ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΑΤΡΩΝ - ΠΟΛΥΤΕΧΝΙΚΗ ΣΧΟΛΗ
Τμήμα Μηχανικών Ηλεκτρονικών Υπολογιστών και Πληροφορικής

Διδακτορική Διατριβή

ΑΠΟΔΟΤΙΚΕΣ ΤΕΧΝΙΚΕΣ ΑΝΤΙΣΤΟΙΧΙΣΗΣ ΚΑΙ ΨΗΦΙΑΚΗΣ
ΥΔΑΤΟΓΡΑΦΗΣΗΣ ΕΙΚΟΝΩΝ

Ειρήνη Γ. Καρύμπαλη

Πάτρα, Οκτώβριος 2005

Στον πατέρα μου, που έφυγε νωρίς...

...και στους ανθρώπους που είναι πάντα δίπλα μου και με στηρίζουν.

Στη μητέρα μου, στον αδερφό μου και στον αρραβωνιαστικό μου.

Στην Αμαλία, στον Γιάννη και στον Μανόλη.

Περιεχόμενα

Περιεχόμενα	i
Ευχαριστίες	v
Περίληψη	vii
1 Εισαγωγή	1
1.1 Αντιστοίχιση Εικόνων	1
1.1.1 Αντιστοίχιση με βάση τα χαρακτηριστικά της εικόνας	2
1.1.2 Αντιστοίχιση με βάση ολόκληρες περιοχές της εικόνας	4
1.1.3 Υπολογισμός παραμέτρων μετασχηματισμού	7
1.1.4 Επαναδειγματοληψία και μετασχηματισμός της εικόνας	10
1.1.5 Ερευνητικά αποτελέσματα στα πλαίσια της διδακτορικής διατριβής . . .	11
1.2 Η σχέση της αντιστοίχισης εικόνων με την ψηφιακή υδατογράφηση εικόνων . .	12
1.3 Ψηφιακή υδατογράφηση εικόνων	14
1.3.1 Μοντέλα υδατογράφησης και επικοινωνίες	17
1.3.2 Υδατογράφηση βασισμένη στη συσχέτιση	19
1.3.3 Υδατογράφηση με επιπρόσθετη πληροφορία	22
1.3.4 Χρήση αντιληπτικών μοντέλων	25
1.3.5 Ερευνητικά αποτελέσματα στα πλαίσια της διδακτορικής διατριβής . . .	27
2 Μια νέα τεχνική αντιστοίχισης εικόνων βασισμένη στη συσχέτιση	29
2.1 Εισαγωγή	30
2.2 Η προτεινόμενη τεχνική	31
2.2.1 Προλεύκανση της εικόνας	31
2.2.2 Η νέα τεχνική αντιστοίχισης	33
2.3 Αποδοτικές υλοποιήσεις	36
2.4 Πειραματικά αποτελέσματα	37
2.5 Εφαρμογή της ICC-PW σε ακολουθίες ιατρικών εικόνων	38
2.5.1 Πειραματικά αποτελέσματα	40

2.6	Συμπεράσματα, μελλοντικές ερευνητικές κατευθύνσεις	43
3	Αποδοτικοί αλγόριθμοι για τον υπολογισμό μετρικών συσχέτισης	45
3.1	Αποδοτικός υπολογισμός της ετεροσυσχέτισης	45
3.2	Αποδοτικό επαναληπτικό σχήμα για αντιστοίχιση με ακρίβεια εικονοστοιχείου .	48
3.2.1	Περιγραφή του αλγορίθμου	49
3.2.2	Πειραματική μελέτη της σύγκλισης του αλγορίθμου	52
3.3	Συμπεράσματα, μελλοντικές ερευνητικές κατευθύνσεις	57
4	Αντιστοίχιση εικόνων με ακρίβεια υποδιαίρεσης εικονοστοιχείου	59
4.1	Εισαγωγή	59
4.2	Το νέο σχήμα αντιστοίχισης	61
4.2.1	Διατύπωση του προβλήματος	61
4.2.2	Η προτεινόμενη μετρική για ακρίβεια υποδιαίρεσης εικονοστοιχείου . .	62
4.2.3	Λύση κλειστού τύπου	64
4.3	Πειραματικά αποτελέσματα	64
4.4	Συμπεράσματα, μελλοντικές ερευνητικές κατευθύνσεις	68
5	Προσαρμοστική υδατογράφηση με “τυφλή” ανίχνευση	69
5.1	Εισαγωγή	70
5.2	Διατύπωση του προβλήματος	71
5.3	Αντιληπτική μάσκα βασισμένη στο σφάλμα πρόβλεψης	72
5.4	Νέο σχήμα “τυφλής” ανίχνευσης (Ιδανική περίπτωση)	73
5.4.1	Ανίχνευση όταν δεν υπάρχει επίθεση	74
5.4.2	Ανίχνευση μετά από πρόσθεση θορύβου	75
5.4.3	Ανίχνευση μετά από γραμμικό φιλτράρισμα	76
5.5	Νέο σχήμα “τυφλής” ανίχνευσης (Περίπτωση πραγματικών δεδομένων)	77
5.5.1	Ντετερμινιστική προσέγγιση	78
5.5.2	Στοχαστική προσέγγιση	79
5.6	Πειραματικά αποτελέσματα	80
5.7	Συμπεράσματα, μελλοντικές ερευνητικές κατευθύνσεις	84
6	Αποδοτική υδατογράφηση στο χωρικό πεδίο μέσω μιας νέας αντιληπτικής μάσκας και ενός νέου “τυφλού” σχήματος ανίχνευσης	87
6.1	Εισαγωγή	88
6.2	Η προτεινόμενη ένθεση υδατογραφήματος	90
6.2.1	Διατύπωση προβλήματος	90
6.2.2	Αποθορυβοποίηση εικόνας και αντιληπτική μάσκα	91
6.2.3	Παραγωγή της νέας αντιληπτικής μάσκας	92
6.2.4	Μελέτη της απόδοσης της προτεινόμενης μάσκας	94
6.3	Το προτεινόμενο σχήμα ανίχνευσης	103

6.3.1	Ντετερμινιστική προσέγγιση	105
6.3.2	Στοχαστική προσέγγιση	107
6.4	Καταλληλότητα του προτεινόμενου ανιχνευτή	110
6.5	Πειραματικά αποτελέσματα	113
6.6	Συμπεράσματα, μελλοντικές ερευνητικές κατευθύνσεις	117
7	Βελτιωμένη ένθεση πολλαπλασιαστικών υδατογραφημάτων μέσω χωρο- χρονικής κωδικοποίησης μπλοκ	119
7.1	Εισαγωγή	119
7.2	Διατύπωση του προβλήματος και μοντέλο καναλιού	121
7.3	Προτεινόμενος αλγόριθμος	122
7.3.1	Ένθεση υδατογραφήματος	123
7.3.2	Ανίχνευση υδατογραφήματος	125
7.4	Πειραματικά αποτελέσματα	127
7.5	Συμπεράσματα, μελλοντικές ερευνητικές κατευθύνσεις	129
8	Γενικά συμπεράσματα. Μελλοντικές ερευνητικές κατευθύνσεις	131
A	Απόδειξη Θεωρήματος 1 (Κεφάλαιο 4)	135
B	Απόδειξη της σχέσης (6.3.5)	137
Γ	Παραγωγή μέσων τιμών και διασπορών των συσχετίσεων (Κεφάλαιο 6)	141
	Βιβλιογραφία	147

Ευχαριστίες

Τώρα που τελειώνω τη διδακτορική μου διατριβή¹ θέλω να ευχαριστήσω κάποιους ανθρώπους. Πρώτα απ' όλους, θα ήθελα να ευχαριστήσω τον επιβλέποντα καθηγητή μου, κ. Κωνσταντίνο Μπερμπερίδη για τη συνεργασία και τις εποικοδομητικές συζητήσεις που είχαμε όλα αυτά τα χρόνια.

Ευχαριστώ, επίσης, τον κ. Εμμανουήλ Ψαράκη για τη συνεργασία που είχαμε το τελευταίο διάστημα.

Τέλος, ευχαριστώ τα μέλη της τριμελούς συμβουλευτικής επιτροπής της διδακτορικής μου διατριβής, κ. Σπυρίδωνα Λυκοθανάση και κ. Αθανάσιο Σκόδρα, καθώς και τα υπόλοιπα μέλη της επταμελούς εξεταστικής επιτροπής κ. Θεόδωρο Παπαθεοδώρου, κ. Αθανάσιο Στουραίτη και κ. Σπυρίδωνα Φωτόπουλο.

¹ Η διδακτορική διατριβή στηρίχθηκε οικονομικά από το ΕΠΕΑΕΚ II / Πρόγραμμα ΗΡΑΚΛΕΙΤΟΣ.

Περίληψη

Δύο πολύ σημαντικά προβλήματα της ψηφιακής επεξεργασίας εικόνας, με μεγάλο πεδίο εφαρμογών, είναι η αντιστοίχιση και η ψηφιακή υδατογράφηση εικόνων. Η αντιστοίχιση εικόνων έχει σαν σκοπό την εύρεση γεωμετρικών και άλλων διαφορών ανάμεσα σε δύο ή περισσότερες εικόνες μιας σκηνής, οι οποίες έχουν ληφθεί σε διαφορετικές χρονικές στιγμές, από διαφορετικούς αισθητήρες ή από διαφορετικές θέσεις. Η ψηφιακή υδατογράφηση εικόνων έρχεται να προσφέρει κατοχύρωση των πνευματικών δικαιωμάτων των ψηφιακών δεδομένων και λειτουργεί συμπληρωματικά με την κρυπτογραφία. Η κρυπτογραφία προστατεύει τα ψηφιακά δεδομένα κατά τη διάρκεια της μετάδοσής τους, αλλά στο δέκτη τα δεδομένα είναι πανομοιότυπα με τα αυθεντικά και απροστάτευτα. Η υδατογράφηση έρχεται να λύσει αυτό το πρόβλημα, εισάγοντας στα δεδομένα ένα ανεπαίσθητο σήμα, ένα υδατογράφημα, με τέτοιο τρόπο ώστε να είναι δύσκολο να αφαιρεθεί.

Το πρόβλημα της αντιστοίχισης μπορεί να αποτελέσει τμήμα της ψηφιακής υδατογράφησης, στη φάση της ανίχνευσης του υδατογραφήματος (συγχρονισμός). Επιπλέον, για την ανίχνευση του υδατογραφήματος χρησιμοποιούνται παρόμοιες ή και ίδιες μετρικές με αυτές που χρησιμοποιούνται στην αντιστοίχιση εικόνων (συσχέτιση, κανονικοποιημένη συσχέτιση, κ.τ.λ.). Έτσι, τα δύο προβλήματα δεν είναι ξένα μεταξύ τους και οποιαδήποτε βελτίωση αφορά την αντιστοίχιση εικόνων μπορεί να έχει θετικές επιδράσεις και στην ψηφιακή υδατογράφηση.

Η έρευνα που έγινε στα πλαίσια της διδακτορικής διατριβής σε σχέση με το πρόβλημα της αντιστοίχισης εικόνων αφορά τις τεχνικές που βασίζονται στη συσχέτιση των εικόνων. Ένα από τα μειονεκτήματα της χωρικής συσχέτισης είναι ότι η περιοχή γύρω από τη μέγιστη τιμή της μπορεί να έχει μεγάλο εύρος και να επηρεάζει την ακρίβεια εκτίμησης του σωστού μετασχηματισμού που πρέπει να εφαρμοστεί σε μια εικόνα προκειμένου να αντιστοιχιστεί με μια άλλη. Για την αντιμετώπιση αυτού του προβλήματος, προτείνεται στο κεφάλαιο 2 μια διαδικασία προ-λεύκανσης των εικόνων, βασισμένη στο φίλτρο σφάλματος πρόβλεψης, καθώς και ένας επαναληπτικός αλγόριθμος αντιστοίχισης, ο οποίος υπολογίζει μετατοπίσεις και περιστροφές [12]. Ο αλγόριθμος αυτός εφαρμόζεται στη συνέχεια σε ακολουθίες ιατρικών εικόνων από

τον τράχηλο της μήτρας με σκοπό τη διάγνωση δυσπλασιών και κακοηθειών [64].

Ένα δεύτερο μειονέκτημα των τεχνικών που βασίζονται στη χωρική συσχέτιση είναι το μεγάλο υπολογιστικό τους κόστος. Στο κεφάλαιο 3 προτείνεται ένα γρήγορο σχήμα υπολογισμού της συσχέτισης, το οποίο βασίζεται σε κατάλληλη τμηματοποίηση της εικόνας και στη χρήση του μετασχηματισμού Fourier [12, 64]. Επιπλέον, το πιο απαιτητικό κομμάτι στη διαδικασία της αντιστοίχισης είναι ο υπολογισμός της χρησιμοποιούμενης μετρικής σαν συνάρτηση της σχετικής θέσης των εικόνων. Έτσι, στο ίδιο κεφάλαιο, για την περίπτωση των μετατοπίσεων, προτείνεται ένα αποδοτικό επαναληπτικό σχήμα, το οποίο μειώνει σημαντικά τον αριθμό των ψαξιμάτων που απαιτούνται για την εύρεση του μεγίστου του συντελεστή συσχέτισης στο χωρικό πεδίο και παρέχει ακρίβεια εικονοστοιχείου [68, 67].

Πολλές εφαρμογές απαιτούν αντιστοίχιση με ακρίβεια υποδιαίρεσης εικονοστοιχείου. Μια τεχνική η οποία παρέχει τέτοια ακρίβεια προτείνεται στο κεφάλαιο 4 και βασίζεται στη μεγιστοποίηση του συντελεστή συσχέτισης. Η τεχνική αυτή δεν απαιτεί ανακατασκευή των τιμών της έντασης και παρέχει μια λύση κλειστού τύπου για την εκτίμηση της μετατόπισης [68, 67].

Όσο αφορά τώρα το πρόβλημα της υδατογράφησης, η έρευνα που διεξάχθηκε στα πλαίσια της διδακτορικής διατριβής είχε σαν στόχο την ένθεση ισχυρών υδατογραφημάτων στο χωρικό πεδίο και τη βελτίωση της ανίχνευσης υδατογραφημάτων και τα αποτελέσματά της παρουσιάζονται στο κεφάλαιο 5. Καταρχήν, προτάθηκε μια χωρική αντιληπτική μάσκα, η οποία βασίζεται στην τοπική διασπορά του σφάλματος πρόβλεψης της αρχικής εικόνας. Παράλληλα, αναπτύχθηκε ένα “τυφλό” σύστημα ανίχνευσης και η βελτιωμένη απόδοσή του σε σχέση με υπάρχοντες ανιχνευτές αποδείχθηκε θεωρητικά [62, 63]. Αρχικά, μελετήθηκε η ιδανική περίπτωση, σύμφωνα με την οποία το υδατογράφημα και η εικόνα ή διαφορετικά υδατογραφήματα θεωρούνται εντελώς ασυσχέτιστα μεταξύ τους και κάποιες ποσότητες μηδενίζονται [62], ενώ στη συνέχεια μελετήθηκε η περίπτωση των πραγματικών δεδομένων, όπου η ιδανική υπόθεση δεν ισχύει [63].

Στο κεφάλαιο 6, προτείνεται μια νέα χωρική μάσκα η οποία επιτρέπει την ένθεση υδατογραφημάτων με εξαιρετικά μεγάλη ενέργεια, διατηρώντας ταυτόχρονα την ποιότητα της εικόνας σε πολύ καλό επίπεδο. Η απόδοσή της συγκρίνεται με πολύ γνωστές και ευρέως χρησιμοποιούμενες μάσκες και αποδεικνύεται σημαντικά καλύτερη. Επίσης, παρουσιάζεται ένα βελτιωμένο σχήμα ανίχνευσης, το οποίο σε συνδυασμό με την προτεινόμενη μάσκα έχει πολύ καλή απόδοση [65].

Τέλος, στο κεφάλαιο 7, προτείνουμε την εισαγωγή ενός υδατογραφήματος στην εικόνα με πολλαπλασιαστικό τρόπο, χρησιμοποιώντας χωρο-χρονική κωδικοποίηση μπλοκ και ειδικότερα μια 4×4 πραγματική, ορθογώνια διάταξη συμβόλων [66]. Η εικόνα χωρίζεται σε τέσσερα

τμήματα, τα οποία με κατάλληλες πράξεις μπορούν να θεωρηθούν σαν τέσσερα διαφορετικά κανάλια με επίπεδη εξασθένιση. Αποδεικνύεται ότι το σχήμα αυτό έχει πολύ καλύτερη απόδοση σε σχέση με την επαναληπτική υδατογράφηση.

Στο κεφάλαιο 8 παρουσιάζονται γενικά συμπεράσματα και κάποιες κατευθύνσεις για μελλοντική έρευνα.

Κεφάλαιο 1

Εισαγωγή

Δύο πολύ σημαντικά προβλήματα της ψηφιακής επεξεργασίας εικόνας είναι η αντιστοίχιση και η ψηφιακή υδατογράφηση εικόνων. Η αντιστοίχιση εικόνων αποτελεί προαπαιτούμενη επεξεργασία σε πολλές εφαρμογές, όπως για παράδειγμα η απόκτηση πληροφορίας από εικόνες που έχουν ληφθεί από αεροσκάφη ή από δορυφόρους και η ιατρική διάγνωση. Η ψηφιακή υδατογράφηση εικόνων δίνει τη δυνατότητα κατοχύρωσης των πνευματικών δικαιωμάτων ψηφιακών δεδομένων. Τα δύο αυτά προβλήματα έχουν κάποια κοινά χαρακτηριστικά (π.χ. μετρικές) και η αντιστοίχιση (συγχρονισμός) αποτελεί προϋπόθεση για την ανίχνευση υδατογραφημάτων σε εικόνες. Θα δούμε στη συνέχεια τα δύο αυτά προβλήματα με περισσότερες λεπτομέρειες.

1.1 Αντιστοίχιση Εικόνων

Αντιστοίχιση εικόνων (image registration) είναι η διαδικασία “ευθυγράμμισης” δύο ή περισσότερων εικόνων της ίδιας σκηνής, οι οποίες έχουν ληφθεί σε διαφορετικές χρονικές στιγμές, από διαφορετικά σημεία ή από διαφορετικούς αισθητήρες, [18, 162]. Η “ευθυγράμμιση” έχει την έννοια της αποκατάστασης γεωμετρικών παραμορφώσεων που μπορεί να υπάρχουν ανάμεσα στις συγκρινόμενες εικόνες. Γνωρίζοντας τη σχέση που υπάρχει ανάμεσα στη θέση των εικονοστοιχείων ή κάποιων χαρακτηριστικών μιας εικόνας και στη θέση των αντίστοιχων εικονοστοιχείων ή χαρακτηριστικών σε μια άλλη εικόνα, μας δίνεται η δυνατότητα να συγκρίνουμε την ένταση των εικόνων στις συγκεκριμένες θέσεις.

Η αντιστοίχιση εικόνων αποτελεί το κρίσιμο στάδιο σε πολλές διαδικασίες ανάλυσης εικόνων. Αποτελεί προαπαιτούμενο στο *remote sensing* για την παρακολούθηση του περιβάλλοντος, την ανίχνευση αλλαγών, την πρόβλεψη του καιρού, την κατασκευή εικόνων υψηλής ανάλυσης και την ενσωμάτωση πληροφορίας σε γεωγραφικά πληροφοριακά συστήματα, καθώς και στην ιατρική, [50, 56, 90, 143], για το συνδυασμό δεδομένων από διαφορετικές πηγές

(π.χ. CT-*Computed Tomography* και MRI-*Magnetic Resonance Imaging*) προκειμένου να υπάρχει πιο πλήρης πληροφορία για τον ασθενή, η οποία θα οδηγεί σε ασφαλέστερη διάγνωση. Χρησιμοποιείται ακόμα στη χαρτογράφηση, καθώς και στην τεχνητή όραση (*computer vision*) για εντοπισμό στόχων ή για αυτόματο έλεγχο ποιότητας.

Εξαιτίας της μεγάλης ποικιλίας των εικόνων και των διαφορετικού τύπου υποβαθμίσεων στις οποίες αυτές μπορεί να υπόκεινται, είναι αδύνατος ο σχεδιασμός μιας γενικής τεχνικής, η οποία να είναι εφαρμόσιμη σε όλα τα προβλήματα αντιστοίχισης εικόνων. Κάθε τεχνική πρέπει να λαμβάνει υπόψη της όχι μόνο τον τύπο της γεωμετρικής παραμόρφωσης ανάμεσα στις εικόνες, αλλά επίσης τις διαφοροποιήσεις στην ένταση των εικονοστοιχείων, την ύπαρξη θορύβου, την επιθυμητή ακρίβεια (*accuracy*) αντιστοίχισης και διάφορα άλλα χαρακτηριστικά που εξαρτώνται από τις επιμέρους εφαρμογές.

Ωστόσο, η πλειοψηφία των τεχνικών αντιστοίχισης ακολουθούν κάποια βασικά βήματα. Αρχικά, αν η αντιστοίχιση δε γίνεται με βάση μια ολόκληρη περιοχή της εικόνας, πρέπει να ανιχνευθούν *χαρακτηριστικά* (*features*) των εικόνων κατάλληλα για αντιστοίχιση, όπως ακμές, περιγράμματα, γωνίες και σημεία τομής ανάμεσα σε γραμμές. Στη συνέχεια πρέπει να καθοριστεί η αντιστοιχία ανάμεσα στα χαρακτηριστικά των συγκρινόμενων εικόνων. Υπάρχουν διάφοροι περιγραφείς σχήματος και μετρικές ομοιότητας που χρησιμοποιούνται για αυτό το σκοπό. Το επόμενο βήμα είναι ο υπολογισμός των παραμέτρων του μετασχηματισμού που πρέπει να εφαρμοστεί σε μια εικόνα, προκειμένου να “ευθυγραμμιστεί” με μία άλλη. Τέλος, ο μετασχηματισμός πρέπει να εφαρμοστεί πάνω στην εικόνα. Σε αυτό το σημείο οι τιμές της εικόνας που αντιστοιχούν σε θέσεις με μη ακέραιες συντεταγμένες υπολογίζονται με κατάλληλες τεχνικές *παρεμβολής* (*interpolation*).

1.1.1 Αντιστοίχιση με βάση τα χαρακτηριστικά της εικόνας

Όταν η αντιστοίχιση γίνεται με βάση κάποια χαρακτηριστικά των εικόνων [30, 48, 81], η επιλογή των χαρακτηριστικών που θα χρησιμοποιηθούν είναι πολύ σημαντικό βήμα. Τα χαρακτηριστικά πρέπει να είναι ευδιάκριτα, κατανεμημένα σε όλη την εικόνα και εύκολα ανιχνεύσιμα και στις δύο συγκρινόμενες εικόνες. Πρέπει να είναι σταθερά στο χρόνο, να παραμένουν δηλαδή σε συγκεκριμένες θέσεις, σε όλη τη διάρκεια λήψης εικόνων. Επίσης, προκειμένου να είναι συγκρίσιμα, πρέπει να παραμένουν *ανεπηρέαστα* (*invariant*) από τις αλλαγές στη γεωμετρία της εικόνας ή την παρουσία θορύβου.

Υπάρχουν χαρακτηριστικά που μοιάζουν με περιοχές (*region-like features*), τα οποία μπορεί να είναι οι προβολές περιοχών κατάλληλου μεγέθους, με κλειστά όρια και μεγάλη αντίθεση [40, 47], όπως λίμνες [46], κτίρια [57], αστικές περιοχές [119], κ.τ.λ. Οι περιοχές αναπαριστώνται συχνά με τα κέντρα βάρους τους, τα οποία δεν επηρεάζονται από περιστροφές, κλιμάκωση,

ασυμμετρίες, θόρυβο ή διαφοροποιήσεις στη φωτεινότητα. Αυτού του είδους τα χαρακτηριστικά ανιχνεύονται με τη χρήση τεχνικών *κατάτμησης (segmentation)*, [98], της οποίας η ακρίβεια έχει σημαντική επίδραση στην απόδοση της αντιστοίχισης. Άλλα χαρακτηριστικά μπορεί να είναι οι *γραμμές (line features)*, οι οποίες μπορεί να είναι γενικές αναπαραστάσεις γραμμών [57], περιγράμματα αντικειμένων [81], ακτές [126], δρόμοι, κ.τ.λ. Η αντιστοιχία των γραμμών συνήθως εκφράζεται από ζεύγη σημείων του τέλους ή της μέσης των γραμμών. Για την ανίχνευση αυτού του είδους των χαρακτηριστικών χρησιμοποιούνται τεχνικές *ανίχνευσης ακμών (edge detection)*, [161]. Τέλος, άλλου είδους χαρακτηριστικά είναι τα *σημεία (point features)*, τα οποία μπορεί να είναι τομές γραμμών [128], διασταυρώσεις δρόμων [119], κεντροειδή υδάτινων περιοχών [141], τοπικές ασυνέχειες καμπύλωσης, οι οποίες ανιχνεύονται με τη χρήση Gabor wavelets [159], ή γωνίες [150]. Οι *γωνίες (corners)* αποτελούν μια ιδιαίτερη κατηγορία χαρακτηριστικών, καθώς είναι δύσκολο να οριστούν μαθηματικά. Διαισθητικά, οι γωνίες γίνονται αντιληπτές σαν σημεία με μεγάλη καμπυλότητα πάνω στα όρια των περιοχών. Για περισσότερες πληροφορίες σχετικά με τους ανιχνευτές γωνιών, ο αναγνώστης μπορεί να ανατρέξει στις εργασίες [118] και [160].

Έχοντας στη διάθεσή μας δύο ομάδες χαρακτηριστικών για δύο συγκρινόμενες εικόνες, το επόμενο βήμα είναι να βρούμε τη μεταξύ τους αντιστοιχία. Αυτό μπορεί να γίνει είτε χρησιμοποιώντας τη χωρική τους σχέση, είτε χρησιμοποιώντας διάφορους περιγραφείς σχήματος. Οι τεχνικές που βασίζονται στη χωρική σχέση ανάμεσα στα χαρακτηριστικά εφαρμόζονται συνήθως σε περιπτώσεις που τα χαρακτηριστικά που έχουν ανιχνευτεί είναι ασαφή ή οι γειτονιές τους είναι τοπικά αλλοιωμένες. Οι τεχνικές αυτές εκμεταλλεύονται την πληροφορία που προκύπτει από την απόσταση ανάμεσα στα χαρακτηριστικά και από την κατανομή τους στο χώρο. Στην εργασία [46], η αντιστοίχιση πραγματοποιείται με βάση έναν αλγόριθμο αντιστοίχισης γραφημάτων. Στην [128], μια τεχνική *ομαδοποίησης (clustering)* προσπαθεί να αντιστοιχίσει σημεία που συνδέονται με ακμές ή τμήματα γραμμών. Τέλος, στην [16], η αντιστοίχιση βασίζεται στην ελαχιστοποίηση του μέσου *RMS-Root Mean Square* ανάμεσα σε χαρακτηριστικά γραμμές.

Η αντιστοιχία των χαρακτηριστικών μπορεί επίσης να εκτιμηθεί χρησιμοποιώντας την περιγραφή τους, η οποία πρέπει να παραμένει ανεπηρέαστη από τις αναμενόμενες παραμορφώσεις της εικόνας. Η περιγραφή πρέπει να ικανοποιεί κάποιες συνθήκες. Οι πιο σημαντικές είναι α) η περιγραφή να είναι ανεπηρέαστη από παραμορφώσεις (*invariance*), που σημαίνει ότι οι περιγραφές αντίστοιχων χαρακτηριστικών από δύο συγκρινόμενες εικόνες πρέπει να είναι ίδιες, β) η *μοναδικότητα (uniqueness)*, που σημαίνει ότι δύο διαφορετικά χαρακτηριστικά πρέπει να έχουν διαφορετική περιγραφή, γ) η *σταθερότητα (stability)*, που σημαίνει ότι αν ένα χαρακτηριστικό είναι ελαφρώς παραμορφωμένο σε σχέση με ένα άλλο κατά άγνωστο τρόπο, οι

περιγραφές τους πρέπει να είναι κοντινές, και δ) η *ανεξαρτησία* (*independence*), που σημαίνει ότι αν η περιγραφή είναι ένα διάνυσμα, τότε τα στοιχεία του διανύσματος πρέπει να είναι ανεξάρτητα. Στην πράξη, δεν είναι απαραίτητο ή δε συμβαίνει να ικανοποιούνται όλες οι παραπάνω συνθήκες.

Τα χαρακτηριστικά που έχουν τις πιο κοντινές περιγραφές θεωρούνται να είναι αντίστοιχα. Η επιλογή του τύπου της περιγραφής εξαρτάται από τα χαρακτηριστικά και την υποτιθέμενη γεωμετρική παραμόρφωση των εικόνων. Για την αντιστοίχιση των περιγραφών χρησιμοποιείται συνήθως ο κανόνας της ελάχιστης απόστασης με τη χρήση κατωφλίου. Αν απαιτείται ένας αλγόριθμος με καλύτερη απόδοση, μπορούμε να χρησιμοποιήσουμε τους συντελεστές της πιθανότητας αντιστοίχισης [39], οι οποίοι συμπεριφέρονται καλύτερα στην περίπτωση που υπάρχουν αμφισβητήσιμες καταστάσεις. Ένας άλλος τρόπος επιλογής των χαρακτηριστικών είναι σύμφωνα με την αξιοπιστία των πιθανών αντιστοιχίσεων τους [49].

Η πιο απλή περιγραφή χαρακτηριστικών είναι η ένταση της εικόνας σε μια μικρή γειτονιά τους. Για την εκτίμηση της αντιστοιχίας υπολογίζεται ο συντελεστής συσχέτισης ανάμεσα σε τέτοιου είδους γειτονιές [159]. Η *αμοιβαία πληροφορία* (*mutual information*) έχει χρησιμοποιηθεί για την αντιστοίχιση χαρακτηριστικών [163], ενώ οποιοσδήποτε ανεπηρέαστος περιγραφέας σχήματος μπορεί επίσης να χρησιμοποιηθεί, [100]. Πολλές μέθοδοι χρησιμοποιούν *moment-based invariants* για την περιγραφή χαρακτηριστικών που βασίζονται σε περιοχές με κλειστά όρια. Περισσότερες πληροφορίες μπορεί να βρει κανείς στην εργασία [162].

Σε αντίθεση με τις τεχνικές που βασίζονται σε ολόκληρες περιοχές της εικόνας, οι τεχνικές που βασίζονται στα χαρακτηριστικά της δε δουλεύουν απευθείας με τις τιμές της έντασης της εικόνας. Τα χαρακτηριστικά αναπαριστούν πληροφορία σε υψηλότερο επίπεδο και η χρήση τους είναι κατάλληλη για περιπτώσεις που υπάρχουν αλλαγές στη φωτεινότητα της εικόνας ή όταν τα δεδομένα προέρχονται από διαφορετικούς αισθητήρες. Επίσης, δεδομένου ότι η αντιστοίχιση γίνεται μεταξύ διανυσμάτων μικρού μήκους σε σχέση με την αντιστοίχιση ανάμεσα σε ολόκληρες περιοχές, το υπολογιστικό κόστος αυτών των τεχνικών είναι μικρότερο. Παρόλα αυτά, το πρόβλημα καθορισμού χαρακτηριστικών προς αντιστοίχιση είναι ιδιαίτερα πολύπλοκο και οι τεχνικές αυτές είναι πολύ ευαίσθητες στην ακρίβεια εξαγωγής των χαρακτηριστικών.

1.1.2 Αντιστοίχιση με βάση ολόκληρες περιοχές της εικόνας

Οι τεχνικές που βασίζονται σε περιοχές των εικόνων και συνεπώς στην ένταση των εικόνων είναι πιο ακριβές υπολογιστικά, αλλά αποφεύγουν τις δυσκολίες που έχει το στάδιο της εξαγωγής χαρακτηριστικών. Σύμφωνα με αυτές τις τεχνικές παράθυρα προκαθορισμένου μεγέθους ή ακόμα και ολόκληρες εικόνες χρησιμοποιούνται για την εκτίμηση των παραμέτρων αντιστοίχισης, [2, 7, 111]. Η χρήση παραθύρου μπορεί σε κάποιες περιπτώσεις να είναι προβληματική,

καθώς αν χρησιμοποιηθεί για παράδειγμα ένα παραθύρο που περιέχει μια ομαλή περιοχή χωρίς φανερές λεπτομέρειες μπορεί να οδηγήσει σε λανθασμένη αντιστοίχιση. Ωστόσο, το πρόβλημα αυτό αντιμετωπίζεται με τη χρήση παραθύρων μεγαλύτερου μεγέθους, ανάλογα με την περίπτωση που εξετάζουμε. Επίσης, καθώς κάποιες από τις τεχνικές αυτές βασίζονται στην ένταση των εικόνων, είναι ευαίσθητες σε αλλαγές της έντασης που εισάγονται στις εικόνες είτε από θόρυβο, είτε από αλλαγές στη φωτεινότητα, είτε από τη χρήση διαφορετικών αισθητήρων.

Μια χαρακτηριστική τεχνική αυτής της κατηγορίας είναι η *ετεροσυσχέτιση* (*cross-correlation*) και οι παραλλαγές της [112]. Ο *συντελεστής συσχέτισης* (*correlation coefficient*) είναι μία από αυτές τις παραλλαγές, και λόγω της κανονικοποίησης που υπεισέρχεται στον υπολογισμό του είναι μια μετρική ανθεκτική σε διαφοροποιήσεις της έντασης των εικόνων. Αυτή η μετρική υπολογίζεται για ζεύγη παραθύρων των συγκρινόμενων εικόνων και η θέση της μέγιστης τιμής της δηλώνει τη μετατόπιση που πρέπει να υποστεί η μια εικόνα σε σχέση με την άλλη, προκειμένου να βρεθούν τα αντίστοιχα τμήματά τους. Αν η ακρίβεια που απαιτείται είναι υποδιαίρεσης *εικονοστοιχείου* (*subpixel*), τότε η συνήθης διαδικασία είναι να χρησιμοποιηθεί παρεμβολή στις τιμές του συντελεστή συσχέτισης [139].

Η αντιστοίχιση που βασίζεται στη συσχέτιση των εικόνων δουλεύει με ακρίβεια στην περίπτωση που η γεωμετρική παραμόρφωση που υπάρχει ανάμεσα στις εικόνες είναι η *μετατόπιση* (*translation*). Ωστόσο, μπορεί να εφαρμοστεί με επιτυχία και στην περίπτωση που υπάρχει μικρή περιστροφή και κλιμάκωση. Υπάρχουν πιο γενικευμένες εκδόσεις για εικόνες με μεγαλύτερες γεωμετρικές παραμορφώσεις, όπου η συσχέτιση υπολογίζεται για κάθε γεωμετρικό μετασχηματισμό της εικόνας. Αυτό φυσικά έχει ως συνέπεια να αυξάνεται το υπολογιστικό κόστος.

Παρόμοιος με τις τεχνικές συσχέτισης είναι και ο *αλγόριθμος ακολουθιακής ανίχνευσης ομοιότητας* (*SSDA-Sequential Similarity Detection Algorithm*) [7]. Χρησιμοποιεί μια προσέγγιση ακολουθιακού ψαξίματος και ένα μέτρο απόστασης πιο απλό από τη συσχέτιση. Υπολογίζει το άθροισμα των απόλυτων διαφορών ανάμεσα στις τιμές της έντασης των εικόνων και εφαρμόζει ένα κατώφλι. Αν η τιμή του αθροίσματος ξεπερνά το δεδομένο κατώφλι, τότε το υποψήφιο ζεύγος προς αντιστοίχιση απορρίπτεται και εξετάζεται το επόμενο ζεύγος. Ο αλγόριθμος αυτός είναι πιο γρήγορος από τη συσχέτιση, αλλά λιγότερο ακριβής. Μια πιο πρόσφατη μετρική που χρησιμοποιείται για αντιστοίχιση εικόνων που προέρχονται από διαφορετικές πηγές είναι ο *λόγος συσχέτισης* (*correlation ratio*), [117]. Η μετρική αυτή μπορεί να αντιμετωπίσει διαφοροποιήσεις στην ένταση των εικόνων, σε αντίθεση με την κλασική συσχέτιση.

Δύο μειονεκτήματα των τεχνικών που βασίζονται στη χωρική συσχέτιση είναι ότι η περιοχή γύρω από το μέγιστο της εκάστοτε μετρικής έχει μεγάλο εύρος και ότι το υπολογιστικό τους

κόστος είναι μεγάλο. Το μέγιστο μπορεί να γίνει πιο οξύ με κατάλληλη προεπεξεργασία ή με κατάλληλο φιλτράρισμα, ενώ το υπολογιστικό κόστος αυτών των τεχνικών μπορεί να μειωθεί με τη χρήση γρήγορων αλγορίθμων. Τέτοιου είδους τεχνικές χρησιμοποιούνται ευρέως και εξαιτίας του ότι είναι εύκολα υλοποιήσιμες σε υλικό είναι χρήσιμες για εφαρμογές πραγματικού χρόνου.

Στην περίπτωση που επιθυμούμε μεγαλύτερη ταχύτητα ή οι εικόνες έχουν ληφθεί κάτω από διαφορετικές συνθήκες ή υπάρχει θόρυβος που εξαρτάται από τις συχνότητες, προτιμάται η χρήση τεχνικών που εφαρμόζονται στο πεδίο των συχνοτήτων. Μια τέτοια τεχνική είναι η *συσχέτιση συχνότητας* (*phase correlation*), η οποία βασίζεται στο *θεώρημα ολίσθησης του Fourier* (*Fourier Shift Theorem*) και η οποία προτάθηκε για την αντιστοίχιση μετατοπισμένων εικόνων. Υπολογίζει το *cross-power spectrum* των εικόνων

$$\frac{\mathcal{F}(f)\mathcal{F}(f_t)^*}{|\mathcal{F}(f)\mathcal{F}(f_t)^*|} = e^{2\pi(ix+jy)} \quad (1.1.1)$$

όπου f είναι η εικόνα αναφοράς, f_t μια μετατοπισμένη έκδοσή της, $\mathcal{F}$ είναι ο μετασχηματισμός Fourier και το σύμβολο $*$ δηλώνει συζυγή μιγαδικό. Στη συνέχεια, παίρνει τον αντίστροφο μετασχηματισμό Fourier (που μοιάζει με τη συνάρτηση Δέλτα) και ψάχνει για τη θέση της *αιχμής* (*peak*). Η τεχνική αυτή είναι ανθεκτική στο συσχετισμένο θόρυβο και στις μη ομοιόμορφες διαφοροποιήσεις στη φωτεινότητα. Επίσης, το υπολογιστικό της κόστος είναι μικρότερο σε σχέση με τις τεχνικές στο χωρικό πεδίο, ειδικότερα για μεγάλες εικόνες. Μια επέκταση της συσχέτισης φάσης παρουσιάστηκε στην εργασία [135], για την περίπτωση που υπάρχει και περιστροφή. Αν υπάρχει και κλιμάκωση, τότε η αντιστοίχιση μπορεί να επιτευχθεί συνδυάζοντας τη συσχέτιση φάσης με τη μετατροπή του μέτρου του FFT σε *πολικό-λογαριθμικό* (*polar-log*), που παραπέμπει στο μετασχηματισμό Fourier-Mellin [21, 116, 154]. Τέλος, στην εργασία [4] προτείνεται ο υπολογισμός της συσχέτισης στο πεδίο των συχνοτήτων.

Μια άλλη κατηγορία τεχνικών περιοχής που έχουν εμφανιστεί πρόσφατα χρησιμοποιούν ως μετρική την αμοιβαία πληροφορία και φαίνεται να επικρατούν στο πρόβλημα της αντιστοίχισης εικόνων που προέρχονται από διαφορετικές πηγές. Το πρόβλημα αυτό είναι αρκετά δύσκολο, αλλά πολλές φορές η λύση του είναι απαραίτητη, κυρίως όταν πρόκειται για ιατρικές εικόνες. Η σύγκριση τέτοιων εικόνων μπορεί να οδηγήσει σε διάγνωση, που σε άλλη περίπτωση θα ήταν αδύνατη. Η αμοιβαία πληροφορία προέρχεται από τη θεωρία πληροφορίας και είναι ένα μέτρο της στατιστικής εξάρτησης ανάμεσα σε δύο ομάδες δεδομένων. Η αμοιβαία πληροφορία ανάμεσα σε δύο τυχαίες μεταβλητές f και f_t , που στην περίπτωσή μας είναι οι συγκρινόμενες εικόνες, δίνεται από τη σχέση

$$MI(f, f_t) = H(f_t) - H(f_t|f) = H(f) + H(f_t) - H(f, f_t) \quad (1.1.2)$$

όπου το $H(f) = -E_f(\log(P(f)))$ αναπαριστά την εντροπία της τυχαίας μεταβλητής f και $P(f)$ είναι η κατανομή της πιθανότητας της f . Οι τεχνικές αντιστοίχισης βασίζονται στη μεγιστοποίηση της αμοιβαίας πληροφορίας. Στα πρώτα άρθρα που προτείνονται τέτοιου είδους τεχνικές [146, 152] η αμοιβαία πληροφορία μεγιστοποιείται χρησιμοποιώντας τον αλγόριθμο βελτιστοποίησης του gradient-descent. Περισσότερες αναφορές για αντιστοίχιση εικόνων με χρήση της αμοιβαίας πληροφορίας βρίσκονται στο άρθρο [108].

Η εύρεση της ελάχιστης τιμής ενός μέτρου ανομοιότητας ή η εύρεση της μέγιστης τιμής ενός μέτρου ομοιότητας είναι ένα πολυδιάστατο πρόβλημα βελτιστοποίησης, όπου ο αριθμός των διαστάσεων αντιστοιχεί στους βαθμούς ελευθερίας του αναμενόμενου γεωμετρικού μετασχηματισμού. Η μόνη μέθοδος που δίνει λύση ολικού ακρότατου είναι το εξαντλητικό ψάξιμο πάνω στην εικόνα. Αν και αυτή η διαδικασία είναι υπολογιστικά ακριβή, χρησιμοποιείται συχνά, κυρίως στην περίπτωση που υπάρχουν μόνο μετατοπίσεις.

