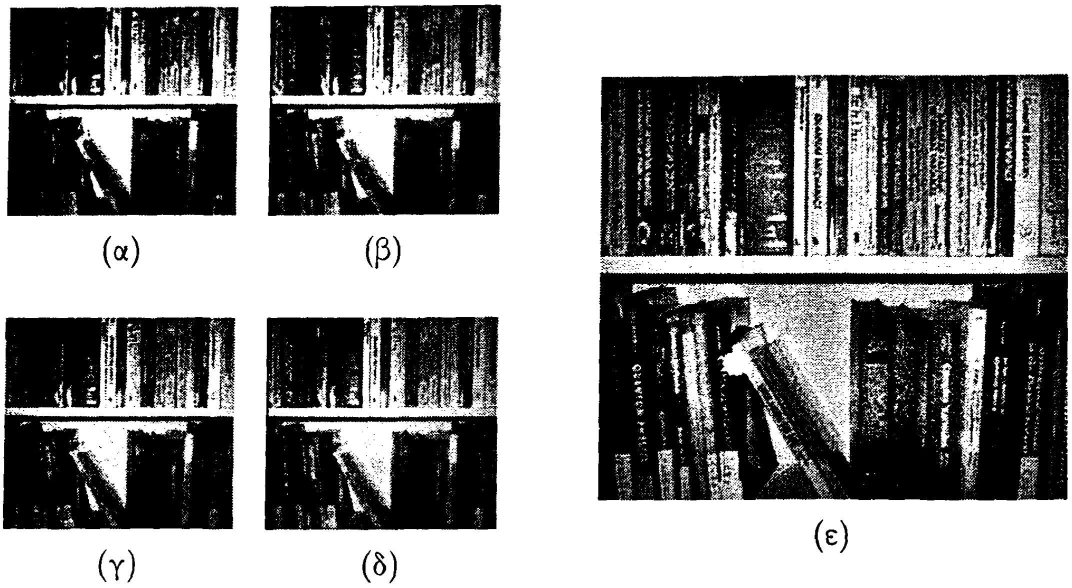
Σχήμα 4.4: *Αυτοκίνητο.* (α) – (δ) Εικόνες χαμηλής ανάλυσης. (ε) Ανακατασκευ υψηλής ανάλυσης. Ο αριθμός των χαμηλής ανάλυσης εικόνων που χρησιμοποιή ανακασκευή της (ε) είναι 4.



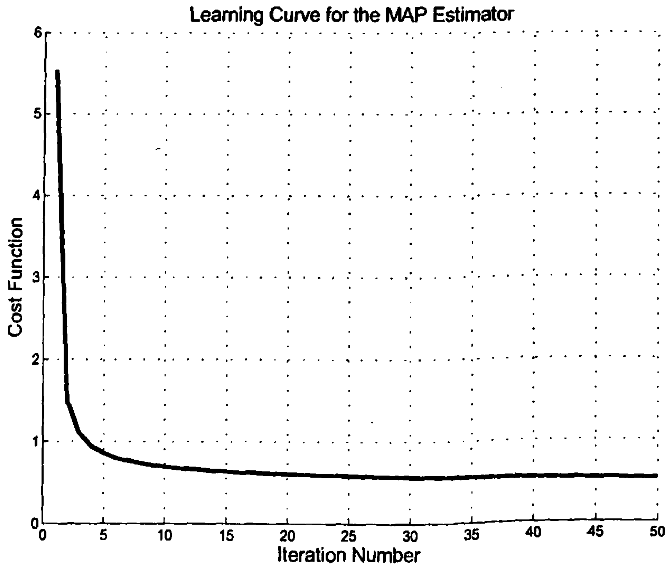
Σχήμα 4.5: Η συνάρτηση κόστους $L(\mathbf{z}, \mathbf{s})$ σε σχέση με τον αριθμό των επαναλήψεων για την εικόνα *Αυτοκίνητο* του σχήματος 4.4.



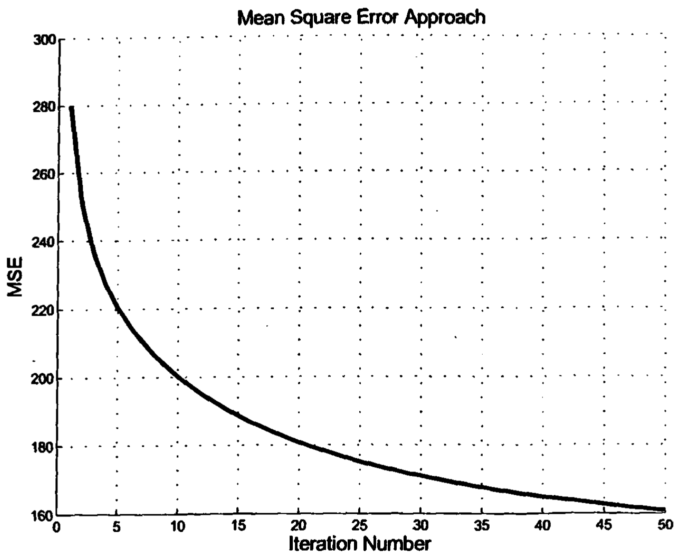
Σχήμα 4.6: Καμπύλη μέσου τετραγωνικού σφάλματος μεταξύ της πραγματικής εικόνας υψηλής ανάλυσης και της εκτίμησης της εικόνας υψηλής ανάλυσης για την εικόνα *Αυτοκίνητο* του σχήματος 4.4. Η ανακατασκευασμένη εικόνα υψηλής ανάλυσης έχει PSNR 28.13 dB.



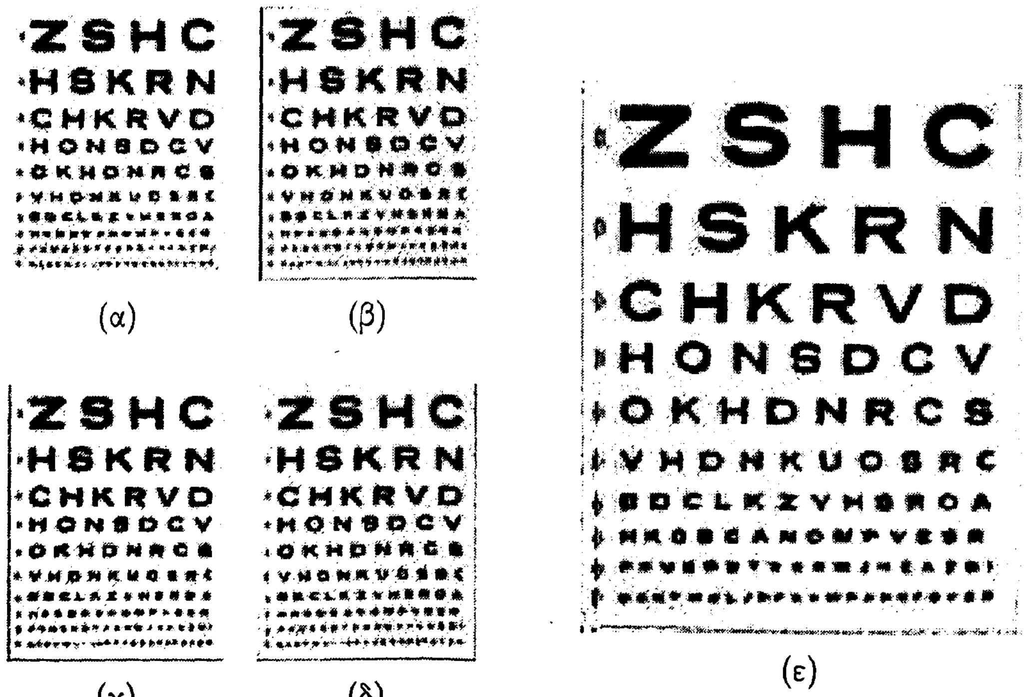
Σχήμα 4.7: *Βιβλία.* (α) – (δ) Εικόνες χαμηλής ανάλυσης. (ε) Ανακατασκευασμένη εικόνα υψηλής ανάλυσης. Ο αριθμός των χαμηλής ανάλυσης εικόνων που χρησιμοποιήθηκαν για την ανακασκευή της (ε) είναι 4.



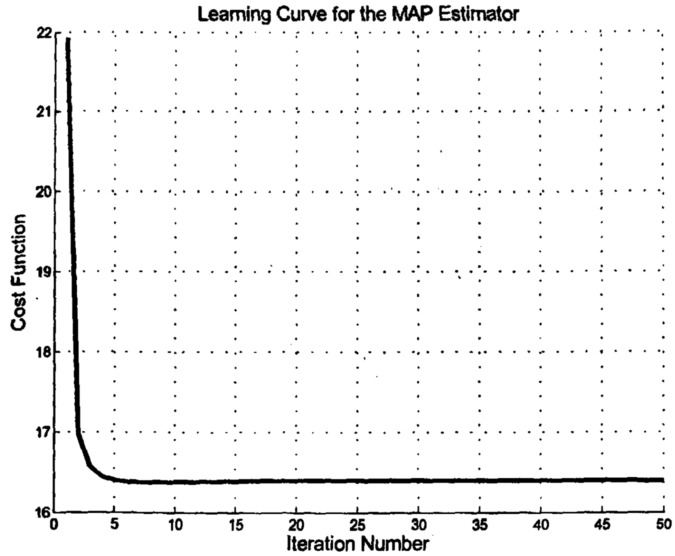
Σχήμα 4.8: Η συνάρτηση κόστους $L(\mathbf{z}, \mathbf{s})$ σε σχέση με τον αριθμό των επαναλήψεων για την εικόνα *Βιβλία* του σχήματος 4.7.



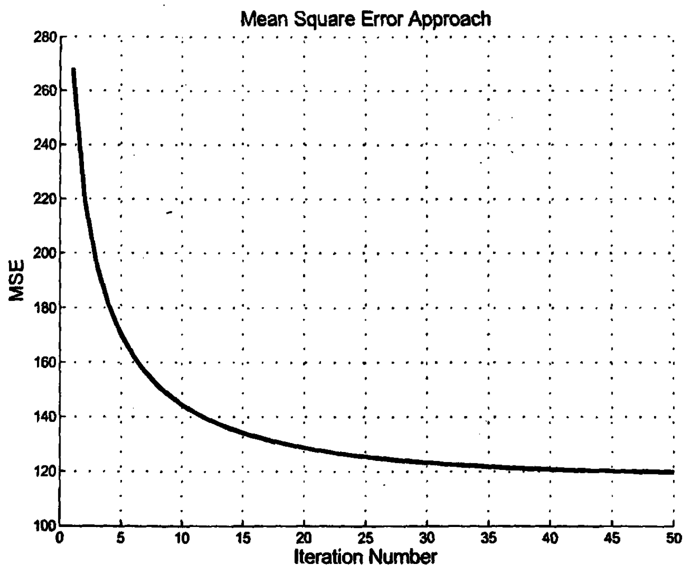
Σχήμα 4.9: Καμπύλη μέσου τετραγωνικού σφάλματος μεταξύ της πραγματικής εικόνας υψηλής ανάλυσης και της εκτίμησης της εικόνας υψηλής ανάλυσης για την εικόνα *Βιβλία* του σχήματος 4.7. Η ανακατασκευασμένη εικόνα υψηλής ανάλυσης έχει PSNR 26.06 dB.



Σχήμα 4.10: *Πίνακας Οφθαλμίατρου.* (α) – (δ) Εικόνες χαμηλής ανάλυσης. (ε) Ανακατασκευασμένη εικόνα υψηλής ανάλυσης. Ο αριθμός των χαμηλής ανάλυσης εικόνων που χρησιμοποιήθηκαν για την ανακασκευή της (ε) είναι 4.

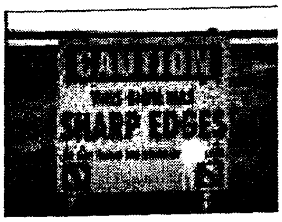


Σχήμα 4.11: Η συνάρτηση κόστους $L(\mathbf{z}, \mathbf{s})$ σε σχέση με τον αριθμό των επαναλήψεων για την εικόνα *Πίνακας Οφθαλμίατρου* του σχήματος 4.10.

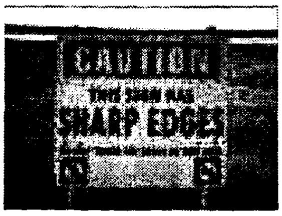


Σχήμα 4.12: Καμπύλη μέσου τεραγωνικού σφάλματος μεταξύ της πραγματικής εικόνας υψηλής ανάλυσης και της εκτίμησης της εικόνας υψηλής ανάλυσης για την εικόνα *Πίνακας Οφθαλμίατρου* του σχήματος 4.10. Η ανακατασκευασμένη εικόνα υψηλής ανάλυσης έχει PSNR 27.33 dB.