Οι τεχνικές αντιστοίχισης που βασίζονται σε ολόκληρες περιοχές των εικόνων (και όχι στα χαρακτηριστικά τους) είναι συνήθως αυτές που μπορούν να παρέχουν ακρίβεια σε υποδιαίρεση του εικονοστοιχείου. Οι πιο συχνά χρησιμοποιούμενες από αυτές βασίζονται στην παρεμβολή, [139]. Κάποια παραδείγματα είναι η παρεμβολή στην ένταση των εικονοστοιχείων της εικόνας, στη συσχέτιση και στη συσχέτιση φάσης. Ωστόσο, η ακρίβεια αυτών των τεχνικών εξαρτάται από την ποιότητα των αλγορίθμων παρεμβολής. Άλλες τεχνικές βασίζονται στις διαφορικές ιδιότητες που έχουν οι ακολουθίες των εικόνων [139] ή διατυπώνουν το πρόβλημα της αντιστοίχισης με ακρίβεια υποδιαίρεσης εικονοστοιχείου σαν ένα πρόβλημα βελτιστοποίησης [71, 138, 142]. Ένας μεγάλος αριθμός τεχνικών εφαρμόζεται στο πεδίο των συχνοτήτων, το οποίο δίνει τη δυνατότητα για πιο εύκολη περιγραφή και χειρισμό της ψευδωνυμίας (*aliasing*). Οι τεχνικές αυτές βασίζονται στην ακύρωση των συχνοτήτων της εικόνας που περιέχουν ψευδωνυμία [71, 129, 142, 145]. Το πεδίο των συχνοτήτων όμως περιορίζεται σε καθολικά μοντέλα κίνησης, σε αντίθεση με το χωρικό πεδίο, το οποίο επιτρέπει τη χρήση πιο γενικών μοντέλων [2, 70, 104].

1.1.3 Υπολογισμός παραμέτρων μετασχηματισμού

Μετά την εύρεση των χαρακτηριστικών των εικόνων και της μεταξύ τους αντιστοιχίας, το επόμενο βήμα είναι η εκτίμηση των παραμέτρων του μετασχηματισμού που πρέπει να υποστεί η μια εικόνα σε σχέση με την άλλη προκειμένου να βρεθούν τα αντίστοιχα τμήματά τους. Αρχικά, πρέπει να επιλεγεί μια κατάλληλη *συνάρτηση αντιστοίχισης (mapping function)*, η οποία να λαμβάνει υπόψη της τον υποτιθέμενο γεωμετρικό μετασχηματισμό, τη διαδικασία λήψης της εικόνας και την απαιτούμενη ακρίβεια αντιστοίχισης. Σε ειδικές περιπτώσεις που ο γεωμετρικός μετασχηματισμός είναι μερικώς γνωστός, όπως για παράδειγμα όταν είναι γνωστή

η παραμόρφωση που εισάγεται από τη συσκευή λήψης ή η γεωμετρία της σκηνής, μπορεί να εφαρμοστεί ο αντίστροφος μετασχηματισμός.

Τα μοντέλα των συναρτήσεων αντιστοίχισης μπορούν να χωριστούν σε δύο κατηγορίες, με βάση την περιοχή υποστήριξης που χρησιμοποιούν, το μέγεθος δηλαδή της περιοχής της εικόνας που χρησιμοποιούν. Τα καθολικά μοντέλα χρησιμοποιούν όλα τα σημεία ελέγχου (*control points*) για την εκτίμηση παραμέτρων που αφορούν ολόκληρη την εικόνα. Τα τοπικά μοντέλα αντιμετωπίζουν την εικόνα σαν ένα σύνολο από υποεικόνες και υπολογίζουν παραμέτρους για κάθε υποεικόνα ξεχωριστά. Μια άλλη κατηγοριοποίηση των συναρτήσεων αντιστοίχισης μπορεί να γίνει με βάση την ακρίβεια της αντιστοίχιας των σημείων ελέγχου που χρησιμοποιούνται για την εκτίμηση των παραμέτρων. Συναρτήσεις παρεμβολής προσπαθούν να βρουν την ακριβή αντιστοίχια ανάμεσα στα σημεία ελέγχου των συγκρινόμενων εικόνων, ενώ προσεγγιστικές συναρτήσεις προσπαθούν να βρουν την καλύτερη δυνατή σχέση (*trade-off*) ανάμεσα στην ακρίβεια της αντιστοίχισης και σε άλλες απαιτήσεις που επιβάλλονται στο χαρακτήρα της συνάρτησης. Καθώς οι συντεταγμένες των σημείων ελέγχου συνήθως δεν είναι ακριβείς, τα προσεγγιστικά μοντέλα είναι αυτά που χρησιμοποιούνται περισσότερο.

Όσο αφορά τα καθολικά μοντέλα συναρτήσεων, το πιο απλό είναι αυτό που αποτελείται από μετατόπιση, περιστροφή και κλιμάκωση και περιγράφεται από τις ακόλουθες σχέσεις

$$\begin{aligned} u &= sc(i \cos \varphi - j \sin \varphi) + x \\ v &= sc(i \sin \varphi + j \cos \varphi) + y \end{aligned} \quad (1.1.3)$$

όπου (i, j) είναι οι αρχικές συντεταγμένες, (x, y) η μετατόπιση, φ η γωνία περιστροφής, sc η παράμετρος κλιμάκωσης και (u, v) οι νέες συντεταγμένες. Το μοντέλο (1.1.3) ονομάζεται συγχνά μοντέλο διατήρησης σχήματος, γιατί διατηρεί γωνίες και καμπυλώσεις. Για τον καθορισμό του αρκούν δύο σημεία ελέγχου.

Ένα ελαφρώς πιο γενικό μοντέλο, το οποίο όμως είναι γραμμικό, είναι ο μετασχηματισμός *affine*

$$\begin{aligned} u &= \alpha_0 + \alpha_1 i + \alpha_2 j \\ v &= \beta_0 + \beta_1 i + \beta_2 j \end{aligned} \quad (1.1.4)$$

ο οποίος μπορεί να αντιστοιχίσει ένα παραλληλόγραμμο σε ένα τετράγωνο. Αυτό το μοντέλο ορίζεται από τρία σημεία ελέγχου, τα οποία δεν πρέπει να βρίσκονται στην ίδια ευθεία, και διατηρεί τις ευθείες γραμμές και την παραλληλία τους. Για να χρησιμοποιηθεί, υποτίθεται ότι η απόσταση της κάμερας από τη σκηνή είναι μεγάλη σε σχέση με το μέγεθος της σκηνής, ότι η σκηνή είναι επίπεδη και ότι δεν υπάρχουν τοπικές γεωμετρικές παραμορφώσεις.

Στην περίπτωση που δεν ικανοποιείται η συνθήκη για την απόσταση της κάμερας από τη σκηνή χρησιμοποιείται ένα μοντέλο προοπτικής προβολής (*perspective projection model*)

$$\begin{aligned} u &= \frac{\alpha_0 + \alpha_1 i + \alpha_2 j}{1 + \gamma_1 i + \gamma_2 j} \\ v &= \frac{\beta_0 + \beta_1 i + \beta_2 j}{1 + \gamma_1 i + \gamma_2 j}. \end{aligned} \quad (1.1.5)$$

Το μοντέλο αυτό περιγράφει το μετασχηματισμό μιας επίπεδης σκηνής, η οποία φωτογραφίζεται από μία κάμερα της οποίας ο οπτικός άξονας δεν είναι κάθετος στη σκηνή. Μπορεί να αντιστοιχίσει ένα τετράπλευρο σε ένα τετράγωνο, διατηρώντας τις ευθείες γραμμές. Για τον καθορισμό του χρειάζονται τέσσερα ανεξάρτητα σημεία ελέγχου.

Αν οι παραπάνω υποθέσεις δεν ικανοποιούνται, τότε είναι απαραίτητη η χρήση πολυωνυμικών μοντέλων δεύτερης ή τρίτης τάξης. Μοντέλα μεγαλύτερης τάξης συνήθως δε χρησιμοποιούνται σε πρακτικές εφαρμογές γιατί μπορεί να επιφέρουν παραμόρφωση (*warping*) σε περιοχές της εικόνας που βρίσκονται μακριά από τα σημεία ελέγχου.

Γενικά, ο αριθμός των σημείων ελέγχου είναι μεγαλύτερος από τον ελάχιστο αριθμό σημείων που είναι απαραίτητος για τον καθορισμό της χρησιμοποιούμενης συνάρτησης. Οι παράμετροι των συναρτήσεων καθορίζονται με τη χρήση ελαχίστων τετραγώνων, έτσι ώστε τα πολυώνυμα να ελαχιστοποιούν το άθροισμα των τετραγωνικών σφαλμάτων μεταξύ των σημείων ελέγχου.

Τα καθολικά μοντέλα αντιστοίχισης έχουν το μειονέκτημα ότι δεν μπορούν να χειριστούν κατάλληλα εικόνες στις οποίες υπάρχουν τοπικές παραμορφώσεις, όπως για παράδειγμα οι ιατρικές εικόνες. Η τεχνική των ελαχίστων τετραγώνων έχει ως αποτέλεσμα να μη λαμβάνονται υπόψη οι τοπικές γεωμετρικές παραμορφώσεις, και να προκύπτει ένας καθολικός μετασχηματισμός για ολόκληρη την εικόνα, κάτι το οποίο δεν είναι επιθυμητό. Οι τοπικές περιοχές της εικόνας πρέπει να αντιστοιχίζονται λαμβάνοντας υπόψη την πληροφορία που υπάρχει για τις τοπικές γεωμετρικές παραμορφώσεις. Έτσι, στις περιπτώσεις εικόνων με τοπικές παραμορφώσεις, οι τοπικές μέθοδοι αντιστοίχισης ή τουλάχιστον μέθοδοι αντιστοίχισης που είναι ευαίσθητες στις τοπικές παραμορφώσεις έχουν καλύτερη απόδοση, [38, 45]. Για παράδειγμα στην εργασία [45], παρουσιάζεται μια τεχνική ελαχίστων τετραγώνων με βάρη, η οποία δίνει τη δυνατότητα να αντιστοιχηθούν τοπικά οι εικόνες, διαφοροποιώντας ελαφρώς την αρχική τεχνική των ελαχίστων τετραγώνων.

Μια άλλη κατηγορία συναρτήσεων αντιστοίχισης είναι οι *συναρτήσεις ακτινικής βάσης* (*radial basis functions*). Οι συναρτήσεις αυτές είναι αντιπροσωπευτικές των καθολικών μεθόδων αντιστοίχισης, αλλά μπορούν να χρησιμοποιηθούν και για τοπικές γεωμετρικές παραμορφώσεις. Μια τέτοια συνάρτηση έχει τη μορφή ενός γραμμικού συνδυασμού ακτινικά

μετατοπισμένων συναρτήσεων συν ένα πολυώνυμο μικρού βαθμού, όπως φαίνεται στη συνέχεια

$$u = \alpha_0 + \alpha_1 i + \alpha_2 j + \sum_{k=1}^N \gamma_k d(\mathbf{x}, \mathbf{x}_k). \quad (1.1.6)$$

Η συνάρτηση για τη συντεταγμένη u ορίζεται με αντίστοιχο τρόπο. Αρχικά, αυτού του είδους οι συναρτήσεις αναπτύχθηκαν για την παρεμβολή ανώμαλων επιφανειών. Το όνομά τους “ακτινικές” παραπέμπει στην εξής σημαντική ιδιότητα της τιμής μιας τέτοιας συνάρτησης για κάθε σημείο: εξαρτάται από την απόσταση από το σημείο ελέγχου και όχι από την ακριβή θέση του. Οι πιο συχνά χρησιμοποιούμενες συναρτήσεις ακτινικής βάσης είναι γνωστές σαν *thin-plate splines* (TPS) [15], όπου οι ακτινικοί όροι είναι της μορφής

$$d(\mathbf{x}, \mathbf{x}_k) = \|\mathbf{x} - \mathbf{x}_k\|^2 \ln(\|\mathbf{x} - \mathbf{x}_k\|). \quad (1.1.7)$$

Ένας άλλος τύπος συναρτήσεων που βασίζονται σε splines είναι οι *elastic body splines* (EBS), [31].

Μια άλλη προσέγγιση για αντιστοίχιση εικόνων με ιδιαίτερα πολύπλοκους και/ή τοπικούς μετασχηματισμούς είναι να γίνει χρήση παραμετρικών συναρτήσεων, όπου η εκτίμηση του γεωμετρικού μετασχηματισμού ισοδυναμεί με την εύρεση των “καλύτερων” παραμέτρων. Σύμφωνα με αυτή την προσέγγιση, η οποία λέγεται *ελαστική αντιστοίχιση* (*elastic registration*) [5], η εικόνα θεωρείται σαν ένα *ελαστικό σεντόνι* (*rubber sheet*), πάνω στο οποίο ασκούνται εξωτερικές δυνάμεις που προκαλούν το *τέντωμα* (*stretching*) της εικόνας ή εσωτερικές δυνάμεις που προκαλούν το *λύγισμά* της. Οι δυνάμεις αυτές ορίζονται με βάση κάποιους περιορισμούς *ακαμψίας* (*stiffness*) ή *ομαλότητας* (*smoothness*) με σκοπό την αντιστοίχιση με το μικρότερο δυνατό *λύγισμα* (*bending*) ή *τέντωμα*. Τα βήματα της αντιστοίχισης των χαρακτηριστικών και της εύρεσης των παραμέτρων των συναρτήσεων αντιστοίχισης γίνονται ταυτόχρονα. Αυτό είναι ένα από τα πλεονεκτήματα της ελαστικής αντιστοίχισης, γιατί δεν υπάρχουν περιγραφείς χαρακτηριστικών που να είναι ανεπηρέαστοι από πολύπλοκους γεωμετρικούς μετασχηματισμούς. Η ελαστική αντιστοίχιση μειονεκτεί όταν οι παραμορφώσεις της εικόνας είναι “πολύ τοπικές”.

1.1.4 Επαναδειγματοληψία και μετασχηματισμός της εικόνας

Το τελευταίο βήμα στη διαδικασία της αντιστοίχισης εικόνων είναι να μετασχηματιστεί κατάλληλα η μια εικόνα προκειμένου να αντιστοιχηθεί με την άλλη. Στη μετασχηματισμένη εικόνα μπορεί να υπάρχουν κάποια κενά ή κάποιες επικαλύψεις. Για να αποφευχθούν αυτά τα προβλήματα πρέπει να χρησιμοποιηθεί παρεμβολή. Η παρεμβολή πραγματοποιείται μέσω

της συνέλιξης της εικόνας με έναν κατάλληλο πυρήνα παρεμβολής. Ο βέλτιστος *παρεμβολέας* (*interpolant*) είναι μια δισδιάστατη συνάρτηση *sinc*, η οποία όμως είναι δύσκολο να υλοποιηθεί λόγω της άπειρης έκτασής της. Άλλοι παρεμβολείς που έχουν χρησιμοποιηθεί στη βιβλιογραφία είναι η συνάρτηση *κοντινότερου γείτονα* (*nearest neighbor*), η *διγραμμική* (*bilinear*) και η *δικυβική* (*bicubic*) συνάρτηση, *τετραγωνικά splines*, *κυβικά B-splines*, κ.τ.λ. [99]. Έχει αποδειχτεί ότι οι μεγάλης τάξης πολυωνυμικοί πυρήνες παρουσιάζουν μια πολύ μικρή βελτίωση σε σχέση με την κυβική παρεμβολή, ενώ αυξάνουν σημαντικά το υπολογιστικό κόστος. Την καλύτερη δυνατή σχέση ανάμεσα σε ακρίβεια και υπολογιστικό κόστος την προσφέρει η διγραμμική παρεμβολή και αυτή είναι που χρησιμοποιείται συνήθως στην πράξη. Η χρήση της κυβικής παρεμβολής προτείνεται σε περιπτώσεις που ο γεωμετρικός μετασχηματισμός περιλαμβάνει σημαντική μεγέθυνση της εικόνας. Τέλος, η παρεμβολή κοντινότερου γείτονα πρέπει να αποφεύγεται γιατί προκαλεί αλλοιώσεις στην εναπαδειγματοληπτημένη εικόνα. Η μόνη εξαίρεση είναι οι εικόνες που περιέχουν λίγες εντάσεις και στις οποίες δε θέλουμε να εισάγουμε επιπλέον χρώματα χρησιμοποιώντας μια παρεμβολή μεγαλύτερης τάξης.

1.1.5 Ερευνητικά αποτελέσματα στα πλαίσια της διδακτορικής διατριβής

Η έρευνα που έγινε στα πλαίσια της διδακτορικής διατριβής σε σχέση με το πρόβλημα της αντιστοίχισης εικόνων επικεντρώθηκε σε τεχνικές που βασίζονται στη συσχέτιση των εικόνων. Ένα από τα μειονεκτήματα της χωρικής συσχέτισης είναι ότι η περιοχή γύρω από το μέγιστο έχει μεγάλο εύρος. Αυτό μπορεί να έχει επίπτωση στην εκτίμηση του σωστού μετασχηματισμού που πρέπει να εφαρμοστεί στην εικόνα κατά τη διαδικασία της αντιστοίχισης. Στο κεφάλαιο 2, προκειμένου να αντιμετωπιστεί αυτό το πρόβλημα, προτείνουμε την εφαρμογή μιας κατάλληλης *προ-λεύκανσης* (*pre-whitening*) των εικόνων πριν τον υπολογισμό της ετεροσυσχέτισής τους. Η προ-λεύκανση γίνεται με τη χρήση του *φίλτρου σφάλματος πρόβλεψης* (*prediction error filter*), το οποίο υπολογίζεται από τα δεδομένα. Η ετεροσυσχέτιση των εικόνων μετά την προ-λεύκανση έχει σαν αποτέλεσμα πιο οξύ μέγιστο. Η τεχνική αντιστοίχισης που προτείνεται στο ίδιο κεφάλαιο βρίσκει με επαναληπτικό τρόπο περιστροφές και μετατοπίσεις [12]. Στη συνέχεια, η προτεινόμενη τεχνική αντιστοίχισης εφαρμόζεται με επιτυχία σε ακολουθίες ιατρικών εικόνων από τον τράχηλο της μήτρας με σκοπό τη βελτίωση στη διάγνωση δυσπλασιών και κακοηθειών [64].

Ένα δεύτερο μειονέκτημα των τεχνικών που βασίζονται στη χωρική συσχέτιση είναι το μεγάλο υπολογιστικό τους κόστος. Έτσι, στο κεφάλαιο 3, προτείνουμε ένα γρήγορο σχήμα υπολογισμού της συσχέτισης που βασίζεται σε κατάλληλη τμηματοποίηση της εικόνας και στη χρήση του μετασχηματισμού Fourier [12, 64]. Επιπλέον, το πιο απαιτητικό κομμάτι

στη διαδικασία της αντιστοίχισης είναι ο υπολογισμός της χρησιμοποιούμενης μετρικής σαν συνάρτηση της σχετικής θέσης των εικόνων. Στο ίδιο κεφάλαιο, για την περίπτωση που υπάρχουν μόνο μετατοπίσεις ανάμεσα στις συγκρινόμενες εικόνες, προτείνουμε ένα αποδοτικό επαναληπτικό σχήμα για τον υπολογισμό του συντελεστή συσχέτισης στο χωρικό πεδίο, το οποίο μειώνει σημαντικά τον αριθμό των ψαξιμάτων που απαιτούνται για την εύρεση του μεγίστου και παρέχει ακρίβεια εικονοστοιχείου [68, 67].

Πολλές εφαρμογές απαιτούν αντιστοίχιση με ακρίβεια υποδιαίρεσης εικονοστοιχείου. Οι πιο συχνά χρησιμοποιούμενες τεχνικές που παρέχουν τέτοια ακρίβεια βασίζονται σε παρεμβολή. Στο κεφάλαιο 4, προτείνουμε μια νέα τεχνική αντιστοίχισης με ακρίβεια υποδιαίρεσης εικονοστοιχείου, η οποία βασίζεται στη μεγιστοποίηση του συντελεστή συσχέτισης. Η τεχνική αυτή δεν απαιτεί την ανακατασκευή των τιμών της έντασης και παρέχει μια λύση κλειστού τύπου για την εκτίμηση της μετατόπισης [68, 67]. Επιπλέον, έχει καλύτερη απόδοση σε σχέση με τις υπάρχουσες τεχνικές.

1.2 Η σχέση της αντιστοίχισης εικόνων με την ψηφιακή υδατογράφηση εικόνων

Μια υδατογραφημένη εικόνα μπορεί να υποστεί διάφορες γεωμετρικές παραμορφώσεις, καθολικές ή τοπικές. Αυτό έχει σαν αποτέλεσμα να δημιουργείται πρόβλημα στο στάδιο της ανίχνευσης του υδατογραφήματος. Έτσι, η αντιστοίχιση εικόνων αποτελεί πολλές φορές προϋπόθεση για την ανίχνευση υδατογραφήματων σε εικόνες και είναι ισοδύναμη διαδικασία με το συγχρονισμό, ο οποίος έχει μελετηθεί εκτενώς στις τηλεπικοινωνίες. Επιπλέον, για την ανίχνευση του υδατογραφήματος χρησιμοποιούνται πολλές φορές ίδιες μετρικές ομοιότητας σαν αυτές που χρησιμοποιούνται στην αντιστοίχιση εικόνων, όπως είναι η συσχέτιση και οι παραλλαγές της. Συνεπώς, οποιαδήποτε βελτίωση αφορά τις μετρικές αντιστοίχισης εικόνων, έχει άμεση επίδραση και στην υδατογράφηση εικόνων, στη φάση της ανίχνευσης.

Η πιο απλή προσέγγιση για ανίχνευση υδατογραφήματος μετά από κάποια γεωμετρική παραμόρφωση είναι το εξαντλητικό ψάξιμο [28]. Καθορίζεται ένα πιθανό εύρος για τις τιμές των παραμέτρων της παραμόρφωσης και η επιθυμητή ανάλυση ψαξίματος και εξετάζεται κάθε δυνατός συνδυασμός. Κάθε συνδυασμός αναπαριστά έναν υποθετικό γεωμετρικό μετασχηματισμό που μπορεί να έχει εφαρμοστεί στην εικόνα μετά την ένθεση του υδατογραφήματος. Το ψάξιμο μπορεί να περιλαμβάνει την εφαρμογή του αντίστροφου μετασχηματισμού στην υδατογραφημένη εικόνα και μετά την εφαρμογή του ανιχνευτή υδατογραφήματος για κάθε πιθανό πρότυπο αναφοράς. Εναλλακτικά, μπορεί ο μετασχηματισμός να εφαρμόζεται σε κάθε πιθανό υδατογράφημα, πριν την ανίχνευση. Χρησιμοποιώντας το εξαντλητικό ψάξιμο, όσο αυξάνεται

ο αριθμός των δυνατών παραμορφώσεων, τόσο αυξάνεται και το υπολογιστικό κόστος και η πιθανότητα λανθασμένης ανίχνευσης (*false positive probability*).

Αυτό το πρόβλημα μπορεί να αντιμετωπιστεί αν η υδατογραφημένη εικόνα και το πρότυπο αναφοράς αντιστοιχηθούν σε ένα στάδιο πριν τον ανιχνευτή (*detector*). Αυτή η διαδικασία ονομάζεται *συγχρονισμός* στη βιβλιογραφία που σχετίζεται με ήχο και *αντιστοίχιση* στη βιβλιογραφία που σχετίζεται με εικόνα. Όταν η αρχική εικόνα (πριν την υδατογράφηση) είναι διαθέσιμη στον ανιχνευτή, τότε μπορούμε να χρησιμοποιήσουμε κάποια τεχνική αντιστοίχισης εικόνων [18, 50, 56, 90, 108, 143, 162], για την αντιστοίχιση της αρχικής εικόνας με την υδατογραφημένη εικόνα που φτάνει στον ανιχνευτή, η οποία μπορεί να έχει υποστεί κάποια γεωμετρική παραμόρφωση. Στην περίπτωση που ο ανιχνευτής είναι “τυφλός” (*blind*), δεν έχει δηλαδή γνώση της αρχικής εικόνας, μια συχνά χρησιμοποιούμενη προσέγγιση συγχρονισμού είναι η ένθεση στην εικόνα ενός προτύπου συγχρονισμού μαζί με το υδατογράφημα [140]. Ακολουθώντας μια τέτοια διαδικασία, η ανίχνευση του υδατογραφήματος περιλαμβάνει πρώτα μια φάση εύρεσης του προτύπου συγχρονισμού μέσα στην εξεταζόμενη εικόνα. Ο γεωμετρικός μετασχηματισμός που μπορεί να έχει εφαρμοστεί πάνω σε αυτή την εικόνα μπορεί να καθοριστεί στη συνέχεια, από τη σύγκριση του εξαγόμενου προτύπου συγχρονισμού και του υδατογραφήματος που έχει εισαχθεί στην εικόνα. Στη συνέχεια, ο μετασχηματισμός αντιστρέφεται, και η διαδικασία συνεχίζεται με την ανίχνευση του υδατογραφήματος. Συμπερασματικά, τα πρότυπα συγχρονισμού μπορούν να απλοποιήσουν το ψάξιμο. Αυτό έχει ως αποτέλεσμα να μην αυξάνεται ιδιαίτερα η πιθανότητα λανθασμένης ανίχνευσης και το υπολογιστικό κόστος να παραμένει μικρό σε σχέση με το εξαντλητικό ψάξιμο. Ωστόσο, με τη χρήση τους εισάγονται δύο περιπτώσεις αποτυχίας στην ανίχνευση, αποτυχία στη σωστή ανίχνευση του προτύπου συγχρονισμού και αποτυχία στην ανίχνευση του υδατογραφήματος μετά την αντιστοίχιση. Προκειμένου να υπάρξει σωστή ανίχνευση, πρέπει και το υδατογράφημα και το πρότυπο συγχρονισμού να εισαχθούν στην εικόνα και να ανιχνευθούν με επιτυχία.

Σε κάποιες περιπτώσεις το πρότυπο συγχρονισμού μπορεί να λειτουργήσει και σαν κανονικό υδατογράφημα που φέρει πληροφορία. Αυτό σημαίνει ότι το υδατογράφημα πρέπει να έχει κάποιες ιδιότητες που να επιτρέπουν συγχρονισμό. Μια τέτοια ιδιότητα είναι η περιοδικότητα και μπορεί να χρησιμοποιηθεί σε μια προσέγγιση βασισμένη στην *αυτοσυσχέτιση* (*autocorrelation*). Η αυτοσυσχέτιση μιας εικόνας έχει μια μεγάλη αιχμή στο μηδέν, η οποία μειώνεται δραστηρικά για μη μηδενικές μετατοπίσεις. Όταν στην εικόνα εισάγεται ένα περιοδικό, λευκό πρότυπο συγχρονισμού, η αυτοσυσχέτιση που προκύπτει θα έχει μια περιοδική ακολουθία από αιχμές, η οποία θα καθορίζει την περιοδικότητα του προτύπου. Έτσι, με βάση αυτές τις αιχμές μπορεί να γίνει η απαραίτητη αντιστοίχιση.

Όσο αφορά την ανθεκτικότητα μιας υδατογραφημένης εικόνας σε γεωμετρικές παραμορφώσεις, μπορεί να ακολουθηθεί μια διαφορετική προσέγγιση, η οποία δεν προϋποθέτει τη διαδικασία της αντιστοίχισης. Η προσέγγιση αυτή βασίζεται στην κατασκευή υδατογραφημάτων ανθεκτικών σε γεωμετρικές παραμορφώσεις. Μια βασική κατηγορία τέτοιων τεχνικών κατασκευάζει υδατογραφήματα ανθεκτικά σε μετατοπίσεις, περιστροφές και κλιμακώσεις, χρησιμοποιώντας τις ιδιότητες του μετασχηματισμού Fourier. Συνήθως, το μέτρο του μετασχηματισμού Fourier μετατρέπεται στην πολική-λογαριθμική αναπαράσταση του και η ένθεση του υδατογραφήματος γίνεται σε αυτό το πεδίο. Αυτό επιτυγχάνεται με τη χρήση του μετασχηματισμού Fourier-Mellin, [85, 121]. Ωστόσο, αυτή η διαδικασία περιλαμβάνει αναγκαστικά παρεμβολή, η οποία επιφέρει αλλοιώσεις στην εικόνα, άρα και στο υδατογράφημα, και μπορεί να δημιουργήσει πρόβλημα στην ανίχνευση του υδατογραφήματος.

Μια άλλη κατηγορία τεχνικών υδατογράφησης, όπου γίνεται συγχρονισμός πριν την ανίχνευση του υδατογραφήματος, είναι αυτή που αντί για πρότυπα συγχρονισμού, χρησιμοποιούνται τα χαρακτηριστικά της εικόνας. Η ένθεση του υδατογραφήματος γίνεται με βάση τη γεωμετρία που ακολουθούν τα χαρακτηριστικά της εικόνας. Ωστόσο, η απόδοση μιας τέτοιας διαδικασίας εξαρτάται σε σημαντικό βαθμό από την ακρίβεια εξαγωγής των χαρακτηριστικών της εικόνας.

Φαίνεται λοιπόν από την παραπάνω ανάλυση ότι η αντιστοίχιση εικόνων συνδέεται με την υδατογράφηση εικόνων στο στάδιο της ανίχνευσης του υδατογραφήματος. Επίσης, οι ανιχνευτές που κυρίως χρησιμοποιούνται βασίζονται σε μετρικές συσχέτισης. Τέλος, η πιο απλή προσέγγιση ανίχνευσης υδατογραφήματος μετά από γεωμετρικές παραμορφώσεις είναι το εξαντλητικό ψάξιμο. Άρα, οποιαδήποτε βελτίωση αφορά τις τεχνικές αντιστοίχισης, τον υπολογισμό των μετρικών συσχέτισης ή τον αριθμό των απαιτούμενων ψαξίμάτων έχει άμεσες θετικές επιδράσεις και στη φάση της ανίχνευσης υδατογραφήματος.

1.3 Ψηφιακή υδατογράφηση εικόνων

Ψηφιακή υδατογράφηση εικόνων (*image watermarking*) είναι η διαδικασία ένθεσης ενός αδιόρατου σήματος, ενός υδατογραφήματος, σε μια εικόνα, με τέτοιο τρόπο ώστε να είναι δύσκολο να αφαιρεθεί και να μη δημιουργεί προβλήματα στην ποιότητα της εικόνας [28, 79]. Δεδομένης της μαζικής διανομής ψηφιακών δεδομένων μέσω του διαδικτύου, η ψηφιακή υδατογράφηση έχει γίνει αναγκαία, προκειμένου να προστατεύονται τα πνευματικά δικαιώματα (*copyright*) των ιδιοκτητών των δεδομένων. Η υδατογράφηση εικόνων λειτουργεί συμπληρωματικά με την

κρυπτογραφία. Η κρυπτογραφία παρέχει προστασία κατά τη διάρκεια μετάδοσης των δεδομένων, αλλά στο δέκτη τα δεδομένα είναι όμοια με τα αυθεντικά και απροστάτευτα. Η υδατογράφηση επιδρά στο περιεχόμενο των δεδομένων, εισάγοντας ένα υδατογράφημα το οποίο παραμένει στα δεδομένα και μετά τη μετάδοσή τους.

Η υδατογράφηση έχει αρκετές εφαρμογές, όπως για παράδειγμα η *προστασία των πνευματικών δικαιωμάτων* (*copyright protection*). Ο ιδιοκτήτης μπορεί να εισάγει ένα υδατογράφημα στα δεδομένα, το οποίο να δηλώνει την ταυτότητά του. Επιπλέον, ο ιδιοκτήτης μπορεί να ακολουθήσει μια τεχνική “*δακτυλικών αποτυπωμάτων*” (*fingerprinting*), προκειμένου να βρεθεί η πηγή που παράγει παράνομα αντίγραφα των αυθεντικών δεδομένων. Μπορεί, δηλαδή, να εισάγει διαφορετικά υδατογραφήματα στα αντίγραφα που παρέχει σε διαφορετικούς πελάτες, το οποίο του δίνει τη δυνατότητα να εντοπίσει τους πελάτες που παρέχουν παράνομα αντίγραφα των δεδομένων σε τρίτους. Μια άλλη εφαρμογή της υδατογράφησης είναι η *προστασία αντιγραφής* (*copy protection*). Η πληροφορία που περιέχεται σε ένα υδατογράφημα μπορεί να χρησιμοποιηθεί για τον έλεγχο συσκευών αντιγραφής δεδομένων, προκειμένου να μην επιτρέπεται η αντιγραφή των δεδομένων. Η *παρακολούθηση εκπομπής* (*broadcast monitoring*) των δεδομένων είναι επίσης μια εφαρμογή της υδατογράφησης [3]. Μπορεί για παράδειγμα να εισαχθεί ένα υδατογράφημα σε μια διαφήμιση, το οποίο να επιτρέπει την επιβεβαίωση του τύπου και του χρόνου εκπομπής της διαφήμισης. Επίσης, η υδατογράφηση μπορεί να χρησιμοποιηθεί για την *πιστοποίηση των δεδομένων* (*data authentication*) [89]. Ένα “*εύθραυστο*” (*fragile*) υδατογράφημα μπορεί να εισαχθεί στα δεδομένα, το οποίο δείχνει αν υπάρχει κάποια τροποποίηση στα δεδομένα και ποιο τμήμα των δεδομένων έχει τροποποιηθεί.

Κάθε εφαρμογή της υδατογράφησης έχει τις δικές της απαιτήσεις και η καταλληλότητα ενός συστήματος υδατογράφησης για την εκάστοτε εφαρμογή κρίνεται με βάση κάποιες ιδιότητες του συστήματος [28, 51, 156]. Μια ιδιότητα που πρέπει να έχει ένα σύστημα υδατογράφησης είναι η *πιστότητα* (*fidelity*) [26, 110], η οποία ορίζεται σαν την ομοιότητα ανάμεσα στα μη υδατογραφημένα και στα υδατογραφημένα δεδομένα. Στις πιο πολλές εφαρμογές, το υδατογράφημα πρέπει να εισαχθεί στα δεδομένα με τέτοιο τρόπο, έτσι ώστε να μην επηρεάζεται η ποιότητά τους. Μια δεύτερη ιδιότητα που χαρακτηρίζει ένα σύστημα υδατογράφησης είναι το *ωφέλιμο φορτίο* (*payload*), το οποίο είναι το ποσό της πληροφορίας (σε bits) που μπορεί να αποθηκευτεί σε ένα υδατογράφημα και το οποίο ποικίλει ανάλογα με την εφαρμογή [77]. Η *ανθεκτικότητα* (*robustness*) είναι μια άλλη πολύ σημαντική ιδιότητα. Το υδατογράφημα πρέπει να παραμένει στα δεδομένα ακόμα κι αν η ποιότητα των δεδομένων υποβαθμισθεί μετά από κάποια επεξεργασία, σκόπιμη ή μη. Μια άλλη ιδιότητα είναι η *ασφάλεια* (*security*), η οποία δηλώνει την ικανότητα του υδατογραφήματος να παραμένει στα δεδομένα ύστερα από εχθρικές επιθέσεις που έχουν ως σκοπό την ανίχνευσή του και την αφαίρεσή του από τα δεδομένα

[131]. Ένα επιπλέον χαρακτηριστικό των συστημάτων υδατογράφησης είναι το πώς γίνεται η ανίχνευση του υδατογραφήματος. Ο ανιχνευτής μπορεί να είναι “πληροφορημένος” (*informed detector*), να έχει δηλαδή γνώση των αρχικών δεδομένων ή “τυφλός” (*blind detector*), να μην έχει δηλαδή κάποια πληροφορία για τα αρχικά δεδομένα. Κάποιες επιπρόσθετες ιδιότητες ενός συστήματος υδατογράφησης είναι η *αποτελεσματικότητα της ένθεσης*, η πιθανότητα δηλαδή η ένθεση του υδατογραφήματος να γίνει με επιτυχία, η *πιθανότητα εσφαλμένης ανίχνευσης*, η πιθανότητα δηλαδή να ανιχνευθεί υδατογράφημα σε δεδομένα τα οποία δεν έχουν υδατογραφηθεί, η *χρήση κρυπτογραφικών κλειδιών* για τον έλεγχο της απόκρυψης κάποιου μηνύματος ή *κλειδιών υδατογράφησης* για τον έλεγχο της ένθεσης και της απόκρυψης (*cipher, watermark keys*) [28], η πιθανότητα αλλαγής του υδατογραφήματος που έχει εισαχθεί στα δεδομένα ή η πιθανότητα ένθεσης πολλαπλών υδατογραφημάτων και τέλος το υπολογιστικό κόστος της ένθεσης και της ανίχνευσης.

Οι παραπάνω ιδιότητες δεν είναι ανεξάρτητες μεταξύ τους. Για παράδειγμα, ένα πολύ ανθεκτικό υδατογράφημα μπορεί να προκύψει αν τροποποιήσουμε σε μεγάλο βαθμό τα δεδομένα. Αυτό όμως θα έχει επίπτωση στην πιστότητα των υδατογραφημένων δεδομένων, καθώς οι τροποποιήσεις θα γίνονται αντιληπτές. Έτσι, για κάθε εφαρμογή πρέπει να βρίσκουμε την καλύτερη δυνατή σχέση ανάμεσα στις ιδιότητες που πρέπει να έχει το σύστημα υδατογράφησης για να ικανοποιεί τις απαιτήσεις της εφαρμογής.

Προκειμένου να αξιολογηθεί ένα σύστημα υδατογράφησης με δίκαιο αλλά και αυτόμα-το τρόπο, χρειάζεται να δοκιμαστούν οι επιδόσεις του με βάση κάποιο σημείο αναφοράς (*benchmarking*). Η πρώτη ερευνητική δουλειά προς αυτή την κατεύθυνση είχε σαν αποτέλεσμα το StirMark [61, 105]. Το StirMark εισήγαγε τις τυχαίες διγραμμικές γεωμετρικές παραμορφώσεις σαν μια καινούρια επίθεση κατά της υδατογραφημένης εικόνας. Ένα άλλο σύστημα για benchmarking είναι το OPTIMARK [127], το οποίο παρέχει στατιστικά χαρακτηριστικά για το λογισμικό υδατογράφησης που ελέγχει. Το CHECKMARK [60, 101] είναι ένα τρίτο σύστημα, το οποίο βρίσκεται πολύ κοντά στο Ευρωπαϊκό σύστημα CERTIMARK. Περιέχει όλες τις επιθέσεις του StirMark 3.1 και ένα μεγάλο αριθμό από επιπρόσθετες επιθέσεις. Για κάθε επίθεση ελέγχει αν ένα μήνυμα έχει αποκωδικοποιηθεί σωστά ή όχι. Το CERTIMARK [59] έχει χρηματοδοτηθεί από την Ευρωπαϊκή Ένωση και περιλαμβάνει όλα τα χαρακτηριστικά του CHECKMARK και του OPTIMARK. Τα αποτελέσματά του όμως δεν είναι διαθέσιμα μέσω του διαδικτύου, ούτε ο κώδικάς του είναι ελεύθερα προσβάσιμος (*open source*). Μέχρι τώρα, κανένα benchmark δεν έχει αποτελέσει μια οριστική και καθολική λύση στο πρόβλημα της αξιολόγησης των συστημάτων υδατογράφησης, για προφανείς λόγους, και χρειάζεται πολύ δουλειά ακόμα προς αυτή την κατεύθυνση [88].

1.3.1 Μοντέλα υδατογράφησης και επικοινωνίες

Η υδατογράφηση είναι υπό μία έννοια μια μορφή επικοινωνίας. Ένα μήνυμα μεταδίδεται από το σημείο ένθεσης του υδατογραφήματος (*watermark embedder*) στο δέκτη του υδατογραφήματος (*watermark receiver*). Θα περιγράψουμε τρία μοντέλα συστημάτων υδατογράφησης, τα οποία βασίζονται στο κλασικό μοντέλο καναλιού επικοινωνίας (*communication channel*) [28, 29]. Η διαφορά ανάμεσα στα τρία μοντέλα καθορίζεται από τον τρόπο με τον οποίο ενσωματώνονται τα αρχικά δεδομένα, η αρχική εικόνα στην περίπτωση μας, στο μοντέλο επικοινωνίας.

Στο βασικό μοντέλο, η αρχική εικόνα θεωρείται σαν θόρυβος που προστίθεται κατά τη διάρκεια μετάδοσης του σήματος υδατογράφησης. Η διαδικασία ένθεσης του υδατογραφήματος αποτελείται από δύο βήματα. Καταρχήν, το μήνυμα αντιστοιχίζεται με κάποιο τρόπο στο υδατογράφημα που θα προστεθεί στην εικόνα. Για παράδειγμα, ο κωδικοποιητής του υδατογραφήματος πρέπει να παράγει ένα σήμα που να έχει ίδιες διαστάσεις με την αρχική εικόνα. Στη συνέχεια, το σήμα αυτό προστίθεται στην αρχική εικόνα και έτσι παράγεται η υδατογραφημένη εικόνα. Επειδή κατά τη διάρκεια της ένθεσης δε λαμβάνεται υπόψη η αρχική εικόνα, λέμε ότι η ένθεση είναι “τυφλή” (*blind embedding*). Μετά την ένθεση του υδατογραφήματος, η υδατογραφημένη εικόνα μπορεί να υποστεί κάποιου είδους επεξεργασία ή κάποια επίθεση. Η επίδραση αυτής της επεξεργασίας ή της επίθεσης μπορεί να μοντελοποιηθεί σαν πρόσθεση θορύβου. Αν χρησιμοποιηθεί ένας “πληροφορημένος” ανιχνευτής, η διαδικασία της ανίχνευσης γίνεται σε δύο βήματα. Πρώτα αφαιρείται η αρχική εικόνα από την εικόνα που φτάνει στο δέκτη και στη συνέχεια το ενθόρυβο υδατογράφημα που απομένει αποκωδικοποιείται με τη βοήθεια ενός κλειδιού υδατογράφησης. Επειδή η πρόσθεση της αρχικής εικόνας στη φάση της ένθεσης ακυρώνεται με την αφαίρεσή της στον ανιχνευτή, η μόνη διαφορά ανάμεσα στο αρχικό υδατογράφημα και σε αυτό που φτάνει στο δέκτη είναι ο θόρυβος. Αυτό σημαίνει ότι η πρόσθεση της αρχικής εικόνας μπορεί να αγνοηθεί και ότι ο κωδικοποιητής του υδατογραφήματος, η διαδικασία θορύβου και ο ανιχνευτής αποτελούν ένα σύστημα εντελώς αντίστοιχο με ένα επικοινωνιακό σύστημα. Αν ο ανιχνευτής είναι “τυφλός” (*blind detector*), τότε η αρχική εικόνα δεν είναι διαθέσιμη και δεν μπορεί να αφαιρεθεί από την εικόνα που φτάνει στο δέκτη πριν από τη φάση της αποκωδικοποίησης. Σε αυτή την περίπτωση η λαμβανόμενη εικόνα θεωρείται να είναι μια αλλοιωμένη έκδοση του υδατογραφήματος και ο ανιχνευτής του υδατογραφήματος θεωρείται ότι είναι ο αποκωδικοποιητής καναλιού (*channel decoder*). Τα μοντέλα υδατογράφησης αυτού του είδους μπορούν να χρησιμοποιηθούν σε εφαρμογές που απαιτούν ανθεκτικότητα, όπου το ζητούμενο είναι να μεγιστοποιηθεί η πιθανότητα ανίχνευσης του σωστού υδατογραφήματος. Δεν μπορούν όμως να χρησιμοποιηθούν για τον έλεγχο

αυθεντικότητας, γιατί εκεί ο στόχος δεν είναι η ανίχνευση του υδατογραφήματος, αλλά το να γίνει γνωστό αν η υδατογραφημένη εικόνα έχει τροποποιηθεί και με ποιο τρόπο.

Ένα άλλο μοντέλο είναι η υδατογράφηση σαν επικοινωνία με *επιπρόσθετη πληροφορία* (*side information*) στον πομπό. Δεδομένου ότι η αρχική εικόνα είναι γνωστή στη φάση της ένθεσης, δεν υπάρχει κανένας λόγος να επιβάλλουμε τον περιορισμό το υδατογράφημα που θα προστεθεί να είναι ανεξάρτητο από την αρχική εικόνα, όπως γίνεται στο βασικό μοντέλο υδατογράφησης. Πιο αποτελεσματικοί αλγόριθμοι υδατογράφησης μπορούν να προκύψουν αν ληφθεί υπόψη η πληροφορία της αρχικής εικόνας πριν την κωδικοποίηση του υδατογραφήματος. Αν συνεχίσουμε να θεωρούμε την αρχική εικόνα σαν κομμάτι της διαδικασίας θορύβου στο κανάλι μετάδοσης, τότε το μοντέλο αυτό αντιστοιχεί σε ένα επικοινωνιακό σύστημα με επιπρόσθετη πληροφορία στον πομπό, [123]. Με άλλα λόγια, υπάρχει κάποια γνώση για τον θόρυβο του καναλιού, η οποία μπορεί να χρησιμοποιηθεί κατά τη διαδικασία της ένθεσης. Οι επικοινωνίες με επιπρόσθετη πληροφορία έχουν μελετηθεί από αρκετούς ερευνητές [25, 53] και αποτελέσματά τους έχουν αρχίσει να εφαρμόζονται στο πρόβλημα της υδατογράφησης [19, 20, 29]. Θα δούμε περισσότερες λεπτομέρειες για το θέμα αυτό στην ενότητα 1.3.3.

Στο τρίτο μοντέλο υδατογράφησης η αρχική εικόνα δεν θεωρείται πλέον τμήμα του καναλιού μετάδοσης, αλλά ένα δεύτερο μήνυμα που μεταδίδεται μαζί με το υδατογράφημα στο ίδιο σήμα. Τα δύο *πολυπλεγμένα* (*multiplexed*) μηνύματα, η αρχική εικόνα και το υδατογράφημα, ανιχνεύονται και αποκωδικοποιούνται από δύο διαφορετικούς δέκτες, έναν άνθρωπο και έναν ανιχνευτή υδατογραφήματος, αντίστοιχα. Ο συνδυασμός της αρχικής εικόνας και του υδατογραφήματος σε ένα σήμα είναι παρόμοιος με τη μετάδοση πολλαπλών μηνυμάτων πάνω από μια απλή γραμμή στις επικοινωνίες, όπου γίνεται χρήση *πολυπλεξίας με διαίρεση χρόνου, συχνότητας ή κώδικα* (*time-division, frequency-division, code-division multiplexing*). Μια διαφορά είναι ότι στις επικοινωνίες χρησιμοποιείται η ίδια βασική τεχνολογία για τα διαφορετικά μηνύματα, τα οποία διαχωρίζονται από μια απλή παράμετρο (χρόνος, συχνότητα, κώδικας). Αντίθετα, στην υδατογράφηση τα δύο μηνύματα διαχωρίζονται από διαφορετικές τεχνολογίες. Όταν το σήμα περνά από το κανάλι μετάδοσης, καταλήγει σε ένα σύστημα ανθρώπινης όρασης ή σε έναν ανιχνευτή υδατογραφήματος. Ο άνθρωπος πρέπει να δει κάτι που να μοιάζει πολύ με την αρχική εικόνα, χωρίς *παρεμβολή* (*interference*) από το υδατογράφημα. Ο ανιχνευτής πρέπει να ανακτήσει το μήνυμα υδατογράφησης, χωρίς παρεμβολή από την αρχική εικόνα. Αν παρουσιάζονται προβλήματα πιστότητας, το σύστημα υδατογράφησης μπορεί να βελτιωθεί χρησιμοποιώντας ένα *αντιληπτικό* (*perceptual*) μοντέλο και ρυθμίζοντας κατάλληλα την ενέργεια του υδατογραφήματος. Περισσότερες λεπτομέρειες για τη χρήση αντιληπτικών μοντέλων θα δούμε στην ενότητα 1.3.4.

1.3.2 Υδατογράφηση βασισμένη στη συσχέτιση

Στην προηγούμενη ενότητα είδαμε κάποια γενικά μοντέλα υδατογράφησης που μπορούν να χρησιμοποιηθούν για κάθε σύστημα υδατογράφησης. Ένας μεγάλος αριθμός τεχνικών υδατογράφησης, η πλειοψηφία ίσως, χρησιμοποιεί κάποιου είδους συσχέτιση στη φάση της ανίχνευσης. Θα μπορούσαμε λοιπόν να πούμε ότι μια βασική κατηγορία τεχνικών υδατογράφησης είναι αυτές που βασίζονται στη συσχέτιση.

Υπάρχουν αρκετοί τύποι συσχέτισης που χρησιμοποιούνται στην ανίχνευση υδατογραφήματος. Ο πιο βασικός είναι η γραμμική συσχέτιση (*linear correlation*). Κάποιοι άλλοι τύποι συσχέτισης διακρίνονται με βάση την κανονικοποίηση που γίνεται στα διανύσματα πριν τον υπολογισμό του εσωτερικού γινομένου τους. Αν κανονικοποιήσουμε τα δύο διανύσματα έτσι ώστε τα μέτρα τους να είναι ίσα με τη μονάδα, τότε υπολογίζουμε την *κανονικοποιημένη συσχέτισή τους* (*normalized correlation*). Αν αφαιρέσουμε τις μέσες τιμές των διανυσμάτων πριν τον υπολογισμό της κανονικοποιημένης συσχέτισης, τότε υπολογίζουμε το *συντελεστή συσχέτισής τους* (*correlation coefficient*).

Σε μερικά συστήματα, οι αλγόριθμοι υδατογράφησης χτίζονται γύρω από τη χρήση τέτοιων μετρικών συσχέτισης. Σε άλλα συστήματα, η χρήση της συσχέτισης μπορεί να είναι λιγότερο προφανής ή η συσχέτιση να αποτελεί ένα μικρό τμήμα του αλγορίθμου ανίχνευσης. Μπορεί επίσης, κάποια συστήματα υδατογράφησης να χρησιμοποιούν κάποια τεχνική ανίχνευσης η οποία να είναι μαθηματικά ισοδύναμη με τη χρήση κάποιου τύπου συσχέτισης. Θα δούμε στη συνέχεια τις μετρικές συσχέτισης και κάποιες από τις αντίστοιχες ισοδύναμες τεχνικές ανίχνευσης [28].

Η γραμμική συσχέτιση ανάμεσα σε δύο διανύσματα υπολογίζεται από το εσωτερικό τους γινόμενο το οποίο διαιρείται με το πλήθος των στοιχείων του διανύσματος. Στις επικοινωνίες ο υπολογισμός της γραμμικής συσχέτισης ανάμεσα στο μεταδιδόμενο και στο λαμβανόμενο σήμα και η σύγκρισή της με ένα κατώφλι είναι συνηθισμένη τακτική, προκειμένου να ελεγχθεί η παρουσία του μεταδιδόμενου σήματος. Αυτή η διαδικασία είναι γνωστή σαν *matched filtering* και θεωρείται να είναι η βέλτιστη τεχνική ανίχνευσης σημάτων στην παρουσία λευκού, προσθετικού, Gaussian θορύβου με την εξής έννοια: για κάθε δεδομένη *πιθανότητα λανθασμένης ανίχνευσης* (*false positive probability*), το *matched filtering* ελαχιστοποιεί την πιθανότητα να μην ανιχνευθεί ένα υπάρχον σήμα (*false negative*). Αυτό μπορεί να αποδειχτεί, δείχνοντας ότι το *matched filtering* είναι ισοδύναμο με το κριτήριο Neyman-Pearson (Neyman-Pearson hypothesis testing) [69].

Γενικά, κάθε αλγόριθμος ανίχνευσης που υπολογίζει μια γραμμική συνάρτηση των δειγμάτων και συγκρίνει αυτή τη συνάρτηση με ένα κατώφλι μπορεί να θεωρηθεί ότι χρησιμοποιεί

γραμμική συσχέτιση. Ένα παράδειγμα ανίχνευσης ισοδύναμης με τη γραμμική συσχέτιση είναι το ακόλουθο: τα εικονοστοιχεία μιας εικόνας χωρίζονται σε δύο ομάδες και το άθροισμα των εικονοστοιχείων της μιας ομάδας αφαιρείται από το άθροισμα των εικονοστοιχείων της άλλης. Το αποτέλεσμα συγκρίνεται με ένα κατώφλι, προκειμένου να καθοριστεί αν το υδατογράφημα είναι παρόν ή όχι. Αυτή η διαδικασία είναι παρόμοια με τον υπολογισμό της συσχέτισης ανάμεσα στην εικόνα και ένα υδατογράφημα που αποτελείται από 1 και -1 .

Ένα λιγότερο προφανές παράδειγμα, υλοποιείται στο πεδίο του μετασχηματισμού DCT. Υπολογίζεται ο μετασχηματισμός DCT για κάθε τμήμα της εικόνας διαστάσεων 8×8 και επιλεγμένοι συντελεστές ομαδοποιούνται σαν ζεύγη σε διάταξη. Η βασική ιδέα είναι να κωδικοποιηθεί ένα bit πληροφορίας σε κάθε ζευγάρι, σύμφωνα με το αν ο πρώτος συντελεστής στο ζευγάρι είναι μεγαλύτερος από το δεύτερο. Για παράδειγμα, αν ο πρώτος συντελεστής είναι μεγαλύτερος θα κωδικοποιηθεί ένα bit 1, ενώ αν είναι μικρότερος θα κωδικοποιηθεί ένα bit 0. Για να υλοποιηθεί ένας τέτοιος αλγόριθμος με γραμμική συσχέτιση, ορίζεται ένα πρότυπο (*pattern*) για κάθε bit. Όλοι οι συντελεστές του DCT είναι 0, εκτός από το ζευγάρι των συντελεστών που χρησιμοποιούνται για την κωδικοποίηση του bit. Ο πρώτος συντελεστής του ζεύγους θα είναι ίσος με 1 και ο δεύτερος ίσος με -1 . Έτσι αν υπολογίσουμε τη συσχέτιση ανάμεσα στο μετασχηματισμό DCT ενός τέτοιου προτύπου και στο μετασχηματισμό DCT της εικόνας, το πρόσημό της θα δείξει αν ο πρώτος συντελεστής είναι μεγαλύτερος από το δεύτερο ή όχι. Επειδή ο μετασχηματισμός DCT είναι γραμμικός, μπορούμε να πάρουμε το ίδιο αποτέλεσμα αν υπολογίσουμε τη συσχέτιση του προτύπου με την εικόνα στο χωρικό πεδίο.

Ένα από τα προβλήματα της γραμμικής συσχέτισης είναι ότι οι τιμές της ανίχνευσης εξαρτώνται από τα μέτρα των διανυσμάτων που συμμετέχουν στον υπολογισμό της. Αυτό σημαίνει ότι ένα υδατογράφημα δε θα είναι ανθεκτικό σε κάποια αλλαγή στη φωτεινότητα της εικόνας. Αυτό το πρόβλημα μπορεί να αντιμετωπιστεί κανονικοποιώντας τα διανύσματα που συμμετέχουν στον υπολογισμό της συσχέτισης έτσι ώστε να έχουν μέτρο ίσο με τη μονάδα, υπολογίζοντας δηλαδή την κανονικοποιημένη συσχέτιση. Μερικές φορές η γραμμική συσχέτιση μπορεί να χρησιμοποιηθεί σαν μετρική ανίχνευσης, αλλά το κατώφλι του ανιχνευτή πρέπει να προσαρμοστεί κατάλληλα, λαμβάνοντας υπόψη το μέτρο του εξαγόμενου υδατογραφήματος. Αυτή η διαδικασία είναι ισοδύναμη με το να χρησιμοποιηθεί η κανονικοποιημένη συσχέτιση σαν μετρική ανίχνευσης.

Αν αφαιρεθεί η μέση τιμή από τα δύο διανύσματα πριν τον υπολογισμό της κανονικοποιημένης συσχέτισής τους, τότε υπολογίζεται ο συντελεστής συσχέτισης. Ο συντελεστής συσχέτισης παρέχει ανθεκτικότητα σε αλλαγές του DC όρου, όπως για παράδειγμα η πρόσθεση μιας σταθερής έντασης σε όλα τα εικονοστοιχεία της εικόνας. Αν η τιμή της γραμμικής

συσχέτισης προσαρμόσκει κατάλληλα σε σχέση με την τυπική απόκλιση του εξαγόμενου υδατογραφήματος, τότε στην ουσία η μετρική που χρησιμοποιείται για ανίχνευση είναι ο συντελεστής συσχέτισης.

Οι μετρικές συσχέτισης μπορούν να χρησιμοποιηθούν για την ανίχνευση υδατογραφημάτων σε διάφορα πεδία. Ένας μεγάλος αριθμός τεχνικών υδατογράφησης που έχουν προταθεί στη βιβλιογραφία υλοποιούνται στο χωρικό πεδίο. Ο πιο άμεσος τρόπος να προστεθεί ένα υδατογράφημα σε μια εικόνα στο χωρικό πεδίο είναι να προστεθεί ένα ψευδοτυχαίο πρότυπο στις τιμές της φωτεινότητας των εικονοστοιχείων [43, 106, 155, 158]. Το πρότυπο παράγεται με τη χρήση ενός κλειδιού και οι μόνοι περιορισμοί είναι η ενέργειά του να είναι ομοιόμορφα κατανομημένη και να είναι ασυσχέτιστο με την εικόνα στην οποία πρόκειται να προστεθεί. Η ανίχνευσή του μπορεί να γίνει με τη χρήση κάποιας μετρικής συσχέτισης. Η αξιοπιστία του ανιχνευτή μπορεί να βελτιωθεί αν πριν από τον υπολογισμό της συσχέτισης εφαρμοστεί *matched filtering* [33]. Το ωφέλιμο φορτίο μιας τεχνικής υδατογράφησης μπορεί εύκολα να αυξηθεί, μπορεί δηλαδή να αποθηκευτεί στην εικόνα περισσότερη πληροφορία από ένα bit. Ο πιο απλός τρόπος για να γίνει αυτό είναι να χωριστεί η εικόνα σε υποεικόνες και σε κάθε μία από αυτές να προστεθεί ένα bit υδατογραφήματος. Ένας άλλος τρόπος είναι να χρησιμοποιηθούν τεχνικές *απλωμένου φάσματος* (*spread spectrum*) [121]. Για κάθε bit υδατογραφήματος παράγεται ένα διαφορετικό στοχαστικά ανεξάρτητο πρότυπο, το οποίο έχει ίδιες διαστάσεις με την εικόνα. Το άθροισμα όλων των ανεξάρτητων προτύπων αποτελεί το υδατογράφημα που θα εισαχθεί στην εικόνα.

Αντίστοιχες τεχνικές υδατογράφησης μπορούν να εφαρμοστούν και στο πεδίο των συχνοτήτων. Για παράδειγμα, ένα υδατογράφημα μπορεί να προστεθεί στη φάση του μετασχηματισμού FFT, η οποία περιέχει τη σημαντικότερη πληροφορία για την εικόνα. Αυτό έχει σαν αποτέλεσμα το υδατογράφημα να είναι πιο ανθεκτικό, γιατί αν κάποιος προσπαθήσει να το αφαιρέσει, η ποιότητα της εικόνας θα υποβαθμιστεί [120]. Το υδατογράφημα μπορεί επίσης να εισαχθεί στο μέτρο του μετασχηματισμού FFT, το οποίο δεν επηρεάζεται από μετατοπίσεις, με αποτέλεσμα και το υδατογράφημα να είναι ανθεκτικό στις μετατοπίσεις [121]. Επιπλέον, το υδατογράφημα μπορεί να προστεθεί μόνο στη μεσαία μπάντα συχνοτήτων, γιατί η τροποποίηση των χαμηλών συχνοτήτων μπορεί να έχει επιπτώσεις στην ποιότητα της εικόνας, και οι υψηλές συχνότητες είναι ευαίσθητες στο θόρυβο, στο φιλτράρισμα και στη συμπίεση με απώλειες (*lossy compression*).

Ένα άλλο πεδίο το οποίο χρησιμοποιείται αρκετά για την ένθεση υδατογραφημάτων είναι το πεδίο του μετασχηματισμού DCT [8, 27, 107, 109, 156]. Χρησιμοποιώντας το μετασχηματισμό DCT, μια εικόνα χωρίζεται εύκολα σε μπάντες συχνοτήτων, με αποτέλεσμα να διευκολύνεται η ένθεση του υδατογραφήματος στις μεσαίες συχνότητες. Επιπλέον, έχει μελετηθεί εκτενώς

η ευαισθησία της ανθρώπινης όρασης στο μετασχηματισμό DCT και έχει προκύψει ο πίνακας κβαντισμού JPEG, ο οποίος μπορεί να χρησιμοποιηθεί για την ελαχιστοποίηση των πιθανών αλλοιώσεων που μπορεί να προκαλέσει στην εικόνα η ένθεση του υδατογραφήματος. Ένα ακόμα πλεονέκτημα του μετασχηματισμού DCT είναι ότι χρησιμοποιείται για τη συμπίεση της εικόνας. Έτσι, μπορούμε να εισάγουμε ένα υδατογράφημα ανθεκτικό στη συμπίεση, αφού γνωρίζουμε ποιοι είναι οι συντελεστές που θα απορριφθούν. Η απόδοση των τεχνικών υδατογράφησης που βασίζονται στη συσχέτιση και υλοποιούνται στο πεδίο του DCT μπορεί να βελτιωθεί αν χρησιμοποιηθούν κατάλληλοι ανιχνευτές, οι οποίοι υποθέτουν ότι το αποτέλεσμα της συσχέτισης ακολουθεί μια γενικευμένη *Gaussian* (*generalized Gaussian*) κατανομή, αντί για μια καθαρά *Gaussian* κατανομή [54].

Οι τεχνικές υδατογράφησης εικόνων μπορούν να εκμεταλλευτούν τα χαρακτηριστικά της ανθρώπινης όρασης, έτσι ώστε να είναι δυνατή η ένθεση υδατογραφημάτων με μεγαλύτερη ενέργεια, τα οποία θα είναι πιο ανθεκτικά σε επιθέσεις. Από αυτή την άποψη, ο διακριτός μετασχηματισμός κυματιδίου (*DWT-Discrete Wavelet Transform*) είναι ιδιαίτερα ελκυστικός, γιατί μπορεί να χρησιμοποιηθεί σαν μια υπολογιστικά αποδοτική εκδοχή των συχνοτικών μοντέλων του συστήματος ανθρώπινης όρασης (*HVS-Human Visual System*) [9, 10]. Επίσης, ο μετασχηματισμός DWT χρησιμοποιείται για συμπίεση στο JPEG2000 και η ένθεση υδατογραφήματος στο πεδίο του DWT μπορεί να εξασφαλίσει ανθεκτικότητα στη συμπίεση. Σε αρκετές τεχνικές που έχουν προταθεί, οι βασικές προσεγγίσεις στην υδατογράφηση εφαρμόζονται στις μπάντες υψηλής ανάλυσης του μετασχηματισμού DWT [74, 157].

1.3.3 Υδατογράφηση με επιπρόσθετη πληροφορία

Όπως αναφέραμε στην ενότητα 1.3.1, ένα μοντέλο υδατογράφησης είναι αυτό που χρησιμοποιεί επιπρόσθετη πληροφορία στον πομπό, την πληροφορία της αρχικής εικόνας η οποία δε θεωρείται πλέον άγνωστος θόρυβος. Η πληροφορία αυτή μπορεί να χρησιμοποιηθεί μετά την παραγωγή του μηνύματος και τότε μιλάμε για “πληροφορημένη” ένθεση (*informed embedding*). Αυτή η διαδικασία μπορεί να αντιμετωπιστεί σαν ένα πρόβλημα βελτιστοποίησης, όπου πρέπει να μεγιστοποιηθεί η ανθεκτικότητα ενώ διατηρείται σταθερή η πιστότητα ή να μεγιστοποιηθεί η πιστότητα ενώ παραμένει σταθερή η ανθεκτικότητα. Αν η επιπρόσθετη πληροφορία χρησιμοποιηθεί κατά τη διάρκεια επιλογής του μηνύματος που θα εισαχθεί στην εικόνα, τότε μιλάμε για “πληροφορημένη” κωδικοποίηση (*informed coding*). Αυτό μπορεί να γίνει ορίζοντας έναν “κώδικα βρώμικου χαρτιού” (*dirty-paper code*) στον οποίο κάθε μήνυμα μπορεί να αναπαρασταθεί από πολλά εναλλακτικά διανύσματα. Από αυτά τα διανύσματα επιλέγεται εκείνο που αναπαριστά το επιθυμητό μήνυμα και είναι πιο κοντινό στην αρχική εικόνα. Συνδυάζοντας την “πληροφορημένη” κωδικοποίηση με την “πληροφορημένη” ένθεση είναι δυνατό

να έχουμε τεχνικές υδατογράφησης με σημαντικά βελτιωμένη απόδοση.

Αρχικά, η “πληροφορημένη” ένθεση χρησιμοποίησε αντιληπτικά μοντέλα, προκειμένου να προσαρμοστεί κατάλληλα η ενέργεια του υδατογραφήματος τοπικά και να διατηρηθεί σε καλά επίπεδα η ποιότητα της εικόνας μετά την πρόσθεσή του. Κάποια τέτοια μοντέλα θα παρουσιάσουμε στην επόμενη ενότητα. Ωστόσο, στα πρώτα συστήματα αυτής της μορφής δε γίνεται κάποια σαφής θεώρηση για τη φάση της ανίχνευσης και η τοπική προσαρμογή που γίνεται στις τιμές του υδατογραφήματος θεωρείται σαν μια μορφή θορύβου. Ο σχεδιασμός της “πληροφορημένης” ένθεσης όμως μπορεί να θεωρηθεί σαν ένα πρόβλημα βελτιστοποίησης. Η ενέργεια του υδατογραφήματος μπορεί να καθοριστεί με βάση τη μετρική ανίχνευσης ή μπορεί να οριστεί μια διαφορετική μετρική, η οποία να προσπαθεί να προβλέψει την ανθεκτικότητα του υδατογραφήματος. Η πρώτη προσέγγιση είναι κατάλληλη για απλά συστήματα που βασίζονται στη γραμμική συσχέτιση, ενώ για συστήματα που χρησιμοποιούν την κανονικοποιημένη συσχέτιση απαιτείται η δεύτερη προσέγγιση.

Γενικά, οι αλγόριθμοι “πληροφορημένης” ένθεσης έχουν καλύτερη απόδοση από τους αντίστοιχους αλγορίθμους “τυφλής” ένθεσης. Επιπλέον, η πληροφορία της αρχικής εικόνας μπορεί να χρησιμοποιηθεί και στη φάση της κωδικοποίησης, προκειμένου να επιλεγεί εκείνο το διάνυμα που θα προκαλέσει τη λιγότερη αλλοίωση στην εικόνα που θα εισαχθεί. Συνδυάζοντας αυτή τη στρατηγική με την “πληροφορημένη” ένθεση, είναι πιθανό να εξασφαλιστεί σημαντικά καλύτερη απόδοση.

Η “πληροφορημένη” κωδικοποίηση βασίζεται στο θεωρητικό αποτέλεσμα του Costa, στην εργασία “Writing on Dirty Paper” [25]. Το αποτέλεσμα αυτό αφορά τη χωρητικότητα ενός Gaussian καναλιού, όπου υπάρχει παρεμβολή (*interference*) η οποία είναι γνωστή στον πομπό. Το σήμα που φτάνει στο δέκτη, y , είναι

$$y = s + i + n^1 \quad (1.3.1)$$

όπου $i \sim N(0, Q)$ είναι η παρεμβολή, $n \sim N(0, N)$ είναι Gaussian θόρυβος και το σήμα s χρησιμοποιείται για τη μετάδοση μιας κωδικής λέξης u . Αποδεικνύεται ότι η χωρητικότητα αυτού του συστήματος είναι ίδια με την περίπτωση που δεν υπάρχει παρεμβολή. Αν το σήμα s που στέλνει ο πομπός έχει ένα περιορισμό στην ισχύ του, $|s|^2 \leq P$, τότε η χωρητικότητα του είναι

$$C = \frac{1}{2} \log \left(1 + \frac{P}{N} \right) \quad (1.3.2)$$

¹Πρέπει να σημειώσουμε ότι ο συμβολισμός που χρησιμοποιούμε σχετικά με την υδατογράφηση εικόνων δε συμπίπτει με αυτόν της αντιστοίχισης, προκειμένου να συμβαδίζουμε με τη βιβλιογραφία της κάθε περιοχής. Για παράδειγμα, στην αντιστοίχιση εικόνων τα x, y χρησιμοποιούνται για μετατοπίσεις ανάμεσα στις συγκρινόμενες εικόνες, ενώ στην υδατογράφηση χρησιμοποιούνται για την αρχική και την υδατογραφημένη (ή τη λαμβανόμενη) εικόνα, αντίστοιχα.

ανεξάρτητα από την τιμή του Q . Ο τίτλος της εργασίας προκύπτει από την αναλογία με το πρόβλημα της γραφής σε ένα βρώμικο χαρτί, όπου ο αναγνώστης δεν μπορεί να διακρίνει τη βρωμιά από το μελάνι. Στην περίπτωση της υδατογράφησης, το βρώμικο χαρτί (η γνωστή παρεμβολή) είναι η εικόνα, το μεταδιδόμενο σήμα είναι το υδατογράφημα και ο περιορισμός στην ισχύ του εκφράζει την απαίτηση να διατηρηθεί η ποιότητα της εικόνας. Ο θόρυβος μπορεί να προέλθει από κάποια επεξεργασία της εικόνας ή από κάποια επίθεση με σκοπό την αφαίρεση του υδατογραφήματος.

Στην πράξη ωστόσο υπάρχουν αρκετές διαφορές ανάμεσα στο κανάλι που περιγράφει ο Costa και σε ένα σύστημα υδατογράφησης. Στην υδατογράφηση, η αρχική εικόνα (παρεμβολή) και ο θόρυβος σπάνια έχουν Gaussian κατανομή. Επιπλέον, εκφράζοντας τον περιορισμό που αφορά την πιστότητα σαν ένα περιορισμό ισχύος, σημαίνει ότι για τον υπολογισμό των αλλοιώσεων χρησιμοποιείται το μέσο τετραγωνικό σφάλμα, το οποίο όμως δεν είναι ένα καλό μοντέλο για την περιγραφή των αλλοιώσεων. Στην εργασία [95] έχουν μελετηθεί πιο κατάλληλα κανάλια για το πρόβλημα της υδατογράφησης. Ο θόρυβος αντικαθίσταται από μια διαδικασία αφαίρεσης του υδατογραφήματος. Για αυτό το κανάλι, αν η αρχική εικόνα έχει Gaussian κατανομή και ο περιορισμός πιστότητας εκφραστεί σαν περιορισμός ισχύος, το αποτέλεσμα του Costa συνεχίζει να ισχύει. Αυτό σημαίνει ότι η χωρητικότητα του καναλιού είναι ανεξάρτητη από τη διασπορά της αρχικής εικόνας. Για το ίδιο κανάλι, αν η κατανομή της αρχικής εικόνας δεν είναι Gaussian, το αποτέλεσμα του Costa είναι κατά προσέγγιση αληθές, δεδομένου ότι το όριο που τίθεται στην ισχύ είναι μικρό σε σχέση με τη διασπορά της αρχικής εικόνας. Επειδή συνήθως οι περιορισμοί που αφορούν την πιστότητα εξαρτώνται από τη μεταβλητότητα της εικόνας, φαίνεται ότι το αποτέλεσμα του Costa μπορεί να ισχύει για πραγματικά συστήματα υδατογράφησης.

Η βασική ιδέα στη χρησιμοποίηση της πληροφορίας της αρχικής εικόνας κατά τη διάρκεια της κωδικοποίησης του μηνύματος είναι μια κλάση κωδίκων, οι οποίοι είναι γνωστοί σαν *dirty-paper codes*. Σε αντίθεση με τη συνηθισμένη κωδικοποίηση, όπου σε κάθε μήνυμα ανατίθεται μια μοναδική κωδική λέξη, οι κώδικες αυτοί αναθέτουν αρκετές εναλλακτικές κωδικές λέξεις σε κάθε μήνυμα και επιλέγεται τελικά εκείνη η κωδική λέξη, η οποία είναι πιο κοντινή στην αρχική εικόνα. Στην περίπτωση που το ωφέλιμο φορτίο που θέλουμε να κωδικοποιήσουμε είναι αρκετό, απαιτούνται δομημένοι *dirty-paper* κώδικες, όπου οι κωδικές λέξεις ταξινομούνται με τέτοιο τρόπο ώστε να είναι δυνατό το γρήγορο ψάξιμο κοντινότερου γείτονα. Ένα παράδειγμα τέτοιων κωδίκων είναι οι *κώδικες πλέγματος* (*lattice codes*), όπου οι κωδικές λέξεις αναπαρίστανται σαν σημεία πάνω σε ένα συμμετρικό πλέγμα. Τα διαφορετικά μηνύματα αναπαρίστανται με *υποπλέγματα* (*sublattices*) και οι κοντινότερες κωδικές λέξεις βρίσκονται με στρογγυλοποίηση. Οι κωδικοί πλέγματος είναι εύκολα υλοποιήσιμοι, αλλά είναι ευαίσθητοι

σε αλλοιώσεις όπως η αλλαγή στην αντίθεση της εικόνας. Η ένθεση υδατογραφήματος στα λιγότερα σημαντικά bits της εικόνας είναι μια υποπερίπτωση των κωδίκων πλέγματος, ενώ κάποιοι άλλοι δομημένοι κώδικες που χρησιμοποιούνται στην “πληροφορημένη” κωδικοποίηση είναι γνωστοί σαν *syndrome codes* [28].

Η “πληροφορημένη” κωδικοποίηση, η “πληροφορημένη” ένθεση και γενικότερα οι τεχνικές υδατογράφησης που αξιοποιούν την πληροφορία της αρχικής εικόνας στον πομπό, έχουν αρχίσει να κερδίζουν έδαφος σε σχέση με τις κλασικές τεχνικές υδατογράφησης. Κάποιοι πιστεύουν ότι θα κυριαρχήσουν, κάποιοι θεωρούν ότι έχουν αρκετά μειονεκτήματα σε σχέση με τις υπάρχουσες τεχνικές με βασικότερο ίσως το μεγάλο υπολογιστικό κόστος, ενώ άλλοι υποστηρίζουν ότι σε συνδυασμό με τις υπάρχουσες τεχνικές θα οδηγήσουν σε συστήματα υδατογράφησης με βελτιωμένη απόδοση [42]. Αρκετές δουλειές εμφανίζονται στη βιβλιογραφία σχετικά με την “πληροφορημένη” κωδικοποίηση και ένθεση [20, 34, 35, 36, 37, 72, 93, 103], ενώ παράλληλα προτείνονται σημαντικές βελτιώσεις των κλασικών τεχνικών υδατογράφησης [73].

1.3.4 Χρήση αντιληπτικών μοντέλων

Τα υδατογραφήματα που εισάγονται σε μια εικόνα υποτίθεται ότι πρέπει είναι αόρατα. Αυτή η υπόθεση όμως γεννά δύο ερωτήματα. Το πρώτο είναι το πώς μπορεί να μετρηθεί αν ένα υδατογράφημα είναι αντιληπτό ή όχι και το δεύτερο ποιός είναι ο τρόπος με τον οποίο πρέπει να εισάχθει ένα υδατογράφημα σε μια εικόνα, προκειμένου να μην είναι αντιληπτό. Ένα υδατογράφημα μπορεί να γίνεται λιγότερο ή περισσότερο αντιληπτό από έναν παρατηρητή, και υπάρχουν κάποιες βασικές μεθοδολογίες με τις οποίες μπορεί να μετρηθεί η πιθανότητα το υδατογράφημα να είναι αντιληπτό. Επίσης, έχουν κατασκευαστεί διάφορα *αντιληπτικά μοντέλα* (*perceptual models*), τα οποία μπορούν να ενσωματωθούν σε ένα σύστημα υδατογράφησης και να ελέγχουν πόσο ορατό είναι το υδατογράφημα που εισάγεται στην εικόνα [110, 148, 151, 156].

Το πόσο αντιληπτό είναι ένα υδατογράφημα μπορεί να μετρηθεί με δύο τρόπους, την *πιστότητα* και την *ποιότητα*. Η πιστότητα είναι ένα μέτρο ομοιότητας ανάμεσα στην αρχική και στην υδατογραφημένη εικόνα. Η ποιότητα είναι ένα απόλυτο μέτρο, το οποίο απλά δείχνει πόσο καλά φαίνεται η εικόνα. Σε κάποιες εφαρμογές υδατογράφησης το σημαντικό μέτρο είναι η πιστότητα. Για παράδειγμα, όταν πρόκειται για κάποιο πίνακα ζωγραφικής ή για κάποια ιατρική εικόνα είναι σημαντικό η υδατογραφημένη εικόνα να είναι πιστό αντίγραφο της αρχικής. Σε άλλες εφαρμογές αυτό που είναι σημαντικό είναι απλά η υδατογραφημένη εικόνα να είναι καλής ποιότητας, αφού στις περισσότερες περιπτώσεις η αρχική εικόνα δεν είναι διαθέσιμη για να γίνει η σύγκριση μαζί της.

Σε πολλές εργασίες που εμφανίζονται στη βιβλιογραφία γίνονται εκτιμήσεις για την αντιληπτότητα του υδατογραφήματος με αυτόματο τρόπο ή με παρατηρήσεις μικρού αριθμού ανθρώπων για ελάχιστο αριθμό πειραμάτων. Αυτές οι εμπειρικές μετρήσεις δεν είναι κατάλληλες για την σωστή εκτίμηση της αντιληπτότητας ή για τη σύγκριση τεχνικών υδατογράφησης. Στις μελέτες που εμπλέκονται άνθρωποι πρέπει να λαμβάνεται υπόψη ότι οι ικανότητες της όρασης διαφέρουν σημαντικά από άνθρωπο σε άνθρωπο ή ακόμα και στον ίδιο άνθρωπο σε διαφορετικές χρονικές στιγμές. Έτσι, τέτοιου είδους μελέτες πρέπει να χρησιμοποιούν ένα μεγάλο αριθμό ανθρώπων και ένα μεγάλο αριθμό πειραμάτων. Επειδή διαφορετικοί παρατηρητές συμπεριφέρονται με διαφορετικό τρόπο, τα πειράματα είναι στην ουσία στατιστικά. Ένα επίπεδο διαφοροποιήσεων που γίνονται αντιληπτές στο 50% των πειραμάτων αναφέρεται συχνά σαν *μόλις αισθητή διαφορά (JND-Just Noticeable Difference)*.

Οι πειραματικές τεχνικές που περιλαμβάνουν την εκτίμηση πολλών ανθρώπων για μεγάλο αριθμό πειραμάτων μπορούν να δώσουν ακριβείς πληροφορίες για την πιστότητα των υδατογραφημένων δεδομένων. Ωστόσο, αυτές οι τεχνικές είναι πολύ ακριβές και δεν μπορούν να επαναληφθούν εύκολα. Μια εναλλακτική προσέγγιση είναι η χρήση μιας αλγοριθμικής μέτρησης της ποιότητας των υδατογραφημένων δεδομένων, η οποία θα βασίζεται σε κάποιο αντιληπτικό μοντέλο. Ο σκοπός είναι το μοντέλο να προβλέπει την απάντηση του παρατηρητή. Ένα τέτοιο σύστημα είναι πιο φτηνό, υλοποιείται ευκολότερα και οι εκτιμήσεις του μπορούν να επαναληφθούν, καθιστώντας το ικανό να χρησιμοποιηθεί για τη σύγκριση διαφορετικών μεθόδων.

Στην πράξη είναι πολύ δύσκολο να υπάρξει ακριβής εκτίμηση της απάντησης ενός ανθρώπου σχετικά με την πιστότητα ή την ποιότητα μιας υδατογραφημένης εικόνας. Ωστόσο, για τη σύγκριση της πιστότητας διαφορετικών τεχνικών υδατογράφησης, αρκεί το σύστημα να παρέχει μια εκτίμηση που να έχει μονότονη σχέση με τα αποτελέσματα των πειραμάτων που συμμετέχουν άνθρωποι, να εκτιμά δηλαδή τη σχετική απόδοση των διαφορετικών αλγορίθμων σε αυτά τα πειράματα. Ένα τέτοιο αντιληπτικό σύστημα ή μοντέλο είναι στην ουσία μια συνάρτηση, η οποία δίνει ένα μέτρο για την απόσταση ανάμεσα στα αρχικά και στα υδατογραφημένα δεδομένα. Η πιο απλή συνάρτηση που χρησιμοποιείται σαν μετρική απόστασης είναι το *μέσο τετραγωνικό σφάλμα (MSE-Mean Squared Error)*, το οποίο όμως δε δίνει μια καλή εκτίμηση της πιστότητας, καθώς αντιμετωπίζει τις αλλαγές σε ολόκληρη την εικόνα με τον ίδιο τρόπο.

Το σύστημα της ανθρώπινης όρασης όμως δεν αντιλαμβάνεται τις αλλαγές που μπορεί να έχει υποστεί μια εικόνα με ομοιόμορφο τρόπο. Η απόκρισή του διαφέρει ανάλογα με τη συχνότητα, τη φωτεινότητα ή το χρώμα του ερεθίσματος που λαμβάνει. Αυτό σημαίνει ότι το υδατογράφημα δε γίνεται αντιληπτό με τον ίδιο τρόπο σε όλη την επιφάνεια της εικόνας. Έτσι, στα σημεία που είναι λιγότερο ορατό, η ενέργειά του μπορεί να ενισχυθεί, προκειμένου να γίνει

πιο ανθεκτικό. Στη βιβλιογραφία έχουν προταθεί πολλές μάσκες, οι οποίες λαμβάνουν υπόψη τους τις ιδιότητες του συστήματος της ανθρώπινης όρασης. Οι μάσκες αυτές επιτρέπουν την προσαρμοστική ένθεση ενός ανθεκτικού υδατογραφήματος σε μια εικόνα, χωρίς να επηρεάζεται η ποιότητα της εικόνας.

Η ανθρώπινη όραση είναι λιγότερο ευαίσθητη σε αλλαγές τις εικόνες γύρω από ακμές ή σε περιοχές με πολλές λεπτομέρειες. Στην εργασία [148], έχει προταθεί μια μάσκα η οποία εκμεταλλεύεται τις τοπικές ιδιότητες της εικόνας. Στην εργασία [76], το υδατογράφημα προστίθεται στο μπλε κανάλι, λαμβάνοντας υπόψη το γεγονός ότι το μάτι είναι λιγότερο ευαίσθητο σε αυτό το συγκεκριμένο κανάλι. Στην [26], το υδατογράφημα προστίθεται σε μερικούς συντελεστές του μετασχηματισμού DCT από τις χαμηλές συχνότητες και οι τιμές του προσαρμόζονται κατάλληλα σε σχέση με τις τιμές των συντελεστών. Σε μια άλλη εργασία, την [110], για την ένθεση υδατογραφήματος στα πεδία μετασχηματισμών, χρησιμοποιείται το μοντέλο Watson [151]. Οι συγγραφείς του άρθρου [102] προτείνουν την ένθεση υδατογραφήματος στο πεδίο των συχνοτήτων, λαμβάνοντας υπόψη κάποιους περιορισμούς από το χωρικό πεδίο. Στην εργασία [32] χρησιμοποιείται μια μάσκα για θόρυβο περιορισμένου εύρους ζώνης (*bandlimited*) σε περιοχές με πολλές λεπτομέρειες και γύρω από περιγράμματα. Πιο πρόσφατα, οι συγγραφείς του άρθρου [78] παρουσίασαν ένα αντιληπτικό μοντέλο το οποίο περιγράφει την ευαισθησία του ανθρώπινου ματιού μέσω μιας τοπικής ισοτροπικής μετρικής αντίθεσης και μιας μάσκας. Περισσότερες αναφορές σχετικά με τη χρήση αντιληπτικών μοντέλων σε τεχνικές υδατογράφησης εικόνων μπορεί να βρει κανείς στα [28, 79, 156].

1.3.5 Ερευνητικά αποτελέσματα στα πλαίσια της διδακτορικής διατριβής

Η έρευνα που έγινε στα πλαίσια της διδακτορικής διατριβής σχετικά με το πρόβλημα της υδατογράφησης εικόνων είχε σαν στόχο την ένθεση ισχυρών υδατογραφημάτων στο χωρικό πεδίο, χωρίς να επηρεάζεται η ποιότητα της εικόνας. Επιπλέον, προτάθηκαν κατάλληλα σχήματα “τυφλής” ανίχνευσης με εξαιρετικά καλή απόδοση.