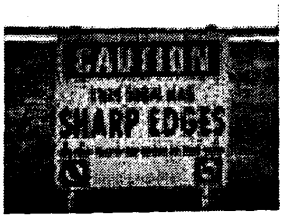
(α)



(β)



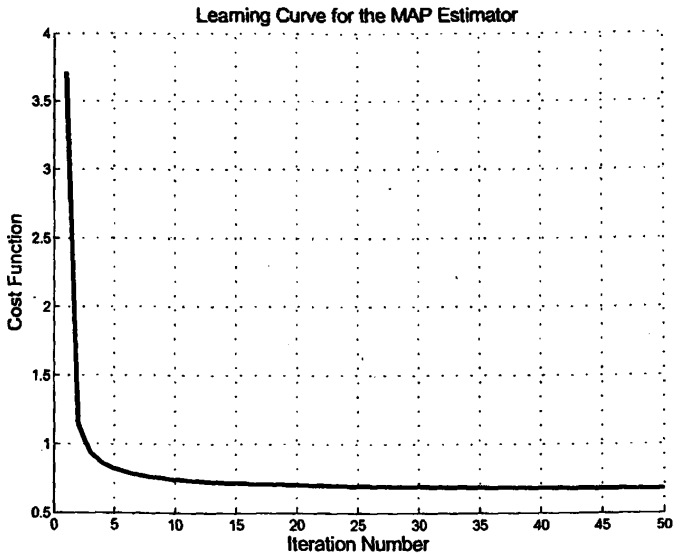
(γ)



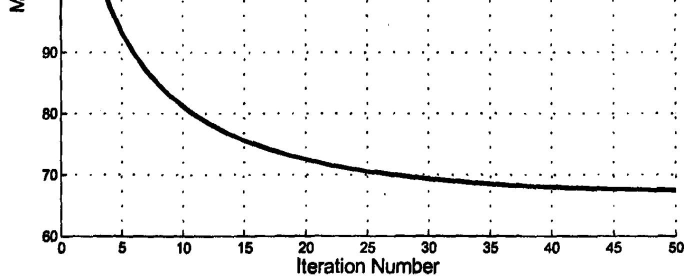
(δ)

(ε)

Σχήμα 4.13: *Πινακίδα.* (α) - (δ) Εικόνες χαμηλής ανάλυσης. (ε) Ανακατασκευασμένη εικόνα υψηλής ανάλυσης. Ο αριθμός των χαμηλής ανάλυσης εικόνων που χρησιμοποιήθηκαν για την αναχασκευή της (ε) είναι 4.



Σχήμα 4.14: Η συνάρτηση κόστους $L(\mathbf{z}, \mathbf{s})$ σε σχέση με τον αριθμό των επαναλήψεων για την εικόνα *Πινακίδα* του σχήματος 4.13.



Σχήμα 4.15: Καμπύλη μέσου τετραγωνικού σφάλματος μεταξύ της πραγματικής εικόνας υψηλής ανάλυσης και της εκτίμησης της εικόνας υψηλής ανάλυσης για την εικόνα *Πινακίδα* του σχήματος 4.13. Η ανακατασκευασμένη εικόνα υψηλής ανάλυσης έχει PSNR 29.80 dB.

Πίνακας 4.3: Αριθμητικά αποτελέσματα για το PSNR που παρουσιάζονται στα σχήματα 4.1, 4.4, 4.7, 4.10 και 4.13 με μέγεθος θορύβου 30 dB.

Εικόνα	**PSNR**		
	mean	**std**	**median**
Εξώφυλλο (σχ. 4.1)	26.78	0.41	26.52
Αυτοκίνητο (σχ. 4.4)	27.80	0.59	27.69
Βιβλία (σχ. 4.7)	25.34	0.42	25.45
Πίνακας Οφθαλμίατρου (σχ. 4.10)	27.01	0.11	27.06
Πινακίδα (σχ. 4.13)	29.11	0.57	29.03

Πίνακας 4.4: Αριθμητικά αποτελέσματα για τις παραμέτρους υπέρθεσης που παρουσιάζονται στα σχήματα 4.1, 4.4, 4.7, 4.10 και 4.13 με μέγεθος θορύβου 20 dB.

Εικόνα	**Σφάλμα δ^x**		**Σφάλμα δ^y**		**Σφάλμα θ^o**	
	mean	**std**	**mean**	**std**	**mean**	**std**
Εξώφυλλο (σχ. 4.1)	-0.31	0.05	0.27	0.03	0.51	0.07
Αυτοκίνητο (σχ. 4.4)	0.35	0.18	0.39	0.02	0.38	0.31
Βιβλία (σχ. 4.7)	-0.05	0.03	0.06	0.01	0.18	0.21
Πίνακας Οφθαλμίατρου (σχ. 4.10)	0.39	0.09	0.16	0.06	0.51	0.16
Πινακίδα (σχ. 4.13)	0.41	0.19	-0.02	0.02	0.11	0.21

Πίνακας 4.5: Αριθμητικά αποτελέσματα για το PSNR που παρουσιάζονται στα σχήματα 4.1, 4.4, 4.7, 4.10 και 4.13 με μέγεθος θορύβου 20 dB.

Εικόνα	**PSNR**		
	mean	**std**	**median**
Εξώφυλλο (σχ. 4.1)	24.83	0.36	24.85
Αυτοκίνητο (σχ. 4.4)	21.56	0.19	21.47
Βιβλία (σχ. 4.7)	23.64	0.31	23.52
Πίνακας Οφθαλμίατρου (σχ. 4.10)	24.59	0.26	24.64
Πινακίδα (σχ. 4.13)	24.48	0.27	24.39

Κεφάλαιο 5

Υπερανάλυση Εικόνας με Χρήση Ευρωστων Εκτιμητών

5.1 Εισαγωγή

Στα προηγούμενα κεφάλαια παρουσιάστηκε και αναλύθηκε το πρόβλημα της υπερανάλυσης εικόνας από την σκοπιά της εύρεσης των παραμέτρων υπέρθεσης. Στο κεφάλαιο αυτό, θα παρουσιάσουμε μια μέθοδο για υπερανάλυση εικόνας εξετάζοντας ασυνέχειες στις χαμηλής ανάλυσης εικόνες, οι οποίες "παραβιάζουν" την υπόθεση για διατήρηση των δεδομένων και τη χωρική συνάφεια. Πιο συγκεκριμένα, τέτοιου είδους περιορισμοί στο σύνολο δεδομένων των χαμηλής ανάλυσης εικόνων προκαλούνται από μετρήσεις που δεν ακολουθούν το κυρίαρχο μοντέλο (*outliers*). Πολλά προβλήματα της Υπολογιστικής Όρασης περιέχουν τέτοιου είδους δεδομένα.

Για το λόγο αυτό, επιλέγεται η χρήση εύρωστων εκτιμητών (*robust M-estimators*) προσαρμόζοντας κατάλληλα την διαδικασία εκτίμησης σε κάθε πλαίσιο χαμηλής ανάλυσης. Η μέθοδος που προτείνεται στην παρούσα εργασία, εξαλείφει τις "κακές" μετρήσεις (*outliers*).

Παρόμοιες μελέτες πάνω στον ίδιο τομέα έχουν προταθεί από τους N. El-Yamany και P. Papamichalis [8]. Στην εργασία τους εφαρμόζουν έναν εύρωστο εκτιμητή πάνω στην αντικειμενική συνάρτηση της παραγώγου και προσαρμόζουν τη διαδικασία της εκτίμησης σε κάθε χαμηλής ανάλυσης πλαίσιο. Η μέθοδος αυτή εξαλείφει αποκλίνοντα σημεία με αποτελεσματικό τρόπο και μπορεί εύκολα να εφαρμοστεί σε υπερανάλυση έγχρωμων εικόνων με έντονες λεπτομέρεις, χωρίς την ανάγκη ομολοποίησης. Στην εργασία των V. Patanavijit και S. Jitapunkul [25] παρουσιάζεται ένα παραπλήσιο μοντέλο υπερανάλυσης εικόνας, το οποίο βασίζεται σε στοχαστικές τεχνικές ελαχιστοποιώντας την συνάρτηση κόστους. Για περισσότερες λεπτομέριες παραπέμπουμε τον αναγνώστη στην αναφορά αυτή. Μία ακόμη ενδιαφέρουσα προσέγγιση στο πρόβλημα της υπερανάλυσης εικόνας παρουσιάζεται στην εργασία των Zomet et al. [39]. Η τεχνική αυτή, εκτιμά την ενδιάμεση τιμή της παραγώγου για όλες τις χαμηλής ανάλυσης εικόνες εικόνες και στη συνέχεια, εκτελεί το βήμα της ενημέρωσης για την εκίμηση της υψηλής ανάλυσης εικόνας.

Κύριος σκοπός ενός εύρωστου εκτιμητή είναι να:

1. Περιγράψει τη δομή που ταιριάζει καλύτερα στα δεδομένα.

2. Αναγνωρίσει αποκλίνοντα δεδομένα (*outliers*).

Ο καλύτερος εκτιμητής είναι εκείνος που συμπεριφέρεται καλύτερα στην χειρότερη κατανομή για παραμετρικά μοντέλα. Αυτό είναι ένα κριτήριο που μπορεί να χρησιμοποιηθεί για να παράγουμε μια πλειάδα εκτιμητών.

Για να θέσουμε το θέμα σε μία πιο στέρεα βάση, ένας εύρωστος εκτιμητής διευθετεί το πρόβλημα της εύρεσης των βέλτιστων παραμέτρων $\theta = [\theta_0, \ldots, \theta_n]^T$, του μοντέλου, για ένα σύνολο δεδομένων $\mathbf{x} = \{x_0, x_1, \ldots, x_S\}$. Για να ταιριάξουμε ένα μοντέλο, σκοπός μας είναι να βρούμε τις τιμές των παραμέτρων θ που ελαχιστοποιούν τα υπολοιπόμενα σφάλματα $r_i(x_i, \theta_i)$:

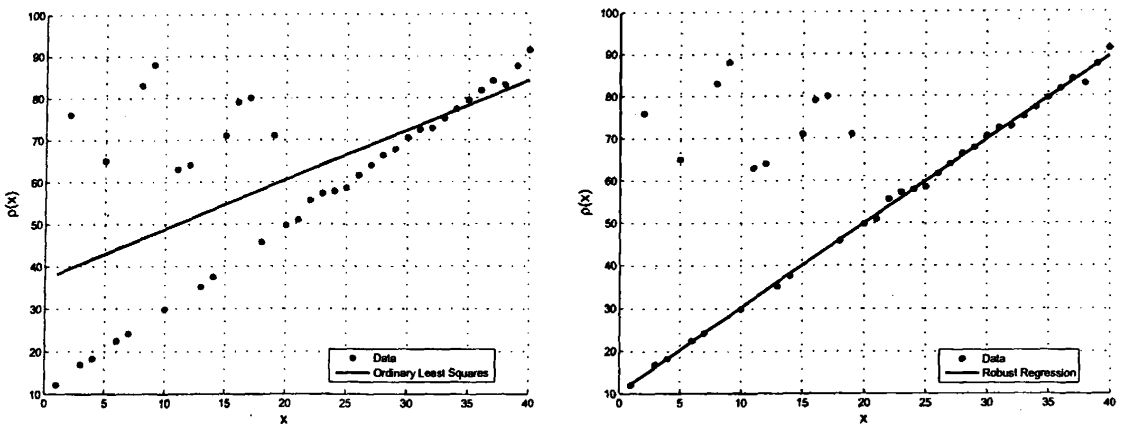
$$\min_{\theta} \sum_{s \in S} \rho(r_s(x_s, \theta); \sigma_s)$$

όπου σ_s είναι ένας παράγοντας κλίμακας και ελέγχει το σημείο στο οποίο ο εκτιμητής θεωρεί τις μετρήσεις αποκλίνουσες και ρ είναι ο εκτιμητής. Η συνάρτηση ρ ονομάζεται και *M-estimator* μίας και αντιστοιχεί στην εκτίμηση της μεγιστοποίησης της πιθανοφάνειας (*Maximum Likelihood*). Η ευρωστία ενός εκτιμητή έχει να κάνει με την ευαισθησία του σε μετρήσεις που αποκλίνουν από το επικρατόν μοντέλο.

5.2 Εύρωστοι Εκτιμητές

Ο πιο κοινός εκτιμητής που χρησημοποιείται συνήθως είναι ο εκτιμητής ελάχιστων τετραγώνων L_2.

Το σχήμα 5.1 δείχνει ένα παράδειγμα ταιριάσματος μίας ευθείας σε δεδομένα υπό την παρουσία αποκλίνοντων σημείων. Η εικόνα 5.1(α) παρουσιάζει πως η ευθεία για το ταίριασμα των δεδομένων με την χρήση ελαχίστων τετραγώνων αποκλίνει από τη σωστή λύση,



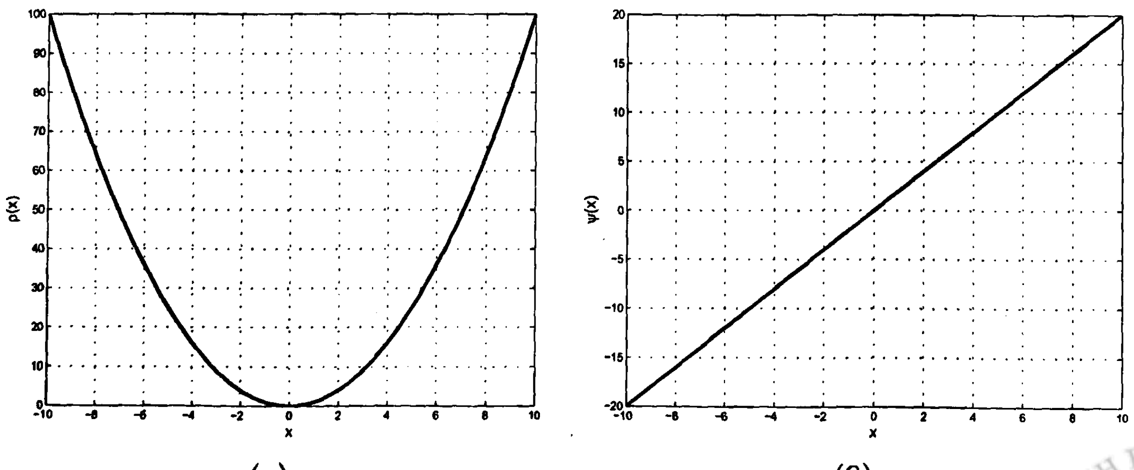
(α) (β)

Σχήμα 5.1: Ταιριάζοντας μία ευθεία γραμμή. Το μοντέλο για την πλειονότητα των δεδομένων είναι η ευθεία $y(x) = 2x + 10$. Υπάρχει ένας αριθμός από αποκλίνοντα σημεία που δεν έχουν καλή συμπεριφορά σε σχέση με το μοντέλο. (α) Ταίριασμα του μοντέλου στα δεδομένα με χρήση ελαχίστων τετραγώνων. (β) Ταίριασμα του μοντέλου στα δεδομένα με χρήση Lorentzian εκτιμητή.

υπό την παρουσία σημείων που δεν ακολοθούν το μοντέλο. Το ταίριασμα στο σχήμα 5.1(β) είναι πιο εύρωστο υπό την έννοια ότι απορρίπτει τα "κακά" σημεία και έτσι ανακτά ένα καλύτερο ταίριασμα για την πλειοψηφία των δεδομένων.

Το πρόβλημα με την προσέγγιση των ελαχίστων τετραγώνων είναι ότι τα αποκλίνοντα αυτά σημεία συνεισφέρουν αρκετά στην τελική λύση. Για να αναλύσουμε την συμπεριφορά ενός εκτιμητή πρέπει να λάβουμε υπ' όψη μας την συνάρτηση επιρροής (*influence function*). Η συνάρτηση επιρροής καθορίζει την πόλωση που μία μέτρηση έχει πανω στην λύση και ορίζεται ως η παράγωγος ψ, του εκτιμητή ($\psi(X) = d\rho(x)/dx$). Αν θεωρήσουμε για παράδειγμα των τετραγωνικό εκτιμητή τότε ισχύει:

$$\rho(x) = x^2, \quad \psi(x) = 2x \tag{5.1}$$



(α) (β)

Σχήμα 5.2: (α) Τετραγωνικός εκτιμητής $\rho(x)$. (β) Συνάρτηση επιρροής $\psi(x)$.

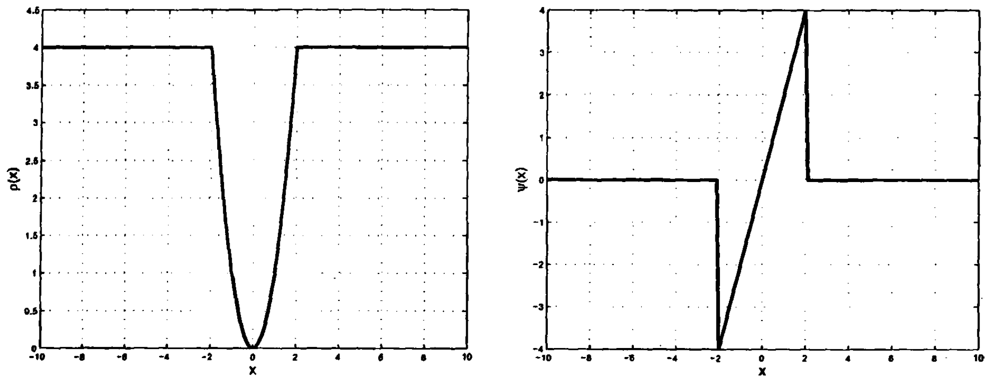
Η συνάρτηση επιρροής ψ, μετρά την επιρροή ενός στοιχείου για την τιμή της παραμέτρου εκτίμησης. Στο παράδειγμα του τετραγωνικού εκτιμητή η συνάρτηση επιρροής αυξάνεται γραμμικά χωρίς όριο, πράγμα που δεν κατωχυρώνει την ευρωστία του εκτιμητή. Το σχήμα 5.2 παρουσιάζει παρουσιάζει τον τετραγωνικό εκτιμητή και την συνάρτηση επιρροής του.

Υπάρχουν κάποιοι συγκεκριμένοι περιορισμοί που ένας εύρωστος εκτιμητής πρέπει να πληρεί:

- Πρέπει να έχει φραγμένη συνάρτηση επιρροής.
- Ο εκτιμητής πρέπει να είναι μοναδικός. Αυτό σημαίνει ότι η αντικειμενική συνάρτηση του διανύσματος παραμέτρων θ πρέπει να έχει μοναδικό ελάχιστο. Αυτό απαιτεί η συνάρτηση ρ να είναι κυρτή για τις παραμέτρους θ.

Για να αυξήσουμε την ευρωστία, ένας εκτιμητής πρέπει να είναι όσο το δυνατόν πιο αυστηρός σε ότι έχει να κάνει με αποκλίνοντα σημεία. Το επόμενο βήμα είναι να χρησιμοποιήσουμε τον εκτιμητή των κατατετμημένων ελαχίστων τετραγώνων (*Truncated Least Squares*):

$$\rho(x) = \begin{cases} x^2, & \text{αν } |x| \le \sigma, \\ \sigma^2, & \text{αν } |x| > \sigma, \end{cases} \qquad \psi(x) = \begin{cases} 2x, & \text{αν } |x| \le \sigma \\ 0, & \text{αν } |x| > \sigma \end{cases} \tag{5.2}$$



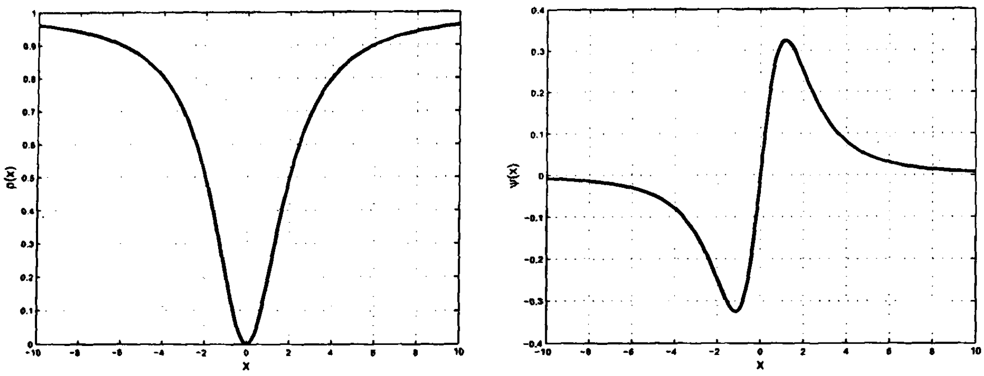
(α) (β)

Σχήμα 5.3: (α) Truncated Least Squares εκτιμητής $\rho(x)$. (β) Συνάρτηση επιρροής $\psi(x)$.

Στο σχήμα 5.3(α) απεικονίζεται ο Truncated Least Squares εκτιμητής. Στη συνάρτηση αυτή υπεισέρχεται η ρυθμιστική παράμετρος σ, που καθορίζει το σημείο εκείνο πέρα από το οποίο τα δεδομένα θεωρούνται ότι δεν ακολουθούν το μοντέλο. Το σχήμα 5.3(β) δείχνει την συνάρτηση επιρροής για τον εκτιμητή αυτόν. Η συνάρτηση επιρροής είναι φραγμένη από το σ. Ο εκτιμητής αυτός μειώνει τις επιρροές για μεγάλα σφάλματα.

Για να αυξήσουμε την ευρωστία ακόμα περισσότερο, μπορούμε να χρησιμοποιήσουμε εκτιμητές για τους οποίους η επιρροή των αποκλίνοντων σημείων τείνει στο μηδέν, πιο ομαλά από τον Truncated Least Squares εκτιμητή. Ένα παράδειγμα τέτοιου εκτιμητή είναι ο Geman-McClure εκτιμητής:

$$\rho(x) = \frac{x^2}{\sigma^2 + x^2}, \quad \psi(x) = \frac{2\sigma^2 x}{(\sigma^2 + x^2)^2} \tag{5.3}$$



(α) (β)

Σχήμα 5.4: (α) Geman-McClure εκτιμητής $\rho(x)$. (β) Συνάρτηση επιρροής $\psi(x)$.

Παρατηρώντας το σχήμα 5.4 και εξετάζοντας την συνάρτηση ψ βλέπουμε ότι η επιρροή των αποκλίνοντων σημείων τείνει στο μηδέν. Παρατηρούμε ότι η συνάρτηση ρ (σχήμα 5.4(α)) είναι συνεχής και διαφορήσιμη και η συνάρτηση επιρροής της ψ (σχήμα 5.4(β)) έχει μια αρκετά απλή μορφή. Αυτό ακρίβώς είναι που κάνει την χρήση τους αρκετά διαδεδομένη.

Είναι αρκετά δύσκολο να επιλέξει κανείς μια συνάρτηση ρ για γενική χρήση σε όλες τις περιπτώσεις χωρίς να είναι κάπως αυθαίρετος. Η χρήση του κατάλληλου εκτιμητή εξαρτάται από το πρόβλημα. 'Ενα αρκετά δύσκολο σημείο που απαιτεί αρκετή προσοχή είναι η επιλογή της ρυθμιστικής παραμέτρου σ.

5.3 Εύρωστη Υπερανάλυση

Στην βιβλιογραφία έχουν προταθεί αρκετά μοντέλα για το πρόβλημα της ανακατασκευής μιας εικόνας υψηλής ανάλυσης. Το μαθηματικό μοντέλο που ακολουθούμε στην διατριβή αυτή και έχει παρουσιαστεί αναλυτικά στο κεφάλαιο 2 είναι αυτο που φαίνεται παρακάτω:

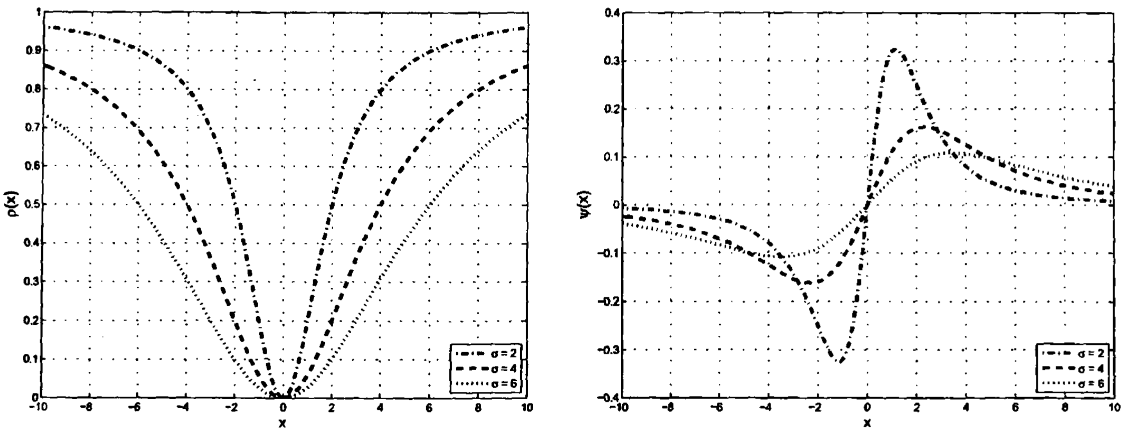
$$\mathbf{y}_k = \mathbf{W}_k \mathbf{z} + \mathbf{n}_k \tag{5.4}$$