Καταρχήν, προτείναμε τη χρήση μιας χωρική μάσκας για την ένθεση υδατογραφήματος σε εικόνες, η οποία κατασκευάζεται με βάση την τοπική διασπορά του σφάλματος πρόβλεψης της εικόνας. Επιπλέον, αναπτύξαμε ένα αποδοτικό σχήμα ανίχνευσης. Το σχήμα αυτό είναι “τυφλό”, δεν απαιτεί δηλαδή τη γνώση της αρχικής εικόνας, και χρησιμοποιεί σαν μετρική ανίχνευσης την κανονικοποιημένη συσχέτιση ανάμεσα στο υδατογράφημα και το σφάλμα πρόβλεψης της εικόνας που φτάνει στο δέκτη. Η βελτίωση στην απόδοση του ανιχνευτή έχει αποδειχτεί θεωρητικά, αρχικά στην περίπτωση που υδατογράφημα και εικόνα, αλλά και διαφορετικά υδατογραφήματα μεταξύ τους, θεωρούνται εντελώς ασυσχέτιστα [62]. Στη συνέχεια,

μελετήθηκε η περίπτωση που αυτή η ιδανική υπόθεση δεν ισχύει, όπως άλλωστε συμβαίνει στην πράξη, και αποδείχτηκε θεωρητικά ότι ο προτεινόμενος ανιχνευτής έχει βελτιωμένη απόδοση, ακόμα και όταν η υδατογραφημένη εικόνα έχει υποστεί επίθεση με γραμμικό φίλτρο και θόρυβο. Επίσης, υπολογίστηκαν κατάλληλες εκφράσεις που επιτρέπουν τον υπολογισμό κατωφλίων ανίχνευσης, προσαρμοσμένων στις συνθήκες των επιθέσεων [63]. Τα παραπάνω ερευνητικά αποτελέσματα παρουσιάζονται στο κεφάλαιο 5.

Τα αποτελέσματα αυτά αποτέλεσαν τη βάση για περαιτέρω έρευνα πάνω στο ίδιο αντικείμενο, την εύρεση μιας αποδοτικής χωρικής μάσκας και του αντίστοιχου σχήματος ανίχνευσης, τα οποία παρουσιάζονται στο κεφάλαιο 6. Προτείναμε τη χρήση μιας νέας χωρικής μάσκας, η οποία επιτρέπει την ένθεση υδατογραφημάτων με εξαιρετικά μεγάλη ενέργεια, τα οποία κατά συνέπεια είναι πολύ ανθεκτικά. Η μάσκα αυτή βασίζεται στο σφάλμα πρόβλεψης της αρχικής εικόνας (και όχι στη διασπορά του σφάλματος πρόβλεψης, όπως στην προηγούμενη μάσκα που προτείναμε). Η απόδοσή της συγκρίθηκε με πολύ γνωστές και ευρέως χρησιμοποιούμενες μάσκες, [78, 148], και προέκυψε σημαντικά καλύτερη. Επιπλέον, στη φάση της ανίχνευσης του υδατογραφήματος προτάθηκε η χρήση μιας κατάλληλης διαδικασίας προ-λεύκανσης τόσο της εικόνας που φτάνει στο δέκτη, όσο και του υδατογραφήματος μετά την εφαρμογή της μάσκας, η οποία εκτιμάται στο δέκτη από τη λαμβανόμενη εικόνα. Η ανίχνευση έχει σημαντικά βελτιωμένη απόδοση, η οποία αποδεικνύεται θεωρητικά. Η θεωρητική ανάλυση είναι ανεξάρτητη της μάσκας και τα αποτελέσματά της μπορούν να χρησιμοποιηθούν σε κάθε τεχνική υδατογράφησης βασισμένη σε χωρικές μάσκες. Πρέπει να τονίσουμε όμως ότι τα πειραματικά αποτελέσματα έδειξαν ότι ο συνδυασμός του προτεινόμενου ανιχνευτή με την προτεινόμενη μάσκα οδηγεί στην καλύτερη απόδοση [65].

Η χρήση μιας αντιληπτικής μάσκας σε συνδυασμό με την επαναληπτική ένθεση του υδατογραφήματος σαν μια διαδικασία κωδικοποίησης αυξάνουν σημαντικά την ανθεκτικότητα του υδατογραφήματος. Ωστόσο, υπάρχουν κάποια μειονεκτήματα από πλευράς αποδοτικότητας της κωδικοποίησης και πρέπει να βρεθεί ο κατάλληλος συμβιβασμός ανάμεσα στις επαναλήψεις του υδατογραφήματος και στην αποδοτικότητα της κωδικοποίησης. Πιθανές λύσεις σε αυτό το πρόβλημα είναι η χρήση τεχνικών *διαφορικότητας (diversity)* ή σύνθετων σχημάτων κωδικοποίησης [147]. Εμείς προτείναμε ένα σχήμα υδατογράφησης στο χωρικό πεδίο, βασισμένο σε χωρο-χρονική κωδικοποίηση μπλοκ [66], το οποίο παρουσιάζεται στο κεφάλαιο 7. Ειδικότερα, χρησιμοποιούμε μια 4×4 πραγματική, ορθογώνια διάταξη, για την ένθεση ενός πολλαπλασιαστικού υδατογραφήματος στην εικόνα. Η εικόνα χωρίζεται σε τέσσερα τμήματα τα οποία με κατάλληλες πράξεις μπορούν να θεωρηθούν σαν τέσσερα διαφορετικά κανάλια με *επίπεδη εξασθένιση (flat fading)*. Αποδεικνύεται ότι το προτεινόμενο σχήμα έχει πολύ καλύτερη απόδοση σε σχέση με την επαναληπτική υδατογράφηση.

Κεφάλαιο 2

Μια νέα τεχνική αντιστοίχισης εικόνων βασισμένη στη συσχέτιση

Σε πολλές εφαρμογές η αντιστοίχιση των εικόνων αποτελεί βασική προϋπόθεση για περαιτέρω επεξεργασία και ανάλυση των δεδομένων. Η έρευνα που έγινε στα πλαίσια της διδακτορικής διατριβής αφορά τεχνικές αντιστοίχισης που βασίζονται στη συσχέτιση των εικόνων. Όπως αναφέρθηκε και στο πρώτο κεφάλαιο ένα μειονέκτημα της χωρικής συσχέτισης είναι ότι η περιοχή γύρω από τη μέγιστη τιμή έχει μεγάλο εύρος. Αυτό μπορεί να έχει επίπτωση στην ακρίβεια της αντιστοίχισης. Σε αυτό το κεφάλαιο προτείνεται μια νέα αποδοτική τεχνική αντιστοίχισης εικόνων, η οποία λύνει το παραπάνω πρόβλημα και είναι κατάλληλη για την αποκατάσταση μετατοπίσεων και περιστροφών. Η τεχνική αυτή αποτελείται από δύο βασικά μέρη. Καταρχήν, υπάρχει ένα στάδιο προλεύκανσης, το οποίο υλοποιείται με αποδοτικό τρόπο, και στη συνέχεια ένα επαναληπτικό σχήμα, βασισμένο στη συσχέτιση των εικόνων, το οποίο υπολογίζει τη μετατόπιση και την περιστροφή της μιας εικόνας σε σχέση με την άλλη σε λίγες επαναλήψεις [12]. Η προλεύκανση χρησιμοποιείται πριν τον υπολογισμό της συσχέτισης, προκειμένου να προκύψει πιο οξύ μέγιστο και γίνεται με το *φίλτρο πρόβλεψης σφάλματος* (*prediction error filter*), το οποίο υπολογίζεται από τα δεδομένα. Οι πράξεις ετεροσυσχέτισης που περιλαμβάνονται στο επαναληπτικό τμήμα του αλγορίθμου, εκτελούνται αποδοτικά μέσω κατάλληλης τμηματοποίησης της εικόνας και χρήσης του μετασχηματισμού Fourier. Λεπτομέρειες για τον αποδοτικό υπολογισμό της ετεροσυσχέτισης δίνονται στο κεφάλαιο 3. Τέλος, η προτεινόμενη τεχνική αντιστοίχισης εφαρμόστηκε με επιτυχία σε μια ακολουθία ιατρικών εικόνων από τον τράχηλο της μήτρας, με σκοπό τη διάγνωση δυσπλασιών και κακοηθειών [64].

2.1 Εισαγωγή

Πολλές φορές είναι απαραίτητη η σύγκριση εικόνων σημείο προς σημείο, προκειμένου να βρεθεί η μεταξύ τους αντιστοιχία. Οι διαφορετικές εικόνες μπορεί να είναι εικόνες του ίδιου αντικειμένου οι οποίες έχουν ληφθεί σε διαφορετικές χρονικές στιγμές ή από διαφορετικούς αισθητήρες. Έτσι, μπορεί να υπάρχουν κάποιες γεωμετρικές διαφοροποιήσεις ανάμεσά τους, όπως μετατόπιση, περιστροφή, κλιμάκωση και άλλες [112]. Υπάρχουν αρκετές εφαρμογές που απαιτούν αντιστοίχιση εικόνων και δύο τέτοια παραδείγματα είναι η σύγκριση εικόνων της γης, οι οποίες έχουν ληφθεί από δορυφόρους και οι ιατρικές εφαρμογές, οι οποίες στοχεύουν στη βελτίωση της διάγνωσης.

Ένας μεγάλος αριθμός τεχνικών αντιστοίχισης έχει αναπτυχθεί, για διάφορους τύπους δεδομένων και προβλημάτων [18, 136]. Μια διάκριση που μπορεί να γίνει στη βιβλιογραφία είναι οι τεχνικές που βασίζονται στη συσχέτιση των εικόνων (στο χωρικό πεδίο ή στο πεδίο των συχνοτήτων) και οι τεχνικές που βασίζονται στην αντιστοίχιση κατάλληλα επιλεγμένων χαρακτηριστικών ή μοντέλων των εικόνων. Οι διαφορετικές τεχνικές συγκρίνονται με βάση κάποια κριτήρια όπως η ακρίβεια, η εφαρμοσιμότητα, οι διάφοροι περιορισμοί στην υλοποίηση, το επίπεδο αυτοματισμού και η υπολογιστική πολυπλοκότητα. Δύο γνωστές τεχνικές αντιστοίχισης εικόνων είναι η χωρική *ετεροσυσχέτιση* (*cross-correlation*) και η *συσχέτιση φάσης* (*phase correlation*). Αν μια εικόνα έχει λευκό θόρυβο, δηλαδή θόρυβο που εξαπλώνεται σε όλες τις συχνότητες, η θέση της μέγιστης τιμής της συσχέτισης φάσης δε θα είναι ιδιαίτερα ακριβής. Σε τέτοιες περιπτώσεις προτιμάται ο υπολογισμός της χωρικής συσχέτισης [18].

Η τεχνική αντιστοίχισης που παρουσιάζεται σε αυτό το κεφάλαιο βασίζεται στη χωρική συσχέτιση. Αποκαθιστά τη μετατόπιση και την περιστροφή που μπορεί να υπάρχουν ανάμεσα σε δύο συγκρινόμενες εικόνες και αποτελείται από δύο βασικά μέρη. Στο πρώτο μέρος εφαρμόζεται στις εικόνες ένα κατάλληλο φιλτράρισμα, το οποίο αυξάνει την απόδοση της όλης τεχνικής. Ειδικότερα, υπολογίζεται ένα δισδιάστατο φίλτρο σφάλματος πρόβλεψης από την εικόνα αναφοράς. Στη συνέχεια, το φίλτρο αυτό εφαρμόζεται και στις δύο εικόνες, την αρχική και τη μετατοπισμένη, και η διαδικασία της συνέλιξης υλοποιείται στο πεδίο των συχνοτήτων χρησιμοποιώντας ένα γρήγορο αλγόριθμο. Στο δεύτερο μέρος, μια επαναληπτική διαδικασία αντιστοίχισης, βασισμένη στη χωρική συσχέτιση, εφαρμόζεται στις φιλτραρισμένες εικόνες και επιστρέφει τη συνολική μετατόπιση και περιστροφή μετά από λίγες επαναλήψεις. Το πιο “ακριβό” κομμάτι του αλγορίθμου είναι ο υπολογισμός της ετεροσυσχέτισης, ο οποίος όμως μπορεί να επιταχυνθεί, χρησιμοποιώντας ένα αποδοτικό σχήμα, το οποίο θα παρουσιαστεί στο επόμενο κεφάλαιο.

Η προτεινόμενη τεχνική αντιστοίχισης θα αναφέρεται στο εξής σαν *Επαναληπτική Ετεροσυσχέτιση με Προλεύκανση* (*Iterative Cross-Correlation with Pre-Whitening - ICC-PW*). Παραλείποντας το στάδιο της προλεύκανσης, προκύπτει μια διαφορετική έκδοση της παραπάνω τεχνικής, η οποία θα αναφέρεται σαν ICC. Η ICC έχει μικρότερο κόστος από την ICC-PW, αλλά και χειρότερη απόδοση. Η ICC-PW συγκρίνεται επιπλέον και με μια επέκταση της συσχέτισης φάσης, η οποία αποκαθιστά μετατοπίσεις και περιστροφές [135]. Η τεχνική αυτή είναι πολύ γνωστή, έχει καλή απόδοση και θα αναφέρεται στη συνέχεια σαν EPC (Enhanced Phase-Correlation). Τα πειραματικά αποτελέσματα έδειξαν ότι η ICC-PW έχει καλύτερη απόδοση και από τις ICC και EPC.

Στη συνέχεια, η ICC-PW χρησιμοποιήθηκε για την αντιστοίχιση ακολουθιών ιατρικών εικόνων με σκοπό την ανίχνευση δυσπλασιών και κακοηθειών στον τράχηλο της μήτρας [64]. Η λήψη αυτών των εικόνων έγινε με τη χρήση μιας πολυφασματικής κάμερας, η οποία περιγράφεται στην εργασία [6]. Η διάγνωση βασίζεται στην ποσοτική αποτίμηση των διαφοροποιήσεων της έντασης των εικονοστοιχείων της εξεταζόμενης περιοχής, η οποία μεταβάλλεται μετά την τοπική εφαρμογή διαλύματος οξικού οξέος.

2.2 Η προτεινόμενη τεχνική

Έστω f μια εικόνα αναφοράς και f_{tr} μια μετατοπισμένη και περιστραμμένη έκδοση της f . Ο σκοπός είναι να αποκατασταθούν οι διαφορές ανάμεσα στην f και στην f_{tr} . Τα βασικά βήματα του αλγορίθμου που επιτυγχάνει αυτή την αποκατάσταση περιγράφονται στη συνέχεια.

2.2.1 Προλεύκανση της εικόνας

Όπως έχει ήδη αναφερθεί, μια κλασική τεχνική για την αποκατάσταση μετατοπίσεων ανάμεσα σε εικόνες βασίζεται στη χωρική ετεροσυσχέτιση [112]. Η τεχνική αυτή όμως έχει κάποια μειονεκτήματα. Καταρχήν, η πιθανή παρουσία μιας δεύτερης αιχμής (η οποία εμφανίζεται κυρίως όταν στις συγκρινόμενες εικόνες υπάρχουν περιοδικότητες) δημιουργεί πρόβλημα στην επιλογή του πραγματικού μεγίστου. Επιπλέον, η περιοχή γύρω από το μέγιστο μπορεί να έχει μεγάλο εύρος με αποτέλεσμα ο ακριβής εντοπισμός της θέσης του μεγίστου να μην είναι εφικτός.

Προκειμένου να ενισχύσουμε τη μέγιστη τιμή της ετεροσυσχέτισης και να τη κάνουμε πιο ευδιάκριτη, προτείνουμε μια κατάλληλη προλεύκανση των αρχικών εικόνων. Η μέθοδος προλεύκανσης που χρησιμοποιείται βασίζεται στο φιλτράρισμα των εικόνων με ένα δισδιάστατο φίλτρο πρόβλεψης σφάλματος, το οποίο υπολογίζεται από μια κατάλληλα επιλεγμένη περιοχή της εικόνας αναφοράς. Η επίδραση που έχει το φίλτρο πάνω στις εικόνες είναι να απομακρύνει

την πλεονάζουσα πληροφορία, με αποτέλεσμα οι τελικές εικόνες να είναι λιγότερο συσχετισμένες.

Η πρόβλεψη της τιμής ενός εικονοστοιχείου $f(k, l)$ δίνεται από τη σχέση [137]

$$\tilde{f}(k, l) = \sum_i \sum_j a(i, j) f(k - i, l - j) \quad (2.2.1)$$

όπου το $a(i, j)$ αντιστοιχεί σε ένα δισδιάστατο συντελεστή γραμμικής πρόβλεψης. Για τον πίνακα γραμμικής πρόβλεψης, επιλέγεται μια μη αιτιατή περιοχή υποστήριξης, διαστάσεων $p \times p$, και οι συντελεστές που ελαχιστοποιούν τη συνάρτηση κόστους

$$J_{LP} = E\{|f(k, l) - \tilde{f}(k, l)|^2\} \quad (2.2.2)$$

υπολογίζονται από τις ακόλουθες κανονικές εξισώσεις

$$\sum_{i,j \in \Omega} a(i, j) r_{ff}(k - i, l - j) = \begin{cases} 1, & (k, l) = (0, 0) \\ 0, & \text{αλλού} \end{cases} \quad (2.2.3)$$

όπου r_{ff} είναι ο συντελεστής αυτοσυσχέτισης της δεδομένης εικόνας.

Το φίλτρο σφάλματος πρόβλεψης είναι ένας πίνακας διαστάσεων $p \times p$, ο οποίος έχει 1 στο κέντρο του και τους συντελεστές $-a(i, j)$ στις υπόλοιπες θέσεις. Το φίλτρο αυτό επιλέγεται να είναι μη αιτιατό, καθώς δεν υπάρχει φυσικό ισοδύναμο για την έννοια της αιτιατότητας στις δύο διαστάσεις. Οι συντεταγμένες των δεδομένων είναι χωρικές και οποιαδήποτε αιτιατότητα αναφέρεται σε σχέση με την εικόνα, έχει να κάνει με τον τρόπο ανάγνωσης ή λήψης της εικόνας. Εντούτοις, μπορεί να χρησιμοποιηθεί αιτιατή πρόβλεψη, χωρίς να επηρεάσει σημαντικά την απόδοση του αλγορίθμου.

Αν η διαφοροποίηση ανάμεσα στις συγκρινόμενες εικόνες ήταν απλά μετατόπιση ή περιστροφή, τότε η διαδικασία αντιστοίχισης θα είχε ως εξής. Καταρχήν, υπολογίζεται το φίλτρο πρόβλεψης a_f για την εικόνα αναφοράς f και εφαρμόζεται και στις δύο εικόνες f και f_{tr} . Οι εικόνες που προκύπτουν είναι η f' και η f'_{tr} . Στην περίπτωση που υπάρχει μόνο μετατόπιση η κλασική τεχνική ετεροσυσχέτισης εφαρμόζεται για ένα κατάλληλα επιλεγμένο παράθυρο της f' και μια περιοχή ψαξίματος της f'_{tr} . Στην περίπτωση που υπάρχει μόνο περιστροφή, ένα παράθυρο της f' περιστρέφεται σε σχέση με το αντίστοιχο παράθυρο της f'_{tr} . Μετά από κάθε περιστροφή υπολογίζεται το μέτρο ετεροσυσχέτισης ανάμεσα στα συγκρινόμενα παράθυρα. Και στις δύο περιπτώσεις η άγνωστη μετατόπιση είναι αυτή που αντιστοιχεί στο βέλτιστο ταίριασμα των εικόνων, το οποίο συμβαίνει όταν η μετρική συσχέτισης παίρνει τη μέγιστη τιμή της. Αν ανάμεσα στις συγκρινόμενες εικόνες υπάρχει και μετατόπιση και περιστροφή, τότε η αντιστοίχιση θα γίνει σύμφωνα με τον προτεινόμενο επαναληπτικό αλγόριθμο.

2.2.2 Η νέα τεχνική αντιστοίχισης

Η προτεινόμενη τεχνική αντιστοίχισης αποτελείται από τρία βήματα, τα οποία περιγράφονται στη συνέχεια. Τα δύο πρώτα εφαρμόζονται μία μόνο φορά και αποτελούν το μη επαναληπτικό τμήμα του αλγορίθμου, ενώ το τρίτο βήμα είναι αυτό που υπολογίζει με επαναληπτικό τρόπο τη μετατόπιση και την περιστροφή ανάμεσα στις συγκρινόμενες εικόνες.

1. Υπολογισμός του φίλτρου σφάλματος πρόβλεψης, a_f , για την εικόνα αναφοράς f , από τη λύση των κανονικών εξισώσεων της σχέσης (2.2.3).
2. Υπολογισμός της συνέλιξης f' και f'_{tr} , του φίλτρου a_f με τις εικόνες f και f_{tr} , αντίστοιχα, και αρχικοποιήσεις: $i = 1$, $f'_0 = f'$, $f'_{tr0} = f'_{tr}$.
3. Επαναληπτικό βήμα

3.1 Μετατόπιση

- Υπολογισμός ετεροσυσχέτισης των f'_{i-1} και f'_{tri-1} για όλες τις δυνατές μετατοπίσεις. Εύρεση της θέσης (x_i, y_i) της βέλτιστης αντιστοίχισης.
- Μετατόπιση της μιας εικόνας σε σχέση με την άλλη σύμφωνα τη (x_i, y_i) και εξαγωγή των επικαλυπτόμενων υποεικόνων, $\tilde{f}_i$ και $\tilde{f}_{tri}$.

3.2 Περιστροφή

- Περιστροφή της εικόνας $\tilde{f}_i$ κατά $-\varphi$, $-\varphi + d\varphi$, $\dots$, φ μοίρες και υπολογισμός της ετεροσυσχέτισης ανάμεσα στις κοινές περιοχές της περιστραμμένης $\tilde{f}_i$ και της $\tilde{f}_{tri}$. Εύρεση της γωνίας που δίνει τη βέλτιστη αντιστοίχιση.
- Περιστροφή της μιας εικόνας σε σχέση με την άλλη σύμφωνα με τη γωνία φ_i και εξαγωγή των επικαλυπτόμενων υποεικόνων f'_i , f'_{tri} .

3.3 Αν x_i ή y_i ή $\varphi_i \neq 0$ (κάποια από τις μετατοπίσεις ή η γωνία περιστροφής είναι διαφορετική του μηδέν), τότε τίθεται $i = i + 1$ και επαναλαμβάνεται το βήμα 3.

Το επαναληπτικό τμήμα του αλγορίθμου (δηλαδή το βήμα 3) σταματά, όταν πάψουν να ανιχνεύονται μετατοπίσεις ή περιστροφές από τη μια επανάληψη στην άλλη. Η συνολική μετατόπιση και περιστροφή μεταξύ των συγκρινόμενων εικόνων υπολογίζεται από το αλγεβρικό άθροισμα των επιμέρους μετατοπίσεων και περιστροφών που προκύπτουν σε κάθε επανάληψη. Το διάγραμμα ροής του αλγορίθμου φαίνεται στο Σχήμα 2.1.

Στο Σχήμα 2.2 φαίνεται σχηματικά το επαναληπτικό τμήμα του αλγορίθμου. Έστω f'_{i-1} και f'_{tri-1} οι εικόνες που πρόκειται να αντιστοιχηθούν στην i -οστή επανάληψη, με την f'_{i-1} να

Σχήμα 2.1: Διάγραμμα ροής του αλγορίθμου

Σχήμα 2.2: Σχηματική αναπαράσταση της επαναληπτικής διαδικασίας

είναι η εικόνα αναφοράς. Αρχικά υπολογίζεται η ετεροσυσχέτιση ανάμεσα στις f'_{i-1} και $f'_{tr,i-1}$ για κάθε δυνατή μετατόπιση. Από την εικόνα $f'_{tr,i-1}$ επιλέγεται ένα κατάλληλο παράθυρο και υπολογίζεται η ετεροσυσχέτισή του με όλα τα πιθανά παράθυρα μέσα στην περιοχή ψαξίματος της f'_{i-1} . Η θέση (x_i, y_i) για την οποία η μετρική της ετεροσυσχέτισης μεγιστοποιείται είναι η θέση της βέλτιστης αντιστοίχισης. Έτσι, η μια εικόνα μετατοπίζεται σε σχέση με την άλλη κατά (x_i, y_i) , όπως φαίνεται στο Σχήμα 2.2, και εξάγονται οι επικαλυπτόμενες υποεικόνες $\tilde{f}_i$ και $\tilde{f}_{tr,i}$. Αυτές οι εικόνες αποτελούν τις αρχικές εικόνες για το επόμενο βήμα του αλγορίθμου, που αφορά την εύρεση της περιστροφής. Σε αυτό το βήμα, η εικόνα $\tilde{f}_i$ περιστρέφεται σε σχέση με την $\tilde{f}_{tr,i}$, για διαφορετικές γωνίες, και υπολογίζεται η ετεροσυσχέτιση για τις κοινές περιοχές των δύο εικόνων. Η θέση της βέλτιστης αντιστοίχισης είναι ξανά αυτή που μεγιστοποιεί τη

μετρική της ετεροσυσχέτισης. Στη συνέχεια, η μια εικόνα περιστρέφεται σε σχέση με την άλλη σύμφωνα με την εκτιμώμενη περιστροφή, όπως φαίνεται στο Σχήμα 2.2, και εξάγονται οι επικαλυπτόμενες περιοχές τους, f'_i και f'_{tr_i} , οι οποίες θα αποτελέσουν τις αρχικές εικόνες για την επόμενη επανάληψη. Στην πράξη, ο αλγόριθμος συγκλίνει σε μικρό αριθμό επαναλήψεων.

2.3 Αποδοτικές υλοποιήσεις

Κάποια τμήματα του προτεινόμενου αλγορίθμου έχουν μεγάλο υπολογιστικό κόστος. Ωστόσο, μπορούμε να χρησιμοποιήσουμε αποδοτικούς αλγορίθμους, οι οποίοι μειώνουν αυτό το κόστος.

Στο πρώτο βήμα του αλγορίθμου υπολογίζεται το φίλτρο σφάλματος πρόβλεψης από την επίλυση ενός συστήματος κανονικών εξισώσεων (2.2.3). Χρησιμοποιώντας αντιστροφή πίνακα για την επίλυση του συστήματος, η πολυπλοκότητα είναι $O(p^6)$, όπου p^2 είναι ο αριθμός των στοιχείων του φίλτρου. Ωστόσο, χρησιμοποιώντας αλγορίθμους τύπου Levinson η πολυπλοκότητα μπορεί να μειωθεί σε $O(p^4)$ [137].

Στο δεύτερο βήμα του αλγορίθμου, στο στάδιο της προλεύκανσης, πρέπει να υπολογιστεί η συνέλιξη του φίλτρου σφάλματος πρόβλεψης με τις εικόνες f και f_{tr} . Υποθέτουμε ότι η υποπεριοχή της εικόνας f_{tr} που θα χρησιμοποιηθεί στη διαδικασία της αντιστοίχισης έχει διαστάσεις $m \times m$. Τότε, πρέπει να υπολογιστεί η συνέλιξη ανάμεσα στο φίλτρο, που έχει διαστάσεις $p \times p$, και στις περιοχές των εικόνων f και f_{tr} , που έχουν διαστάσεις $n \times n$ και $m \times m$ ($m > n$), αντίστοιχα. Οι απαιτούμενες συνέλιξεις μπορούν να υπολογιστούν αποδοτικά χρησιμοποιώντας το μετασχηματισμό Fourier και τη δισδιάστατη έκδοση της μεθόδου overlap-save [84]. Έτσι, ο υπολογιστικός φόρτος αυτού του βήματος του αλγορίθμου είναι κατά προσέγγιση $O(n^2 \log p)$.

Τέλος, το τρίτο βήμα του αλγορίθμου και ιδιαίτερα το κομμάτι που αφορά την εύρεση της μετατόπισης (εξαντλητικό ψάξιμο στις δύο διαστάσεις) έχει το μεγαλύτερο υπολογιστικό κόστος. Αν $N = m - n + 1$ είναι η μέγιστη δυνατή μετατόπιση της μιας εικόνας σε σχέση με την άλλη, τότε πρέπει να υπολογίσουμε τη συσχέτιση μιας περιοχής διαστάσεων $n \times n$ από τη μια εικόνα με μια περιοχή διαστάσεων $m \times m$ από την άλλη. Σε πολλές περιπτώσεις ισχύει ότι $N \ll n$ και για αυτό το λόγο ο συμβατικός υπολογισμός της συσχέτισης με χρήση του FFT δεν είναι αποδοτικός. Για το λόγο αυτό προτείνουμε ένα αποδοτικό σχήμα υπολογισμού της συσχέτισης, το οποίο βασίζεται σε κατάλληλη τμηματοποίηση της εικόνας και έχει πολυπλοκότητα $O(n^2 \log N)$. Το σχήμα αυτό περιγράφεται αναλυτικά στο επόμενο κεφάλαιο.

(α)

(β)

Σχήμα 2.3: (α) *Lenna* (256×256) και (β) αεροφωτογραφία (290×445)

2.4 Πειραματικά αποτελέσματα

Η απόδοση του προτεινόμενου αλγορίθμου αντιστοίχισης μελετήθηκε για μεγάλο αριθμό εικόνων. Παρουσιάζουμε εδώ τα αποτελέσματα για δύο εικόνες, τη γνωστή *Lenna* και μια αεροφωτογραφία, διαστάσεων 256×256 και 290×445 αντίστοιχα. Οι εικόνες αυτές φαίνονται στο Σχήμα 2.3. Για κάθε μια, θεωρήσαμε ένα ζευγάρι υποεικόνων, διαστάσεων 200×200 , με τη δεύτερη υποεικόνα να είναι μια μετατοπισμένη και περιστραμμένη έκδοση της πρώτης. Για κάθε μια από τις εξεταζόμενες εικόνες, κάναμε 20 διαφορετικά πειράματα και η προτεινόμενη τεχνική, η ICC-PW, συγκρίνεται με την ICC.

Προκειμένου να αποτιμήσουμε την απόδοση της τεχνικής μας χρησιμοποιήσαμε δύο κριτήρια. Το πρώτο είναι ο αριθμός των επιτυχημένων αντιστοιχίσεων (αριθμός επιτυχιών) και το δεύτερο το σφάλμα αντιστοίχισης, το οποίο ορίζεται από τη σχέση

$$D = \frac{1}{N_1 \times N_2} \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} |sf(i, j) - sf_{tr}(i, j)| \quad (2.4.1)$$

όπου sf και sf_{tr} είναι οι υποεικόνες που προκύπτουν μετά την αντιστοίχιση των εικόνων f και f_{tr} , αντίστοιχα, ενώ $N_1 \times N_2$ είναι ο αριθμός των εικονοστοιχείων για κάθε μία από αυτές τις υποεικόνες.

Το σφάλμα αντιστοίχισης φαίνεται στο Σχήμα 2.4(α) για τη *Lenna* και στο Σχήμα 2.4(β) για την αεροφωτογραφία. Η διακεκομμένη καμπύλη αντιστοιχεί στην ICC, ενώ η συνεχής καμπύλη αντιστοιχεί στην προτεινόμενη ICC-PW. Φαίνεται καθαρά ότι το σφάλμα αντιστοίχισης είναι μικρότερο για την προτεινόμενη τεχνική.

Τέλος, στον Πίνακα 2.1 φαίνεται ο αριθμός των επιτυχημένων αντιστοιχίσεων, καθώς και η μέση τιμή και η διασπορά του σφάλματος αντιστοίχισης για το σύνολο των πειραμάτων. Η

Σχήμα 2.4: Σφάλμα αντιστοίχισης για (α) τη Lenna και (β) την αεροφωτογραφία

μέθοδός μας έχει περισσότερες επιτυχίες, ενώ η μέση τιμή και η διασπορά του σφάλματος αντιστοίχισης είναι μικρότερα. Πρέπει να σημειωθεί ότι όταν η προτεινόμενη τεχνική αποτυγχάνει να βρει τη σωστή μετατόπιση ή περιστροφή, το σφάλμα είναι πολύ μικρό, σε αντίθεση με την κλασική τεχνική ετεροσυσχέτισης, της οποίας οι εκτιμήσεις μπορεί να διαφέρουν αρκετά από τις πραγματικές τιμές.

2.5 Εφαρμογή της ICC-PW σε ακολουθίες ιατρικών εικόνων

Η τεχνική αντιστοίχισης που περιγράφηκε παραπάνω εφαρμόστηκε σε ακολουθίες εικόνων του τραχήλου της μήτρας, με σκοπό τη βελτίωση στη διάγνωση δυσπλασιών και κακοηθειών. Η ανάπτυξη μεθόδων πρόληψης του καρκίνου του τραχήλου της μήτρας έχει σαν σκοπό την

Πίνακας 2.1: ΑΡΙΘΜΟΣ ΕΠΙΤΥΧΗΜΕΝΩΝ ΑΝΤΙΣΤΟΙΧΙΣΕΩΝ, ΜΕΣΗ ΤΙΜΗ ΚΑΙ ΔΙΑΣΠΟΡΑ ΣΦΑΛΜΑΤΟΣ ΑΝΤΙΣΤΟΙΧΙΣΗΣ

Lenna	Επιτυχίες	Μέση τιμή	Διασπορά
Cross-Correlation	3	15.5055	126.9687
Προτεινόμενη τεχνική	8	7.0421	11.3229
Αεροφωτογραφία	Επιτυχίες	Μέση τιμή	Διασπορά
Cross-Correlation	6	6.9658	14.3385
Προτεινόμενη τεχνική	11	4.6507	3.4834

ανίχνευση και την αναγνώριση νεοπλασιών (Cervical Intraepithelial Neoplasia - CIN) σε πρώιμο στάδιο. Μια τέτοια μέθοδος είναι το τεστ ΠΑΠ, το οποίο χρησιμεύει στη διάγνωση καρκίνου ή αλλοιώσεων του ιστού σε προ-καρκινικό στάδιο. Αν το τεστ ΠΑΠ δείξει ότι υπάρχουν μη κανονικά κύτταρα, τότε ακολουθεί κολποσκόπια, όπου ο τράχηλος εξετάζεται με ένα μικροσκόπιο χαμηλής ισχύος [87]. Η διαδικασία εξέτασης περιλαμβάνει την τοπική εφαρμογή διαλύματος οξικού οξέος (3% – 5%), η οποία επιφέρει μια παροδική μεταβολή στις ιδιότητες διασκόρπισης του φωτός του καρκινικού ιστού. Στην ουσία δηλαδή παρατηρείται μια παροδική μεταβολή στο χρώμα του προβληματικού ιστού. Το φαινόμενο διαρκεί από μερικά δευτερόλεπτα μέχρι το πολύ 10 λεπτά και κατά τη διάρκεια αυτής της χρονικής περιόδου, οι μη κανονικές περιοχές γίνονται πιο ορατές και μοιάζουν με άσπρα μπαλώματα [125]. Το επιλεκτικό βάψιμο του μη κανονικού ιστού βασίζεται στην ιδιότητα του οξικού οξέος να αλληλεπιδρά με τα διαφοροποιημένα μεταβολικά ή δομικά χαρακτηριστικά (τα οποία σχετίζονται με την πρόοδο της ασθένειας) και να αυξάνει την αντίθεση ανάμεσα στον υγιή και τον παθολογικό ιστό.

Αν και τα δυναμικά χαρακτηριστικά αυτού του φαινομένου καθορίζονται από το είδος και το στάδιο της αλλοίωσης, η οπτική αποτίμηση αυτών των γρήγορα μεταβαλλόμενων χαρακτηριστικών δεν μπορεί να είναι ακριβής και επιπλέον ο εξεταστής δεν μπορεί να εκμεταλλευτεί πλήρως το διαγνωστικό περιεχόμενο αυτού του τεστ. Άρα, στη κολποσκόπια υπάρχει το πρόβλημα της χαμηλής ευαισθησίας, ενώ το οξικό οξύ δε χρησιμοποιείται για διαγνωστικούς σκοπούς, αλλά μόνο για την ένδειξη προβληματικών περιοχών, από όπου πρέπει να ληφθεί δείγμα για βιοψία. Το δείγμα αυτό υποβάλλεται σε ιστολογική εξέταση, η οποία θεωρείται έως σήμερα η καλύτερη διαγνωστική τεχνική. Ωστόσο, η ιστολογία, καθώς είναι ποιοτική και υποκειμενική, πάσχει από αρκετούς περιορισμούς και μειονεκτήματα. Ο ιστός λαμβάνεται με δειγματοληψία (η οποία μπορεί να μην είναι επιτυχημένη) και σφάλματα ανάγνωσης μπορεί να συμβούν. Έτσι, υπάρχει ανάγκη για ανάπτυξη νέων, πιο αποδοτικών διαγνωστικών μεθόδων και συστημάτων.

Σχετικά πρόσφατα, αναπτύχθηκε μια διαγνωστική μέθοδος βασισμένη στην απεικόνιση, η οποία είναι κατάλληλη για την ανίχνευση καρκινικών και άλλων νεοπλασιών ή κακοηθειών σε ζωντανό οργανισμό (*in vivo*) και την ταξινόμηση τους σε στάδια εξέλιξης της νόσου [6]. Η μέθοδος αυτή βασίζεται στην ποιοτική αποτίμηση των μεταβολών στο χρώμα του κάθε χωρικού σημείου της εξεταζόμενης περιοχής, μετά την τοπική εφαρμογή διαλύματος οξικού οξέος 3%. Υπολογίζονται οι χαμπύλες των τιμών της έντασης σε σχέση με το χρόνο για κάθε σημείο σε μια ακολουθία εικόνων, οι οποίες λαμβάνονται διαδοχικά, κατά τη διάρκεια εξέλιξης του φαινομένου. Κάποια πρώτα πειράματα έδειξαν ότι η μορφή αυτών των χαμπυλών περιέχει διαγνωστική πληροφορία, η οποία δίνει τη δυνατότητα για διάγνωση καρκινικών και προ-καρκινικών βλαβών και την ταξινόμησή τους σε στάδια, χωρίς να χρειάζεται αφαίρεση

Σχήμα 2.5: Εικόνες του τραχήλου της μήτρας από διαφορετικές ασθενείς

ιστού.

Η λήψη των εικόνων γίνεται με μια πολυφασματική κάμερα, η οποία περιγράφεται στην εργασία [6]. Αυτή η κάμερα παρέχει τη δυνατότητα λήψης εικόνων σε διαφορετικές μπάντες του φάσματος. Η μπάντα που επιλέγεται για τη λήψη των εικόνων του τραχήλου της μήτρας είναι τα $525 \pm 25nm$. Η φασματική απεικόνιση και ανάλυση του τραχήλου έχει δείξει ότι σε αυτή τη μπάντα επιτυγχάνεται η μέγιστη αντίθεση ανάμεσα στις περιοχές που ανταποκρίνονται στο οξικό οξύ και σε αυτές που δεν ανταποκρίνονται. Έτσι, η ακολουθία των εικόνων λαμβάνεται στην ίδια φασματική μπάντα, αλλά σε διαφορετικές χρονικές στιγμές.

Κατά τη διάρκεια λήψης των εικόνων, μπορεί να υπάρξουν κάποιες μικρο-μετακινήσεις του εξεταζόμενου ιστού σε σχέση με την κάμερα, λόγω της αναπνοής ή κάποιας ενόχλησης της ασθενούς. Οι μετακινήσεις αυτές περιλαμβάνουν μετατοπίσεις και περιστροφές. Άρα, ένας καθοριστικός παράγοντας που επηρεάζει την ακρίβεια της μεθόδου στη μέτρηση των κινητικών χαρακτηριστικών του φαινομένου και στον καθορισμό των ορίων της βλάβης είναι η αντιστοίχιση των ακολουθιών των εικόνων.

2.5.1 Πειραματικά αποτελέσματα

Ο νέος αλγόριθμος αντιστοίχισης, ο ICC-PW, εφαρμόστηκε σε ακολουθίες εικόνων και η απόδοσή του συγκρίθηκε με τις τεχνικές ICC και EPC. Οι τρεις τεχνικές εφαρμόστηκαν σε ακολουθίες εικόνων διαστάσεων 576×768 . Δύο τέτοιες εικόνες, οι οποίες ανήκουν σε διαφορετική ακολουθία η κάθε μία (ανήκουν δηλαδή σε διαφορετικές ασθενείς), φαίνονται στο Σχήμα 2.5.

Για κάθε ζευγάρι εικόνων θεωρούμε δύο υποεικόνες διαστάσεων 200×200 , με τη δεύτερη εικόνα να είναι μια μετατοπισμένη και περιστραμμένη έκδοση της πρώτης. Διεξάγονται 40 διαφορετικά πειράματα αυτής της κατηγορίας, και προκειμένου να συγκρίνουμε τις τρεις μεθόδους υπολογίζουμε τον αριθμό των επιτυχημένων μετατοπίσεων για κάθε περίπτωση και το

Σχήμα 2.6: Σφάλμα αντιστοίχισης

σφάλμα αντιστοίχισης. Το σφάλμα αντιστοίχισης φαίνεται στο Σχήμα 2.6. Η καμπύλη με τις τελείες αντιστοιχεί στο σφάλμα αντιστοίχισης της EPC, η διακεκομμένη καμπύλη στο σφάλμα αντιστοίχισης της ICC, ενώ η συνεχής καμπύλη αντιστοιχεί στην προτεινόμενη τεχνική, ICC-PW. Οι καμπύλες αυτές δείχνουν ότι η προτεινόμενη τεχνική δουλεύει καλύτερα από τις άλλες δύο. Αυτό επιβεβαιώνεται και από τον Πίνακα 2.2, όπου φαίνεται ότι η προτεινόμενη τεχνική έχει περισσότερες επιτυχίες, ενώ η μέση τιμή και η διασπορά του σφάλματος είναι μικρότερα. Ένας μεγάλος αριθμός πειραμάτων που διεξάχθηκαν στη συνέχεια επιβεβαίωσαν αυτά τα αποτελέσματα.