όπου $k = 1, 2, \ldots, p$ ο αριθμός των χαμηλής ανάλυσης εικόνων, $\mathbf{y}_k$ και $\mathbf{z}$ η k-οστή χαμηλής ανάλυσης εικόνα μεγέθους $N_1 \times N_2$ με $M = N_1 N_2$ και η επιθυμητή εικόνα υψηλής ανάλυσης μεγέθους $L_1 N_1 \times L_2 N_2$ με $N = L_1 N_1 L_2 N_2$, $\mathbf{W}_k$ είναι ο πίναχας υποδειγματισμού για το k-οστό χαμηλής ανάλυσης πλαίσιο και τέλος $\mathbf{n}_k$ είναι ο προσθετικός θόρυβος. Ακολουθώντας επομένως αυτό το μοντέλο παρατήρησης και συνδιάζοντας το πρόβλημα της υπερανάλυσης εικόνας με την χρήση των εύρωστων εκτιμητών, προκύπτει το ακόλουθο πρόβλημα ελαχιστοποίησης:

$$\begin{aligned} \mathbf{z}^* &= \arg\min_{\mathbf{z}} \sum_{k=1}^{pM} \rho\left(\sum_{m=1}^{N} \mathbf{W}_{k,m}\mathbf{z}_m - \mathbf{y}_k; \sigma_k\right) \\ &= \arg\min_{\mathbf{z}} \sum_{k=1}^{pM} \rho\left(\mathbf{E}_k; \sigma_k\right) \end{aligned} \tag{5.5}$$

όπου $\mathbf{E}_k = \sum_{m=1}^{N} (\mathbf{W}_{k,m}\mathbf{z}_m - \mathbf{y}_k)$ είναι το διάνυσμα σφαλμάτων που αντιστιχούν στο k-οστό χαμηλής ανάλυσης πλαίσιο, σ είναι η παράμετρος του εκτιμητή και ρ είναι η συνάρτηση του εύρωστου εκτιμητή, για να ελαχιστοποιηθεί:

$$\frac{\partial}{\partial \mathbf{z}} \sum_{k=1}^{pM} \rho\left(\mathbf{E}_k; \sigma_k\right) = 0 \Rightarrow \sum_{k=1}^{pM} \left(\mathbf{W_k}\right)^T \psi\left(\mathbf{E}_k; \sigma_k\right) = 0. \tag{5.6}$$



(α) (β)

Σχήμα 5.5: (α) Geman-McClure εκτιμητής $\rho(x)$. (β) Συνάρτηση επιρροής $\psi(x)$. Οι συναρτήσεις απεικονίζονται για διάφορες τιμές της ρυθμιστικής παραμέτρου σ.

Παραβιάσεις στο μαθηματικό μοντέλο της εξίσωσης (5.4) οδηγούν σε μεγάλα σφάλματα ($\mathbf{E}_k$), τα οποία αναφέρονται ως *outliers*, και βλάπτουν σε πολυ μεγάλο βαθμό την διαδικασία της ανακατασκευής αν η διαδικασία της εκτίμησης δεν εξαλείψει την συνεισφορά που έχουν στην εκτίμηση της υψηλής ανάλυσης εικόνας. Όπως φαίνεται και στο σχήμα 5.5, η επιλογή της ρυθμιστικής παραμέτρου σ για έναν εκτιμητή έχει πολύ μεγάλη σημασία για το χειρισμό μετρήσεων που δεν ανταποκρίνονται στο μοντέλο. Στα σφάλματα που ξεπερνούν την τιμή της παραμέτρου σ ανατίθεται μικρότερο βάρος καθώς το σφάλμα μεγαλώνει, έτσι μετρήσεις που δεν ακολουθούν το μοντέλο εξαλείφονται. Επίσης, για μικρότερες τιμές του σ η συνάρτηση επιρροής φθίνει πιο γρήγορα, αναθέτοντας έτσι μικρότερα βάρη σε σφάλματα που ξεπερνούν την τιμή της παραμέτρου σ. Αν επιλέξουμε την τιμή της παραμέτρου αυτής να είναι μικρή, η συνεισφορά απ' όλα τα χαμηλής ανάλυσης πλαίσια θα απορριφθεί, οδηγώντας έτσι σε κακές εκτιμήσεις της εικόνας υψηλής ανάλυσης, λόγω της ανεπαρκούς πληροφορίας που διαθέτουμε από τις χαμηλής ανάλυσης εικόνες. Από την άλλη μεριά, αν επιλέξουμε μεγάλη τιμή για την παράμετρο σ, μετρήσεις που δεν ανταποκρίνονται στο μοντέλο θα συνεισφέρουν στην εκτίμηση της υψηλής ανάλυσης εικόνας, έχοντας ως αποτέλεσμα η έκτιμηση της εικόνας να περιέχει ανακριβείς μετρήσεις. Στο [31] αναφέρεται ένας τρόπος υπολογισμού της παραμέτρου αυτής λαμβάνοντας υπ' όψην τα υπολοιπόμενα σφάλματα, ένω στην εργασία των El-Yamany και Papamichalis [8] ακολουθείται μια διαφορετική προσέγγιση για την εκτίμηση της παραμέτρου αυτής.

5.3.1 Υπολογισμός Παραμέτρου Κατωφλίωσης

Από την ανάλυση που παρουσιάστηκε στην προηγούμενη ενότητα, είναι κατανοητό ότι είναι απαραίτητη μία διαδικασία για τον υπολογισμό της ρυθμιστικής παραμέτρου σ, που να αποφασίζει ποια είναι εκείνα τα στοιχεία που θα λαμβάνονται υπ' όψη στους υπολογισμούς. Από την διαδικασία αυτή, ανατίθενται διαφορετικές τιμές στην παράμετρο σ για κάθε χαμηλής ανάλυσης πλαίσιο, ανάλογα με την ομοιότητά του με την εικόνα υψηλής ανάλυσης.

Η ενδιάμεση απόλυτη απόκλιση (*Median Absolute Deviation, MAD*) είναι ένα μέτρο στατιστικής διασποράς.

$$MAD = \underset{k}{\text{median}} \left\{ |r_k^n(\mathbf{W}_k \mathbf{z}^{n-1}; \mathbf{y}_k) - \underset{m}{\text{median}}(r_m^n(\mathbf{W}_m \mathbf{z}^{n-1}; \mathbf{y}_m))| \right\}$$

Είναι πίο εύρωστος εκτιμητής της παραμέτρου σ από την τυπική απόκλιση. Για παράδειγμα, η ενδιάμεση απόλυτη απόκλιση είναι ένα πιο εύρωστο στατιστικό εργαλείο, που είναι πιο ανθεκτικό στην παρουσία αποκλίνοντων δεδομένων σε σχέση με την τυπική απόκλιση. 'Οταν χρησιμοποιούμε την τυπική απόκλιση, οι αποστάσεις από την μέση τιμή είναι τετραγωνικές, έτσι σε μεγάλες αποκλίσεις ανατίθεται μεγάλο βάρος και επομένως, οι ακοκλίνουσες τιμές επιρρεάζουν το αποτέλεσμα. Στο κριτήριο της ενδιάμεσης απόλυτης απόκλισης από την άλλη, οι αποστάσεις των αποκλίνοντων δεδομένων τείνουν στη μέση τιμή.

'Οταν τα δεδομένα ακολουθούν κανονική κατανομή, μπορούμε να χρησιμοποιήσουμε το κριτήριο της ενδιάμεσης απόλυτης απόκλισης για την εκτίμηση της παραμέτρου κατωφλίωσης σ με την μορφή:

$$\sigma = K \cdot MAD$$

όπου K είναι μία σταθερά που εξαρτάται από την κατανομή.

Για μια συμμετρική κατανομή, το κριτήριο ενδιάμεσης τυπικής απόκλισης είναι η απόσταση μεταξύ του πρώτου και του δεύτερου (ισοδύναμα δεύτερου και τρίτου) τεταρτημόριου. Η παράμετρος κατωφλίωσης σ, όταν χρησιμοποιούμε το κριτήριο της ενδιάμεσης απόλυτης απόκλισης για κανονική κατανομή βρίσκεται στο 75% της κανονικής κατανομής με τυπική απόκλιση 1. 'Ετσι, για δεδομένα που ακολουθούν κανονική κατανομή τυπικής απόκλισης 1, η σταθερά K επιλέγεται ως $1/\Phi^{-1}(3/4) \approx 1.4826$, όπου Φ^{-1} είναι η αντίστροφη συνάρτηση κατανομής.

Η εκτίμηση της παραμέτρου σ γίνεται σε κάθε βήμα του επαναληπτικού αλγορίθμου υπερανάλυσης εικόνας. Η πιο δημοφιλής εκτίμηση της παραμέτρου αυτής [31] δίνεται από:

$$\sigma_k^n = 1.4826 \cdot MAD \tag{5.7}$$

για $k = 1, \ldots, p$ και $n = 1, 2, \ldots$ είναι το πλήθος των επαναλήψεων. 'Οπου $r_k^n(\mathbf{W}_k \mathbf{z}^{n-1}; \mathbf{y}_k) = (\sum_{m=1}^{N} \mathbf{W}_{k,m} \mathbf{z}_m^{n-1} - \mathbf{y}_k)^2$ είναι το υπολοιπόμενο σφάλμα για το k-οστό χαμηλής ανάλυσης πλαίσιο.

Με βάση την εξίσωση (5.7), ο υπολογισμός της ρυθμιστικής παραμέτρου σ για τον εύρωστο εκτιμητή, γίνεται με έναν αυτόματο τρόπο από όλες τις χαμηλής ανάλυσης εικόνες. Χρησιμοποιώντας ένα μέτρο ομοιότητας μεταξύ των χαμηλής ανάλυσης πλαισίων και της εκτίμησης της υψηλής ανάλυσης εικόνας, η τιμή της ρυθμιστικής παραμέτρου υπολογίζεται με τέτοιο τρόπο έτσι ώστε όταν η ομοιότητα είναι μικρή, δηλαδή το υπολοιπόμενο σφάλμα είναι μεγάλο, τότε το σ υπολογίζεται ως η ενδιάμεση τιμή των σφαλμάτων αυτών, έτσι ώστε να μειωθεί η επίδραση των μετρήσεων που δεν ακολουθούν το μοντέλο.

5.3.2 Εκτίμηση της Υψηλής Ανάλυσης Εικόνας

Για να βρούμε τη λύση στην εξίσωση (5.5) χρησιμοποιούμε τη μέθοδο της πτώσης της παραγώγου. Η συνάρτηση κόστους, χρησιμοποιώντας εύρωστους εκτιμητές, μπορεί να εκφραστεί πλέον ως εξής:

$$L(\mathbf{z}, \mathbf{s}) = \frac{1}{2\sigma_\eta^2} \sum_{m=1}^{pM} \rho\left(y_m - \sum_{r=1}^{N} w_{m,r}(\mathbf{s}) z_r;\ \sigma_m \right) + \frac{1}{2\lambda} \sum_{i=1}^{N} \left(\sum_{j=1}^{N} d_{i,j} z_j \right)^2 \tag{5.8}$$

όπου $\mathbf{s}$ είναι οι άγνωστες παράμετροι υπέρθεσης, οι οποίες εκτιμώνται με τη χρήση των περιγραφέων SIFT (βλ. κεφάλαιο 3).

Για να βρούμε την ενημέρωση της παραγώγου για την εκτίμηση της υψηλής ανάλυσης εικόνας, αρκεί να παραγωγίσουμε την συνάρτηση κόστους (5.8) ως προς το εικονοστοιχείο z_k για $k = 1, 2, \ldots, N$. Η μερική παράγωγος δίνεται από:

$$\begin{aligned} g_k(\mathbf{z}, \mathbf{s}) &= \frac{\partial L(\mathbf{z}, \mathbf{s})}{\partial z_k} \\ &= \frac{1}{\sigma_\eta^2} \sum_{m=1}^{pM} \psi\left(\sum_{r=1}^{N} w_{m,r}(\mathbf{s}) z_r - y_m;\ \sigma_m \right) w_{m,k}(\mathbf{s}) + \frac{1}{\lambda} \sum_{i=1}^{N} d_{i,k} \left(\sum_{j=1}^{N} d_{i,j} z_j \right). \end{aligned} \tag{5.9}$$

Στην εξίσωση (5.9), η παράγωγος εκτιμάται από την συνάρτηση επιρροής για το σφάλμα μεταξύ της υποβαθμισμένης εικόνας υψηλής ανάλυσης και της χαμηλής ανάλυσης εικόνας.

Η ενημέρωση της λύσης με χρήση της παραγώγου για κάθε εκτίμηση του εικονοστοιχείου υψηλής ανάλυσης δίνεται από:

$$\hat{z}_k^{n+1} = \hat{z}_k^n - \varepsilon^n g_k(\hat{\mathbf{z}}^n, \hat{\mathbf{s}}^n) \tag{5.10}$$

για $n = 0, 1, 2, \ldots$ και $k = 1, 2, \ldots, N$.