Η ICC-PW εφαρμόστηκε στις πραγματικές ακολουθίες των εικόνων του τραχήλου της μήτρας και υπολογίστηκαν οι καμπύλες της χρωματικής απόκρισης των εικονοστοιχείων σε σχέση με το χρόνο. Κάποιες χαρακτηριστικές καμπύλες, πριν και μετά την αντιστοίχιση (διακεκομμένες και συνεχείς καμπύλες, αντίστοιχα), φαίνονται στο Σχήμα 2.7. Μετά την

Πίνακας 2.2: ΑΡΙΘΜΟΣ ΕΠΙΤΥΧΗΜΕΝΩΝ ΑΝΤΙΣΤΟΙΧΙΣΕΩΝ, ΜΕΣΗ ΤΙΜΗ ΚΑΙ ΔΙΑΣΠΟΡΑ ΣΦΑΛΜΑΤΟΣ ΑΝΤΙΣΤΟΙΧΙΣΗΣ ΓΙΑ ΤΙΣ ΙΑΤΡΙΚΕΣ ΕΙΚΟΝΕΣ.

Ιατρικές εικόνες	Επιτυχίες	Μέση τιμή	Διασπορά
<i>EPC</i>	23	3.1928	2.5162
<i>ICC</i>	9	5.5481	9.7750
<i>ICC-PW</i>	25	1.8168	0.5811

Σχήμα 2.7: Χρωματικές αποκρίσεις πριν και μετά την αντιστοίχιση

αντιστοίχιση οι καμπύλες έχουν μικρότερη διασπορά και σχήμα κατάλληλο για αποτίμηση. Για παράδειγμα, η καμπύλη πριν την αντιστοίχιση στο πάνω αριστερά σχήμα είναι ακατάλληλη για διάγνωση. Μετά την αντιστοίχιση όμως η συγκεκριμένη καμπύλη παραμένει σχεδόν στο ίδιο επίπεδο, δηλώνοντας ότι ο ιστός που αντιστοιχεί στο συγκεκριμένο σημείο είναι μη καρκινικός. Στο σχήμα πάνω δεξιά, στην καμπύλη πριν την αντιστοίχιση υπάρχουν κάποιες αιχμές, οι οποίες δηλώνουν ότι πρέπει να υπάρχει κάποια μεγάλη μετατόπιση ανάμεσα σε διαδοχικές εικόνες. Μετά την αντιστοίχιση αυτές οι μετατοπίσεις αποκαθίστανται και προκύπτει μια καμπύλη κατάλληλη για διάγνωση. Η γρήγορη μετάβαση της έντασης των εικονοστοιχείων από χαμηλές σε υψηλές τιμές δηλώνει κακοήθειες. Στη συγκεκριμένη περίπτωση, ο τύπος της κακοήθειας είναι CIN II, καθώς μετά από κάποια ώρα η ένταση προσεγγίζει την αρχική της τιμή. Τέλος, στο κάτω σχήμα, η καμπύλη πριν την αντιστοίχιση δηλώνει CIN II, ενώ μετά την αντιστοίχιση, CIN III. Στην περίπτωση της CIN III χρειάζεται περισσότερη ώρα για να επανέλθει η ένταση στην αρχική της τιμή.

Όπως, φαίνεται λοιπόν από τα παραπάνω αποτελέσματα, οι καμπύλες μετά την αντιστοίχιση δεν περιλαμβάνουν “θόρυβο” λόγω των μετατοπίσεων και διευκολύνουν τη διάγνωση. Μπορεί να γίνει ακόμα και διάκριση ανάμεσα στα διαφορετικά στάδια των νεοπλασιών από το σχήμα των καμπυλών.

2.6 Συμπεράσματα, μελλοντικές ερευνητικές κατευθύνσεις

Σε αυτό το κεφάλαιο παρουσιάστηκε μια νέα τεχνική αντιστοίχισης, η οποία αποκαθιστά μετατοπίσεις και περιστροφές και έχει καλύτερη απόδοση σε σχέση με άλλες τεχνικές. Η τεχνική αυτή εφαρμόστηκε σε ακολουθίες εικόνων του τραχήλου της μήτρας με σκοπό τη διάγνωση δυσπλασιών και κακοηθειών. Το γεγονός ότι το χρώμα του προβληματικού ιστού αλλάζει μετά την τοπική εφαρμογή διαλύματος οξικού οξέος στην εξεταζόμενη περιοχή εισάγει άλλη μια παράμετρο στο πρόβλημα της αντιστοίχισης, εκτός από αυτές που αφορούν τις γεωμετρικές παραμορφώσεις. Στα πειράματά μας θεωρήσαμε ότι η αλλαγή αυτή είναι γραμμική και άρα μπορούμε να αντιμετωπίσουμε το πρόβλημα αυτό χρησιμοποιώντας σαν μετρική αντιστοίχισης την κανονικοποιημένη συσχέτιση ανάμεσα σε ακολουθίες οι οποίες έχουν υποστεί προ-λεύκανση. Μια καλύτερη προσέγγιση θα ήταν η υπόθεση αυτή να γίνεται για μικρότερα παράθυρα των εικόνων, να εξετάζεται δηλαδή το πρόβλημα τοπικά. Επίσης, θα μπορούσε να γίνει έλεγχος για περισσότερα είδη γεωμετρικών παραμορφώσεων, αν και εξετάζοντας το πρόβλημα τοπικά, η μετατόπιση και η περιστροφή συνήθως αρκούν για να αποκαταστήσουν τις γεωμετρικές παραμορφώσεις. Τα θέματα αυτά αποτελούν κάποιες αρχικές κατευθύνσεις για μελλοντική έρευνα.

Κεφάλαιο 3

Αποδοτικοί αλγόριθμοι για τον υπολογισμό μετρικών συσχέτισης

Το πιο ακριβό κομμάτι της διαδικασίας αντιστοίχισης εικόνων είναι ο υπολογισμός της χρησιμοποιούμενης μετρικής ομοιότητας (της χωρικής συσχέτισης στην περίπτωση μας) σαν συνάρτηση της σχετικής θέσης των συγκρινόμενων εικόνων. Ο υπολογισμός αυτός είναι στην ουσία ένα ψάξιμο στις δύο διαστάσεις, προκειμένου να βρούμε για ποια θέση μεγιστοποιείται η μετρική συσχέτισης. Σε αυτό το κεφάλαιο προτείνουμε τη χρήση ενός γρήγορου σχήματος υπολογισμού της συσχέτισης, το οποίο βασίζεται σε κατάλληλη τμηματοποίηση της εικόνας και στη χρήση του μετασχηματισμού Fourier [12, 64]. Προτείνουμε επίσης ένα αποδοτικό επαναληπτικό σχήμα, το οποίο μειώνει σημαντικά τον αριθμό των ψαξιμάτων που απαιτούνται για την εύρεση του μεγίστου του συντελεστή συσχέτισης στο χωρικό πεδίο και παρέχει ακρίβεια εικονοστοιχείου [68, 67]. Και οι δύο αλγόριθμοι αφορούν την περίπτωση της μετατόπισης (*translation*).

3.1 Αποδοτικός υπολογισμός της ετεροσυσχέτισης

Έστω f μια εικόνα αναφοράς και f_t μια μετατοπισμένη έκδοσή της. Έστω, επίσης ένα παραθύρο της εικόνας f , διαστάσεων $n \times n$, και μια περιοχή ψαξίματος μέσα στη εικόνα f_t , διαστάσεων $m \times m$ (με $m > n$), όπου $N = m - n + 1$ είναι η μέγιστη δυνατή μετατόπιση σε κάθε διάσταση. Άρα, η μετρική συσχέτισης πρέπει να υπολογιστεί για N^2 διαφορετικές μετατοπίσεις. Αν υπολογίζουμε τη δισδιάστατη συσχέτιση κατευθείαν από τον ορισμό της, τότε η πολυπλοκότητα είναι της τάξης του $O(n^2 N^2)$. Εναλλακτικά, η συσχέτιση μπορεί να υπολογιστεί στο πεδίο των συχνοτήτων μετά από zero-padding του παραθύρου της εικόνας f σε διαστάσεις $m \times m$. Σε αυτή την περίπτωση, η πολυπλοκότητα είναι $O(m^2 \log m)$ [84].

Πρέπει να σημειωθεί, όμως, ότι στην πράξη σε πολλές περιπτώσεις, μπορεί να γίνει η υπόθεση ότι το N είναι πολύ μικρότερο από το n ($N \ll n$). Σε πολλές εφαρμογές οι εικόνες λαμβάνονται με ρυθμό τέτοιο ώστε μεταξύ διαδοχικών εικόνων να υπάρχουν πολύ μικρές μετατοπίσεις και οπωσδήποτε αρκετά μικρότερες από το μέγεθος του παραθύρου και της περιοχής ψαξίματος που θα χρησιμοποιηθούν στη διαδικασία της αντιστοίχισης. Έτσι, μπορούμε να θεωρήσουμε ότι η χωρική συσχέτιση θα είναι ειδικού τύπου, αφού απαιτείται ο υπολογισμός της συσχέτισης για μικρό αριθμό μετατοπίσεων σε σχέση με το μέγεθος του παραθύρου. Σε αυτή την περίπτωση, ο υπολογισμός της συσχέτισης στο πεδίο των συχνοτήτων, καθώς και οι δισδιάστατες εκδοχές των τεχνικών overlap-save ή overlap-add δεν αποτελούν αποδοτικά σχήματα, καθώς το παράθυρο της εικόνας f και η περιοχή ψαξίματος της εικόνας f_t έχουν συγκρίσιμα μεγέθη.

Στη συνέχεια, παρουσιάζουμε μια νέα μέθοδο υπολογισμού της συσχέτισης, η οποία βασίζεται σε μια ιδέα που παρουσιάστηκε στην εργασία [13] για την περίπτωση της μίας διάστασης. Ειδικότερα, στις εικόνες f και f_t εφαρμόζεται μια κατάλληλη τμηματοποίηση και το πρόβλημα χωρίζεται σε έναν αριθμό από υπο-προβλήματα συσχέτισης. Αυτά τα μικρά υπο-προβλήματα υπολογίζονται αποδοτικά στο πεδίο των συχνοτήτων. Εξαιτίας της τμηματοποίησης, το προτεινόμενο σχήμα έχει μικρότερη πολυπλοκότητα σε σχέση με τις προαναφερθείσες τεχνικές. Ακολουθεί η περιγραφή του, όπου για ευκολία υποθέτουμε ότι τα n και N είναι δυνάμεις του 2.

Το προτεινόμενο σχήμα

1. Έστω $f_t(k, l)$, με $k, l = 1, 2, \dots, m$, τα εικονοστοιχεία της περιοχής ψαξίματος της εικόνας f_t , διαστάσεων $m \times m$, η οποία θα πάρει μέρος στον υπολογισμό της δισδιάστατης συσχέτισης.
2. Η περιοχή ψαξίματος της εικόνας f_t , διαστάσεων $m \times m$ διαιρείται σε n^2/N^2 επικαλυπτόμενα τμήματα, διαστάσεων $2N \times 2N$. Το κάθε τμήμα $f_{t_{i,j}}$ περιέχει τα εικονοστοιχεία $f_t[(i-1)N + u, (j-1)N + v]$, όπου $u, v = 1, 2, \dots, 2N$.
3. Το $n \times n$ παράθυρο της εικόνας f διαιρείται σε n^2/N^2 μη επικαλυπτόμενα και συνεχόμενα τμήματα, διαστάσεων $N \times N$. Έτσι, το τμήμα $f_{i,j}$ περιέχει τα εικονοστοιχεία $f[(i-1)N + k, (j-1)N + l]$, όπου $k, l = 1, 2, \dots, N$. Για κάθε $f_{i,j}$ φτιάχνουμε τον αντίστοιχο πίνακα (χάνοντας zero-padding), διαστάσεων $2N \times 2N$, τον $f_{i,j}^e$, ως εξής:

$$f_{i,j}^e = \begin{pmatrix} Jf_{i,j}J & 0 \\ 0 & 0 \end{pmatrix}$$

Σχήμα 3.1: Σχηματική αναπαράσταση της νέας τεχνικής

όπου J είναι ένας πίνακας διαστάσεων $N \times N$, ο οποίος έχει άσσους στην αντιδιαγώνιο και μηδενικά στις υπόλοιπες θέσεις του.

4. Για κάθε (i, j) , με $i, j = 1, 2, \dots, n^2/N^2$, υπολογίζονται οι δισδιάστατοι μετασχηματισμοί Fourier των αντίστοιχων $f_{t,i,j}$ και $f_{i,j}^e$ και στη συνέχεια το γινόμενο τους σημείο-προς-σημείο, $C_{i,j} = \mathbf{f}_{t,i,j} \mathbf{f}_{i,j}^e$.
5. Υπολογίζεται το άθροισμα:

$$C_{tot} = \sum_{i=1}^{n/N} \sum_{j=1}^{n/N} C_{i,j}$$

6. Υπολογίζεται ο αντίστροφος δισδιάστατος μετασχηματισμός Fourier του C_{tot} και από τον πίνακα διαστάσεων $2N \times 2N$ που προκύπτει, κρατάμε μόνο το τέταρτο που αντιστοιχεί στη συσχέτιση που μας ενδιαφέρει.

Η συνολική υπολογιστική πολυπλοκότητα του νέου σχήματος είναι περίπου $O(n^2 \log N)$. Έτσι, εκτός από τους μικρούς FFT που εμπλέκονται στους υπολογισμούς, επιλέον μείωση σε υπολογισμούς μπορεί να επιτευχθεί, ανάλογα με τις τιμές των n και N . Στο Σχήμα 3.1 φαίνεται ένα απλοποιημένο παράδειγμα της όλης διαδικασίας, για $N = n/2$. Η περιοχή ψαξίματος της εικόνας f_t χωρίζεται σε τέσσερα επικαλυπτόμενα τμήματα, διαστάσεων $2N \times 2N$. Ένα τέτοιο τμήμα φαίνεται στο σχήμα κάτω αριστερά. Το παράθυρο της εικόνας f χωρίζεται επίσης σε τέσσερα μη επικαλυπτόμενα τμήματα και σε κάθε τμήμα γίνεται το κατάλληλο zero-padding, όπως φαίνεται στο σχήμα κάτω δεξιά.

3.2 Αποδοτικό επαναληπτικό σχήμα για αντιστοίχιση με ακρίβεια εικονοστοιχείου

Όπως ήδη έχουμε πει, το πιο ακριβό κομμάτι της διαδικασίας αντιστοίχισης είναι ο υπολογισμός της χρησιμοποιούμενης μετρικής ομοιότητας (του συντελεστή συσχέτισης στην περίπτωση μας) σαν συνάρτηση των σχετικών θέσεων των συγκρινόμενων εικόνων. Είναι γνωστό ότι η ετεροσυσχέτιση μπορεί να υλοποιηθεί αποδοτικά στο πεδίο των συχνοτήτων. Ο συντελεστής συσχέτισης, όμως, δεν μπορεί να υπολογιστεί μέσω μιας απλής και αποδοτικής διαδικασίας στο πεδίο των συχνοτήτων και για το λόγο αυτό υπολογίζεται στο χωρικό πεδίο [44, 112].

Εξαιτίας του μεγάλου υπολογιστικού κόστους της συνέλιξης στο χωρικό πεδίο, έχουν αναπτυχθεί αρκετοί γρήγοροι, αλλά προσεγγιστικοί αλγόριθμοι [7]. Στην εργασία [80], μια κατάλληλα κανονικοποιημένη μετρική συσχέτισης υπολογίζεται μέσω συνέλιξης στο πεδίο των συχνοτήτων. Οι συγγραφείς της εργασίας [130] χρησιμοποιούν για τον υπολογισμό της κανονικοποιημένης συσχέτισης ένα σχήμα βασισμένο στο μετασχηματισμό Fourier, μαζί με ένα μη γραμμικό προ-φιλτράρισμα και κατάλληλη κατωφλίωση, προκειμένου να εκτελέσουν γρήγορη αντιστοίχιση. Η μέθοδός τους επιτυγχάνει αντιστοίχιση με ακρίβεια ± 2 εικονοστοιχεία. Εμείς θα παρουσιάσουμε σε αυτή την ενότητα έναν αποδοτικό επαναληπτικό αλγόριθμο για αντιστοίχιση εικόνων με ακρίβεια εικονοστοιχείου.

Αντιστοίχιση με ακρίβεια εικονοστοιχείου μπορεί να επιτευχθεί μεγιστοποιώντας το συντελεστή συσχέτισης για αμέριστε μετατοπίσεις. Ο αριθμός των απαιτούμενων ψαξιμάτων είναι N^2 , όπου N είναι ο αριθμός των ψαξιμάτων σε κάθε διάσταση. Ο σκοπός του προτεινόμενου αλγορίθμου είναι να μειώσει το συνολικό αριθμό ψαξιμάτων, ακολουθώντας μια επαναληπτική διαδικασία ψαξίματος σε γραμμές και στήλες. Σε κάθε επανάληψη αντιστοιχίζουμε τις συγκρινόμενες εικόνες και κρατάμε τις κοινές τους περιοχές, οι οποίες αποτελούν τις αρχικές εικόνες στην επόμενη επανάληψη. Ακολουθώντας αυτή τη στρατηγική ψαξίματος, ο αριθμός των ψαξιμάτων μειώνεται σε $2aN$, όπου $2N$ είναι ο αριθμός των ψαξιμάτων για κάθε επανάληψη και a ο αριθμός επαναλήψεων, ο οποίος, στις περισσότερες περιπτώσεις, είναι πολύ μικρότερος από $N/2$. Αυτός ο αλγόριθμος συγκλίνει στη σωστή λύση αν η αρχική εκτίμηση της μετατόπισης βρίσκεται μέσα στο βασικό λοβό της συνάρτησης του συντελεστή συσχέτισης. Σε αντίθετη περίπτωση, χρειάζεται μια κατάλληλη επαναρχικοποίηση του αλγορίθμου, όπως θα περιγραφεί στη συνέχεια, προκειμένου να επιτευχθεί σύγκλιση. Αν συμβεί κάτι τέτοιο, ο αριθμός των ψαξιμάτων γίνεται $2N \cdot \sum_{i=1}^k a_i$, όπου k είναι ο συνολικός αριθμός των επαναρχικοποιήσεων, δηλαδή ο αριθμός των επαναρχικοποιήσεων συν ένα (η αρχικοποίηση που γίνεται όταν ο αλγόριθμος εκτελείται για πρώτη φορά), και a_i ο αριθμός των επαναλήψεων του αλγορίθμου για κάθε επαναρχικοποίηση i .

3.2.1 Περιγραφή του αλγορίθμου

Έστω f η εικόνα αναφοράς και w ένα παράθυρο της f , με διαστάσεις $n \times n$. Επίσης, έστω g η περιοχή ψαξίματος σε μια εικόνα f_t , με διαστάσεις $m \times m$ (όπου $m > n$), η οποία είναι μια μετατοπισμένη έκδοση της f . Για τα w και g , η άνω αριστερά γωνία τους βρίσκεται στην αρχή ενός καθολικού συστήματος συντεταγμένων. Όλες οι δυνατές θέσεις του παραθύρου w μέσα στην περιοχή ψαξίματος g παίρνουν τιμές στο σύνολο

$$\mathcal{I} = [0, N - 1] \times [0, N - 1], \quad N = m - n + 1 \quad (3.2.1)$$

και ο μέγιστος αριθμός τους φράσσεται άνω από τον αριθμό των στοιχείων του συνόλου $\mathcal{I}$, είναι δηλαδή ίσος με N^2 . Έστω, τέλος, $s(x, y)$ ένα παράθυρο μέσα στην περιοχή ψαξίματος g , ίδιου μεγέθους με το w , και (x, y) οι συντεταγμένες της άνω αριστερά γωνίας. Το πρόβλημα της αντιστοίχισης εικόνων μπορεί να εκφραστεί σαν ένα πρόβλημα ψαξίματος. Ψάχνουμε για μια θέση $(x_0, y_0) \in \mathcal{I}$ η οποία να ικανοποιεί τη σχέση

$$s(x_0, y_0) = w, \quad (x_0, y_0) \in \mathcal{I}. \quad (3.2.2)$$

Προκειμένου να λύσουμε αυτό το πρόβλημα χρειαζόμαστε μια μετρική ομοιότητας και χρησιμοποιούμε το συντελεστή συσχέτισης, C , των παραθύρων w και $s(x, y)$, ο οποίος ορίζεται σαν

$$C(x, y) = \bar{\mathbf{w}}^T \bar{\mathbf{s}}(x, y) \quad (3.2.3)$$

όπου $\bar{\mathbf{w}} = \mathbf{w} / \|\mathbf{w}\|$ και $\bar{\mathbf{s}} = \mathbf{s} / \|\mathbf{s}\|$ είναι οι κανονικοποιημένες (με την Ευκλείδεια νόρμα) εκδόσεις, μέσης τιμής μηδέν, των διανυσμάτων $\mathbf{w} = \text{vec}(w - m_w)$ και $\mathbf{s}(x, y) = \text{vec}[s(x, y) - m_{s(x, y)}]$, όπου m_w και $m_{s(x, y)}$ δηλώνουν τις μέσες τιμές των παραθύρων w και $s(x, y)$ αντίστοιχα, ενώ ο τελεστής vec μετατρέπει τα παράθυρα σε διανύσματα στήλες, μήκους n^2 .

Ο προτεινόμενος επαναληπτικός αλγόριθμος περιγράφεται στη συνέχεια.

```

1  function [Newf, Newft] = iterative_registration(reference_image,
                                           translated_image, threshold)
2  f = reference_image, ft = translated_image, Cfft = threshold - 1
3  while Cfft < threshold
4      Determine w and g
5      for i = 0 : N - 1
6          C(i, 0) = corr2[w, s(i, 0)]
7      end
8      x = find[C(x, 0) == max(C(x, 0))]
```

```

9      Register  $f$  and  $f_t$  according to  $x \rightarrow$  Intermediate images  $f'$  and  $f'_t$ 
10     Determine  $w'$  and  $g'$ 
11     for  $j = 0 : N - 1$ 
12          $C(0, j) = \text{corr2} [w', s'(0, j)]$ 
13     end
14      $y = \text{find} [C(0, y) == \max (C(0, y))]$ 
15     Register  $f'$  and  $f'_t$  according to  $y \rightarrow$  New images  $\text{New}f$  and  $\text{New}f_t$ 
16      $C_{ff_t} = \text{corr2}(\text{New}f, \text{New}f_t)$ 
17     if  $((x \neq 0) \text{ OR } (y \neq 0)) \text{ AND } (C_{ff_t} < \text{threshold})$ 
18          $f = \text{New}f, f_t = \text{New}f_t$ 
19     elseif  $(x == 0) \text{ AND } (y == 0) \text{ AND } (C_{ff_t} < \text{threshold})$ 
20          $f = \text{reference\_image}, f_t = \text{translated\_image}$ 
21         Give a new random value for column translation,  $\text{New}x$ 
22          $x = \text{New}x$ 
23         goto 9
24     end
25 end

```

Η συνάρτηση *iterative_registration* παίρνει σαν είσοδο μια αρχική εικόνα αναφοράς και μια μετατοπισμένη έκδοσή της, καθώς και μια τιμή κατωφλίου. Το κατώφλι αυτό αφορά την τιμή του συντελεστή συσχέτισης ανάμεσα στις αντιστοιχισμένες εικόνες, C_{ff_t} , και στα πειράματά μας το θέσαμε ίσο με 0.99. Οι πρώτες ενέργειες που πραγματοποιεί η συνάρτηση είναι κάποιες αρχικοποιήσεις (γραμμή 2). Οι μεταβλητές f και f_t τίθενται ίσες με τις εικόνες εισόδου, ενώ ο συντελεστής συσχέτισης, C_{ff_t} , τίθεται ίσος με $\text{threshold} - 1$. Αυτό είναι απαραίτητο για την εισαγωγή στο βρόχο *while* που ακολουθεί.

Εφόσον η συνθήκη $C_{ff_t} < \text{threshold}$ είναι αληθής (γραμμή 3), ο συντελεστής συσχέτισης, $C(x, 0)$, υπολογίζεται για N κατακόρυφες μετατοπίσεις του παραθύρου w στην περιοχή ψαξίματος g (γραμμές 4 – 7). Βρίσκουμε την κατακόρυφη μετατόπιση x , για την οποία ο συντελεστής συσχέτισης παίρνει τη μέγιστη τιμή του, αντιστοιχίζουμε τις εικόνες f και f_t σύμφωνα με το x και κρατάμε τις κοινές τους περιοχές f' και f'_t , αντίστοιχα (γραμμές 8 – 9). Στη συνέχεια, καθορίζουμε νέο παράθυρο και νέα περιοχή ψαξίματος, w' και g' , τα οποία χρησιμοποιούνται για τον υπολογισμό του συντελεστή συσχέτισης, $C(0, y)$, για οριζόντιες μετατοπίσεις (γραμμές 10–13). Βρίσκουμε την οριζόντια μετατόπιση y , για την οποία μεγιστοποιείται ο συντελεστής συσχέτισης και αφού αντιστοιχίσουμε τις εικόνες f' και f'_t σύμφωνα με το y , κρατάμε τις κοινές τους περιοχές, $\text{New}f$ και $\text{New}f_t$ (γραμμές 14 – 15). Τέλος,

υπολογίζουμε το συντελεστή συσχέτισης, C_{ff_t} , για τις εικόνες που έχουν προκύψει, δηλαδή τις $Newf$ και $Newf_t$ (γραμμή 16).

Αν κάποια από τις μετατοπίσεις x ή y είναι διαφορετική από το μηδέν και ο συντελεστής συσχέτισης C_{ff_t} είναι μικρότερος από ένα προκαθορισμένο κατώφλι, τότε οι μεταβλητές f και f_t τίθενται ίσες με τις εικόνες $Newf$ και $Newf_t$, αντίστοιχα (γραμμές 17 – 18), και εκτελείται η επόμενη επανάληψη. Αν οι μετατοπίσεις x και y είναι ίσες με μηδέν, αλλά ο συντελεστής συσχέτισης C_{ff_t} είναι μικρότερος από το προκαθορισμένο κατώφλι (γραμμή 19), αυτό σημαίνει ότι ο αλγόριθμος έχει εγκλωβιστεί σε κάποιο τοπικό μέγιστο. Σε αυτή την περίπτωση ο αλγόριθμος χρειάζεται μια “διαταραχή” (επαναρχικοποίηση), προκειμένου να ξεφύγει από την περιοχή του τοπικού μεγίστου.

Η σύγκλιση του αλγορίθμου εξαρτάται από την αρχική εκτίμηση για την κατακόρυφη μετατόπιση x . Αν αυτή βρίσκεται μέσα στην περιοχή (στο λοβό) γύρω από τη μέγιστη τιμή του συντελεστή συσχέτισης (όπως υπολογίζεται στην περίπτωση του εξαντλητικού ψαξίματος), τότε ο αλγόριθμος συγκλίνει στο καθολικό μέγιστο. Αυτή είναι η πιο προφανής περίπτωση, καθώς το σχήμα της συνάρτησης του συντελεστή συσχέτισης γύρω από τη μέγιστη τιμή της θυμίζει κοίλη (*concave*) συνάρτηση. Επιπλέον, πρέπει να σημειωθεί ότι αν η συνάρτηση του συντελεστή συσχέτισης είναι συμμετρική γύρω από το ολικό μέγιστο και η αρχική εκτίμηση της μετατόπισης βρίσκεται μέσα στο βασικό λοβό, τότε αυτή η εκτίμηση αντιστοιχεί στη συνολική κατακόρυφη μετατόπιση και η επόμενη εκτίμηση αντιστοιχεί στη συνολική οριζόντια μετατόπιση. Άρα, σε αυτή την περίπτωση, ο αριθμός των ψαξιμάτων είναι ίσος με $2N$, που αντιστοιχεί στην καλύτερη περίπτωση.

Καθώς η αρχική εκτίμηση της μετατόπισης είναι αυτή που καθορίζει τη σύγκλιση του αλγορίθμου, μπορούμε να ξεπεράσουμε τον πιθανό εγκλωβισμό του αλγορίθμου σε ένα τοπικό μέγιστο επαναρχικοποιώντας τον αλγόριθμο με μια τυχαία κατακόρυφη μετατόπιση. Αρχικά, οι μεταβλητές f και f_t αρχικοποιούνται καταλλήλως (τίθενται ίσες με τις εικόνες εισόδου της συνάρτησης) και στη συνέχεια επιλέγεται με τυχαίο τρόπο μια νέα τιμή για την αρχική κάθετη μετατόπιση x , η $Newx$ (γραμμές 20–22). Στη συνέχεια, ο αλγόριθμος επιστρέφει στη γραμμή 9 και εκτελεί από εκεί και κάτω τις εντολές που υπάρχουν μέσα στο βρόχο *while*. Πρέπει να σημειωθεί ότι η εντολή *goto* στη γραμμή 23 χρησιμοποιείται προκειμένου να παρακάμψουμε τις εντολές που βρίσκονται στις γραμμές 4 – 7, οι οποίες αντιστοιχούν στην εκτίμηση της κάθετης μετατόπισης. Τέλος, αν οι μετατοπίσεις x και y γίνουν μηδέν και ο συντελεστής συσχέτισης, C_{ff_t} , των αντιστοιχισμένων εικόνων ίσος ή μεγαλύτερος από το κατώφλι, τότε ο αλγόριθμος έχει επιτύχει σύγκλιση και σταματά (γραμμές 24 – 25).

Στην περίπτωση που αλγόριθμος χρειάζεται νέες αρχικοποιήσεις, ο αριθμός των ψαξιμάτων είναι ίσος με $2N \cdot \sum_{i=1}^k a_i$, όπου k είναι ο συνολικός αριθμός αρχικοποιήσεων του αλγορίθμου,

δηλαδή ο αριθμός των επαναρχικοποιήσεων συν ένα (η αρχικοποίηση που γίνεται όταν ο αλγόριθμος εκτελείται για πρώτη φορά). Επιπλέον, a_i είναι ο αριθμός των επαναλήψεων για κάθε αρχικοποίηση και $\sum_{i=1}^k a_i$ είναι ο συνολικός αριθμός επαναλήψεων του αλγορίθμου, ο οποίος, τις περισσότερες φορές, είναι πολύ μικρότερος του $N/2$. Κάποια χαρακτηριστικά παραδείγματα που επιβεβαιώνουν την παραπάνω ανάλυση θα παρουσιαστούν στη συνέχεια, στην ενότητα 3.2.2.

Πρέπει να σημειωθεί ότι εξαιτίας των επαναληπτικών βημάτων της διαδικασίας αντιστοίχισης, οι τελικές εικόνες μπορεί να είναι μικρότερες από εκείνες που θα προέκυπταν με τη συμβατική αντιστοίχιση. Εντούτοις, αυτό το πρόβλημα λύνεται πολύ εύκολα. Το αλγεβρικό άθροισμα των μετατοπίσεων που προκύπτουν σε κάθε επανάληψη ισούται με τη συνολική μετατόπιση ανάμεσα στις συγκρινόμενες εικόνες. Έτσι, με βάση αυτά τα αθροίσματα, μπορούμε να εφαρμόσουμε τις αντίστοιχες μετατοπίσεις στις αρχικές εικόνες.

3.2.2 Πειραματική μελέτη της σύγκλισης του αλγορίθμου

Οι ιδιότητες σύγκλισης του επαναληπτικού αλγορίθμου έχουν μελετηθεί μέσω εκτεταμένων πειραμάτων. Εδώ παρουσιάζονται τρία χαρακτηριστικά παραδείγματα, τριών διαφορετικών εικόνων, για τις οποίες οι συναρτήσεις του συντελεστή συσχέτισης έχουν αρκετά διαφορετικό σχήμα. Οι δύο πρώτες εικόνες είναι τεχνητές και η τρίτη είναι η γνωστή εικόνα της Lenna.

Η πρώτη εικόνα μοιάζει με Gaussian συνάρτηση και ορίζεται ως εξής:

$$f(i, j) = 1000 \frac{1}{\sigma \sqrt{2\pi}} e^{-[(i-m)^2 + (j-m)^2]/2\sigma^2} \quad (3.2.4)$$

ενώ η μετατοπισμένη έκδοσή της υπολογίζεται από τη σχέση

$$g(i, j) = f(i - x, j - y). \quad (3.2.5)$$

Οι μετατοπίσεις δηλώνονται με τα x και y , ενώ σ και m είναι η τυπική απόκλιση και η μέση τιμή. Οι εικόνες που χρησιμοποιήθηκαν στα πειράματά μας είχαν διαστάσεις 400×400 . Επιπλέον, τα σ και m τέθηκαν ίσα με 80 και 200, αντίστοιχα. Ένα τέτοιο ζεύγος εικόνων φαίνεται στο Σχήμα 3.2, μαζί με τη συνάρτηση του συντελεστή συσχέτισής τους, όπου το παράθυρο και η περιοχή ψαξίματος που επιλέχθηκαν είχαν μέγεθος 128×128 και 228×228 , αντίστοιχα.

Η μετατόπιση ήταν (23, 17). Πρέπει να σημειώσουμε ότι το σχήμα της συνάρτησης του συντελεστή συσχέτισης, το οποίο θυμίζει Gaussian συνάρτηση, είναι συμμετρικό. Σε αυτή την περίπτωση, ο επαναληπτικός αλγόριθμος αναμένεται να συγκλίνει σε μία μόνο επανάληψη, δηλαδή μετά από $2N$ ψαξίματα, η οποία είναι και η καλύτερη περίπτωση.

Σχήμα 3.2: Μια τεχνητή εικόνα που μοιάζει με *Gaussian* συνάρτηση, (α), η μετατοπισμένη έκδοσή της κατά (23,17), (β), και η συνάρτηση του συντελεστή συσχέτισής τους, (γ).

Σχήμα 3.3: Μια ημιτονοειδής τεχνητή εικόνα, (α), η μετατοπισμένη έκδοσή της κατά (23,17), (β), και η συνάρτηση του συντελεστή συσχέτισής τους, (γ).

Η δεύτερη εικόνα είναι ημιτονοειδής και ορίζεται ως

$$f(i, j) = 120 \frac{\sin K_x(i - 100.1)}{K_x(i - 100.1)} \frac{\sin K_y(j - 100.1)}{K_y(j - 100.1)} \quad (3.2.6)$$

όπου τα K_x και K_y καθορίζουν το σχήμα της συνάρτησης κατά μήκος των δύο αξόνων [139]. Για την εικόνα που φαίνεται στο Σχήμα 3.3, $K_x = 0.2$ και $K_y = 0.2$, και η μετατοπισμένη της έκδοση δίνεται από τη σχέση

$$g(i, j) = f(i + x, j + y). \quad (3.2.7)$$

Η συνάρτηση του συντελεστή συσχέτισης έχει τώρα πολλούς επαναλαμβανόμενους λοβούς, οι οποίοι περιπλέκουν τη διαδικασία εκτίμησης της σωστής μετατόπισης, καθώς ο αλγόριθμος μπορεί να εγκλωβιστεί σε τοπικά μέγιστα.

Η τρίτη εικόνα που χρησιμοποιήθηκε είναι αυτή της Lenna, η οποία φαίνεται στο Σχήμα 3.4. Η συνάρτηση του συντελεστή συσχέτισης έχει πιο “φυσική” μορφή συγκρινόμενη με

Σχήμα 3.4: Η Lenna, (α), η μετατοπισμένη έκδοσή της κατά (23,17), (β), και η συνάρτηση του συντελεστή συσχέτισής τους, (γ).

τις προηγούμενες περιπτώσεις. Προφανώς, αν χρησιμοποιηθεί πιο μικρό παράθυρο, το σχήμα της θα είναι διαφορετικό.

Στα πειράματά μας εξετάσαμε τρεις περιπτώσεις παραθύρων με διαστάσεις, 128×128 , 64×64 και 32×32 . Οι αντίστοιχες περιοχές ψαξίματος ήταν 228×228 , 164×164 και 132×132 , που σημαίνει ότι ο μέγιστος αριθμός ψαξίματος είναι 101×101 , για όλες τις παραπάνω περιπτώσεις. Οι μετατοπίσεις, τις οποίες θα κληθεί να εκτιμήσει ο αλγόριθμός μας, παράγονται από μια γεννήτρια τυχαίων αριθμών κανονικά κατανομημένων, με τυπική απόκλιση ίση με 10. Οι αριθμοί που παράγονται στρογγυλοποιούνται στον κοντινότερο ακέραιο. Χρησιμοποιήσαμε 400 διαφορετικά ζεύγη μετατοπίσεων στα πειράματά μας. Ο προτεινόμενος επαναληπτικός αλγόριθμος συνέκλινε σε όλες τις περιπτώσεις, απλά σε κάποιες από αυτές χρειάστηκαν μερικές επαναρχικοποιήσεις.

Ας δούμε πρώτα τι συμβαίνει με την περίπτωση της μετατόπισης κατά (23,17). Όταν χρησιμοποιείται παράθυρο διαστάσεων 128×128 , ο αλγόριθμος συγκλίνει, όπως αναμενόταν, σε μόνο μία επανάληψη (2×101 ψαξίματα) για την εικόνα που μοιάζει με Gaussian. Για την ημιτονοειδή εικόνα ο αλγόριθμος χρειάζεται μία επαναρχικοποίηση και ο αριθμός των επαναλήψεων είναι 2 και 4 για κάθε αρχικοποίηση του αλγορίθμου (12×101 ψαξίματα), ενώ για τη Lenna ο αλγόριθμος συγκλίνει σε 5 επαναλήψεις (10×101 ψαξίματα). Όταν χρησιμοποιείται παράθυρο μεγέθους 64×64 , ο αλγόριθμος συγκλίνει σε μία επανάληψη (2×101 ψαξίματα) για τη Gaussian εικόνα, σε 3 επαναλήψεις (6×101 ψαξίματα) για την ημιτονοειδή και σε 9 επαναλήψεις (18×101 ψαξίματα) για τη Lenna. Στο Σχήμα 3.5 φαίνονται οι συναρτήσεις του συντελεστή συσχέτισης, μαζί με τις γραφικές παραστάσεις των περιγραμμάτων τους, όπου επίσης φαίνεται και η θέση της πρώτης εκτίμησης. Στην περίπτωση που χρησιμοποιηθεί παράθυρο μεγέθους 32×32 , ο αλγόριθμος συγκλίνει ξανά σε μία επανάληψη (2×101 ψαξίματα)

Σχήμα 3.5: Οι συναρτήσεις των συντελεστών συσχέτισης και τα περιγράμματά τους για τη Gaussian εικόνα, (α) και (δ), για την ημιτονοειδή, (β) και (ε), και για τη Lena, (γ) και (ζ), για παράθυρο μεγέθους 64×64 . Στις (δ), (ε) και (ζ), φαίνονται οι αρχικές εκτιμήσεις της μετατόπισης.

Σχήμα 3.6: Οι συναρτήσεις των συντελεστών συσχέτισης και τα περιγράμματά τους (μαζί με τις αρχικές εκτιμήσεις της μετατόπισης) για τη Gaussian εικόνα, (α) και (δ), για την ημιτονοειδή, (β) και (ε), και για τη Lena, (γ) και (ζ), για παράθυρο μεγέθους 32×32 .

για τη Gaussian εικόνα, σε 2 επαναλήψεις (4×101 ψαξίματα) για την ημιτονοειδή εικόνα και σε 19 επαναλήψεις (38×101 ψαξίματα) για τη Lenna. Οι συναρτήσεις των συντελεστών συσχέτισης και τα περιγράμματά τους με τη θέση της πρώτης εκτίμησης της μετατόπισης φαίνονται στο Σχήμα 3.6. Πρέπει να σημειωθεί ότι για την ημιτονοειδή εικόνα ο αλγόριθμος αποδίδει καλύτερα για μικρό παράθυρο σε αντίθεση με την εικόνα της Lenna. Επιπλέον, η αρχική εκτίμηση για τη μετατόπιση είναι έξω από το λοβό του καθολικού μεγίστου, αλλά ο αλγόριθμος συγκλίνει και μάλιστα χωρίς επαναρχικοποίηση για τα δύο μικρότερα παράθυρα.

Τα αποτελέσματα για το σύνολο των πειραμάτων φαίνονται στους Πίνακες 3.1, 3.2 και 3.3. Για την Gaussian εικόνα, ο αλγόριθμος συγκλίνει σε μια επανάληψη (2×101 ψαξίματα), ανεξάρτητα από το μέγεθος του παραθύρου που χρησιμοποιείται. Όπως έχει ήδη αναφερθεί, στην περίπτωση της ημιτονοειδούς εικόνας ο αλγόριθμος αποδίδει καλύτερα για μικρότερα παράθυρα. Ο μέσος αριθμός επαναλήψεων ανά πείραμα, δηλαδή το $mean\left(\sum_{i=1}^k a_i\right)$, είναι ίσος με 1.6525, 1.5700 και 1.1550 για μεγέθη παραθύρων 128×128 , 64×64 και 32×32 , αντίστοιχα. Στην περίπτωση που χρησιμοποιείται παράθυρο διαστάσεων 128×128 , χρειάζονται συνολικά 85 επαναρχικοποιήσεις για όλα τα 400 πειράματα και ο μέγιστος αριθμός επαναρχικοποιήσεων σε ένα πείραμα ήταν ίσος με 4. Ωστόσο, ο μέσος αριθμός επαναλήψεων ανά πείραμα είναι μικρός (1.6525) και βέβαια πολύ μικρότερος του $N/2$ (το οποίο στα πειράματά μας ήταν ίσο με 50.5), ενώ ο μέσος αριθμός ψαξιμάτων, δηλαδή το $2N \cdot mean\left(\sum_{i=1}^k a_i\right)$, είναι μόνο (3.3050×101). Για τη Lenna ο αλγόριθμος χρειάζεται γενικά περισσότερες επαναλήψεις. Εντούτοις, οι αριθμοί παραμένουν μικροί. Ειδικότερα, για παράθυρα μεγέθους 128×128 και 64×64 χρειάζονται μόνο 3 και 2 επαναρχικοποιήσεις, οι μέσοι αριθμοί επαναλήψεων ανά πείραμα είναι 3.2200 και 4.7075, οι οποίοι είναι σημαντικά μικρότεροι από $N/2$, και οι μέσοι αριθμοί ψαξιμάτων είναι 6.4400×101 και 9.4150×101 , αντίστοιχα. Στην περίπτωση παραθύρου μεγέθους 32×32 , ο αλγόριθμος χρειάζεται να επαναρχικοποιηθεί σε 152 πειράματα, ο μέγιστος αριθμός επαναρχικοποιήσεων για ένα πείραμα ήταν ίσος με 15 και ο συνολικός αριθμός επαναρχικοποιήσεων είναι 597. Πρέπει να σημειωθεί, όμως, ότι αυτή η περίπτωση είναι ιδιαίτερα δύσκολη, καθώς οι εξεταζόμενες μετατοπίσεις μπορεί να έχουν τιμές συγκρίσιμες με το μέγεθος του παραθύρου που χρησιμοποιείται (32×32). Ωστόσο, ο μέσος αριθμός επαναλήψεων ανά πείραμα παραμένει μικρός (23.5675, ο οποίος είναι αρκετά μικρότερος από $N/2$), ενώ ο μέσος αριθμός ψαξιμάτων είναι 47.1350×101 . Τα παραπάνω αποτελέσματα δείχνουν ότι, ακόμα και σε δύσκολες περιπτώσεις, ο προτεινόμενος αλγόριθμος συγκλίνει και ο αριθμός των απαιτούμενων ψαξιμάτων παραμένει σημαντικά μικρότερος σε σχέση με το εξαντλητικό ψάξιμο.

Πίνακας 3.1: 400 ΠΕΙΡΑΜΑΤΑ, ΠΑΡΑΘΥΡΟ: 128×128 - ΠΕΡΙΟΧΗ ΨΑΞΙΜΑΤΟΣ: 228×228

Εικόνες:	Gaussian	Ημιτονοειδής	Lenna
Σωστές Αντιστοιχίσεις	400	400	400
Επαναρχικοποιήσεις	0	85	3
Μέγιστος # Επαναρχικοποιήσεων	0	4	1
Πειράματα με επαναρχικοποίηση	0	53	3
Μέσος # επαναλήψεων	1	1.3629	3.1960
Μέσος # ψαξιμάτων (ανά πείραμα)	2×101	3.3050×101	6.4399×101

Πίνακας 3.2: 400 ΠΕΙΡΑΜΑΤΑ, ΠΑΡΑΘΥΡΟ: 64×64 - ΠΕΡΙΟΧΗ ΨΑΞΙΜΑΤΟΣ: 164×164

Εικόνες:	Gaussian	Ημιτονοειδής	Lenna
Σωστές Αντιστοιχίσεις	400	400	400
Επαναρχικοποιήσεις	0	62	2
Μέγιστος # Επαναρχικοποιήσεων	0	5	1
Πειράματα με επαναρχικοποίηση	0	36	2
Μέσος # επαναλήψεων	1	1.3623	4.6841
Μέσος # ψαξιμάτων (ανά πείραμα)	2×101	3.1469×101	9.4150×101

Πίνακας 3.3: 400 ΠΕΙΡΑΜΑΤΑ, ΠΑΡΑΘΥΡΟ: 32×32 - ΠΕΡΙΟΧΗ ΨΑΞΙΜΑΤΟΣ: 132×132

Εικόνες:	Gaussian	Ημιτονοειδής	Lenna
Σωστές Αντιστοιχίσεις	400	400	400
Επαναρχικοποιήσεις	0	0	597
Μέγιστος # Επαναρχικοποιήσεων	0	0	15
Πειράματα με επαναρχικοποίηση	0	0	152
Μέσος # επαναλήψεων	1	1.1550	9.4839
Μέσος # ψαξιμάτων (ανά πείραμα)	2×101	2.31×101	47.2772×101

3.3 Συμπεράσματα, μελλοντικές ερευνητικές κατευθύνσεις

Στο κεφάλαιο αυτό παρουσιάσαμε δύο αποδοτικούς αλγορίθμους. Ο πρώτος αφορά τον υπολογισμό της συσχέτισης με βάση μια κατάλληλη τμηματοποίηση της εικόνας και τη χρήση του μετασχηματισμού Fourier. Ο δεύτερος μειώνει σημαντικά τον αριθμό των ψαξιμάτων που απαιτούνται για την εύρεση του μεγίστου του συντελεστή συσχέτισης στο χωρικό πεδίο και παρέχει ακρίβεια εικονοστοιχείου. Και οι δύο αλγόριθμοι αφορούν την περίπτωση της μετατόπισης (*translation*). Μια πιθανή κατεύθυνση για μελλοντική έρευνα θα μπορούσε να είναι η προσπάθεια εύρεσης αποδοτικών αλγορίθμων στην περίπτωση διαφορετικών γεωμετρικών παραμορφώσεων, όπως για παράδειγμα η περιστροφή και η κλιμάκωση.

Κεφάλαιο 4

Αντιστοίχιση εικόνων με ακρίβεια υποδιαίρεσης εικονοστοιχείου

Πολλές εφαρμογές επεξεργασίας εικόνων απαιτούν την αντιστοίχιση των εικόνων με ακρίβεια υποδιαίρεσης εικονοστοιχείου. Στο κεφάλαιο αυτό θα παρουσιάσουμε μια τεχνική, η οποία παρέχει τέτοια ακρίβεια και βασίζεται στη μεγιστοποίηση του συντελεστή συσχέτισης. Η τεχνική αυτή δεν απαιτεί ανακατασκευή των τιμών της έντασης και παρέχει μια λύση κλειστού τύπου για την εκτίμηση της μετατόπισης [67].

4.1 Εισαγωγή

Οι τεχνικές αντιστοίχισης που βασίζονται σε ολόκληρες περιοχές των εικόνων (και όχι στα χαρακτηριστικά τους) είναι συνήθως αυτές που μπορούν να παρέχουν ακρίβεια σε υποδιαίρεση του εικονοστοιχείου. Οι πιο συχνά χρησιμοποιούμενες από αυτές βασίζονται στην παρεμβολή, [139]. Κάποια παραδείγματα είναι η παρεμβολή στην ένταση των εικονοστοιχείων της εικόνας, στη συσχέτιση και στη συσχέτιση φάσης. Ωστόσο, η ακρίβεια αυτών των τεχνικών εξαρτάται από την ποιότητα των αλγορίθμων παρεμβολής. Άλλες τεχνικές βασίζονται στις διαφορικές ιδιότητες που έχουν οι ακολουθίες των εικόνων [139] ή διατυπώνουν το πρόβλημα της αντιστοίχισης με ακρίβεια υποδιαίρεσης εικονοστοιχείου σαν ένα πρόβλημα βελτιστοποίησης [71, 138, 142].

Στην εργασία [142], περιγράφεται ένας αλγόριθμος αντιστοίχισης ακολουθίας εικόνων με χρήση μη γραμμικής ελαχιστοποίησης στο πεδίο των συχνοτήτων. Σε αυτόν τον αλγόριθμο οι εικόνες υποτίθεται ότι είναι *περιορισμένου εύρους ζώνης* (*bandlimited*). Οι συγγραφείς της εργασίας [71] προτείνουν μια τεχνική αντιστοίχισης εικόνων στις οποίες υπάρχει *ψευδωνυμία* (*aliasing*), η οποία βασίζεται σε ακύρωση των συχνοτήτων που περιέχουν *ψευδωνυμία*.

Οι Reddy και Chatterji [116] χρησιμοποιούν για αντιστοίχιση τη συσχέτιση συχνότητας του πολικού-λογαριθμικού μετασχηματισμού των εικόνων και ένα υψιπερατό φίλτρο για την ενίσχυση των υψηλών συχνοτήτων. Άλλη μια τεχνική στο πεδίο των συχνοτήτων με ακρίβεια υποδιαίρεσης εικονοστοιχείου παρουσιάζεται στην εργασία [129] και περιορίζεται στην αποκατάσταση μετατοπίσεων με την υπόθεση ότι οι αλλαγές στο φωτισμό των εικόνων είναι ομοιόμορφες. Και σε αυτή την τεχνική απομακρύνονται οι συχνότητες που μπορεί να προκαλέσουν σφάλμα στην εκτίμηση των μετατοπίσεων λόγω ψευδωνυμίας. Στο άρθρο [41], αποδεικνύεται ότι η ισχύς του σήματος στη συσχέτιση φάσης αντιστοιχεί στον πολυφασικό μετασχηματισμό ενός φιλτραρισμένου μοναδιαίου παλμού στο σημείο της αντιστοίχισης. Πιο πρόσφατα, στο [145] παρουσιάζεται μια τεχνική αντιστοίχισης στο πεδίο των συχνοτήτων, η οποία βασίζεται στις χαμηλές συχνότητες της εικόνας. Αρχικά, υπολογίζεται η διαφορά φάσης ανάμεσα στις συγχρινόμενες εικόνες και για το τμήμα των συχνοτήτων που δεν περιλαμβάνει ψευδωνυμία γράφονται οι κατάλληλες γραμμικές εξισώσεις που περιγράφουν ένα επίπεδο. Στη συνέχεια, βρίσκονται οι μετατοπίσεις από τη λύση των γραμμικών εξισώσεων υπό την έννοια των ελάχιστων τετραγώνων (*least squares*).

Γενικά, είναι πιο εύκολο να περιγράψουμε και να χειριστούμε τη ψευδωνυμία στο πεδίο των συχνοτήτων. Το πεδίο των συχνοτήτων όμως περιορίζεται σε καθολικά μοντέλα κίνησης. Από την άλλη μεριά, το χωρικό πεδίο επιτρέπει τη χρήση πιο γενικών μοντέλων κίνησης. Στην εργασία [70] παρουσιάζεται ένα επαναληπτικό σχήμα που αντιμετωπίζει μετατοπίσεις και περιστροφές και βασίζεται σε *αναπτύγματα σειρών Taylor* (*Taylor expansions*). Ένα πυραμιδικό σχήμα χρησιμοποιείται προκειμένου να αυξηθεί η ακρίβεια της μεθόδου για μεγάλες παραμέτρους κίνησης. Οι συγγραφείς της εργασίας [2] χρησιμοποιούν τη συσχέτιση αραιά δειγματοληπτημένων περιοχών και η μέθοδός τους παρέχει ακρίβεια καλύτερη από 0.2 εικονοστοιχεία. Στην εργασία [104], μια συνάρτηση σφάλματος, γραμμική στις παραμέτρους του μοντέλου κίνησης (*local affine*), ελαχιστοποιείται με τη χρήση ελάχιστων τετραγώνων. Στη συνέχεια, στη συνάρτηση σφάλματος προστίθεται ένας περιορισμός για τις μη γραμμικές διαφοροποιήσεις της έντασης των εικόνων και η λύση που έχει προκύψει από τα ελάχιστα τετράγωνα χρησιμοποιείται για την εκκίνηση ενός επαναληπτικού, μη γραμμικού αλγορίθμου ελαχιστοποίησης.

Τέλος, στην εργασία [114] παρουσιάζεται μια βελτιωμένη τεχνική αντιστοίχισης εικόνων stereo που βασίζεται στο συντελεστή συσχέτισης. Η τεχνική αυτή βρίσκει την οριζόντια μετατόπιση ανάμεσα στις stereo εικόνες με ακρίβεια υποδιαίρεσης εικονοστοιχείου. Στο κεφάλαιο αυτό θα παρουσιάσουμε ένα νέο σχήμα αντιστοίχισης εικόνων στο χωρικό πεδίο [67, 68], για μετατοπίσεις και στις δύο διαστάσεις, το οποίο παρέχει ακρίβεια υποδιαίρεσης εικονοστοιχείου και μικρό υπολογιστικό κόστος σε σχέση με υπάρχουσες τεχνικές.

4.2 Το νέο σχήμα αντιστοίχισης

4.2.1 Διατύπωση του προβλήματος

Έστω f μια εικόνα αναφοράς και w ένα παράθυρο της f , με διαστάσεις $n \times n$. Έστω, επίσης, g μια περιοχή ψαξίματος μέσα σε μια μετατοπισμένη έκδοση f_t της εικόνας f , με διαστάσεις $m \times m$ (με $m > n$)¹. Και για το w και για το g , θεωρούμε ότι η πάνω αριστερά γωνία τους βρίσκεται στην αρχή ενός καθολικού συστήματος συντεταγμένων. Όλες οι πιθανές θέσεις του παραθύρου w μέσα στην περιοχή ψαξίματος g παίρνουν τιμές στο σύνολο

$$\mathcal{I} = [0, N - 1] \times [0, N - 1], \quad N = m - n + 1 \quad (4.2.1)$$

και η μέγιστη τιμή τους φράσσεται άνω από το πλήθος των στοιχείων του συνόλου $\mathcal{I}$, δηλαδή από το N^2 .

Έστω τώρα $s(x, y)$ ένα παράθυρο της περιοχής ψαξίματος g , το οποίο έχει το ίδιο μέγεθος με το w και (x, y) οι συντεταγμένες της άνω αριστερά γωνίας του. Το πρόβλημα της αντιστοίχισης εικόνων μπορεί να εκφραστεί σαν ένα πρόβλημα ψαξίματος. Ψάχνουμε για $(x_0, y_0) \in \mathcal{I}$, έτσι ώστε η ακόλουθη σχέση να είναι αληθής

$$s(x_0, y_0) = w, \quad (x_0, y_0) \in \mathcal{I}. \quad (4.2.2)$$

Προκειμένου να λύσουμε το παραπάνω πρόβλημα, χρειαζόμαστε μια μετρική ομοιότητας. Χρησιμοποιούμε το συντελεστή συσχέτισης των παραθύρων w και $s(x, y)$, ο οποίος δίνεται από τη σχέση

$$C(x, y) = \bar{\mathbf{w}}^T \bar{\mathbf{s}}(x, y) \quad (4.2.3)$$

όπου $\bar{\mathbf{w}} = \mathbf{w}/\|\mathbf{w}\|$ και $\bar{\mathbf{s}} = \mathbf{s}/\|\mathbf{s}\|$ είναι οι κανονικοποιημένες (με την Ευκλείδεια νόρμα) εκδοχές, μηδενικής μέσης τιμής, των διανυσμάτων $\mathbf{w} = \text{vec}(w - m_w)$ και $\mathbf{s}(x, y) = \text{vec}[s(x, y) - m_{s(x,y)}]$, όπου m_w και $m_{s(x,y)}$ δηλώνουν τις μέσες τιμές των παραθύρων w και $s(x, y)$, αντίστοιχα, και ο τελεστής vec μετατρέπει τα παράθυρα σε διανύσματα στήλες, μήκους n^2 . Ο συντελεστής συσχέτισης που ορίζεται στη σχέση (4.2.3) έχει το πλεονέκτημα να μένει ανεπηρέαστος από γραμμικές φωτομετρικές διαφοροποιήσεις. Αυτή η ιδιότητα απαιτείται σε πολλές εφαρμογές αντιστοίχισης εικόνων.

Έχοντας ορίσει τη μετρική ομοιότητας στη σχέση (4.2.3) και υποθέτοντας ότι οι μετατοπίσεις ανάμεσα στα συγκρινόμενα παράθυρα είναι σταθερές, πρέπει να υπολογίσουμε το

¹Οι παραπάνω ποσότητες, καθώς και κάποιες από αυτές που ακολουθούν, έχουν οριστεί και στην ενότητα 3.2.1, αλλά αναφέρονται και εδώ για διευκόλυνση του αναγνώστη.

συντελεστή συσχέτισης για όλες τις δυνατές θέσεις του παραθύρου w μέσα στην περιοχή φάξιματος g , προκειμένου να βρούμε τη μετατόπιση (x_0, y_0) που μεγιστοποιεί το (4.2.3). Στην πράξη, έχουμε να λύσουμε το ακόλουθο πρόβλημα μεγιστοποίησης

$$\max_{(x,y) \in \mathcal{I}} C(x, y). \quad (4.2.4)$$

Λύνοντας αυτό το πρόβλημα, μπορούμε να αντιστοιχίσουμε εικόνες με ακρίβεια εικονοστοιχείου. Ωστόσο, σε πολλές εφαρμογές απαιτείται ακρίβεια υποδιαίρεσης εικονοστοιχείου [2, 104, 113, 114, 145]. Για το λόγο αυτό, στην επόμενη ενότητα, επεκτείνουμε στις δύο διαστάσεις τη μετρική ομοιότητας που προτείνεται στις εργασίες [113, 114] και διατυπώνουμε ένα κατάλληλο πρόβλημα μεγιστοποίησης, προκειμένου να επιτευχθεί η επιθυμητή ακρίβεια.

4.2.2 Η προτεινόμενη μετρική για ακρίβεια υποδιαίρεσης εικονοστοιχείου

Για το υπολογισμό μετατοπίσεων με ακρίβεια υποδιαίρεσης εικονοστοιχείου, ο συντελεστής συσχέτισης της σχέσης (4.2.3) πρέπει να επαναοριστεί. Τα παράθυρα που εμπλέκονται στον υπολογισμό του πρέπει να αναδομηθούν καταλλήλως με χρήση παρεμβολής, που σημαίνει ότι ο συντελεστής συσχέτισης θα εξαρτάται πλέον από συνεχείς μεταβλητές. Ο νέος ορισμός του συντελεστή συσχέτισης δίνεται από τη σχέση

$$C(x + t_x, y + t_y) = \bar{\mathbf{w}}^T \bar{\mathbf{s}}(x + t_x, y + t_y) \quad (4.2.5)$$

όπου t_x και t_y είναι συνεχείς μεταβλητές, που αντιστοιχούν σε μη ακέραιες μετατοπίσεις. Το αντίστοιχο πρόβλημα μεγιστοποίησης παίρνει τώρα τη μορφή

$$\max_{(x,y) \in \mathcal{I}} \max_{(t_x, t_y)} C(x + t_x, y + t_y) \quad (4.2.6)$$

η οποία δηλώνει μια μεγιστοποίηση σε σχέση με την ακέραιη μετατόπιση (x, y) και μια μεγιστοποίηση σε σχέση με τη συνεχή μετατόπιση (t_x, t_y) .

Η μορφή του συντελεστή συσχέτισης που ορίζεται στη σχέση (4.2.5), όπως και το υπολογιστικό κόστος του προβλήματος μεγιστοποίησης που ορίζεται στη σχέση (4.2.6), εξαρτάται από τον τύπο της παρεμβολής που θα χρησιμοποιηθεί για την ανακατασκευή του παραθύρου. Επίσης, αν η μεγιστοποίηση της (4.2.6) σε σχέση με το (t_x, t_y) δεν έχει μια λύση κλειστού τύπου, απαιτείται ένας αλγόριθμος μεγιστοποίησης, ο οποίος θα αυξήσει σημαντικά το υπολογιστικό κόστος. Το συνολικό πρόβλημα της αντιστοίχισης μπορεί να επιλυθεί αλλάζοντας τη σειρά των μεγιστοποιήσεων στη σχέση (4.2.6) και δειγματοληπτώντας τις συνεχείς μεταβλητές, αλλά το κόστος θα είναι αυξημένο και η ακρίβεια περιορισμένη. Προκειμένου να

αποφύγουμε αυτά τα προβλήματα, για δεδομένη ακέραιη μετατόπιση, ενσωματώνουμε στη σχέση (4.2.5), τον ακόλουθο πυρήνα πρώτης τάξης

$$\begin{aligned} \mathbf{s}(x + t_x, y + t_y) &\approx \mathbf{s}(x, y) + t_x [\mathbf{s}(x, y) - \mathbf{s}(x - 1, y)] + t_y [\mathbf{s}(x, y) - \mathbf{s}(x, y - 1)] \\ &= \mathbf{s} + t_x \Delta \mathbf{s}_{1,0} + t_y \Delta \mathbf{s}_{0,1} \end{aligned} \quad (4.2.7)$$

που στην ουσία είναι η Taylor προσέγγιση πρώτης τάξης του μετατοπισμένου παραθύρου, $\mathbf{s}(x + t_x, y + t_y)$. Τα διανύσματα $\mathbf{s}(x - 1, y)$ και $\mathbf{s}(x, y - 1)$ αντιστοιχούν στις κατά ένα εικονοστοιχείο μετατοπισμένες εκδόσεις του παραθύρου $s(x, y)$ στους άξονες x και y , αντίστοιχα, ενώ $\Delta \mathbf{s}_{1,0} = \mathbf{s}(x, y) - \mathbf{s}(x - 1, y)$, $\Delta \mathbf{s}_{0,1} = \mathbf{s}(x, y) - \mathbf{s}(x, y - 1)$ είναι προσεγγίσεις των χωρικών παραγώγων στις διαστάσεις x και y .

Ορίζουμε τις ακόλουθες ποσότητες

$$\rho_{k,l} = \bar{\mathbf{w}}^t \bar{\mathbf{s}}(x - k, y - l), \quad r_{k,l} = \bar{\mathbf{s}}^t(x - k, l) \bar{\mathbf{s}}(x, y - l), \quad \lambda_{k,l} = \frac{\|\mathbf{s}(x - k, y - l)\|}{\|\mathbf{s}(x, y)\|} \quad (4.2.8)$$

όπου το $\rho_{k,l}$ αντιστοιχεί στο συντελεστή συσχέτισης του παραθύρου w με το παράθυρο $s(x, y)$ (για $k = l = 0$) ή με τις μετατοπισμένες εκδόσεις του $s(x - k, y - l)$, το $r_{k,l}$ ορίζεται με παρόμοιο τρόπο και το $\lambda_{k,l}$ είναι ο λόγος της νόρμας των μετατοπισμένων εκδοχών του διανύσματος $\mathbf{s}(x, y)$, $\mathbf{s}(x - k, y - l)$, προς τη νόρμα του διανύσματος $\mathbf{s}(x, y)$. Ας ορίσουμε, επίσης, τις ακόλουθες χρήσιμες ποσότητες

$$\begin{aligned} a_0 &= \rho_{0,0} & a_1 &= \rho_{0,0} - \rho_{1,0} \lambda_{1,0} & a_2 &= \rho_{0,0} - \rho_{0,1} \lambda_{0,1} \\ b_0 &= 1 & b_1 &= 2(1 - r_{1,0} \lambda_{1,0}) & b_2 &= 2(1 - r_{0,1} \lambda_{0,1}) \\ b_3 &= 2(1 - r_{1,0} \lambda_{1,0} - r_{0,1} \lambda_{0,1} + r_{1,1} \lambda_{1,0} \lambda_{0,1}) \\ b_4 &= 1 + \lambda_{1,0}^2 - 2r_{1,0} \lambda_{1,0} & b_5 &= 1 + \lambda_{0,1}^2 - 2r_{0,1} \lambda_{0,1}. \end{aligned} \quad (4.2.9)$$

Ενσωματώνοντας την (4.2.7) στην (4.2.5), ο συντελεστής συσχέτισης μετατρέπεται σε συνάρτηση των συνεχών παραμέτρων t_x και t_y . Στη συνέχεια, για δεδομένη ακέραια μετατόπιση (x, y) , πρέπει να βρούμε τη λύση (t'_x, t'_y) , η οποία μεγιστοποιεί την ακόλουθη συνάρτηση

$$C(x + t_x, y + t_y) = \frac{a_0 + a_1 t_x + a_2 t_y}{\sqrt{b_0 + b_1 t_x + b_2 t_y + b_3 t_x t_y + b_4 t_x^2 + b_5 t_y^2}}. \quad (4.2.10)$$

Έχοντας ορίσει το συντελεστή συσχέτισης σαν μια συνάρτηση των συνεχών παραμέτρων μετατόπισης t_x, t_y και για δεδομένο (x_0, y_0) του $(x, y) \in \mathcal{I}$, έχουμε να λύσουμε το ακόλουθο πρόβλημα μεγιστοποίησης

$$\max_{(t_x, t_y)} C(x_0 + t_x, y_0 + t_y). \quad (4.2.11)$$

Αυτή η διαδικασία περιγράφεται στη συνέχεια.

4.2.3 Λύση κλειστού τύπου

Ο συντελεστής συσχέτισης που ορίζεται στη σχέση (4.2.10) είναι μη γραμμικός σε σχέση με τις συνεχείς παραμέτρους μετατόπισης t_x και t_y . Εντούτοις, η μεγιστοποίησή του καταλήγει σε μια λύση κλειστού τύπου, η οποία δίνεται στο επόμενο θεώρημα.

ΘΕΩΡΗΜΑ 1: Έστω ότι ο συντελεστής συσχέτισης ορίζεται όπως στη σχέση (4.2.10) και ότι ο παρανομαστής στη (4.2.10) είναι διάφορος του μηδενός, για δεδομένο $(x_0, y_0) \in \mathcal{I}$. Τότε, το $C(x_0 + t_x, y_0 + t_y)$ παίρνει τη μοναδική ακρότατη τιμή του, στο

$$t'_x = \frac{(a_2b_1 - a_1b_2)(a_1b_2 - a_0b_3) + (2a_1b_0 - a_0b_1)(2a_1b_5 - a_2b_3)}{(a_1b_3 - 2a_2b_4)(a_1b_2 - a_0b_3) - (a_1b_1 - 2a_0b_4)(2a_1b_5 - a_2b_3)} \quad (4.2.12)$$

και

$$t'_y = -\frac{(a_2b_1 - a_1b_2)(a_1b_1 - 2a_0b_4) + (2a_1b_0 - a_0b_1)(a_1b_3 - 2a_2b_4)}{(a_1b_3 - 2a_2b_4)(a_1b_2 - a_0b_3) - (a_1b_1 - 2a_0b_4)(2a_1b_5 - a_2b_3)}. \quad (4.2.13)$$

Το ακρότατο είναι μέγιστο, αν και μόνο αν ο Hessian πίνακας του $C(x_0 + t_x, y_0 + t_y)$, υπολογισμένος στο (t'_x, t'_y) , είναι αρνητικά ορισμένος.

ΑΠΟΔΕΙΞΗ: Η απόδειξη του Θεωρήματος 1 δίνεται στο Παράρτημα Α.

Χρησιμοποιώντας το παραπάνω θεώρημα μπορούμε να βρούμε τη βέλτιστη συνεχή μετατόπιση, με μια μικρή αύξηση στο υπολογιστικό κόστος. Το βασικό κόστος στην εκτίμηση της μετατόπισης με ακρίβεια εικονοστοιχείου οφείλεται στον υπολογισμό των ποσοτήτων που ορίζονται στην (4.2.8) και περιλαμβάνονται στην (4.2.9). Για τις ποσότητες αυτές πρέπει να υπολογιστούν τέσσερεις νόρμες και έξι εσωτερικά γινόμενα, με πολυπλοκότητα $O(n^2)$, όπου n^2 είναι ο αριθμός των στοιχείων του κάθε διανύσματος. Ειδικότερα, χρειάζονται $10n^2 + 6$ πολλαπλασιασμοί, $10(n^2 - 1)$ προσθέσεις, 4 τετραγωνικές ρίζες και 8 διαιρέσεις. Πρέπει να σημειωθεί ότι αν οι συγκρινόμενες εικόνες έχουν αντιστοιχηθεί για τις ακέραιες μετατοπίσεις, τότε οι παραπάνω ποσότητες πρέπει να υπολογιστούν μόνο για ένα μικρό αριθμό ψαξιμάτων. Σε αυτή την περίπτωση, το βασικό κόστος είναι αυτό που απαιτείται για την αντιστοίχιση με ακρίβεια εικονοστοιχείου. Αυτή η αντιστοίχιση μπορεί να γίνει αποδοτικά χρησιμοποιώντας τον αλγόριθμο που παρουσιάστηκε στην ενότητα 3.2.

4.3 Πειραματικά αποτελέσματα

Προκειμένου να μετρήσουμε την απόδοση του προτεινόμενου αλγορίθμου, πραγματοποιήσαμε μεγάλο αριθμό πειραμάτων. Συγκρίναμε τον αλγόριθμο με μια πρόσφατη τεχνική αντιστοίχισης [145], η οποία βασίζεται στο πεδίο των συχνοτήτων και έχει καλύτερη απόδοση σε σχέση

(α)

(β)

(γ)

(δ)

Σχήμα 4.1: Ζεύγη εικόνων που έχουν ληφθεί από το *Camera Tool* της *MATLAB* ((α) και (β) διαστάσεων 560×768) και από μια διαδικασία υπο-δειγματοληψίας ((γ) και (δ) διαστάσεων 288×288).

με υπάρχουσες τεχνικές [70, 86, 92]. Στην εργασία [145], οι εικόνες που χρησιμοποιούνται στα πειράματα προέρχονται από εικόνες υψηλής ανάλυσης (στις οποίες αρχικά γίνεται υπερ-δειγματοληψία με παράγοντα 2) με υπο-δειγματοληψία. Οι ακέραιες μετατοπίσεις στις εικόνες μεγάλης ανάλυσης αντιστοιχούν σε μη ακέραιες μετατοπίσεις στις δειγματοληπτημένες εικόνες που προκύπτουν. Προκειμένου να ελεγχθεί η *ψευδωνυμία* (*aliasing*), χρησιμοποιείται ένα χαμηλοπερατό φίλτρο πριν τη υπο-δειγματοληψία. Επιπλέον, οι αρχικές εικόνες πολλαπλασιάζονται με ένα παράθυρο Tukey, προκειμένου να είναι κυκλικά συμμετρικές και να αποφευχθούν οι οποιεσδήποτε επιδράσεις των ορίων. Παρόμοιες προσεγγίσεις για την παραγωγή εικόνων με μη ακέραιες μετατοπίσεις έχουν χρησιμοποιηθεί και στις εργασίες [41, 71, 124, 129].

Αρκετά ζεύγη αρχικών και μετατοπισμένων εικόνων παράχθηκαν για τα πειράματά μας, χρησιμοποιώντας τους ακόλουθους τρεις τρόπους. Αρχικά, παράχθηκαν ζεύγη εικόνων με χρήση του κώδικα σε Matlab που μπορεί να βρεθεί στην ηλεκτρονική διεύθυνση [58]. Στη συνέχεια, άλλα ζεύγη εικόνων παράχθηκαν με χρήση του *Camera Tool* της Matlab. Ειδικότερα, δημιουργήσαμε ένα επίπεδο στον τρισδιάστατο χώρο και η γωνία από την οποία η κάμερα βλέπει το επίπεδο ορίστηκε με τέτοιο τρόπο, ώστε να έχουμε το ίδιο βάθος σε όλα τα σημεία του επιπέδου. Θέσαμε το αζιμούθιο και τη γωνία κάθετης ανύψωσης ίσα με 0 και 90 μοίρες,

Σχήμα 4.2: Απόδοση της μεθόδου που βασίζεται στο πεδίο των συχνοτήτων

αντίστοιχα. Με αυτό τον τρόπο προσομοιώσαμε ένα μοντέλο επίπεδης κίνησης, μετακινώντας την κάμερα σε δύο διευθύνσεις και λαμβάνοντας εικόνες από διαφορετικά σημεία. Αφού μελετάμε την περίπτωση της καθαρής μετατόπισης, κάναμε χρήση της συνάρτησης *camdolly* της Matlab. Ένα ζευγάρι εικόνων που παράγεται με αυτόν τον τρόπο φαίνεται στο Σχήμα 4.1, μαζί με ένα ζευγάρι εικόνων που παράγονται σύμφωνα με τη διαδικασία που περιγράφεται στην εργασία [145], οι οποίες έχουν υποδειγματοληπτηθεί κατά ένα παράγοντα 4. Τέλος, για εικόνες που παράγονται από δισδιάστατες συναρτήσεις, όπως η Gaussian και η ημιτονοειδής εικόνα που χρησιμοποιήθηκαν στο προηγούμενο κεφάλαιο, παράγουμε τα αντίστοιχα ζεύγη εικόνων κατευθείαν από τον ορισμό τους.

Για όλες τις εικόνες έγιναν πειράματα για 400 διαφορετικά ζεύγη μετατοπίσεων, τα οποία παράχθηκαν από μια γεννήτρια τυχαίων αριθμών, κανονικά κατανομημένων και με τυπική

Πίνακας 4.1: ΑΠΟΔΟΣΗ ΠΡΟΤΕΙΝΟΜΕΝΟΥ ΣΧΗΜΑΤΟΣ ΑΝΤΙΣΤΟΙΧΙΣΗΣ ΜΕ ΑΚΡΙΒΕΙΑ ΕΙΚΟΝΟΣΤΟΙΧΕΙΟΥ (ΑΞΟΝΑΣ Χ, ΑΞΟΝΑΣ Υ)

Εικόνες	Μέσο Σφάλμα	STD	Μέγ. Σφάλμα	Ελ. Σφάλμα
<i>Gaussian</i>	$(-3.6170 \cdot 10^{-5}, -1.6908 \cdot 10^{-4})$	(0.0059, 0.0064)	(0.0208, 0.0222)	$(-0.0245, -0.0290)$
<i>Ημιτονοειδής</i>	$(-9.5206 \cdot 10^{-6}, -3.0403 \cdot 10^{-5})$	(0.0026, 0.0027)	(0.0064, 0.0080)	$(-0.0083, -0.0080)$
<i>Lenna</i>	$(-0.0040, -9.5171 \cdot 10^{-5})$	(0.0174, 0.0141)	(0.0416, 0.0480)	$(-0.0535, -0.0431)$
<i>Δειγμ. Lenna</i>	$(-0.0015, 4.7575 \cdot 10^{-5})$	(0.0317, 0.0248)	(0.1111, 0.1094)	$(-0.0861, -0.0914)$

Πίνακας 4.2: ΑΠΟΔΟΣΗ ΤΗΣ ΜΕΘΟΔΟΥ ΠΟΥ ΠΑΡΟΥΣΙΑΖΕΤΑΙ ΣΤΗΝ ΕΡΓΑΣΙΑ [145] (ΑΞΟΝΑΣ Χ, ΑΞΟΝΑΣ Υ)

Gaussian	Μέσο Σφάλμα	STD	Μέγ. Σφάλμα	Ελ. Σφάλμα
<i>ακτίνα=1</i>	(0.0018, 0.0034)	(0.0268, 0.0336)	(0.1542, 0.1934)	$(-0.1237, -0.1907)$
<i>ακτίνα=2</i>	$(-0.0195, -0.0368)$	(0.2935, 0.3723)	(1.3723, 2.1514)	$(-1.7249, -2.1832)$
Ημιτονοειδής	Μέσο Σφάλμα	STD	Μέγ. Σφάλμα	Ελ. Σφάλμα
<i>ακτίνα=3</i>	$(-0.0200, -0.0128)$	(0.1015, 0.1005)	(0.1092, 0.1092)	$(-0.1753, -0.1753)$
<i>ακτίνα=4</i>	$(-0.0192, -0.0151)$	(0.0601, 0.0599)	(0.0726, 0.0726)	$(-0.2067, -0.2067)$
<i>ακτίνα=5</i>	$(-0.0180, -0.0196)$	(0.4702, 0.4681)	(0.0714, 0.0714)	$(-9.3538, -9.3176)$
Lenna	Μέσο Σφάλμα	STD	Μέγ. Σφάλμα	Ελ. Σφάλμα
<i>ακτίνα=4</i>	$(-0.4577, -0.2020)$	(1.1021, 1.3255)	(1.6981, 3.3907)	$(-3.9820, -5.4883)$
<i>ακτίνα=7</i>	$(-0.5457, 0.1180)$	(1.1800, 3.4451)	(3.1791, 11.5689)	$(-5.2417, -8.9678)$
Δειγμ. Lenna	Μέσο Σφάλμα	STD	Μέγ. Σφάλμα	Ελ. Σφάλμα
<i>ακτίνα=1</i>	$(-0.0109, -0.0044)$	(0.0508, 0.0264)	$(2.1742 \cdot 10^{-5}, 0.0044)$	$(-0.5032, -0.2707)$
<i>ακτίνα=2</i>	(0.0023, -0.0015)	(0.0166, 0.0462)	(0.1722, 0.2928)	$(-0.0934, -0.4896)$
<i>ακτίνα=3</i>	$(4.6096 \cdot 10^{-4}, 7.49 \cdot 10^{-6})$	(0.0159, 0.0287)	(0.1267, 0.1804)	$(-0.1411, -0.3010)$
<i>ακτίνα=4</i>	$(-0.0556, -0.0398)$	(0.7542, 0.8621)	(0.0978, 4.3022)	$(-13.8289, -16.1364)$

απόκλιση ίση με 10. Για τις δύο συγκρινόμενες μεθόδους (την προτεινόμενη και αυτή της εργασίας [145]) υπολογίσαμε το σφάλμα ανάμεσα στις πραγματικές μετατοπίσεις και στις εκτιμήσεις τους. Στον Πίνακα 4.1, φαίνονται για την προτεινόμενη τεχνική οι μέσες τιμές του σφάλματος, οι τυπικές αποκλίσεις, καθώς και οι μέγιστες και ελάχιστες τιμές σφάλματος. Στην εργασία [145], η μετατόπιση υπολογίζεται με βάση μια περιοχή των χαμηλών συχνοτήτων της διαφοράς φάσης ανάμεσα στις συγκρινόμενες εικόνες. Κάναμε πειράματα για πολλά διαφορετικά μεγέθη αυτής της περιοχής χαμηλών συχνοτήτων, με την ακτίνα να κυμαίνεται από 1 έως 50. Υπολογίσαμε το σφάλμα για όλες τις ακτίνες και τα αποτελέσματα φαίνονται στο Σχήμα 4.2. Πρέπει να σημειωθεί ότι το σφάλμα παραμένει σχετικά μικρό για μικρές ακτίνες, ενώ για μεγαλύτερες αυξάνεται σημαντικά. Στον Πίνακα 4.2 φαίνονται κάποια αναλυτικά αποτελέσματα για διαφορετικά, μικρά μεγέθη ακτίνας.

Τα αποτελέσματα δείχνουν ότι η προτεινόμενη τεχνική είναι πιο ακριβής, συγκρινόμενη με αυτή της εργασίας [145], η οποία είναι πολύ ευαίσθητη στο μέγεθος της ακτίνας. Επιπλέον, πρέπει να σημειωθεί ότι στην πράξη δεν ξέρουμε ποτέ ποιο είναι το κατάλληλο μέγεθος αυτής

της ακτίνας. Από την άλλη μεριά, η προτεινόμενη τεχνική είναι πιο ανθεκτική, πολύ ακριβής και αποδίδει πολύ καλά και στην περίπτωση των εικόνων που παράγονται με υπο-δειγματοληψία.

4.4 Συμπεράσματα, μελλοντικές ερευνητικές κατευθύνσεις

Στο κεφάλαιο αυτό παρουσιάσαμε μια τεχνική, η οποία παρέχει ακρίβεια υποδιαίρεσης εικονοστοιχείου και βασίζεται στη μεγιστοποίηση του συντελεστή συσχέτισης. Η τεχνική αυτή δεν απαιτεί ανακατασκευή των τιμών της έντασης και παρέχει μια λύση κλειστού τύπου για την εκτίμηση της μετατόπισης. Η απόδοση της τεχνικής αυτής είναι πολύ καλύτερη σε σχέση με υπάρχουσες τεχνικές και έχει μελετηθεί για την περίπτωση καθολικών μετατοπίσεων. Στο μέλλον, η προτεινόμενη τεχνική θα μπορούσε να μελετηθεί για την περίπτωση που υπάρχουν μετατοπίσεις ανάμεσα στις συγκρινόμενες εικόνες, οι οποίες διαφοροποιούνται τοπικά.

Κεφάλαιο 5

Προσαρμοστική υδατογράφηση με “τυφλή” ανίχνευση

Από αυτό το κεφάλαιο και στο εξής θα ασχοληθούμε με το πρόβλημα της ψηφιακής υδατογράφησης εικόνων. Όπως έχει ήδη αναφερθεί στο κεφάλαιο 1, το ζητούμενο στο πρόβλημα της ψηφιακής υδατογράφησης εικόνων είναι η ένθεση ενός αδιόρατου υδατογραφήματος στην εικόνα, το οποίο ταυτόχρονα πρέπει να είναι ανθεκτικό σε προσπάθειες αφαίρεσής του (επιθέσεις). Στο κεφάλαιο αυτό προτείνουμε μια νέα χωρική μάσκα, η οποία λαμβάνει υπόψη της τις ιδιότητες του συστήματος ανθρώπινης όρασης (*HVS-Human Visual System*). Η μάσκα αυτή κατασκευάζεται με βάση την τοπική διασπορά του σφάλματος πρόβλεψης της αρχικής εικόνας και επιτρέπει την ένθεση υδατογραφημάτων μεγαλύτερης ενέργειας σε σχέση με τις υπάρχουσες τεχνικές. Επιπλέον, προτείνουμε ένα βελτιωμένο σύστημα ανίχνευσης υδατογραφήματος, το οποίο είναι “τυφλό”, υπό την έννοια ότι δε χρειάζεται γνώση της αρχικής εικόνας κατά τη φάση της ανίχνευσης. Η μετρική ομοιότητας που χρησιμοποιείται στον ανιχνευτή είναι η κανονικοποιημένη συσχέτιση ανάμεσα στο αναπαραγόμενο υδατογράφημα και το σφάλμα πρόβλεψης της υδατογραφημένης και πιθανώς αλλοιωμένης εικόνας (αντί για την ίδια την εικόνα). Η βελτίωση που επιτυγχάνεται στον ανιχνευτή έχει αποδειχτεί θεωρητικά και έχει επιβεβαιωθεί από μεγάλο αριθμό πειραμάτων [62].

Στη συνέχεια, εξετάζεται η περίπτωση όπου δε γίνεται η υπόθεση ότι το υδατογράφημα με την εικόνα ή διαφορετικά υδατογραφήματα μεταξύ τους είναι εντελώς ασυσχέτιστα μεταξύ τους. Αυτή άλλωστε είναι μια ιδανική υπόθεση και στην πράξη πάντα υπάρχει μια μικρή συσχέτιση ανάμεσα στις προαναφερθείσες ακολουθίες. Θα αποδείξουμε λοιπόν ότι σε αυτήν την περίπτωση το προτεινόμενο σχήμα ένθεσης και ανίχνευσης υδατογραφημάτων σε εικόνες είναι σημαντικά καλύτερο σε σχέση με υπάρχουσες τεχνικές. Θα εξετάσουμε τις περιπτώσεις που δεν υπάρχει επίθεση, που έχει προστεθεί λευκός Gaussian θόρυβος στην υδατογραφημένη

εικόνα, καθώς και την περίπτωση που αυτή έχει φιλτραριστεί με ένα γραμμικό φίλτρο [63]. Η θεωρητική ανάλυση που αφορά την απόδοση του ανιχνευτή είναι ανεξάρτητη της χρησιμοποιούμενης μάσκας και μπορεί να χρησιμοποιηθεί για κάθε τεχνική υδατογράφησης βασισμένη σε ένθεση με χωρική μάσκα. Επιπλέον, θα υπολογιστούν κατάλληλες εκφράσεις, οι οποίες επιτρέπουν τον καθορισμό κατωφλίων ανίχνευσης, προσαρμοζόμενων στις συνθήκες των επιθέσεων.