Δύο σημαντικά ζητήματα που προκύπτουν είναι η εκτίμηση των παραμέτρων εξομάλυνσης (σ_η^2 και λ) και του μεγέθους του βήματος στην εξίσωση (5.10), καθώς αυτά παίζουν σημαντικό ρόλο στην εκτίμηση της υψηλής ανάλυσης εικόνας. Ξαναγράφοντας την συνάρτηση κόστους δίνοντας περισσότερη σημασία στις παραμέτρους εξομάλυνσης έχουμε:

$$L(\mathbf{z}, \mathbf{s}) = \sum_{m=1}^{pM} \rho\left(y_m - \sum_{r=1}^{N} w_{m,r}(\mathbf{s}) z_r;\ \sigma_m \right) + \alpha \sum_{i=1}^{N} \left(\sum_{j=1}^{N} d_{i,j} z_j \right)^2$$

όπου $\alpha = \frac{\sigma_\eta^2}{\lambda}$. Σκοπός μας είναι να βρούμε μια εκτίμηση για τον συντελεστή α με την χρήση των εύρωστων εκτιμητών. Ακολουθώντας την ίδια ανάλυση που κάναμε στην §2.4 του κεφαλαίου 2, καταλήγουμε στο παρακάτω αποτέλεσμα για την παράμετρο εξομάλυνσης:

$$\alpha_k(\mathbf{z}) = \frac{\sum_{m=1}^{pM} \rho\left(y_m - \sum_{r=1}^{N} w_{m,r}(\mathbf{s}) z_r;\ \sigma_m \right)}{2\sum_{m=1}^{pM} y_m^2 - \sum_{i=1}^{N} \left(\sum_{j=1}^{N} d_{i,j} z_j \right)^2} \tag{5.11}$$

για $k = 1, 2, \ldots, p$. Στο σημείο αυτό να σημειώσουμε ότι για κάθε χαμηλής ανάλυσης πλαίσιο έχουμε διαφορετική τιμή για την παράμετρο εξομάλυνσης α.

Το επόμενο θέμα που αξίζει να σημειωθεί είναι η επιλογή του κατάλληλου βήματος ε^n στην εξίσωση (5.10). Επιλέγοντας ένα σταθερό βήμα είναι η απλούστερη προσέγγιση που μπορούμε να κάνουμε. Ωστόσο, η επιλογή αυτή έχει τα μειονεκτήματά της. Μια τέτοια επιλογή μπορεί να οδηγήσει σε σοβαρές αποκλίσεις του μοντέλου παρατήρησης και η εκτίμηση της υψηλής ανάλυσης εικόνας να είναι χειρότερη από την αναμενόμενη. Η κατάλληλη επιλογή για την παράμετρο του βήματος είναι αυτή που ακολουθήθηκε στην §2.3.3 του κεφαλαίου 2. Εφαρμόζοντας ένα κατάλληλο εκτιμητή καταλήγουμε στο παρακάτω αποτέλεσμα:

$$\varepsilon^n = \frac{\frac{1}{\sigma_\eta^2}\sum_{m=1}^{pM}\psi\Big(\sum_{r=1}^{N} w_{m,r}(\hat{\mathrm{s}}^n)\hat{z}_r^n - y_m;\ \sigma_m\Big)\gamma_m + \frac{1}{\lambda}\sum_{i=1}^{N}\bar{g}_i\Big(\sum_{j=1}^{N} d_{i,j}\hat{z}_j^n\Big)}{\frac{1}{\sigma_\eta^2}\sum_{m=1}^{pM}\gamma_m^2 + \frac{1}{\lambda}\sum_{i=1}^{N}\bar{g}_i^2} \tag{5.12}$$

όπου γ_m και $\bar{g}_i$ δίνονται από τους τύπους:

$$\gamma_m = \sum_{r=1}^{N} w_{m,r}(\hat{\mathrm{s}}^n) g_r(\hat{\mathbf{z}}^n, \hat{\mathrm{s}}^n) \tag{5.13}$$

και

$$\bar{g}_i = \sum_{j=1}^{N} d_{i,j} g_j(\hat{\mathbf{z}}^n, \hat{\mathrm{s}}^n) \tag{5.14}$$

αντίστοιχα.

Παρατηρούμε ότι στην εξίσωση του βήματος (5.12), εκτός από την συνάρτηση επιρροής ψ, υπεισέρχονται και οι τιμές των παραγώγων γ_m και $\bar{g}_i$ οι οποίες έχουν ήδη υπολογιστεί χρησιμοποιώντας εύρωστο εκτιμητή. Αυτό κάνει τον υπολογισμό της τιμής του βήματος αρκετά πολύπλοκο και ευαίσθητο στην επιλογή της ρυθμιστικής παραμέτρου σ του εκτιμητή. Για το λόγο αυτό, μόνο για τον υπολογισμό του βήματος χρησιμοποιήσαμε τον Truncated Least Square εκτιμητή, ο οποίος κάνει τον υπολογισμό λιγότερο πολύπλοκο σε σχέση με τους άλλους εύρωστους εκτιμητές.

5.4 Περιγραφή της Μεθόδου

Στην ενότητα αυτή, θα περιγράψουμε την εφαρμογή των εύρωστων εκτιμητών στο πρόβλημα της υπερανάλυσης εικόνας.

Σκοπός μας είναι να εκτιμήσουμε την υψηλής ανάλυσης εικόνα από αλλοιωμένα δεδομένα χαμηλής ανάλυσης. Σε κάθε περίπτωση, η πρώτη εκτίμηση της υψηλής ανάλυσης εικόνας βρίσκεται κάνοντας μία παρεμβολή στο πρώτο χαμηλής ανάλυσης πλαίσιο. Οι παράμετροι μετασχηματισμού υπέρθεσης (διάνυσμα μετατόπισης $\mathbf{T}$ και πίνακας περιστροφής $\mathbf{R}$) αρχικοποιούνται χρησιμοποιώντας τους περιγραφείς SIFT (βλ. κεφάλαιο 3).

Να σημειώσουμε ότι όσο αφορά την παράμετρο σ του εκτιμητή, χρησιμοποιούμε διαφορετικές εκτιμήσεις για την εύρεση των παραμέτρων ομαλοποίησης από την εκτίμηση της παραγώγου και του βέλτιστου βήματος. Ο αλγόριθμος 5 συνοψίζει περιληπτικά τα βήματα της διαδικασίας που ακολουθήσαμε.

Αλγόριθμος 5 Αλγόριθμος υπερανάλυσης εικόνας με χρήση εύρωστων εκτιμητών.

βήμα **1:** Αρχικοποίηση της παραμέτρου σ και της υψηλής ανάλυσης εικόνας $\hat{\mathbf{z}}^0$ από παρεμβολή με το πρώτο χαμηλής αναλύσης πλαίσιο.

βήμα **2:** Εύρεση παραμέτρων μετασχηματισμού υπέρθεσης ($\mathbf{T}$ και $\mathbf{R}$) με χρήση περιγραφέων SIFT.

while Δεν έχουμε φτάσει τον μέγιστο αριθμό επαναλήψεων.

βήμα **3:** Για $k = 1, 2, \ldots, p$ υπολογίζουμε τις παραμέτρους εξομάλυνσης $\alpha_k(\mathbf{z})$ από την εξίσωση (5.11)

βήμα **4:** Εκτίμηση της παραμέτρου $\sigma_{\alpha,k}$ του εκτιμητή, που υπεισέρχεται στην εξίσωση των συντελεστών εξομάλυνσης, για κάθε χαμηλής αναλύσης πλαίσιο με την εξίσωση (5.7).

βήμα **5:** Εκτιμούμε την παράγωγο από την εξίσωση (5.9).

βήμα **6:** Υπολογίζουμε το βέλτιστο βήμα χρησιμοποιώντας την εξίσωση (5.12).

βήμα **7:** Ενημερώνουμε την υψηλής ανάλυσης εικόνα $\hat{\mathbf{z}}^{n+1}$ από την εξίσωση (5.10).

βήμα **8:** Εκτιμάμε την παράμετρο σ_k για τον εύρωστο εκτιμητή της παραγώγου και του βήματος, για κάθε χαμηλής αναλύσης πλαίσιο.

5.5 Πειραματικά Αποτελέσματα

Στην ενότητα αυτή παρουσιάζουμε τα πειράματα και τα αριθμητικά αποτελέσματα που υπολογίσαμε εφαρμόζοντας την προτεινόμενη μέθοδο για υπερανάλυση εικόνας. Να σημειώσουμε ότι τα πειράματα έχουν εκτελεστεί για δύο ειδών διαφορετικούς εκτιμητές, τον Truncated Least Squares εκτιμητή και τον Geman-McClure εκτιμητή. Ο υπολογισμός του βέλτιστου βήματος του αλγορίθμου, σε κάθε περίπτωση υπολογίζεται χρησιμοποιώντας τον Truncated Least Squares εκτιμητή. Σε όλα τα πειράματα, η αρχική εκτίμηση για την υψηλής ανάλυσης εικόνα βρέθηκε κάνοντας παρεμβολή με το πρώτο χαμηλής ανάλυσης πλαίσιο.

Τα πειράματα έχουν υλοποιηθεί σε MATLAB. Το σύνολο δεδομένων που έχουμε χρησιμοποιήσει έχει παραχθεί τεχτητά. Για να προσομοιώσουμε την επίδραση της PSF της καμεράς, οι χαμηλής ανάλυσης εικόνες συνελίσσονται με ένα συμμετρικό Gaussian χαμηλοπερατό φίλτρο τυπικής απόκλισης 1 και μέγεθος παραθύρου 5×5. Στη συνέχεια, οι εικόνες υποδειγματοληπτούνται κατά ένα παράγοντα $L_1 = L_2 = 2$. Τέλος, σε κάθε χαμηλής ανάλυσης πλαίσιο προσθέτουμε Gaussian λευκό θόρυβο, που κυμαίνεται από 15 dB έως 30 dB. Επιπλέον, για να προσομοιώσουμε την παρουσία μετρήσεων που δεν ακολουθούν το κύριο μοντέλο υπολογισμού, προσθέσαμε διαφορετικά μοντέλα θορύβου, όπως salt & pepper θόρυβο με πυκνότητες 0.01, 0.05 και 0.1 και speckle θόρυβο με διακυμάνσεις 0.01, 0.02 και 0.035.

Ο speckle θόρυβος, είναι πολλαπλασιαστικός θόρυβος που υποβαθμίζει την ποιότητα μίας εικόνας. Έστω I είναι η εικόνα στην οποία προσθέτουμε τον speckle θόρυβο τότε:

$$J = I + n * I$$

όπου n είναι ομοιόμορφα κατανεμημένος τυχαίος θόρυβος με μέση τιμή 0 και διακύμανση σ^2 και J είναι η εικόνα που προκύπτει έχοντας προσθέσει speckle θόρυβο.

Για την ποσοτικοποίηση της ποιότητας των υποβαθμισμένων εικόνων, τα επίπεδα του θορύβου στις υποβαθμισμένες εικόνες χαμηλής ανάλυσης και την ποιότητα των ανακατασκευασμένων αποτελεσμάτων χρησιμοποιούμε τη μέγιστη τιμή του λόγου σήματος προς θόρυβο ($PSNR$).

$$PSNR = 10 \log_{10} \frac{(255)^2}{\|\mathbf{f} - \mathbf{g}\|^2}$$

όπου $\mathbf{f}$ και $\mathbf{g}$ είναι η πραγματική και η ανακατασκευασμένη εικόνα, αντιστοίχως.

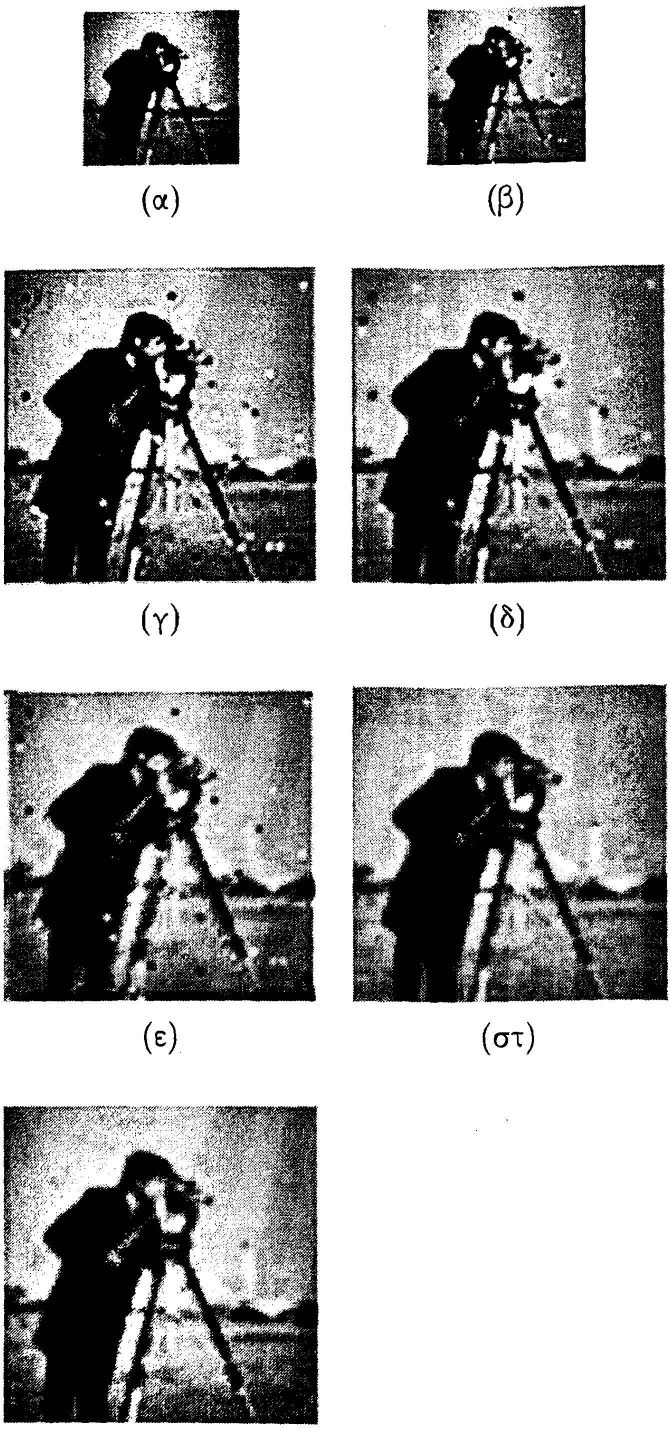
Η δομή με την οποία θα παρουσιάσουμε τα πειραματικά αποτελέσματα είναι η ακόλουθη: για κάθε σύνολο δεδομένων δίνεται ένα υποσύνολο των χαμηλής ανάλυσης εικόνων και η προκύπτουσα υψηλής ανάλυσης εικόνα για τους δύο τύπους των έυρωστων εκτιμητών. Επίσης, παρουσιάζονται οι αντίστοιχοι πίνακες που περιέχουν αριθμητικά αποτελέσματα (μέση τιμή και τυπική απόκλιση) για το PSNR των ανακατασκευασμένων εικόνων υψηλής ανάλυσης.

Στο σχήμα 5.6 συγκρίνουμε τα αποτελέσματα που προκύπτουν από την διαδικασία της υπερανάλυσης εικόνας με χρήση του αλγορίθμου που προτείνουμε στο κεφάλαιο αυτό, σε σχέση με τους αλγορίθμους των κεφαλαίων 2, 3 και 4. Στο πειραμα που παρουσιάζεται στο σχήμα 5.6, χρησιμοποιήσαμε ένα σύνολο από τέσσερις τεχνητά παραχθείσες εικόνες. Οι εικόνες αυτές έχουν μετατοπιστεί στο διάστημα $[-3, 3]$. Κάθε χαμηλής ανάλυσης πλαίσιο έχει λόγο σήματος πρός θόρυβο (SNR) 30 dB. Επιπλέον, στην τέταρτη εικόνα χαμηλής ανάλυσης έχουμε προσθέσει θόρυβο αλατοπίπερου.

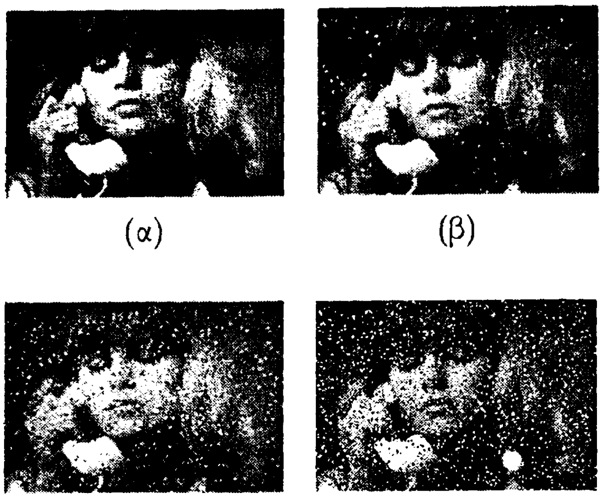
Η εικόνα 5.6(γ) απεικονίζει την ανακατασκευασμένη εικόνα υψηλής ανάλυσης όπως αυτή παράγεται από τον αλγόριθμο 1 το PSNR που προκύπτει είναι 20.03 dB. Η εικόνα 5.6(δ) δειχνει την υψηλής ανάλυσης εικόνα που παράγεται από τον αλγόριμθο 3 χρησιμοποιώντας τους περιγραφείς SIFT. Η εικόνα που προκύτει έχει PSNR 20.54 dB. Η εικόνα 5.6(ε) παρουσιάζει την υψηλής ανάλυσης εικόνα που παράγεται από τον αλγόριθμο 4. Το PSNR για την εικόνα αυτή είναι 16.70 dB. Οι εικόνες 5.6(στ) και 5.6(ζ) είναι οι υψηλής ανάλυσης εικόνες που προκύπτουν από τον προτεινόμενο αλγόριθμο με χρήση του Truncated Least Squares και του Geman-McClure εκτιμητή, αντίστοιχα. Τα αντίστοιχα PSNR που προκύπτουν είναι 21.08 dB για την εικόνα 5.6(στ) και 21.87 dB για την εικόνα 5.6(ζ).

Παρατηρούμε ότι οι κλασικές μέθοδοι υπερανάλυσης δεν παρουσιάζουν καθόλου ικανοποιητικά αποτελέσματα όταν τα δεδομένα εισόδου περιέχουν σφάλματα. Συνεχίζουν να μεταφέρουν τα σφάλματα σε κάθε επανάληψη και το προκύπτον αποτέλεσμα είναι μια κακής ποιότητας εκτίμηση της υψηλής ανάλυσης εικόνας. Αντιθέτως, οι υψηλής ανάλυσης εικόνες που προκύπτουν από τον αλγόριθμο 5 εξαλείφουν πλήρως τα σφάλματα αυτά.

Στο σχήμα 5.7 παρουσιάζονται τέσσερα ενδεικτικά χαμηλής ανάλυσης πλαίσια για την εικονοσειρά *Susie*. Το σύνολο των πλαισίων που χρησιμοποιήσαμε στα πειράματά μας ήταν 20. Σε κάθε πλαίσιο έχει προστεθεί Gaussian λευκός θόρυβος της τάξης των 30 dB. Επιπλέον, σε τυχαία πλαίσια που αντιστοιχούν στο 50% των πλαισίων που χρησιμοποιούμε, έχουμε προσθέσει θόρυβο αλατοπίπερου. Τα πειράματα έχουν εκτελεστεί για 0.01, 0.05 και 0.10 πυκνότητες salt & pepper θορύβου.



Σχήμα 5.6: Ανακατασκευή της εικόνας *Cameraman*. (α)-(β) Χαμηλής ανάλυσης πλαίσια. (γ) Υψηλής ανάλυσης εικόνα με την μέθοδο του κεφαλαίου 2 (PSNR = 20.03), (δ) με χρήση περιγραφέων SIFT (PSNR = 20.54) και (ε) με χρήση αμοιβαίας πληροφορίας (PSNR = 16.70). (στ) Εκτίμηση εικόνας υψηλής ανάλυσης με χρήση Truncated Least Squares εκτιμητή (PSNR = 21.08), (ζ) με χρήση Geman-McClure εκτιμητή (PSNR = 21.87).



(α) (β)

(γ) (δ)

Σχήμα 5.7: *Susie.* (α)-(δ) Ενδεικτικά χαμηλής ανάλυσης πλαίσια. (β) Salt & pepper θόρυβος 1%. (γ) Salt & pepper θόρυβος 5%. (δ) Salt & pepper θόρυβος 10%.

Έχοντας υπ' όψην τα αποτελέσματα για το σχήμα 5.8 και τον πίνακα 5.1, τόσο ο Truncated Least Squares εκτιμητής όσο και ο Geman-McClure εκτιμητής μπορούν επιτυχώς να ανακατασκευάσουν την εικόνα, που έχει αλλοιωθεί από την παρουσία salt & pepper θορύβου.

Πίνακας 5.1: Αριθμητικά αποτελέσματα για την εικόνα *Susie* (σχ. 5.8) με διάφορες πυκνότητες salt & pepper θορύβου για το 50% των πλαισίων με ταυτόχρονη εκτίμηση των παραμέτρων α και ε.

Εκτιμητής	PSNR					
	salt & pepper 1%		**salt & pepper 5%**		**salt & pepper 10%**	
	mean	**std**	**mean**	**std**	**mean**	**std**
Truncated Least Squares	25.82	0.82	25.75	0.43	23.59	0.30
Geman-McClure	24.42	0.47	24.39	0.37	21.45	0.48

Στο σχήμα 5.9 παρουσιάζονται τέσσερα ενδεικτικά χαμηλής ανάλυσης πλαίσια για την εικονοσειρά *Claire.* Το σύνολο των πλαισίων που χρησιμοποιήσαμε στα πειράματά μας ήταν 20. Σε κάθε πλαίσιο έχει προστεθεί Gaussian λευκός θόρυβος της τάξης των 30 dB. Επιπλέον, σε τυχαία πλαίσια που αντιστοιχούν στο 50% των πλαισίων που χρησιμοποιούμε, έχουμε προσθέσει θόρυβο Speckle. Τα πειράματα έχουν εκτελεστεί για 0.01, 0.02 και 0.035 διακυμάνσεις του θορύβου Speckle σε τυχαία πλαίσια της εικονοσειράς.

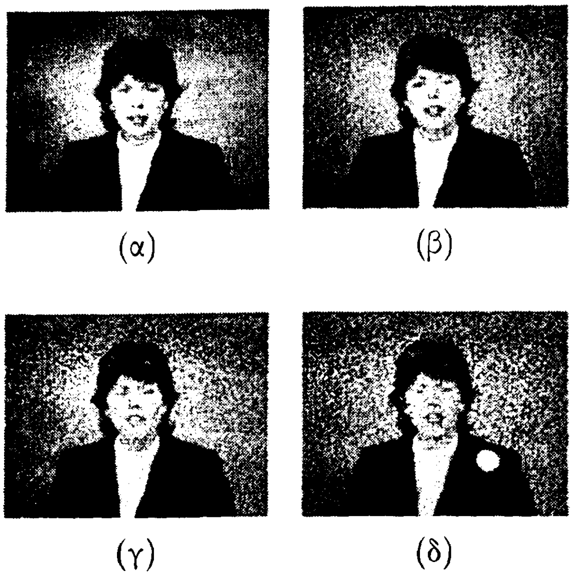
Παρατηρώντας τα αποτελέσματα για το σχήμα 5.10 και τον πίνακα 5.2, ο Geman-McClure εκτιμητής μπορεί επιτυχώς να ανακατασκευάσει την εικόνα, που έχει αλλοιωθεί από την παρουσία του θορύβου Speckle, πολύ καλύτερα απ' ότι ο Truncated Least Squares

(θόρυβος 1%)

(θόρυβος 1%)

(θόρυβος 5%)

(θόρυβος 5%)



Σχήμα 5.9: *Claire*. (α)-(δ) Ενδεικτικά χαμηλής ανάλυσης πλαίσια. (β) Speckle θόρυβος 1%. (γ) Speckle θόρυβος 2%. (δ) Speckle θόρυβος 3.5%.

εκτιμητής, επειδή ο Geman-McClure εκτιμητής είναι πιο εύρωστος για εκτιμήσεις σε μεγάλες αλλοιώσεις της εικόνας.

Πίνακας 5.2: Αριθμητικά αποτελέσματα για την εικόνα *Claire* (σχ. 5.10) με διάφορες διακυμάνσεις θορύβου Speckle για το 50% των πλαισίων με ταυτόχρονη εκτίμηση των παραμέτρων α και ε.

Εκτιμητής	PSNR					
	Speckle 0.01		**Speckle 0.02**		**Speckle 0.035**	
	mean	**std**	**mean**	**std**	**mean**	**std**
Truncated Least Squares	29.50	0.26	29.20	0.32	29.40	0.53
Geman-McClure	31.35	0.14	31.19	0.10	30.58	0.45

Το σχήμα 5.11 απεικονίζει 4 ενδεικτικά χαμηλής ανάλυσης πλαίσια, για δύο διαφορετικές τιμές Gaussian λευκού θορύβου, 30 dB, 20 dB και 15 dB. Η ακολουθία των χαμηλής ανάλυσης πλαισίων αποτελείται από 7 διαφορετικές εικόνες, οι οποίες έχουν παραχθεί ως εξής: πέντε από τα χαμηλής ανάλυσης πλαίσια έχουν μετατοπιστεί κατάλληλα στο διάστημα $[-7, 7]$, ενώ τα υπόλοιπα δύο περιέχουν μόνο περιστροφή στο διαστήμα $[-6, 6]$.