5.1 Εισαγωγή

Η προστασία των πνευματικών δικαιωμάτων και η δυνατότητα απόδειξης της αυθεντικότητας των ψηφιακών δεδομένων έχουν γίνει αντικείμενο εντατικής έρευνας τα τελευταία χρόνια. Η ψηφιακή υδατογράφηση έχει σαν στόχο την ένθεση χρήσιμης πληροφορίας στα δεδομένα με τέτοιο τρόπο, ώστε να είναι δύσκολο να απομακρυνθεί. Αυτή η πληροφορία, η οποία είναι γνωστή ως υδατογράφημα, αποδεικνύει την ταυτότητα του ιδιοκτήτη των δεδομένων. Είναι ένα ψευδοτυχαίο σήμα που παράγεται με βάση ένα κλειδί, το οποίο πρέπει να έχει ασφαλές μήκος, προκειμένου να είναι ανθεκτικό σε κρυπτογραφικές επιθέσεις. Το υδατογράφημα που εισάγεται σε μια εικόνα πρέπει να μην επηρεάζει την ποιότητα της εικόνας, ενώ ταυτόχρονα πρέπει να είναι ανθεκτικό σε επιθέσεις. Προφανώς, ένα υδατογράφημα μεγάλης ενέργειας είναι πιο ανθεκτικό από ένα υδατογράφημα μικρής ενέργειας. Μια αποδεκτή όμως τιμή για την ενέργεια του υδατογραφήματος εξαρτάται από τη χωρητικότητα του καναλιού (της αρχικής εικόνας). Η χωρητικότητα της εικόνας καθορίζεται από το ποσό της πληροφορίας που μπορεί να εισαχθεί σε μια εικόνα, χωρίς να προκαλεί ορατές αλλοιώσεις.

Ένας αποτελεσματικός τρόπος για να αυξηθεί η ανθεκτικότητα ενός υδατογραφήματος, χωρίς να υποβαθμιστεί η ποιότητα της εικόνας, είναι να αυξηθεί η ενέργειά του με ένα χωρικά προσαρμοστικό τρόπο, χρησιμοποιώντας μια μάσκα που να εκμεταλλεύεται τις ιδιότητες του συστήματος ανθρώπινης όρασης [28]. Το σύστημα ανθρώπινης όρασης είναι λιγότερο ευαίσθητο σε αλλοιώσεις γύρω από ακμές ή από περιοχές με πολλές λεπτομέρειες (με έντονη υφή). Στην εργασία [148] προτείνεται η χρήση μιας μάσκας βασισμένης στις τοπικές ιδιότητες της εικόνας. Στην εργασία [76], το υδατογράφημα προστίθεται στο μπλε κανάλι, λαμβάνοντας υπόψη το γεγονός ότι το μάτι είναι λιγότερο ευαίσθητο σε αυτό το χρωματικό κανάλι. Στην [26], το υδατογράφημα προστίθεται σε έναν αριθμό συντελεστών του μετασχηματισμού DCT από τις χαμηλές συχνότητες και η ενέργειά του προσαρμόζεται ανάλογα με τις τιμές των συντελεστών. Στην εργασία [102], προτείνεται ένα εναλλακτικό σχήμα υδατογράφησης στο πεδίο των συχνοτήτων, το οποίο λαμβάνει υπόψη του περιορισμούς του χωρικού πεδίου. Περισσότερες αναφορές σχετικά με τεχνικές υδατογράφησης που βασίζονται σε μάσκες στο

χωρικό πεδίο και στο πεδίο των συχνοτήτων μπορεί να βρει κανείς στα άρθρα [79] και [156].

Στο νέο σχήμα υδατογράφησης που παρουσιάζεται σε αυτό το κεφάλαιο, η μάσκα κατασκευάζεται με βάση την τοπική διασπορά του σφάλματος πρόβλεψης της αρχικής εικόνας. Το σφάλμα πρόβλεψης ταιριάζει πολύ καλά με το σύστημα της ανθρώπινης όρασης, καθώς τα σφάλματα είναι μικρότερα σε ομαλές περιοχές και μεγαλύτερα σε περιοχές με ακμές και έντονη υφή. Όσο αφορά τώρα τη διαδικασία ανίχνευσης, αυτή γίνεται “τυφλά”, δηλαδή δεν απαιτείται η γνώση της αρχικής εικόνας, και το μόνο που χρειάζεται είναι το κλειδί υδατογράφησης. Επιπλέον, η μετρική ομοιότητας που χρησιμοποιείται στον ανιχνευτή είναι η κανονικοποιημένη συσχέτιση ανάμεσα στο υδατογράφημα και στο σφάλμα πρόβλεψης της εικόνας που φτάνει στο δέκτη (της υδατογραφημένης και πιθανώς αλλοιωμένης από επιθέσεις εικόνας). Συνήθως, η ανίχνευση γίνεται υπολογίζοντας τη συσχέτιση ανάμεσα στο υδατογράφημα και την εικόνα που είναι διαθέσιμη στο δέκτη.

Οι τροποποιήσεις που περιλαμβάνει το προτεινόμενο σχήμα έχουν ως αποτέλεσμα σημαντικά βελτιωμένη απόδοση σε σχέση με υπάρχουσες τεχνικές ένθεσης και ανίχνευσης υδατογραφήματος. Η βελτίωση στην απόδοση αποδεικνύεται θεωρητικά για τις περιπτώσεις που δεν υπάρχει επίθεση, που προστίθεται θόρυβος στην υδατογραφημένη εικόνα και που η υδατογραφημένη εικόνα περνάει από ένα γραμμικό φίλτρο. Η ανθεκτικότητα της προτεινόμενης τεχνικής σε διάφορα είδη επιθέσεων καταδεικνύεται μέσω εξαντλητικών πειραμάτων [62]. Στη συνέχεια, για την περίπτωση των πραγματικών δεδομένων, όπου δε γίνεται η ιδανική υπόθεση της μη συσχέτισης των εμπλεκόμενων στο πρόβλημα ακολουθιών (εικόνας, υδατογραφήματων), χρησιμοποιούμε σαν μετρική ομοιότητας την κανονικοποιημένη συσχέτιση ανάμεσα στο υδατογράφημα με τη μάσκα και στο σφάλμα πρόβλεψης της λαμβανόμενης εικόνας. Η ανάλυση αφορά τη γενική περίπτωση επίθεσης, που μοντελοποιείται σαν γραμμικό φιλτράρισμα και πρόσθεση θορύβου. Επιπλέον, υπολογίζουμε κατάλληλες εκφράσεις (μέσες τιμές και διασπορές των εξόδων των συγκρινόμενων ανιχνευτών), οι οποίες μπορούν να χρησιμοποιηθούν για τον καθορισμό κατωφλίων ανίχνευσης, προσαρμοσμένων στις συνθήκες των επιθέσεων [63].

5.2 Διατύπωση του προβλήματος

Έστω x μια αρχική εικόνα και w_0 το υδατογράφημα, το οποίο είναι ένα ψευδοτυχαίο σήμα με μέση τιμή μηδέν και διασπορά $\sigma_{w_0}^2$. Το υδατογράφημα είναι ίδιου μεγέθους με την εικόνα και σε πρώτη φάση θεωρούμε ότι είναι εντελώς ασυσχέτιστο με την εικόνα. Υιοθετώντας το προσθετικό μοντέλο υδατογράφησης, η υδατογραφημένη εικόνα δίνεται από την ακόλουθη σχέση

$$y = x + w_0. \quad (5.2.1)$$

Αν χρησιμοποιηθεί μια χωρική μάσκα για την ένθεση του υδατογραφήματος, τότε η υδατογραφημένη εικόνα μπορεί να γραφτεί σαν

$$y = x + M \odot w_0 \quad (5.2.2)$$

όπου M είναι η χρησιμοποιούμενη μάσκα και ο τελεστής $\odot$ δηλώνει πολλαπλασιασμό σημείο προς σημείο. Η ενέργεια του υδατογραφήματος ενσωματώνεται στο w_0 . Το υδατογράφημα στο οποίο έχει εφαρμοστεί η μάσκα, $w_0 \equiv M \odot w_0$, μπορεί εύκολα ναδειχθεί ότι είναι μια λευκή διαδικασία μέσης τιμής μηδέν, ασυσχέτιστη με την αρχική εικόνα x^1 .

5.3 Αντιληπτική μάσκα βασισμένη στο σφάλμα πρόβλεψης

Όπως ήδη αναφέρθηκε στην ενότητα 5.1, η προτεινόμενη αντιληπτική μάσκα βασίζεται στο σφάλμα πρόβλεψης της αρχικής εικόνας. Καταρχήν, υποθέτοντας *στασιμότητα* (*stationarity*), υπολογίζουμε από την αρχική εικόνα, x , το φίλτρο σφάλματος πρόβλεψης που ελαχιστοποιεί την ακόλουθη συνάρτηση κόστους

$$J_{LP} = E\{|x(i, j) - \hat{x}(i, j)|^2\} \quad (5.3.1)$$

όπου $\hat{x}(i, j)$ είναι η προβλεπόμενη τιμή του εικονοστοιχείου $x(i, j)$, η οποία δίνεται από τη σχέση

$$\hat{x}(i, j) = \tilde{\mathbf{a}}_x^T \tilde{\mathbf{x}}(i, j) \quad (5.3.2)$$

όπου $\tilde{\mathbf{a}}_x$ είναι ένα διάνυσμα μήκους $(p^2 - 1)$, το οποίο περιέχει κατά γραμμές τους συντελεστές γραμμικής πρόβλεψης και το διάνυσμα $\tilde{\mathbf{x}}(i, j)$ περιέχει κατά γραμμές τα αντίστοιχα εικονοστοιχεία της μη αιτιατής γειτονιάς του $x(i, j)$, διαστάσεων $p \times p$, εκτός από το κεντρικό στη θέση (i, j) . Η ελαχιστοποίηση της συνάρτησης κόστους που ορίζεται στη σχέση (5.3.1) σε σχέση με το $\tilde{\mathbf{a}}_x$ οδηγεί στο ακόλουθο σύστημα εξισώσεων

$$R_x \tilde{\mathbf{a}}_x = \mathbf{r}_x \quad (5.3.3)$$

όπου R_x είναι ο πίνακας αυτοσυσχέτισης της εικόνας x και $\mathbf{r}_x$ το αντίστοιχο διάνυσμα ετεροσυσχέτισης. Η λύση του συστήματος αυτού παρέχει το γραμμικό φίλτρο πρόβλεψης $\tilde{\mathbf{a}}_x$. Έτσι, η επιθυμητή ακολουθία σφάλματος πρόβλεψης υπολογίζεται από την ακόλουθη σχέση

$$e_x(i, j) = x(i, j) - \tilde{\mathbf{a}}_x^T \tilde{\mathbf{x}}(i, j). \quad (5.3.4)$$

¹Θυμίζουμε εδώ ότι ο συμβολισμός που χρησιμοποιούμε στην υδατογράφιση εικόνων δε συμπίπτει με αυτόν της αντιστοίχισης εικόνων, προκειμένου να είμαστε συμβατοί με τη βιβλιογραφία της κάθε περιοχής. Για παράδειγμα, τα x , y , στην αντιστοίχιση δηλώνουν μετατόπιση, ενώ στην υδατογράφιση χρησιμοποιούνται για την αρχική και την υδατογραφημένη (ή τη λαμβανόμενη) εικόνα, αντίστοιχα.

Η ακολουθία του σφάλματος πρόβλεψης της αρχικής εικόνας μεταβάλλεται χωρικά με τρόπο που ταιριάζει στις ιδιότητες του συστήματος ανθρώπινης όρασης. Έχει μικρότερες τιμές για ομαλές περιοχές της εικόνας (οι οποίες είναι πιο προβλέψιμες) και μεγαλύτερες για περιοχές με ακμές και έντονη υφή (οι οποίες προβλέπονται δυσκολότερα). Η προτεινόμενη μάσκα ορίζεται από την ακόλουθη σχέση

$$M(i, j) = 1 - \frac{1}{1 + \sigma_{e_x}^2(i, j)} \quad (5.3.5)$$

όπου $\sigma_{e_x}^2(i, j)$ είναι η τοπική διασπορά του σφάλματος πρόβλεψης σε μια γειτονιά του εικονοστοιχείου που βρίσκεται στη θέση (i, j) . Ο παραπάνω ορισμός είναι παρόμοιος με τη γνωστή Noise Visibility Function-NVF, η οποία προτείνεται στην εργασία [148]. Η διαφορά είναι ότι στην $NVF(i, j)$ χρησιμοποιείται η τοπική διασπορά της εικόνας, δηλαδή η $\sigma_x^2(i, j)$ αντί για τη $\sigma_{e_x}^2(i, j)$ και καθώς ισχύει ότι $\sigma_{e_x}^2(i, j) < \sigma_x^2(i, j)$ (αποδεικνύεται εύκολα), συμπεραίνουμε ότι $M(i, j) < NVF(i, j)$. Έτσι, χρησιμοποιώντας την προτεινόμενη μάσκα που βασίζεται στο σφάλμα πρόβλεψης (PE-Prediction Error), δίνεται η δυνατότητα για ένθεση υδατογραφήματων μεγαλύτερης ενέργειας και άρα πιο ανθεκτικών.

Η απόδοση της προτεινόμενης μάσκας ελέγχθηκε πειραματικά και κάποια αντιπροσωπευτικά αποτελέσματα φαίνονται στον Πίνακα 5.1 (για την εικόνα “Boat”), όπου για το ίδιο PSNR (Peak Signal to Noise Ratio), η νέα μάσκα επιτρέπει την ένθεση υδατογραφήματων με μεγαλύτερη ενέργεια (τυπική απόκλιση υδατογραφήματος).

5.4 Νέο σχήμα “τυφλής” ανίχνευσης (Ιδανική περίπτωση)

Συνήθως, σε μια διαδικασία “τυφλής” ανίχνευσης υδατογραφήματος χρησιμοποιείται μια μετρική συσχέτισης βασισμένη στη συσχέτιση του υδατογραφήματος με τη λαμβανόμενη εικόνα. Στο προτεινόμενο σχήμα ανίχνευσης, η ακολουθία που συσχετίζεται με το υδατογράφημα είναι το σφάλμα πρόβλεψης της λαμβανόμενης εικόνας, η οποία είναι υδατογραφημένη και πιθανώς αλλοιωμένη από επιθέσεις. Αποδεικνύεται ότι το προτεινόμενο σχήμα είναι στην ουσία μια

Πίνακας 5.1: ΣΥΓΚΡΙΣΗ ΕΝΕΡΓΕΙΑΣ ΥΔΑΤΟΓΡΑΦΗΜΑΤΟΣ ΓΙΑ ΔΙΑΦΟΡΕΤΙΚΕΣ ΜΕΘΟΔΟΥΣ ΕΝΘΕΣΗΣ

Μέθοδος ένθεσης	Boat: Ενέργεια υδατογραφήματος		
Μη προσαρμοστική	5	10	15
Μάσκα NVF	5.6	11.2	17
Μάσκα PE	7.46	14.9	22.5
PSNR (dB)	34.1532	28.1326	24.5576

επέκταση τεχνικών των επικοινωνιών για ανίχνευση σημάτων παρουσία μη λευκού θορύβου. Στην περίπτωση μας, το σήμα που πρόκειται να ανιχνευθεί είναι το υδατογράφημα και ο μη λευκός θόρυβος είναι η εικόνα. Στη συνέχεια, αποδεικνύουμε ότι το τροποποιημένο μέτρο συσχέτισης παρέχει καλύτερα αποτελέσματα σε σχέση με το συμβατικό. Η ανάλυση έχει γίνει για τις περιπτώσεις α) που δεν υπάρχει επίθεση, β) που έχει προστεθεί λευκός θόρυβος, και γ) της επίθεσης με γραμμικό φίλτρο.

5.4.1 Ανίχνευση όταν δεν υπάρχει επίθεση

Αρχικά, θα εξετάσουμε την περίπτωση που δεν έχει χρησιμοποιηθεί κάποια μάσκα για την ένθεση του υδατογραφήματος, η υδατογραφημένη εικόνα δηλαδή είναι η $y = x + w_0$. Το αντίστοιχο σφάλμα πρόβλεψης της υδατογραφημένης εικόνας y θα είναι

$$e_y(i, j) = y(i, j) - \mathbf{a}_y^T \tilde{\mathbf{y}}(i, j). \quad (5.4.1)$$

Ο σκοπός είναι να συγκρίνουμε τη συσχέτιση της εικόνας y και του υδατογραφήματος w_0 , με τη συσχέτιση του σφάλματος πρόβλεψης e_y και του υδατογραφήματος w_0 . Αυτές οι μετρικές συσχέτισης ορίζονται ως

$$C_{y, w_0} = \frac{E[y(i, j)w_0(i, j)]}{\sqrt{E[y(i, j)y(i, j)]}\sqrt{E[w_0(i, j)w_0(i, j)]}} \quad (5.4.2)$$

και

$$C_{e_y, w_0} = \frac{E[e_y(i, j)w_0(i, j)]}{\sqrt{E[e_y(i, j)e_y(i, j)]}\sqrt{E[w_0(i, j)w_0(i, j)]}}. \quad (5.4.3)$$

Ξεκινώντας από τις σχέσεις (5.4.2) και (5.4.3) και κάνοντας μερικές πράξεις μπορεί να αποδειχτεί ότι

$$\text{αν } P_x \geq \sigma_{e_x}^2 + \sigma_{w_0}^2 \|\tilde{\mathbf{a}}_x\|^2, \quad \text{τότε } C_{e_y, w_0} \geq C_{y, w_0} \quad (5.4.4)$$

όπου P_x είναι η ισχύς της αρχικής εικόνας και $\|\cdot\|$ η Ευκλείδεια νόρμα. Δηλαδή, η έξοδος του προτεινόμενου ανιχνευτή είναι μεγαλύτερη από αυτή του κλασικού αν η συνθήκη (η αριστερή ανισότητα) της σχέσης (5.4.4) είναι αληθής, το οποίο συμβαίνει πάντα.

Ας θεωρήσουμε τώρα την περίπτωση που η υδατογραφημένη εικόνα είναι η

$$y = x + u_0 \quad (5.4.5)$$

όπου $u = M \odot w_0$, έχει δηλαδή εφαρμοστεί στο υδατογράφημα μια κατάλληλη μάσκα M . Τότε, ξεκινώντας πάλι από τις σχέσεις (5.4.2) και (5.4.3), καταλήγουμε στο ακόλουθο συμπέρασμα

$$\text{αν } P_x \geq \sigma_{e_x}^2 + P_M \sigma_{w_0}^2 \|\tilde{\mathbf{a}}_x\|^2, \quad \text{τότε } C_{e_y, w_0} \geq C_{y, w_0} \quad (5.4.6)$$

όπου P_M είναι η ισχύς της μάσκας. Ο προτεινόμενος ανιχνευτής έχει καλύτερη απόδοση, όταν ισχύει η αριστερή ανισότητα της σχέσης 5.4.6, το οποίο συμβαίνει στην πράξη.

Αν στη διαδικασία της ανίχνευσης χρησιμοποιήσουμε το υδατογράφημα με τη μάσκα, δηλαδή το u_0 αντί για το w_0 , οι μετρικές συσχέτισης (5.4.2) και (5.4.3) μεταβάλλονται ως εξής

$$C_{y,u_0} = \frac{E[y(i,j)u_0(i,j)]}{\sqrt{E[y(i,j)y(i,j)]}\sqrt{E[u_0(i,j)u_0(i,j)]}} \quad (5.4.7)$$

και

$$C_{e_y,u_0} = \frac{E[e_y(i,j)u_0(i,j)]}{\sqrt{E[e_y(i,j)e_y(i,j)]}\sqrt{E[u_0(i,j)u_0(i,j)]}}. \quad (5.4.8)$$

Ο προτεινόμενος ανιχνευτής έχει καλύτερη απόδοση αν ισχύει η συνθήκη στην παρακάτω σχέση

$$\text{αν } P_x \geq \sigma_{e_x}^2 + P_M \sigma_w^2 \|\tilde{\mathbf{a}}_x\|^2, \quad \text{τότε } C_{e_y,u_0} \geq C_{y,u_0} \quad (5.4.9)$$

η οποία είναι ίδια με αυτή της σχέσης (5.4.6). Εντούτοις, συγκρίνοντας την (5.4.7) με την (5.4.2) και την (5.4.8) με την (5.4.3), προκύπτει ότι

$$\text{αν } P_M \geq \mu_M^2, \quad \text{τότε } C_{y,u_0} \geq C_{y,w_0} \quad \text{και} \quad C_{e_y,u_0} \geq C_{e_y,w_0} \quad (5.4.10)$$

όπου μ_M είναι η μέση τιμή της μάσκας. Από την (5.4.10) συμπεραίνουμε ότι είναι πάντα προτιμότερο να χρησιμοποιούμε στην ανίχνευση το u_0 αντί το w_0 . Για τον υπολογισμό του u_0 χρειαζόμαστε τη μάσκα που χρησιμοποιήθηκε στην ένθεση, η οποία όμως είναι άγνωστη στη φάση της ανίχνευσης, καθώς δε διαθέτουμε την αρχική εικόνα. Μπορούμε, ωστόσο, να έχουμε μια εκτίμησή της, η οποία υπολογίζεται από τη λαμβανόμενη εικόνα y .

5.4.2 Ανίχνευση μετά από πρόσθεση θορύβου

Έστω ότι στην υδατογραφημένη εικόνα προστίθεται λευκός Gaussian θόρυβος n , με μέση τιμή μηδέν και διασπορά σ_n^2 . Τότε η λαμβανόμενη εικόνα είναι η

$$y = x + w_0 + n \quad (5.4.11)$$

και ακολουθώντας αντίστοιχη διαδικασία με αυτή που περιγράφηκε στην ενότητα 5.4.1, οι σχέσεις που προκύπτουν και αντιστοιχούν στις (5.4.4), (5.4.6) και (5.4.9), είναι οι ακόλουθες,

$$\text{αν } P_x \geq \sigma_{e_x}^2 + (\sigma_{w_0}^2 + \sigma_n^2) \|\tilde{\mathbf{a}}_x\|^2, \quad \text{τότε } C_{e_y,w_0} \geq C_{y,w_0} \quad (5.4.12)$$

$$\text{αν } P_x \geq \sigma_{e_x}^2 + (P_M \sigma_{w_0}^2 + \sigma_n^2) \|\tilde{\mathbf{a}}_x\|^2, \quad \text{τότε } C_{e_y,w_0} \geq C_{y,w_0} \quad (5.4.13)$$

και

$$\text{αν } P_x \geq \sigma_{e_x}^2 + (P_M \sigma_{w_0}^2 + \sigma_n^2) \|\tilde{\mathbf{a}}_x\|^2, \quad \text{τότε } C_{e_y, u_0} \geq C_{y, u_0}. \quad (5.4.14)$$

Τα σχόλια που έγιναν παραπάνω για τη σχέση (5.4.10), σχετικά με τη χρήση του u_0 αντί του w_0 στην ανίχνευση, ισχύουν και σε αυτήν την περίπτωση.

5.4.3 Ανίχνευση μετά από γραμμικό φιλτράρισμα

Έστω ότι στην υδατογραφημένη εικόνα $y = x + w_0$ εφαρμόζεται ένα γραμμικό φίλτρο. Τότε, η λαμβανόμενη εικόνα δίνεται από τη σχέση

$$z = \mathbf{h}^T \mathbf{y}(i, j) \quad (5.4.15)$$

όπου το διάνυσμα $\mathbf{h}$ περιέχει τους συντελεστές του γραμμικού φίλτρου, μεγέθους $l \times l$, κατά γραμμές. Υποτίθεται ότι γενικά $l \geq p$. Το σφάλμα πρόβλεψης της λαμβανόμενης εικόνας είναι τώρα

$$e_z(i, j) = z(i, j) - \tilde{\mathbf{a}}_z^T \tilde{\mathbf{z}}(i, j) \quad (5.4.16)$$

και οι μετρικές συσχέτισης που θα χρησιμοποιηθούν για τη σύγκριση της απόδοσης των ανιχνευτών δίνονται από τις σχέσεις

$$C_{z, w_0} = \frac{E[z(i, j)w_0(i, j)]}{\sqrt{E[z(i, j)z(i, j)]} \sqrt{E[w_0(i, j)w_0(i, j)]}} \quad (5.4.17)$$

και

$$C_{e_z, w_0} = \frac{E[e_z(i, j)w_0(i, j)]}{\sqrt{E[e_z(i, j)e_z(i, j)]} \sqrt{E[w_0(i, j)w_0(i, j)]}}. \quad (5.4.18)$$

Οι (5.4.17) και (5.4.18) μπορούν να γραφούν σαν

$$C_{z, w_0} = \frac{h_0 \sigma_{w_0}^2}{\sqrt{P_z} \sqrt{\sigma_{w_0}^2}} \quad (5.4.19)$$

και

$$C_{e_z, w_0} = \frac{(h_0 - \tilde{\mathbf{a}}_z^T \tilde{\mathbf{h}}) \sigma_{w_0}^2}{\sqrt{\sigma_{e_z}^2} \sqrt{\sigma_{w_0}^2}} \quad (5.4.20)$$

όπου P_z είναι η ισχύς της εικόνας z , και $\tilde{\mathbf{h}}$ είναι το αποκομμένο κεντρικό τμήμα, μήκους $(p^2 - 1)$, του γραμμικού φίλτρου $\mathbf{h}$ (εκτός του κεντρικού του στοιχείου h_0). Το προτεινόμενο σχήμα ανίχνευσης είναι καλύτερο από το κλασικό, στην περίπτωση που η κανονικοποιημένη συσχέτιση C_{e_z, w_0} είναι μεγαλύτερη από τη C_{z, w_0} . Αυτό ισχύει στην περίπτωση φίλτρων με ενισχυμένο κέντρο και υπάρχουν πολλά τέτοιου είδους φίλτρα. Περισσότερες λεπτομέρειες θα δούμε στην ενότητα που ακολουθεί, όπου εξετάζεται η περίπτωση της χρησιμοποίησης του u_0 στην ανίχνευση αντί του w_0 (για πραγματικά δεδομένα).

5.5 Νέο σχήμα “τυφλής” ανίχνευσης (Περίπτωση πραγματικών δεδομένων)

Είδαμε ότι αν $P_M > \mu_M^2$, όπου P_M είναι η ισχύς της μάσκας και μ_M η μέση τιμή της, είναι προτιμότερο να χρησιμοποιούμε στην ανίχνευση το υδατογράφημα με τη μάσκα, το u_0 . Υποθέτοντας ότι η λαμβανόμενη εικόνα διατηρεί κατά προσέγγιση της ιδιότητες προβλεψιμότητας της αρχικής εικόνας, μπορούμε να έχουμε μια ικανοποιητική εκτίμηση της μάσκας. Έτσι, λοιπόν, για το προτεινόμενο σχήμα ανίχνευσης θα υπολογίζεται στο εξής η συσχέτιση του σφάλματος πρόβλεψης της λαμβανόμενης εικόνας με το υδατογράφημα στο οποίο έχει εφαρμοστεί η μάσκα. Υπενθυμίζουμε ότι για τον υπολογισμό του σφάλματος πρόβλεψης χρησιμοποιείται ένα δισδιάστατο γραμμικό φίλτρο σφάλματος πρόβλεψης.

Θα μελετήσουμε στη συνέχεια την απόδοση του προτεινόμενου σχήματος σε σχέση με την κλασική ανίχνευση, για την περίπτωση των πραγματικών δεδομένων. Συνήθως γίνονται ιδανικές υποθέσεις όσο αφορά τη διαδικασία ανίχνευσης, δηλαδή κάθε υδατογράφημα θεωρείται ότι είναι εντελώς ασυσχέτιστο με άλλα υδατογραφήματα, καθώς και με την αρχική εικόνα. Ωστόσο, στην πράξη (δηλαδή στην περίπτωση πραγματικών δεδομένων) οι παραπάνω υποθέσεις είναι αληθείς μόνο κατά προσέγγιση. Κανονικά, όλες οι εμπλεκόμενες ποσότητες αυτοσυσχέτισης και ετεροσυσχέτισης πρέπει να λαμβάνονται υπόψη και οι τελεστές *αναμενόμενης τιμής* (*expectation*) πρέπει να αντικατασταθούν με τους μέσους όρους των δειγμάτων. Επίσης, η συσχέτιση ανάμεσα στη λαμβανόμενη εικόνα (ή το σφάλμα πρόβλεψης της λαμβανόμενης εικόνας) και σε ένα υδατογράφημα διαφορετικό από αυτό που έχει προστεθεί στην αρχική εικόνα δε θα θεωρείται πλέον ίση με μηδέν.

Η ανάλυση αφορά τη γενική περίπτωση που η επίθεση αποτελείται από γραμμικό φιλτράρισμα και πρόσθεση θορύβου. Από τις υπολογιζόμενες εκφράσεις προκύπτουν εύκολα οι υπο-περιπτώσεις α) της μη επίθεσης, β) της πρόσθεσης λευκού θορύβου και γ) του γραμμικού φιλτραρίσματος. Η ανάλυσή μας βασίζεται σε δύο βασικές προσεγγίσεις 1) μια ντετερμινιστική προσέγγιση, όπου υποτίθεται ότι οι μη μηδενικές αυτο- και ετερο-συσχετίσεις έχουν άνω φραγμένα μέτρα, και 2) μια στατιστική προσέγγιση, στην οποία οι ποσότητες συσχέτισης θεωρούνται τυχαίες μεταβλητές με Gaussian κατανομή. Στη δεύτερη προσέγγιση, υπολογίζονται οι εκφράσεις για τις μέσες τιμές και τις διασπορές των εμπλεκόμενων τυχαίων μεταβλητών, οι οποίες μπορούν να χρησιμοποιηθούν για τον υπολογισμό κατωφλίων προσαρμοζόμενων στις συνθήκες των επιθέσεων.

Τρία σενάρια ανίχνευσης έχουν μελετηθεί α) η ανίχνευση του υδατογραφήματος w_0 όταν δεν έχει χρησιμοποιηθεί μάσκα στην ένθεση, β) η ανίχνευση του υδατογραφήματος w_0 όταν έχει χρησιμοποιηθεί μάσκα στην ένθεση και γ) η ανίχνευση του υδατογραφήματος u_0 όταν

έχει χρησιμοποιηθεί μάσκα στην ένθεση.

Το σφάλμα πρόβλεψης της υδατογραφημένης εικόνας είναι $e_y(i, j) = y(i, j) - \tilde{\mathbf{a}}_y^T \tilde{\mathbf{y}}(i, j)$. Ο σκοπός είναι να συγκρίνουμε τα μέτρα συσχέτισης όταν είτε η εικόνα είτε το σφάλμα πρόβλεψης χρησιμοποιούνται στη διαδικασία της ανίχνευσης.

5.5.1 Ντετερμινιστική προσέγγιση

Υπενθυμίζουμε ότι η μετρική της κανονικοποιημένης συσχέτισης δύο δισδιάστατων ακολουθιών x_1 και x_2 ορίζεται από την ακόλουθη σχέση

$$C_{x_1, x_2} = \frac{\sum \sum_{i,j=1}^N x_1(i, j) x_2(i, j)}{\sqrt{\sum \sum_{i,j=1}^N x_1(i, j)^2} \sqrt{\sum \sum_{i,j=1}^N x_2(i, j)^2}} \quad (5.5.1)$$

όπου N είναι ο αριθμός των εικονοστοιχείων. Οι μετρικές συσχέτισης ανάμεσα σε διαφορετικά υδατογραφήματα ή ανάμεσα στην εικόνα και ένα υδατογράφημα διαφορετικό από αυτό που έχει εισαχθεί υποτίθεται ότι έχουν μέτρο άνω φραγμένο από μια μικρή θετική σταθερά ε .

Στην περίπτωση γραμμικού φιλτραρίσματος και προσθήκης θορύβου, η εικόνα περιγράφεται από τη σχέση

$$z = \mathbf{h}^T \mathbf{y}(i, j) + n \quad (5.5.2)$$

όπου το διάνυσμα $\mathbf{h}$ περιλαμβάνει τους συντελεστές ενός γραμμικού φίλτρου, μεγέθους $l \times l$, κατά γραμμές. Γενικά, υποθέτουμε ότι $l \geq p$. Το n δηλώνει λευκό προσθετικό Gaussian θόρυβο, με μέση τιμή μηδέν και διασπορά σ_n^2 . Το σφάλμα πρόβλεψης της λαμβανόμενης εικόνας δίνεται από τη σχέση $e_z(i, j) = z(i, j) - \tilde{\mathbf{a}}_z^T \tilde{\mathbf{z}}(i, j)$. Μετά από κάποιες πράξεις προκύπτει ότι

$$\text{αν } |h_0| P_M \sigma_{w_0}^2 L \geq \|\tilde{\mathbf{a}}_z\| \|\tilde{\mathbf{h}}\| P_M \sigma_{w_0}^2 + c, \quad \text{τότε } C_{e_z, u_0} \geq C_{z, u_0} \quad (5.5.3)$$

όπου $\|\cdot\|$ είναι η Ευκλείδεια νόρμα, $L = 1 - \sqrt{\sigma_{e_z}^2 / P_z}$, P_z η ισχύς της z , και $\tilde{\mathbf{h}}$ είναι το αποκομμένο κεντρικό τμήμα, μήκους $(p^2 - 1)$, του γραμμικού φίλτρου $\mathbf{h}$ (εκτός από το κεντρικό σημείο h_0). Το c είναι ένας πολύ μικρός βαθμωτός, ο οποίος υπολογίζεται σύμφωνα με την ακόλουθη σχέση

$$c = \varepsilon \mu_M \left[\left(\sum_i |h_i| + 1 \right) K \|\tilde{\mathbf{a}}_z\| + (|h_0| + 1 + \|\tilde{\mathbf{h}}\| K) L \right] + \varepsilon^2 \|\tilde{\mathbf{h}}\| L K + (A + B) \|\tilde{\mathbf{a}}_z\| \quad (5.5.4)$$

όπου $K = \sqrt{p^2 - 1}$ και

$$A = \sqrt{\sum_{i \neq 0} [\varepsilon^4 (\sum_{i \neq j} |h_j|)^2]}, \quad B = \sqrt{\sum_{i \neq 0} [2\varepsilon^2 P_M \sigma_{w_0}^2 |h_i| (\sum_{i \neq j} |h_j|)]} \quad (5.5.5)$$

Έτσι, αν ισχύει η συνθήκη της σχέσης (5.5.3), τότε η προτεινόμενη μετρική συσχέτισης αποδίδει καλύτερα σε σχέση με την κλασική. Οι εκφράσεις των ειδικών περιπτώσεων του γραμμικού φιλταρίσματος, του θορύβου και της μη επίθεσης προκύπτουν εύκολα από κατάλληλη επιλογή των εμπλεκόμενων παραμέτρων. Στην περίπτωση επίθεσης με γραμμικό φίλτρο αποδεικνύεται ότι η ανισότητα (5.5.3) είναι αληθής αν $|h_0| \geq \|\tilde{\mathbf{a}}_z\| \|\tilde{\mathbf{h}}\|$. Αυτή η συνθήκη ισχύει για φίλτρα που έχουν ενισχυμένο το κεντρικό σημείο τους, όπως το Laplacian, το Gaussian και το Unsharp, ενώ το υδατογράφημα δεν μπορεί να ανιχνευθεί μετά από εφαρμογή φίλτρου μέσου όρου. Καθώς η ανάλυση είναι ανεξάρτητη από τη χρησιμοποιούμενη μάσκα, το αποτέλεσμα αυτό ισχύει για όλους τους τύπους χωρικής μάσκας.

5.5.2 Στοχαστική προσέγγιση

Στη ντετερμινιστική προσέγγιση συγκρίναμε την απόδοση των ανιχνευτών, υποθέτοντας ότι το υδατογράφημα που εμπλέκεται στον υπολογισμό της συσχέτισης είναι το αυθεντικό (αυτό που έχει προστεθεί στην εικόνα). Ωστόσο, στην περίπτωση των πραγματικών δεδομένων η απόδοση του ανιχνευτή εξαρτάται και από τη συσχέτιση της εικόνας (με ή χωρίς προ-λεύκανση) με ένα υδατογράφημα διαφορετικό από αυτό που έχει προστεθεί στην αρχική εικόνα. Στη στατιστική προσέγγιση που ακολουθεί εξετάζουμε και τις δύο περιπτώσεις και υπολογίζουμε τις αντίστοιχες εκφράσεις για τη μέση τιμή και τη διασπορά των ποσοτήτων συσχέτισης που προκύπτουν.

Το αυθεντικό υδατογράφημα με τη μάσκα συμβολίζεται με u_0 , ενώ κάθε άλλο υδατογράφημα (από μια κατάλληλα σχεδιασμένη τράπεζα υδατογραφημάτων) στο οποίο έχει εφαρμοστεί η μάσκα συμβολίζεται με u_1 . Ας ορίσουμε τις ακόλουθες ποσότητες:

$$k_0 = \sqrt{P_z \sigma_{u_0}^2}, \quad k_1 = \sqrt{P_z \sigma_{u_1}^2}, \quad l_0 = \sqrt{\sigma_{e_z}^2 \sigma_{u_0}^2}, \quad l_1 = \sqrt{\sigma_{e_z}^2 \sigma_{u_1}^2}$$

$$r_{M^2}(i) = E[M^2(n)M^2(n+i)].$$

Επίσης, $\mathbf{a}_z$ είναι ένα διάνυσμα μήκους p^2 το οποίο περιέχει τους συντελεστές γραμμικής πρόβλεψης κατά γραμμές και a_i είναι ένας συντελεστής του φίλτρου σφάλματος πρόβλεψης.

Οι αναλυτικές εκφράσεις για τις μέσες τιμές και τις διασπορές των εξεταζόμενων συσχετίσεων, οι οποίες δίνουν τη δυνατότητα υπολογισμού κατάλληλων κατωφλίων, φαίνονται στη συνέχεια

$$\mu_{C_{z,u_1}}, \mu_{C_{e_z,u_1}} = 0, \quad \mu_{C_{z,u_0}} = \frac{h_0 P_M \sigma_{w_0}^2}{k_0}, \quad \mu_{C_{e_z,u_0}} = \frac{\mathbf{a}_z^T \mathbf{h} P_M \sigma_{w_0}^2}{l_0}$$

$$\sigma_{C_{z,u_1}}^2 = \frac{1}{N k_1} \sigma_{w_1}^2 [P_M \mathbf{h}^T R_x \mathbf{h} + (h_0^2 r_{M^2}(0) + \sum_{i \neq 0} h_i^2 r_{M^2}(i)) \sigma_{w_0}^2 + P_M \sigma_n^2] \quad (5.5.6)$$

$$\sigma_{C_{ez,u_1}}^2 = \frac{1}{Nl_1} \sigma_{w_1}^2 [P_M \mathbf{g}^T R'_x \mathbf{g} + C \sigma_{w_0}^2 + P_M \sigma_n^2 \|\mathbf{a}_z\|^2]$$

$$\sigma_{C_{z,u_0}}^2 = \frac{1}{Nk_0} \sigma_{w_0}^2 [P_M \mathbf{h}^T R_x \mathbf{h} + (3h_0^2 r_{M^2}(0) + \sum_{i \neq 0} h_i^2 r_{M^2}(i) + \frac{1}{N} h_0^2 \sum_{i \neq 0} r_{M^2}(i) - N h_0^2 P_M^2) \sigma_{w_0}^2 + P_M \sigma_n^2]$$

$$\sigma_{C_{ez,u_0}}^2 = \frac{1}{Nl_0} \sigma_{w_0}^2 [P_M \mathbf{g}^T R'_x \mathbf{g} + (D + \frac{1}{N} (\mathbf{a}_z^T \mathbf{h})^2 \sum_{i \neq 0} r_{M^2}(i) - N P_M^2 (\mathbf{a}_z^T \mathbf{h})^2) \sigma_{w_0}^2 + P_M \sigma_n^2 \|\mathbf{a}_z\|^2]$$

όπου R_x είναι ο πίνακας ετεροσυσχέτισης της εικόνας x , $\mathbf{g}$ είναι το αποτέλεσμα της συνέλιξης ανάμεσα στο φίλτρο σφάλματος πρόβλεψης και στο φίλτρο που χρησιμοποιείται για την επίθεση (με μήκος $2p - 1$) και R'_x είναι ο πίνακας ετεροσυσχέτισης αντίστοιχων διαστάσεων. Οι εκφράσεις για τις υποπεριπτώσεις μπορούν να προκύψουν όπως και στη ντετερμινιστική προσέγγιση, παραλείποντας το θόρυβο ($\sigma_n^2 = 0$), ή/και θέτοντας το h_0 ίσο με ένα και τα h_i (για $i \neq 0$) ίσα με το μηδέν. Για παράδειγμα, μετά από επίθεση με θόρυβο, η αντίστοιχη έκφραση της (5.5.6) είναι

$$\sigma_{C_{z,u_1}}^2 = \frac{1}{Nk_1} \sigma_{w_1}^2 (P_M P_x + r_{M^2}(0) \sigma_{w_0}^2 + P_M \sigma_n^2).$$

Οι εκφράσεις για τα C και D εξαρτώνται από το φίλτρο και παραθέτονται εδώ για την περίπτωση που $h_0 = 1$ και $h_i = 0$ (για $i \neq 0$):

$$C = r_{M^2}(0) + \sum_{i \neq 0} a_i^2 r_{M^2}(i), \quad D = 3r_{M^2}(0) + \sum_{i \neq 0} a_i^2 r_{M^2}(i)$$

Με βάση τις μέσες τιμές και τις διασπορές που υπολογίστηκαν μπορούμε να αποδείξουμε ότι ο προτεινόμενος ανιχνευτής είναι καλύτερος από τον κλασικό. Επιπλέον, μπορούν να καθοριστούν κατάλληλες τιμές για τα κατώφλια ανίχνευσης.