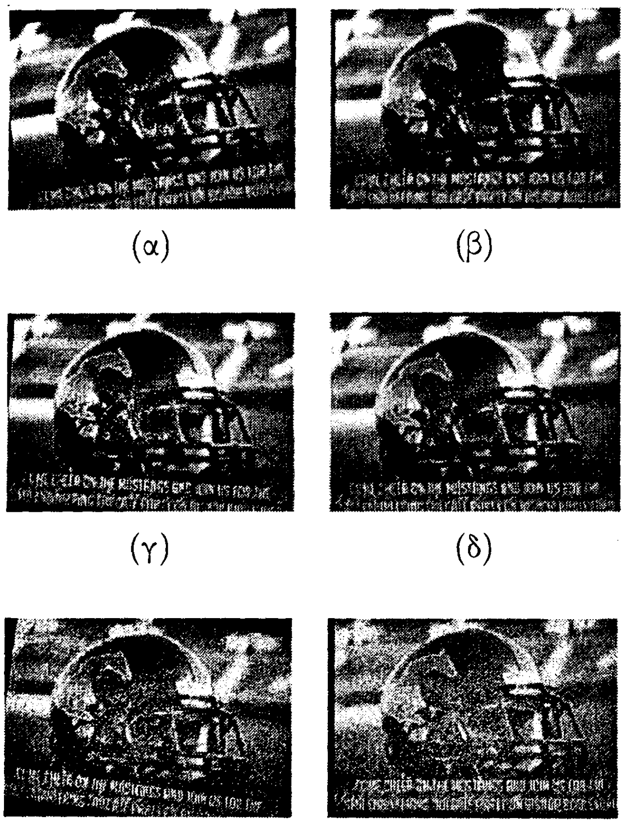
Η συμπεριφορά του Geman-McClure εκτιμητή έναντι του Truncated Least Squares εκτιμητή φαίνεται οπτικά στο σχήμα 5.12. Ο πίνακας 5.3 παρουσιάζει αριθμητικά αποτελέσματα για τα PSNR για τους δύο εκτιμητές. Ο Geman-McClure εκτιμητής φαίνεται να κάνει μία καλύτερη εκτίμηση της υψηλής ανάλυσης εικόνας ακόμα και για μεγαλύτερες τιμές θορύβου. Αυτό οφείλεται στο γεγονός ότι ο Truncated Least Squares εκτιμητής είναι αρκετά ευάλωτος στην παρουσία σφαλμάτων, καθώς η συνάρτηση επιρροής του είναι γραμ-


(θόρυβος 1%)

(θόρυβος 1%)


(θόρυβος 2%)

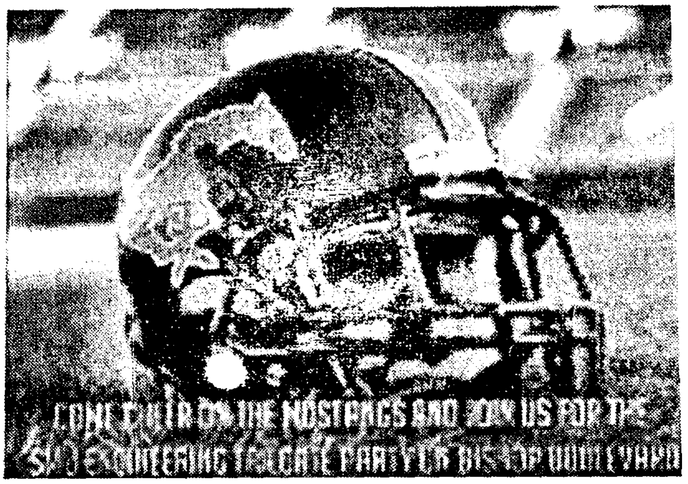
(θόρυβος 2%)

(α) (β) (γ) (δ) (ε) (στ)

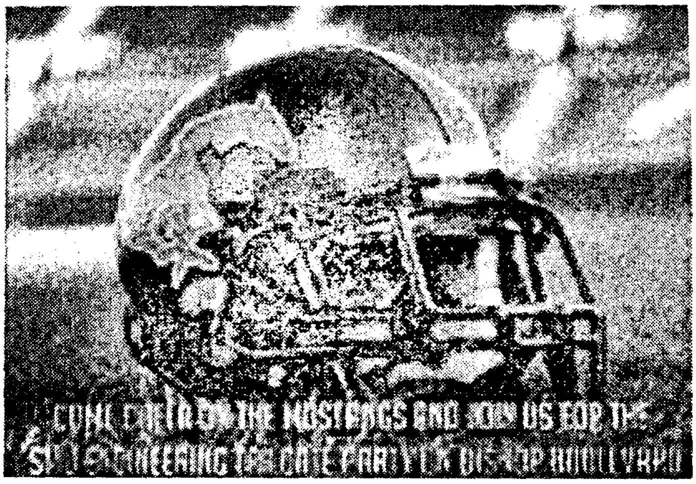
Σχήμα 5.11: *Helmet.* Ενδεικτικά χαμηλής ανάλυσης πλαίσια. (α), (β) Έχουν παραχθεί με Gaussian λευκό θόρυβο 30 dB, (γ), (δ) με 20 dB και (ε), (στ) με 15 dB.

(θόρυβος 30 dB)

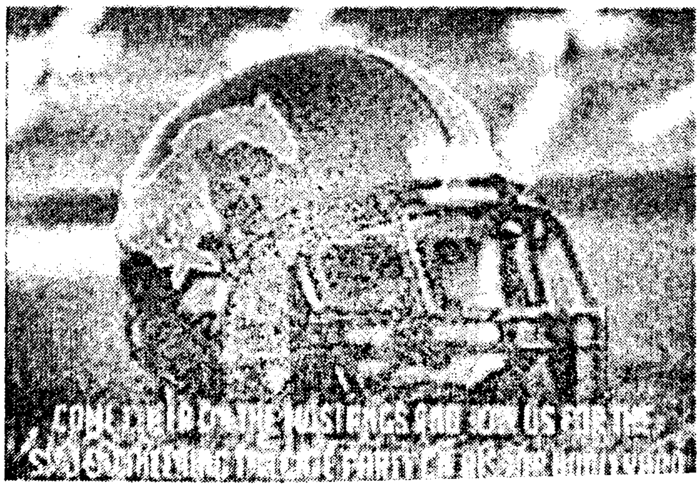


(θόρυβος 30 dB)

(θόρυβος 20 dB)



(θόρυβος 20 dB)



μική και αναθέτει μεγαλύτερα βάρη σε μεγάλα σφάλματα, έτσι ενισχύει την επιρροή τους στην εκτίμηση της υψηλής ανάλυσης εικόνας.

Πίνακας 5.3: Αριθμητικά αποτελέσματα για την εικόνα *Helmet* (σχ. 5.12) για διάφορα μεγέθη Gaussian θορύβου με ταυτόχρονη εκτίμηση των παραμέτρων α και ε.

Εκτιμητής	**PSNR**					
	Gaussian 30 dB		**Gaussian 20 dB**		**Gaussian 15 dB**	
	mean	**std**	**mean**	**std**	**mean**	**std**
Truncated Least Squares	22.50	0.18	21.43	0.36	18.11	0.42
Geman-McClure	22.94	0.31	21.98	0.22	18.39	0.10

'Ενα επιπλέον θέμα που προκύπτει στο σημείο αυτό είναι η αξιολόγηση της προτεινόμενης μεθόθου για υπερανάλυση εικόνας. Η διαδικασία αυτή, βασίζεται στη σύγκριση με την μέθοδο η οποία βασίζεται σε μη εύρωστο υπολογισμό του βήματος και της παραμέτρου εξομάλυνσης. Για το λόγο αυτό, έχουμε εκτελέσει μια σειρά από πειράματα με διαφορετικούς εύρωστους εκτιμητές, συγκρίνοντας την μεθοδο που προτείνουμε (robust ε, α) με εκείνη που δεν χρησιμοποιεί εύρωστη εκτίμηση των παραμέτρων α και ε (no-robust ε, α).

Τα πειράματα έχουν διεξαχθεί για δύο διαφορετικά σύνολα δεδομένων, εικόνες *Susie* και *Claire*. Το σύνολο δεδομένων αποτελείται από 20 χαμηλής ανάλυσης πλαίσια, στο 50% των οποίων έχει προσθεθεί salt & pepper θόρυβος. Τα πειράματα έχουν εκτελεστεί για πυκνότητες θορύβου 5% και 10%.

Πίνακας 5.4: Συγκριτικά αριθμητικά αποτελέσματα για το PSNR για την εικόνα *Susie* με salt & pepper θόρυβο πυκνότητας 5% για διάφορες μεθόδους εύρωστης υπερανάλυσης και ταυτόχρονη εκτίμηση των παραμέτρων α και ε.

Εκτιμητής	**Μέθοδος**			
	no-robust ε, α		**robust** ε, α	
	mean	**std**	**mean**	**std**
Truncated Least Squares	18.10	0.57	25.75	0.43
Geman-McClure	22.10	0.55	24.39	0.37
Lorentzian	20.78	0.26	25.14	0.68

Οι πίνακες 5.4, 5.5, 5.6 και 5.7 παρουσιάζουν τα αριθμητικά αποτελέσματα για το PSNR των ανακατασκευασμένων εικόνων υψηλής ανάλυσης για τις διάφορες μεθόδους υπερανάλυσης εικόνας. 'Οπως μπορούμε εύκολα να αντιληφθούμε, η προτεινόμενη μέθοδος (robust ε, α) παρουσιάζει σαφώς καλύτερα αποτελέσματα, καθώς το PSNR των ανακατασκευασμένων εικόνων υψηλής ανάλυσης για την μέθοδο αυτή ειναι πολύ υψηλότερο από τις υπόλοιπες. Από τα αριθμητικά αποτελέσματα παρατηρούμε ότι η μέθοδος no-robust ε, α παρουσιάζει χαμηλότερα αποτελέσματα σε σχέση με την προτεινόμενη μέθοδο.

Πίνακας 5.5: Συγκριτικά αριθμητικά αποτελέσματα για το PSNR για την εικόνα *Susie* με salt & pepper θόρυβο πυκνότητας 10% για διάφορες μεθόδους εύρωστης υπερανάλυσης και ταυτόχρονη εκτίμηση των παραμέτρων α και ε.

Εκτιμητής	Μέθοδος			
	no-robust ε, α		**robust** ε, α	
	mean	**std**	**mean**	**std**
Truncated Least Squares	15.28	0.92	23.59	0.30
Geman-McClure	19.52	0.38	21.45	0.48
Lorentzian	19.03	0.29	22.93	0.55

Πίνακας 5.6: Συγκριτικά αριθμητικά αποτελέσματα για το PSNR για την εικόνα *Claire* με salt & pepper θόρυβο πυκνότητας 5% για διάφορες μεθόδους εύρωστης υπερανάλυσης και ταυτόχρονη εκτίμηση των παραμέτρων α και ε.

Εκτιμητής	Μέθοδος			
	no-robust ε, α		**robust** ε, α	
	mean	**std**	**mean**	**std**
Truncated Least Squares	13.44	0.07	29.67	0.31
Geman-McClure	28.05	0.89	31.14	0.34
Lorentzian	19.23	0.23	31.34	0.15

Πίνακας 5.7: Συγκριτικά αριθμητικά αποτελέσματα για το PSNR για την εικόνα *Claire* με salt & pepper θόρυβο πυκνότητας 10% για διάφορες μεθόδους εύρωστης υπερανάλυσης και ταυτόχρονη εκτίμηση των παραμέτρων α και ε.

Εκτιμητής	Μέθοδος			
	no-robust ε, α		**robust** ε, α	
	mean	**std**	**mean**	**std**
Truncated Least Squares	11.21	0.73	27.82	0.17
Geman-McClure	22.02	0.53	26.94	0.45
Lorentzian	18.79	0.32	25.77	0.46

Η συνεισφορά μας στην εύρωστη υπερανάλυση εικόνας μπορεί εύκολα να γίνει αντιληπτή συγρίνοντας τις δύο μεθόδους no-robust ϵ, α και robust ϵ, α. Η μέθοδος που προτείνουμε βελτιώνει κατα πολύ το αποτέλεσμα και παράγει καλύτερης ποιότητας εικόνες υψηλής ανάλυσης.

Ο Truncated Least Squares εκτιμητής παρουσιάζει τα χαμηλότερα αποτελέσματα σε σχέση με τους υπόλοιπους δύο εκτιμητές. Από την άλλη, ο Lorentzian και ο Geman-McClure εκτιμητής φαίνεται να παρέχουν μια πολύ καλή συμπεριφορά βελτιώνοντας αρκετά το αποτέλεσμα. Μελετώντας τα πειραματικά αποτελέσματα για τους τρεις αυτούς εκτιμητές, μπορούμε να συμπεράνουμε την ανάγκη για χρήση εύρωστης εκτίμησης των παραμέτρων εξομάλυνσης και του βήματος υπερανάλυσης.

Κεφαλαιο 6

Επιλογος

6.1 Επίλογος

Στην παρούσα εργασία παρουσιάσαμε το πρόβλημα της υπερανάλυσης εικόνας. Η υπερανάλυση εικόνας εστιάζεται στην εύρεση μιας εικόνας υψηλής ανάλυσης από ένα σύνολο εικόνων χαμηλής ανάλυσης. Στο κεφάλαιο 2 ορίσαμε το πρόβλημα της υπερανάλυσης εικόνας και παρουσιάσαμε ένα από τους βασικότερους αλγορίθμους που επιλύουν το συγκεκριμένο πρόβλημα [10]. Ένα από τα σημαντικότερα ζητήματα που έχει μελετηθεί, είναι η εύρεση των άγνωστων παραμέτρων μετασχηματισμού υπερθεσης για τις χαμηλής ανάλυσης εικόνες. Η μοντελοποίηση της άγνωστης υψηλής ανάλυσης εικόνας είναι ένα θέμα που συναντάται σε όλα τα προβλήματα εκτίμησης παραμέτρων. Μέσω της μοντελοποίησης αυτής, όλη η εκ των προτέρων γνώση σχετικά με την αρχική εικόνα ενσωματώνεται στην προσεγγιση της λύσης.

Στο κεφάλαιο 3 κάνουμε λόγο για εξαγωγή χαρακτηριστικών σημείων από μία εικόνα και το πως αυτά σχετίζονται με την διαδικασία της υπερανάλυσης εικόνας. Τα χαρακτηριστικά σημεία αυτά είναι αμετάβλητα στον χώρο κλιμακας της εικόνας και την περιστροφή, και παρουσιάζουν μία ευρωστία απέναντι σε παραμορφωμένα δεδομένα, που περιέχουν θόρυβο ή διαφορές στην φωτεινότητα. Ο υπολογισμός τους είναι αρκετά αποδοτικός, και επομένως εκατοντάδες χαρακτηριστικά σημεία μπορούν να εξαχθούν από μία τυπική εικόνα σε πολύ μικρό χρονικό διάστημα με ένα συνηθισμένο Η/Υ.

Η μέθοδος που προτείνουμε στο κεφάλαιο 4 επιλύει το πρόβλημα της υπερανάλυσης εικόνας βρίσκοντας τις παραμέτρους μετασχηματισμού υπέρθεσης με χρήση της αμοιβαίας πληροφορίας. Η ιδέα είναι να μεγιστοποιήσουμε την αμοιβαία πληροφορία μεταξύ δυό χαμηλής ανάλυσης εικόνων για την εύρεση των μετασχηματισμών υπέρθεσης. Για περισσότερες πληροφορίες σχετικά με την μέθοδο αυτή, παραπέμπουμε τον αναγνώστη στις αναφορές [36] και [20]. Ο αλγόριθμος 4 που προτάθηκε στο κεφάλαιο 4 δίνει καλύτερα αποτελέσματα από τον αλγόριθμο 3 του κεφαλαίου 3. Συγκρίνοντας τους δύο αυτούς αλγορίθμους, το κέρδος που έχουμε για το PSNR στην ανακατασκευασμένη εικόνα υψηλής ανάλυσης φτάνει μέχρι τα 2 dB.

Ένα σημαντικό θέμα για συζήτηση και μελλονική ενασχόληση είναι η εκτέλεση της ελαχιστοποίησης. Για την ελαχιστοποίηση στον αλγόριθμο 4 χρησιμοποιήσαμε την μέθοδο SIMPLEX. Ο αλγόριθμος αυτός είναι ένα αλγόριθμος τοπικής ελαχιστοποίησης και επομένως έχει μεγάλη εξάρτηση από την αρχικοποίηση. Μία λύση θα ήταν να χρησιμοποιήσουμε έναν αλγόριθμο καθολικής ελαχιστοποίησης.

Στο κεφάλαιο 5 παρουσιάσαμε μια εναλλακτική προσέγγιση της μεθόδου που εξετάζουμε στην διατριβή αυτή. Θεωρήσαμε πως το σύνολο δεδομένων μας περιέχει "αλλοιωμένα" δεδομένα, δηλαδή δεδομένα που δεν ακολουθούν το μοντέλο υπολογισμού που προτείνουμε, και για το λόγο αυτό χρησιμοποιήσαμε εύρωστους εκτιμητές. Οι εκτιμητές που χρησιμοποιήσαμε στην εργασία αυτή είναι ο Truncated Least Squares και ο Geman-McClure εκτιμητής. Από τα πειραματικά αποτελέσματα, συμπεράναμε ότι ο αλγόριθμος που προτείνουμε μπορεί να εφαρμοστεί σε πολλά είδη θορύβου όπως λευκό προσθετικό Gaussian θόρυβο, θόρυβο αλατοπίπερου (salt & pepper), θόρυβο Speckle κ.α. Ο προτεινόμενος αλγόριθμος οχι μόνο βελτιώνει το αποτέλεσμα, εξαλείφοντας την επίδραση που έχουν τα ανεπιθύμητα δεδομένα πάνω στο αποτέλεσμα, αλλά δίνει και καλύτερα αποτελέσματα σε σχέση με τους υπόλοιπους αλγορίθμους που μελετήσαμε.

Όπως έχει γίνει κατανοητό μέχρι τώρα, στην παρούσα εργασία επικεντρωθήκαμε στην εξήγηση και ανάλυση των μαθηματικών εργαλείων που απαιτούνται για την επίλυση του προβλήματος της υπερανάλυσης εικόνας. Για τους περισσότερους αλγορίθμους απαιτείται να καθοριστεί ένας αριθμός από παραμέτρους, το οποίο είναι από μόνο του ένα σημαντικό πρόβλημα. Θεωρούμε ότι η κατανόηση των μαθηματικών τύπων και των αλγορίθμων που περιγράψαμε, μπορούν να δώσουν λύση σε πολλά πρόβλήματα υπερανάλυσης εικόνας.

ΒΙΒΛΙΟΓΡΑΦΙΑ

[1] K. V. Arya, P. Gupta, P. K. Kalra, and P. Mitra. *Image Registration Using Robust M-estimators.* Pattern Recognition Letters.

[2] C. Bishop. *Pattern Recognition and Machine Learning.* Springer, 2006.

[3] M. J. Black. *Robust Incremental Optical Flow.* PhD thesis, Yale University, December 1992.

[4] D. Capel and A. Zisserman. Computer vision applied to super-resolution. *IEEE Signal Processing Magazine*, 20(3):75–86, May 2003.

[5] G. K. Chantas, N. P. Galatsanos, and N. A. Woods. Super-resolution based on fast registration and maximum a posteriori reconstruction. *IEEE Transactions on Image Processing*, 16(7):1821–1830, July 2007.

[6] Y. Chen, H. Wang, T. Fang, and J. Tyan. Mutual information regularized Bayesian framework for multiple image restoration. In *IEEE International Conference on Computer Vision (ICCV'05)*, volume 1, 2005.

[7] N. A. El-Yamany and P. E. Papamichalis. An adaptive M-estimation framework for robust image superresolution without regularization. *in Visual Communications and Image Processing*, 6822:1–12, January 2008.

[8] N. A. El-Yamany and P. E. Papamichalis. Robust color image superresolution: An adaptive M-estimation framework. *EURASIP Journal on Image and Video Processing*, 2008. ID 763254.

[9] D. Forsyth and J. Ponce. *Computer Vision. A Modern Approach.* Prentice Hall, 2003.

[10] R. C. Hardie, K. J. Barnard, and E. E. Armstrong. Joint MAP image registration and high-resolution image estimation using a sequence of undersampled images. *IEEE Transactions on Image Processing*, 6(12):1621–1633, December 1997.

[11] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision.* Cambridge University Press, 2000.

[12] H. He and L. P. Kondi. Resolution enhancement of video sequences with simultaneous estimation of the regularization parameter. *SPIE Journal of Electronic Imaging*, 13(3):586–596, July 2004.

[13] H. He and L. P. Kondi. An image super-resolution algorithm for different error levels per frame. *IEEE Transactions on Image Processing*, 15(3):592–603, March 2006.

[14] M. G. Kang and A. K. Katsaggelos. Simultaneous multichannel image restoration and estimation of the regularization parameters. *IEEE Transactions on Image Processing*, 6(5):774–778, May 1997.

[15] A. Katsaggelos, R. Molina, and J. Mateos. *Super Resolution of Images and Video.* Morgan & Claypool Publishers, first edition, 2007.

[16] A. K. Katsaggelos. Iterative image restoration algorithms. *Optical Engineering*, 28(7):735–748, July 1989.

[17] J. C. Lagarias, J. A. Reeds, M. H. Wright, and P. E. Wright. Convergence properties of the nelder-mead simplex method in low dimensions. *SIAM Journal of Optimization*, 9(1):112–147, 1998.

[18] D. G. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2):91–110, 2004.

[19] F. Maes, A. Collignon, D. Delaere, D. Vandermeulen, and P. Suetens. Automated multimodality medical image registration using information theory. *in Proc. 14th International Conference on Information Processing in Medical Imaging IPMI'95*, 3:263–274, June 1995.

[20] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens. Multimodality image registration by maximization of mutual information. *IEEE Transactions on Medical Imaging*, 16(2):187–198, April 1997.

[21] F. Maes, D. Vandermeulen, and P. Suetens. Medical image registration using mutual information. *Proceedings of the IEEE*, 91(10):1699–1722, October 2003.

[22] P. Majorin. *A Survey of Super-Resolution Methods Used in Image Reconstruction*, February 2008.

[23] C. Nikou, F. Heitz, and J. P. Armspach. Robust voxel similarity metrics for the registration of dissimilar single and multimodal images. *Pattern Recognition*, 32(8):1351–1368, 1999.

[24] S. C. Park, M. K. Park, and M. G. Kang. Super-resolution image reconstruction: A technical overview. *IEEE Signal Processing Magazine*, pages 21–36, May 2003.

[25] V. Patanavijit and S. Jitapunkul. A Lorentzian stochastic estimation for a robust iterative multiframe super-resolution reconstruction with Lorentzian-Tikhonov regularization. *EURASIP Journal on Advances in Signal Processing*, 2007(2):21–21, 2007.

[26] V. Patanavijit, S. Tae-O-Sot, and S. Jitapunkul. A robust iterative superresolution reconstruction of image sequences using a Lorentzian Bayesian approach with fast affine block-based registration. *in Proceedings os IEEE International Conference on Image Processing (ICIP '07)*, 5:393–396, 2007.

[27] L. C. Pickup, D. P. Capel, S. J. Roberts, and A. Zisserman. Bayesian image super-resolution, continued. Technical report, Information Engineering Building, Department of Engineering Science, 2003.

[28] L. C. Pickup, D. P. Capel, S. J. Roberts, and A. Zisserman. Overcoming registration uncertainty in image super-resolution: maximize or marginalize? *EURASIP Journal on Advances in Signal Processing*, 2007.

[29] J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever. Mutual information based registration of medical images: a survey. *IEEE Transactions on Medical Imaging*, 22:986–1004, 2003.

[30] J. P. W. Pluim, J. B. A. M.maintz, and M. A. Viergever. Interpolation artefacts in mutual information-based image registration. *Computer Vision and Image Understanding*, 77:211–232, 2000.

[31] P. J. Rousseeuw and A. M. Leory. *Robust Regression and Outlier Detection.* John Wiley & Sons, 1987.

[32] C. Studholme, D. L. G. Hill, and D. J. Hawkes. An overlap invariant entropy measure of 3D medical image alignment. *Pattern Recognition*, 32:71–86, 1998.

[33] M. E. Tipping and C. M. Bishop. Bayesian image super-resolution. In M. Press, editor, *Advances in Neural Information Processing Systems 15*, 2003.

[34] S. Umeyama. Least-squares estimation of transformation parameters between two point patterns. *IEEE Transactions on Pattern Analysis and Intelligence*, 13(4):376–380, 1991.

[35] J. Vajda. Theory of statistical inference and information. *Dordrecht, The Netherlands:Kluwer*, 1989.

[36] P. Viola and W. W. III. Alignment by maximization of mutual information. *International Journal of Computer Vision*, 24(2):137–154, 1997.

[37] I. Vujovic and I. Kuzmanic. Wavelet quasi-superresolution in marine applications. *Multimidia Signal Processing and Cmmunications, 48th International Symposium ELMAR-2006*, pages 65–68, June 2006.

[38] B. Zhang, J. Liu, J. Chu, and J. Qiao. A mutual information based sub-pixel registration method for image super resolution. In *Fifth International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, pages 422–425, 2009.

[39] A. Zomet, A. Rav-Acha, and S. Peleg. Robust super-resolution. *in Proc. of the IEEE Workshop on Applications of Computer Vision*, 1:645–650, 2001.

ΒΙΟΓΡΑΦΙΚΟ

Ο Μιχαήλ Βρίγκας γεννήθηκε στα Ιωάννινα το 1985. Αποφοίτησε το 2003 από το 3ο Ενιαίο Λύκειο Ιωαννίνων. Οι βασικές σπουδές πραγματοποιήθηκαν στο τμήμα Πληροφορικής του Πανεπιστημίου Ιωαννίνων, από όπου και αποφοίτησε το 2008. Τον Σεπτέμβρη της ίδιας χρονιάς έγινε δεκτός στο Μεταπτυχιακό Πρόγραμμα σπουδών της ίδιας σχολής. Στα ερευνητικά του ενδιαφέροντα συγκαταλέγεται η Υπολογιστική Όραση και η Ανάλυση και Επεξεργασία Εικόνων.