5.6 Πειραματικά αποτελέσματα

Καταρχήν θα παρουσιάσουμε κάποια αντιπροσωπευτικά αποτελέσματα για τις εικόνες “Clock” (διαστάσεων 375×500) και Lenna (διαστάσεων 256×256). Πρώτα υπολογίστηκε η προτεινόμενη αντιληπτική μάσκα για κάθε εικόνα και στη συνέχεια πολλαπλασιάστηκε σημείο προς σημείο με το υδατογράφημα που πρόκειται να εισαχθεί στην κάθε εικόνα και οποίο έχει παράγοντα ενέργειας (τυπική απόκλιση) ίσο με 5. Τα βήματα της ένθεσης φαίνονται στο Σχήμα 5.1

(α) Αρχική εικόνα

(β) Σφάλμα Πρόβλεψης

(γ) Μάσκα βασισμένη στο σφάλμα πρόβλεψης

(δ) Υδατογραφημένη εικόνα

Σχήμα 5.1: Τα βήματα της ένθεσης υδατογραφήματος

(α) Συμπίεση Wavelet, QF:10, DD

(β) Συμπίεση Wavelet, QF:10, PED

(γ) Wiener, DD

(δ) Wiener, PED

Σχήμα 5.2: Κλασική και προτεινόμενη ανίχνευση

Πίνακας 5.2: CLOCK - ΑΠΟΚΡΙΣΕΙΣ ΑΝΙΧΝΕΥΤΩΝ ΓΙΑ ΔΙΑΦΟΡΕΤΙΚΕΣ ΕΠΙΘΕΣΕΙΣ (ΧΡΗΣΙΜΟΠΟΙΕΙΤΑΙ ΤΟ u_0). Η ΠΑΥΛΑ (-) ΣΗΜΑΙΝΕΙ ΟΤΙ ΗΤΑΝ ΑΔΥΝΑΤΗ Η ΑΝΙΧΝΕΥΣΗ.

AWGN	30dB	20dB	10dB	0dB
<i>DD</i>	0.0389	0.0389	0.0376	0.0255
<i>PED</i>	0.3879	0.2486	0.1037	0.0312
Συμπίεση JPEG	QF:80	QF:50	QF:15	QF:10
<i>DD</i>	0.0209	0.0140	0.0089	-
<i>(PED)</i>	0.1188	0.0495	0.0190	0.0131
Συμπίεση Wavelet	QF:80	QF:50	QF:20	QF:10
<i>DD</i>	0.0366	0.0282	0.0132	-
<i>PED</i>	0.3832	0.2933	0.0756	0.0232
Μείωση Χρώματος	Dithering	Κατωφλίωση		
<i>DD</i>	0.0452	0.0464		
<i>PED</i>	0.0543	0.1224		
Sampledownup	Περίπτωση 1	Περίπτωση 2		
<i>DD</i>	0.0234	0.0109		
<i>PED</i>	0.1734	0.0280		
Wiener Filtering	3×3	5×5		
<i>DD</i>	0.0138	0.0125		
<i>PED</i>	0.1390	0.1666		
Trimmedmean	3×3	5×5		
<i>DD</i>	-	-		
<i>PED</i>	0.0294	0.0516		
Median	2×2	3×3	4×4	
<i>DD</i>	0.0127	0.0095	-	
<i>PED</i>	0.0710	0.0524	-	
Άλλα Φίλτρα	Laplacian	Gaussian	Unsharp	
<i>DD</i>	0.0273	0.0255	0.1323	
<i>PED</i>	0.3602	0.3940	0.4521	

για την εικόνα “Clock”. Στη συνέχεια η υδατογραφημένη εικόνα δέχεται επιθέσεις. Αρκετές από αυτές προέρχονται από το Checkmark [149]. Τα αποτελέσματα φαίνονται στον Πίνακα 5.2 για την εικόνα “Clock” και στον Πίνακα 5.3 για τη Lenna και έχουν προκύψει εφαρμόζοντας την κλασική διαδικασία ανίχνευσης (Direct Detection-DD) στις λαμβανόμενες εικόνες, καθώς και την προτεινόμενη που βασίζεται στο σφάλμα πρόβλεψης (Prediction Error based Detection-PED). Χρησιμοποιήσαμε μια τράπεζα από 1000 διαφορετικά υδατογραφήματα, με το σωστό να βρίσκεται στη θέση 500. Η ανίχνευση έχει καλύτερα αποτελέσματα για την PED και είναι εφικτή ακόμα και σε περιπτώσεις που η DD δεν μπορεί να βρει το υδατογράφημα. Δύο χαρακτηριστικά παραδείγματα φαίνονται στο Σχήμα 5.2.

Στη συνέχεια, δίνουμε κάποια επιπλέον αποτελέσματα για την εικόνα “Boat”, διαστάσεων

Πίνακας 5.3: LENNA - ΑΠΟΚΡΙΣΕΙΣ ΑΝΙΧΝΕΥΤΩΝ ΓΙΑ ΔΙΑΦΟΡΕΤΙΚΕΣ ΕΠΙΘΕΣΕΙΣ (ΧΡΗΣΙΜΟΠΟΙΕΙΤΑΙ ΤΟ u_0). Η ΠΑΥΛΑ (-) ΣΗΜΑΙΝΕΙ ΟΤΙ ΗΤΑΝ ΑΔΥΝΑΤΗ Η ΑΝΙΧΝΕΥΣΗ.

AWGN	30dB	20dB	10dB	0dB
<i>DD</i>	0.0344	0.0349	0.0348	0.0290
<i>PED</i>	0.4126	0.2438	0.1001	0.0387
Συμπίεση JPEG	QF:80	QF:50	QF:15	QF:10
<i>DD</i>	0.0214	0.0137	-	-
<i>(PED)</i>	0.1893	0.0819	0.0210	-
Συμπίεση Wavelet	QF:80	QF:50	QF:20	QF:10
<i>DD</i>	0.0326	0.0243	0.0131	-
<i>PED</i>	0.4370	0.3070	0.0646	0.0150
Μείωση Χρώματος	Dithering	Κατωφλίωση		
<i>DD</i>	0.0467	0.0441		
<i>PED</i>	0.0615	0.1514		
Sampledownup	Περίπτωση 1	Περίπτωση 2		
<i>DD</i>	0.0204	0.0129		
<i>PED</i>	0.2297	0.0427		
Wiener Filtering	3×3	5×5		
<i>DD</i>	-	-		
<i>PED</i>	0.1283	0.1449		
Trimmedmean	3×3	5×5		
<i>DD</i>	0.0130	0.0128		
<i>PED</i>	0.0546	0.0640		
Median	2×2	3×3	4×4	
<i>DD</i>	-	-	-	
<i>PED</i>	0.0684	0.0694	-	
Άλλα Φίλτρα	Laplacian	Gaussian	Unsharp	
<i>DD</i>	0.0259	0.0226	0.1254	
<i>PED</i>	0.4293	0.5010	0.5127	

470×500 . Αρχικά υπολογίζεται η προτεινόμενη αντιληπτική μάσκα, η οποία πολλαπλασιάζεται με υδατογράφημα με παράγοντα ενέργειας ίσο με 6.3 και προστίθεται στην εικόνα. Ο παράγοντας ενέργειας επιλέχθηκε έτσι ώστε το υδατογράφημα να είναι πρακτικά αόρατο. Για να το πετύχουμε αυτό βασιστήκαμε στη μετρική Watson από το Checkmark [149]. Τα βήματα της ένθεσης για την εξεταζόμενη εικόνα φαίνονται στο Σχήμα 5.3. Στη συνέχεια, η υδατογραφημένη εικόνα δέχτηκε διάφορες επιθέσεις. Τα αποτελέσματα που φαίνονται στον Πίνακα 5.4 προέκυψαν μετά από ανίχνευση με την κλασική μέθοδο (DD) και με την προτεινόμενη (PED). Και για αυτή την περίπτωση χρησιμοποιήθηκε μια τράπεζα από 1000 διαφορετικά υδατογραφήματα, με το σωστό να βρίσκεται στη θέση 500. Είναι προφανές ότι η προτεινόμενη μέθοδος έχει πολύ καλύτερα αποτελέσματα και κάνει εφικτή την ανίχνευση ακόμα και σε περιπτώσεις

(α) Αρχική εικόνα

(β) Σφάλμα πρόβλεψης

(γ) Μάσκα βασισμένη στο σφάλμα πρόβλεψης

(δ) Υδατογραφημένη εικόνα

Σχήμα 5.3: Βήματα υδατογράφησης για την εικόνα “Boat”

που η κλασική μέθοδος αποτυγχάνει.

5.7 Συμπεράσματα, μελλοντικές ερευνητικές κατευθύνσεις

Στο κεφάλαιο αυτό προτάθηκε μια νέα χωρική μάσκα και ένα βελτιωμένο σύστημα ανίχνευσης υδατογραφήματος, το οποίο είναι “τυφλό”, υπό την έννοια ότι δε χρειάζεται γνώση της αρχικής εικόνας κατά τη φάση της ανίχνευσης. Εξετάζεται η ιδανική περίπτωση και η περίπτωση των πραγματικών δεδομένων. Η βελτίωση που επιτυγχάνεται στον ανιχνευτή αποδείχτηκε θεωρητικά και επιβεβαιώθηκε από μεγάλο αριθμό πειραμάτων. Εξετάσαμε τις περιπτώσεις που δεν υπάρχει επίθεση, που έχει προστεθεί λευκός Gaussian θορυβος στην υδατογραφημένη εικόνα και την περίπτωση που αυτή έχει φιλτραριστεί με ένα γραμμικό φίλτρο. Η θεωρητική ανάλυση είναι ανεξάρτητη της χρησιμοποιούμενης μάσκας και μπορεί να χρησιμοποιηθεί για κάθε τεχνική υδατογράφησης βασισμένη σε ένθεση με χωρική μάσκα. Επιπλέον, υπολογίστηκαν κατάλληλες εκφράσεις, οι οποίες επιτρέπουν τον καθορισμό κατωφλίων ανίχνευσης,

Πίνακας 5.4: BOAT - ΑΠΟΚΡΙΣΕΙΣ ΑΝΙΧΝΕΥΤΩΝ ΓΙΑ ΔΙΑΦΟΡΕΤΙΚΕΣ ΕΠΙΘΕΣΕΙΣ (ΧΡΗΣΙΜΟΠΟΙΕΙΤΑΙ ΤΟ u_0). Η ΠΑΥΛΑ (-) ΣΗΜΑΙΝΕΙ ΟΤΙ ΗΤΑΝ ΑΔΥΝΑΤΗ Η ΑΝΙΧΝΕΥΣΗ.

AWGN	30dB	20dB	10dB	0dB
<i>DD</i>	0.0259	0.0258	0.0247	0.0205
<i>PED</i>	0.4033	0.2110	0.0785	0.0291
Συμπίεση JPEG	QF:80	QF:50	QF:15	QF:10
<i>DD</i>	0.0108	-	-	-
<i>(PED)</i>	0.1504	0.0624	0.0437	0.0283
Μείωση Χρώματος	Dithering	Κατωφλίωση		
<i>DD</i>	0.0382	0.0340		
<i>PED</i>	0.0617	0.1553		
Sampledownup	Περίπτωση 1	Περίπτωση 2		
<i>DD</i>	0.0146	-		
<i>PED</i>	0.2405	0.0454		
Wiener Filtering	3 × 3	5 × 5		
<i>DD</i>	-	-		
<i>PED</i>	0.1696	0.1947		
Trimmedmean	3 × 3	5 × 5		
<i>DD</i>	-	-		
<i>PED</i>	0.0540	0.0740		
Median	3 × 3	5 × 5	7 × 7	
<i>DD</i>	-	-	-	
<i>PED</i>	0.0695	0.0234	0.0080	
Άλλα Φίλτρα	Laplacian	Gaussian	Unsharp	
<i>DD</i>	-0.3951	0.0158	0.1099	
<i>PED</i>	-0.4882	0.5198	0.5130	

προσαρμοζόμενων στις συνθήκες των επιθέσεων. Μια πιθανή επέκταση της παραπάνω δουλειάς θα μπορούσε να είναι η θεωρητική μελέτη της απόδοσης του προτεινόμενου σχήματος υδατογράφησης στην περίπτωση μη γραμμικών επιθέσεων. Το πρόβλημα αυτό είναι πολύ δύσκολο καθώς δεν υπάρχει ένα γενικό μοντέλο μη γραμμικής επίθεσης και κάθε είδος τέτοιας επίθεσης πρέπει να μελετηθεί ξεχωριστά.

Κεφάλαιο 6

Αποδοτική υδατογράφηση στο χωρικό πεδίο μέσω μιας νέας αντιληπτικής μάσκας και ενός νέου “τυφλού” σχήματος ανίχνευσης

Σε αυτό το κεφάλαιο θα παρουσιάσουμε μια τεχνική, η οποία έχει ως στόχο τη βελτίωση της υδατογράφησης εικόνων στο χωρικό πεδίο [65]. Ειδικότερα, θα παρουσιάσουμε μια νέα χωρική μάσκα και ένα νέο σχήμα ανίχνευσης υδατογραφημάτων. Η προτεινόμενη χωρική αντιληπτική μάσκα βασίζεται στο σφάλμα πρόβλεψης της αρχικής εικόνας (πριν την υδατογράφηση) και περιγράφει πολύ καλά τις ιδιότητες του συστήματος ανθρώπινης όρασης. Έχει πολύ καλή απόδοση σε σχέση με υπάρχουσες μάσκες, καθώς επιτρέπει την ένθεση υδατογραφημάτων με πολύ μεγάλη ενέργεια, ενώ ταυτόχρονα διατηρεί την ποιότητα της υδατογραφημένης εικόνας σε πολύ καλό επίπεδο. Το νέο “τυφλό” σχήμα ανίχνευσης συνδυάζει μια αποδοτική διαδικασία προ-λεύκανσης και έναν ανιχνευτή βασισμένο στη συσχέτιση. Η προ-λεύκανση γίνεται με το φίλτρο σφάλματος πρόβλεψης *ελαχίστων τετραγώνων* (*Least Squares-LS*) και βελτιώνει σημαντικά την απόδοση του ανιχνευτή. Ο ανιχνευτής που επιλέχθηκε αποδείχτηκε ο πιο κατάλληλος για το πρόβλημά μας. Η βελτιωμένη απόδοση του προτεινόμενου σχήματος ανίχνευσης αποδείχθηκε θεωρητικά και επιβεβαιώθηκε από μεγάλο αριθμό πειραμάτων. Η θεωρητική ανάλυση που αφορά την ανίχνευση είναι ανεξάρτητη από την προτεινόμενη μάσκα και μπορεί να χρησιμοποιηθεί για κάθε τεχνική υδατογράφησης που βασίζεται σε χωρική μάσκα. Πρέπει να σημειωθεί, ωστόσο, ότι ο ανιχνευτής έχει καλύτερη απόδοση αν για την ένθεση

επιλεχθεί η προτεινόμενη μάσκα. Τέλος, υπολογίζουμε αναλυτικούς τύπους, που επιτρέπουν τον υπολογισμό κατωφλίων ανίχνευσης, προσαρμοσμένων στις συνθήκες των επιθέσεων.

6.1 Εισαγωγή

Όπως έχει ήδη αναφερθεί υπάρχουν αντιφατικές απαιτήσεις σε μια διαδικασία υδατογράφησης. Το υδατογράφημα που εισάγεται στην εικόνα δεν πρέπει να επηρεάζει την ποιότητα της εικόνας, ενώ την ίδια στιγμή πρέπει να είναι ανθεκτικό σε επιθέσεις. Προφανώς, ένα υδατογράφημα με μεγάλη ενέργεια είναι πιο ανθεκτικό από ένα υδατογράφημα με μικρή ενέργεια. Το πόσο αποδεκτή είναι όμως η τιμή αυτής της ενέργειας εξαρτάται από τη χωρητικότητα του καναλιού, το οποίο στην περίπτωση μας είναι η αρχική εικόνα. Η χωρητικότητα της εικόνας καθορίζεται από το ποσό της πληροφορίας που μπορεί να εισαχθεί στην εικόνα χωρίς να προκληθούν ορατές αλλοιώσεις. Έτσι, πρέπει να γίνει ένας συμβιβασμός ανάμεσα στην ανθεκτικότητα του υδατογραφήματος και στην ποιότητα της εικόνας. Η χρήση αντιληπτικών масκών, οι οποίες λαμβάνουν υπόψη τους τις ιδιότητες του συστήματος ανθρώπινης όρασης, αποδεικνύεται ότι είναι μια αποδοτική λύση στο πρόβλημα αυτό [26].

Να θυμίσουμε ότι η ανθρώπινη όραση είναι λιγότερο ευαίσθητη σε αλλοιώσεις γύρω από ακμές ή σε περιοχές με έντονη υφή. Έτσι, στην εργασία [148] προτείνεται μια μάσκα, η οποία βασίζεται στις τοπικές ιδιότητες της εικόνας. Στην εργασία [76] το υδατογράφημα προστίθεται στο μπλε κανάλι, όπου το μάτι είναι λιγότερο ευαίσθητο. Στην εργασία [26] το υδατογράφημα προστίθεται σε έναν αριθμό συντελεστών του μετασχηματισμού DCT από τις χαμηλές συχνότητες και προσαρμόζεται στην ενέργειά τους. Στην εργασία [110] χρησιμοποιείται το μοντέλο Watson [151] για ένθεση υδατογραφήματος στο πεδίο των συχνοτήτων. Οι συγγραφείς της εργασίας [102] προτείνουν μια εναλλακτική προσέγγιση για υδατογράφηση στο πεδίο των συχνοτήτων, όπου λαμβάνονται υπόψη περιορισμοί του χωρικού πεδίου. Οι Delaigle, Vleeschouwer και Macq εφαρμόζουν μάσκα στο *θόρυβο περιορισμένου εύρους ζώνης* (*bandlimited noise*) σε περιοχές με έντονη υφή ή γύρω από περιγράμματα, στην εργασία [32]. Πρόσφατα, οι συγγραφείς της εργασίας [78] παρουσίασαν ένα αντιληπτικό μοντέλο, το οποίο λαμβάνει υπόψη του την ευαισθησία της ανθρώπινης όρασης μέσω μιας τοπικής ιστροπικής μετρικής αντίθεσης και ενός μοντέλου μάσκας. Περισσότερες τεχνικές που αφορούν την εφαρμογή масκών μπορούν να βρεθούν στα άρθρα [79] και [156].

Στο κεφάλαιο αυτό θα παρουσιάσουμε μια νέα χωρική αντιληπτική μάσκα, η οποία βασίζεται στο σφάλμα πρόβλεψης ελαχίστων τετραγώνων της αρχικής εικόνας. Το σφάλμα πρόβλεψης ελαχίστων τετραγώνων ταιριάζει πολύ καλά στα χαρακτηριστικά της ανθρώπινης όρασης, καθώς τα σφάλματα αναμένονται να είναι μικρότερα στις ομαλές περιοχές και μεγαλύτερα στις

περιοχές με έντονη υφή και ακμές. Η νέα μάσκα παράγεται έτσι ώστε να βελτιώνει την ανθεκτικότητα του εισερχόμενου υδατογραφήματος στην επίθεση αποθορυβοποίησης (*denoising*). Η επίθεση αποθορυβοποίησης υποτίθεται ότι γίνεται με ένα προσαρμοστικό φίλτρο σφάλματος πρόβλεψης ελαχίστων τετραγώνων. Μια τέτοια επίθεση βασίζεται κατευθείαν στα δεδομένα και δε χρειάζεται να επικαλεστούμε υποθέσεις σχετικά με τις στατιστικές ιδιότητες της εικόνας [52, 137]. Η προτεινόμενη μάσκα έχει πολύ καλύτερη απόδοση σε σχέση με υπάρχουσες μάσκες, οι οποίες χρησιμοποιούνται ευρέως. Επιτρέπει σημαντική αύξηση στην ενέργεια του υδατογραφήματος, ενώ την ίδια στιγμή διατηρεί την ποιότητα της εικόνας. Η πολύ καλή απόδοσή της επιβεβαιώνεται και από τη διαδικασία ανίχνευσης, αφού το υδατογράφημα που εισάγεται χρησιμοποιώντας την προτεινόμενη μάσκα εμφανίζεται να είναι πιο ανθεκτικό.

Όσο αφορά τώρα τη φάση της ανίχνευσης, πρέπει να πούμε ότι έχουν προταθεί πολλοί ανιχνευτές στη βιβλιογραφία, με τον ανιχνευτή που βασίζεται στη συσχέτιση να είναι ο πιο συχνά χρησιμοποιούμενος. Το αν ένας ανιχνευτής είναι βέλτιστος εξαρτάται από τις υποθέσεις που γίνονται για την κατανομή των δεδομένων της αρχικής εικόνας. Για παράδειγμα, οι συντελεστές του μετασχηματισμού DCT (εκτός του DC όρου) μπορούν να προσεγγιστούν αρκετά καλά από τη γενικευμένη Gaussian κατανομή, όπως φαίνεται στις εργασίες [14, 24]. Αυτή η κατανομή χρησιμοποιείται για το σχεδιασμό ανιχνευτών προσθετικών [22, 54], αλλά και πολλαπλασιαστικών υδατογραφημάτων [23]. Στην εργασία [11], όπου ένα πολλαπλασιαστικό υδατογράφημα εισάγεται στο πεδίο του μετασχηματισμού Fourier, χρησιμοποιείται μια Weibull κατανομή για την περιγραφή του πλάτους των συντελεστών του μετασχηματισμού Fourier και την κατασκευή του αντίστοιχου ανιχνευτή. Αυτοί οι ανιχνευτές παρουσιάζουν καλύτερη απόδοση σε σχέση με τον κλασικό ανιχνευτή συσχέτισης, ο οποίος είναι βέλτιστος για δεδομένα που ακολουθούν Gaussian κατανομή.

Σε πολλές περιπτώσεις, όμως, είναι δύσκολο να βρεθεί ένα ικανοποιητικό στατιστικό μοντέλο για τα δεδομένα [55, 96]. Για παράδειγμα, για πραγματικές εικόνες στο χωρικό πεδίο είναι λογικό να υποθέσουμε ότι οι ομαλές περιοχές έχουν Gaussian κατανομή, αλλά οι περιοχές με έντονη υφή ή οι περιοχές που περιέχουν ακμές θα έχουν μια άλλη κατανομή με πιο οξύ *peak* (αιχμή) κοντά στο μηδέν. Σε τέτοιες περιπτώσεις, η πιο συνηθισμένη τακτική είναι να θεωρούμε ένα Gaussian μοντέλο για το θόρυβο (την αρχική εικόνα στην περίπτωση μας) και να χρησιμοποιούμε έναν ανιχνευτή που βασίζεται στη συσχέτιση [55]. Αυτή η τακτική ακολουθείται αρκετές φορές και στις επικοινωνίες απλωμένου φάσματος (*spread spectrum communications*). Εδώ θα παρουσιάσουμε ένα βελτιωμένο ανιχνευτή ο οποίος βασίζεται στη συσχέτιση και στον οποίο ενσωματώνεται ένα αποδοτικό σχήμα που λευκαίνει το θόρυβο (δηλαδή την εικόνα). Ειδικότερα, στη διαδικασία ανίχνευσης υπολογίζεται η κανονικοποιημένη συσχέτιση ανάμεσα στο σφάλμα πρόβλεψης ελαχίστων τετραγώνων της λαμβανόμενης εικόνας

και στο φιλτραρισμένο (με το φίλτρο σφάλματος πρόβλεψης) υδατογράφημα, στο οποίο πρώτα έχει εφαρμοστεί η μάσκα.

Η βελτιωμένη απόδοση του προτεινόμενου ανιχνευτή έχει αποδειχτεί θεωρητικά για τις περιπτώσεις που δεν υπάρχει επίθεση, που προστίθεται θόρυβος και που εφαρμόζεται κάποιο γραμμικό φίλτρο στην υδατογραφημένη εικόνα. Επιπλέον, επιβεβαιώθηκε από μεγάλο αριθμό πειραμάτων. Η θεωρητική ανάλυση είναι ανεξάρτητη της προτεινόμενης μάσκας και οι εκφράσεις που έχουν εξαχθεί μπορούν να χρησιμοποιηθούν για οποιαδήποτε τεχνική βασίζεται σε χωρική μάσκα. Πρέπει να σημειωθεί ωστόσο, ότι η απόδοση του ανιχνευτή είναι καλύτερη για την προτεινόμενη μάσκα. Επιπλέον, για κάθε περίπτωση επίθεσης, υπολογίστηκαν αναλυτικές εκφράσεις για τη μέση τιμή και τη διασπορά του μέτρου συσχέτισης. Αυτές οι εκφράσεις μπορούν να χρησιμοποιηθούν για τον καθορισμό κατωφλίων προσαρμοσμένων στις συνθήκες των επιθέσεων. Εκτεταμένα πειράματα έδειξαν ότι το προτεινόμενο σχήμα υδατογράφησης δεν είναι ανθεκτικό μόνο σε πρόσθεση θορύβου και εφαρμογή γραμμικού φίλτρου αλλά και σε πολλές άλλες επιθέσεις (μη γραμμικές επιθέσεις, συμπίεση, κ.τ.λ.).

6.2 Η προτεινόμενη ένθεση υδατογραφήματος

6.2.1 Διατύπωση προβλήματος

Έστω x η αρχική εικόνα και w_0 το υδατογράφημα, το οποίο είναι ένα ψευδοτυχαίο σήμα με Gaussian κατανομή, μέση τιμή μηδέν και διασπορά $\sigma_{w_0}^2$ ¹. Το υδατογράφημα έχει ίδιες διαστάσεις με την εικόνα και σε πρώτη φάση θεωρούμε ότι είναι εντελώς ασυσχέτιστο με αυτή. Στην πιο απλή περίπτωση, όπου το περιεχόμενο της αρχικής εικόνας δε λαμβάνεται υπόψη κατά τη φάση της ένθεσης υδατογραφήματος (δηλαδή δε χρησιμοποιείται κάποια αντιληπτική μάσκα), η υδατογραφημένη εικόνα γράφεται σαν

$$y = x + w_0. \quad (6.2.1)$$

Αν χρησιμοποιηθεί μια χωρική αντιληπτική μάσκα, M , η υδατογραφημένη εικόνα μπορεί να γραφεί σαν

$$y = x + M \odot w_0 \quad (6.2.2)$$

όπου ο τελεστής $\odot$ δηλώνει πολλαπλασιασμό σημείο προς σημείο. Η ενέργεια του υδατογραφήματος (δηλαδή, η τυπική του απόκλιση) ενσωματώνεται στο w_0 . Καθώς η μάσκα M εξαρτάται μόνο από την εικόνα x , μπορεί εύκολα ναδειχθεί ότι το υδατογράφημα στο οποίο έχει εφαρμοστεί η μάσκα, το $u_0 \equiv M \odot w_0$, είναι μια λευκή διαδικασία μέσης τιμής μηδέν, ασυσχέτιστη με την αρχική εικόνα x .

¹Κάποιοι ορισμοί που υπάρχουν στο κεφάλαιο 5 αναφέρονται και εδώ για διευκόλυνση του αναγνώστη.

Σχήμα 6.1: Γενικό σχήμα αποθορυβοποίησης

6.2.2 Αποθορυβοποίηση εικόνας και αντιληπτική μάσκα

Ας θεωρήσουμε τώρα το πρόβλημα της εκτίμησης του υδατογραφήματος. Το υδατογράφημα στη σχέση (6.2.1) μπορεί να εκτιμηθεί αφαιρώντας από την υδατογραφημένη εικόνα, y , μια εκτίμηση της αρχικής εικόνας, x_{est} . Η εκτίμηση της αρχικής εικόνας αντιστοιχεί στο πρόβλημα της αποθορυβοποίησης της υδατογραφημένης εικόνας. Για την αποθορυβοποίηση μπορεί να χρησιμοποιηθεί ένα γραμμικό φίλτρο, του οποίου η χρουστική απόκριση μπορεί να επιλεχθεί έτσι ώστε να ελαχιστοποιεί κάποια κατάλληλα επιλεγμένη μετρική απόστασης ανάμεσα στην αρχική εικόνα και στην εκτίμησή της. Ένα συνηθισμένο σχήμα αποθορυβοποίησης εικόνας φαίνεται στο Σχήμα 6.1. Η μέση τιμή της ενθόρυβης εικόνας, m_y , αφαιρείται, προκειμένου να εξασφαλιστεί ότι το x_{est} θα είναι μια αμερόληπτη εκτίμηση της αρχικής εικόνας x , [84]. Η εκτίμηση της αρχικής εικόνας δίνεται από την ακόλουθη σχέση

$$x_{est}(i, j) = m_y(i, j) + \mathbf{h}^T[\mathbf{y}(i, j) - \mathbf{m}_y(i, j)] \quad (6.2.3)$$

όπου $\mathbf{h}$ είναι το φίλτρο αποθορυβοποίησης σε μορφή διανύσματος, $\mathbf{y}(i, j)$ είναι ένα διάνυσμα που περιέχει τα αντίστοιχα εικονοστοιχεία μιας μη αιτιατής γειτονιάς του $y(i, j)$, διαστάσεων $p \times p$, κατά γραμμές, και το διάνυσμα $\mathbf{m}_y(i, j)$ περιέχει τις αντίστοιχες τοπικές μέσες τιμές.

Παρατηρούμε στη σχέση (6.2.3) ότι η εικόνα αποσυντίθεται σε ένα τμήμα χαμηλών συχνοτήτων, το m_y , και σε ένα τμήμα υψηλών συχνοτήτων, το $y - m_y$. Το φίλτρο αποθορυβοποίησης επηρεάζει κυρίως τις υψηλές συχνότητες της εικόνας, γεγονός που πρέπει να ληφθεί υπόψη κατά την ένθεση του υδατογραφήματος. Προκειμένου το υδατογράφημα να είναι ανθεκτικό στην επίθεση της αποθορυβοποίησης, πρέπει να ενισχυθεί στις υψηλές συχνότητες, οι οποίες αντιστοιχούν σε περιοχές με έντονη υφή και ακμές. Αυτή η στρατηγική ταιριάζει στις ιδιότητες του συστήματος ανθρώπινης όρασης, καθώς ο θόρυβος είναι λιγότερο αισθητός σε τέτοιες περιοχές. Έτσι, το επόμενο βήμα επικεντρώνεται στην εύρεση μιας κατάλληλης αντιληπτικής μάσκας που να αντισταθμίζει την αποθορυβοποίηση.

6.2.3 Παραγωγή της νέας αντιληπτικής μάσκας

Όταν μια μάσκα σχεδιάζεται έτσι ώστε να κάνει το υδατογράφημα ανθεκτικό στην αποθορυβοποίηση, ο τρόπος που συμπεριφέρεται ένα φίλτρο αποθορυβοποίησης πρέπει να λαμβάνεται υπόψη. Στην εργασία [148] η αποθορυβοποίηση γίνεται με τη χρήση ενός προσαρμοστικού φίλτρου Wiener [84], δηλαδή η αρχική εικόνα εκτιμάται εικονοστοιχείο προς εικονοστοιχείο μέσω της σχέσης

$$x_{est}(i, j) = m_y(i, j) + \frac{\sigma_x^2(i, j)}{\sigma_x^2(i, j) + \sigma_w^2(i, j)} [y(i, j) - m_y(i, j)] \quad (6.2.4)$$

όπου $\sigma_x^2(i, j)$ και $\sigma_w^2(i, j)$ είναι οι τοπικές διασπορές της αρχικής εικόνας και του υδατογραφήματος, αντίστοιχα. Υποθέτοντας ένα μη στάσιμο Gaussian μοντέλο, η αντιληπτική μάσκα ορίζεται από τη σχέση που ακολουθεί

$$M_{NVF}(i, j) = 1 - \frac{1}{1 + \sigma_x^2(i, j)} \quad (6.2.5)$$

όπου το NVF σημαίνει *Noise Visibility Function* (Συνάρτηση Ορατότητας Θορύβου).

Εμείς θα υποθέσουμε ότι η επίθεση της αποθορυβοποίησης γίνεται με τη χρήση ενός προσαρμοστικού μη αιτιατού φίλτρου σφάλματος πρόβλεψης ελαχίστων τετραγώνων. Μια τέτοια επίθεση βασίζεται απευθείας στα δεδομένα, χωρίς να πρέπει να γίνουν διάφορες υποθέσεις για τις στατιστικές ιδιότητες της εικόνας [52, 137], όπως γίνεται στην περίπτωση του προσαρμοστικού φίλτρου Wiener, όπου η αρχική εικόνα υποτίθεται ότι είναι μια τοπικά λευκή διαδικασία.

Το φίλτρο σφάλματος πρόβλεψης ελαχίστων τετραγώνων υπολογίζεται από την ελαχιστοποίηση της ακόλουθης συνάρτησης κόστους

$$J_{LP} = \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} |x(i, j) - \hat{x}(i, j)|^2 \quad (6.2.6)$$

όπου το LP σημαίνει *Linear Prediction* (Γραμμική Πρόβλεψη), $N_1 \times N_2 = N$ είναι ο αριθμός των εικονοστοιχείων και $\hat{x}(i, j)$ είναι η εκτιμώμενη τιμή του $x(i, j)$, η οποία δίνεται από τη σχέση

$$\hat{x}(i, j) = \tilde{\mathbf{a}}_x^T \tilde{\mathbf{x}}(i, j) \quad (6.2.7)$$

όπου $\tilde{\mathbf{a}}_x$ είναι ένα διάνυσμα μήκους $(p^2 - 1)$, το οποίο περιέχει τους συντελεστές γραμμικής πρόβλεψης κατά γραμμές, και το διάνυσμα $\tilde{\mathbf{x}}(i, j)$ περιέχει κατά γραμμές τα αντίστοιχα εικονοστοιχεία της μη αιτιατής γειτονιάς του $x(i, j)$, διαστάσεων $p \times p$, (εκτός από το κεντρικό στη

θέση (i, j)). Η ελαχιστοποίηση της συνάρτησης κόστους (6.2.6) σε σχέση με το $\tilde{\mathbf{a}}_x$ οδηγεί στο ακόλουθο σύστημα κανονικών εξισώσεων

$$R_x \tilde{\mathbf{a}}_x = \mathbf{r}_x \quad (6.2.8)$$

όπου R_x είναι ο πίνακας αυτοσυσχέτισης της εικόνας x και $\mathbf{r}_x$ είναι το αντίστοιχο διάνυσμα ετεροσυσχέτισης. Η λύση του παραπάνω συστήματος παρέχει το γραμμικό προβλέπτη (*predictor*) $\tilde{\mathbf{a}}_x$ [137]. Το φίλτρο σφάλματος πρόβλεψης είναι ένας πίνακας διαστάσεων $p \times p$ που έχει 1 στο κέντρο του και τους συντελεστές $-\tilde{a}_x(i, j)$ στις άλλες θέσεις. Μετασχηματίζεται σε διάνυσμα και συμβολίζεται στο εξής με $\mathbf{a}_x$. Στη συνέχεια, το σφάλμα πρόβλεψης είναι

$$e_x(i, j) = x(i, j) - \tilde{\mathbf{a}}_x^T \tilde{\mathbf{x}}(i, j) \quad (6.2.9)$$

ή

$$e_x(i, j) = \mathbf{a}_x^T \mathbf{x}(i, j). \quad (6.2.10)$$

Χρησιμοποιώντας το φίλτρο σφάλματος πρόβλεψης στη διαδικασία αποθορυβοποίησης, η εκτίμηση της αρχικής εικόνας δίνεται από τη σχέση

$$\begin{aligned} x_{est}(i, j) &= m_y(i, j) + \mathbf{a}_y^T [\mathbf{y}(i, j) - \mathbf{m}_y(i, j)] = m_y(i, j) + e_y(i, j) \\ &= m_y(i, j) + [y(i, j) - m_y(i, j)] - \tilde{\mathbf{a}}_y^T [\tilde{\mathbf{y}}(i, j) - \tilde{\mathbf{m}}_y(i, j)] \end{aligned} \quad (6.2.11)$$

όπου $\mathbf{a}_y$ είναι το σφάλμα πρόβλεψης ελαχίστων τετραγώνων της ενθόρυβης εικόνας, και τα $\tilde{\mathbf{y}}(i, j)$, $\tilde{\mathbf{m}}_y(i, j)$ ορίζονται παρόμοια με το $\tilde{\mathbf{x}}(i, j)$.

Το φίλτρο σφάλματος πρόβλεψης ελαχίστων τετραγώνων μπορεί επίσης να υπολογιστεί προσαρμοστικά χρησιμοποιώντας έναν αναδρομικό αλγόριθμο, όπως για παράδειγμα ένα δισδιάστατο RLS [122] ή ένα δισδιάστατο LMS [97]. Στη συνέχεια, η εκτίμηση της αρχικής εικόνας δίνεται από την ακόλουθη σχέση

$$\begin{aligned} x_{est}(i, j) &= m_y(i, j) + \mathbf{a}_y^T(i, j) [\mathbf{y}(i, j) - \mathbf{m}_y(i, j)] = m_y(i, j) + e_y(i, j) \\ &= m_y(i, j) + [y(i, j) - m_y(i, j)] - \tilde{\mathbf{a}}_y^T(i, j) [\tilde{\mathbf{y}}(i, j) - \tilde{\mathbf{m}}_y(i, j)] \end{aligned} \quad (6.2.12)$$

όπου το $\mathbf{a}_y(i, j)$ περιέχει τους συντελεστές του φίλτρου για το εικονοστοιχείο της θέσης (i, j) . Μια πιο αποδοτική διαδικασία θα ήταν ο υπολογισμός των συντελεστών του φίλτρου για τμήματα της εικόνας, αντί για κάθε εικονοστοιχείο.

Ας μελετήσουμε τώρα τον τρόπο με τον οποίο το φίλτρο σφάλματος πρόβλεψης (προσαρμοστικό ή όχι) ενεργεί πάνω στην εικόνα προκειμένου να αφαιρέσει το θόρυβο. Να σημειώσουμε ότι οι όροι $\mathbf{a}_y^T [\mathbf{y}(i, j) - \mathbf{m}_y(i, j)]$ και $\tilde{\mathbf{a}}_y^T(i, j) [\tilde{\mathbf{y}}(i, j) - \tilde{\mathbf{m}}_y(i, j)]$ στις σχέσεις (6.2.11)

και (6.2.12) αντιστοιχούν στην ακολουθία σφάλματος πρόβλεψης στο μη προσαρμοστικό και στο προσαρμοστικό προβλέπτη, αντίστοιχα. Ωστόσο, ανεξάρτητα από τον τρόπο με τον οποίο υπολογίζονται οι συντελεστές του φίλτρου, το σφάλμα πρόβλεψης που προκύπτει είναι προσαρμοσμένο στην εικόνα. Πράγματι, αν η εικόνα είναι ομαλή, είναι πιο προβλέψιμη και το σφάλμα πρόβλεψης παίρνει τιμές κοντά στο μηδέν. Έτσι, από τη σχέση (6.2.11) [ή την (6.2.12)], μπορούμε εύκολα να συμπεράνουμε ότι το εκτιμώμενο εικονοστοιχείο $x_{est}(i, j)$ θα έχει τιμή κοντά στην τοπική μέση τιμή. Αν η εικόνα έχει έντονη υφή ή πολλές ακμές, τότε δεν είναι ιδιαίτερα προβλέψιμη και το σφάλμα πρόβλεψης θα έχει μεγάλες τιμές. Σε τέτοιες περιοχές τα εικονοστοιχεία σε μια γειτονιά θα είναι (ιδανικά) ασυσχέτιστα, το r_x στη σχέση (6.2.8) θα έχει μηδενικές τιμές, ο R_x θα είναι διαγώνιος πίνακας και η λύση του συστήματος θα είναι μηδέν, δηλαδή οι συντελεστές του προβλέπτη $\tilde{a}_y^T$ θα είναι ίσοι με μηδέν. Έτσι, η σχέση (6.2.11) [ή η (6.2.12)] υποδηλώνει ότι το εκτιμώμενο εικονοστοιχείο είναι στην ουσία το ενθόρυβο εικονοστοιχείο $y(i, j)$.

Η παραπάνω ανάλυση δείχνει ότι το σφάλμα πρόβλεψης περιγράφει αρκετά καλά τις τοπικές ιδιότητες της εικόνας και άρα μπορεί να χρησιμοποιηθεί σαν αντιληπτική μάσκα. Μεταβάλλεται χωρικά με τρόπο που ταιριάζει στις ιδιότητες του συστήματος ανθρώπινης όρασης, γεγονός που επιβεβαιώνεται και από τη μελέτη που ακολουθεί στην ενότητα 6.2.4 και αφορά την ποιότητα της υδατογραφημένης εικόνας. Στην πράξη, η προτεινόμενη μάσκα είναι μια κανονικοποιημένη έκδοση του μέτρου του σφάλματος πρόβλεψης, έτσι ώστε να παίρνει τιμές στο διάστημα $[0,1]$, και υπολογίζεται από τη σχέση

$$M_e(i, j) = \frac{|e_x(i, j)|}{\max |e_x(i, j)|}. \quad (6.2.13)$$

Αν και αυτή η μάσκα αυξάνει σημαντικά την ενέργεια του υδατογραφήματος, το κάνει λιγότερο ορατό σε σχέση με άλλες μάσκες, όπως η M_{NMF} που ορίστηκε στη σχέση (6.2.5) και μια μάσκα η οποία κατασκευάζεται με βάση τις ιδιότητες της ανθρώπινης όρασης [78], την οποία θα αποκαλούμε στο εξής M_{HVS} . Επιπλέον, η καλύτερη απόδοση που επιτυγχάνει σε σχέση με τη M_{HVS} σημαίνει ότι το σφάλμα πρόβλεψης ελαχίστων τετραγώνων παρέχει μια καλή εκτίμηση των τοπικών δυνατοτήτων κάλυψης (*masking capabilities*) της εικόνας. Πρέπει να σημειώσουμε εδώ ότι ο τρόπος υπολογισμού του φίλτρου σφάλματος πρόβλεψης, το οποίο είναι απαραίτητο για τον υπολογισμό της μάσκας, καθορίζεται από το χρήστη. Στις υλοποιήσεις μας, για λόγους απλότητας, δεν υπολογίσαμε το φίλτρο προσαρμοστικά.

6.2.4 Μελέτη της απόδοσης της προτεινόμενης μάσκας

Ελέγξαμε την απόδοση της προτεινόμενης μάσκας μέσω εκτεταμένων πειραμάτων. Θα παρουσιάσουμε εδώ μερικά πειράματα που εφαρμόστηκαν στις πολύ γνωστές εικόνες Lenna

Σχήμα 6.2: Αρχικές εικόνες, Lenna και Mandrill

(512×512) και Mandrill (478×500), οι οποίες φαίνονται στο Σχήμα 6.2. Η Lenna είναι μια εικόνα με μεγάλες ομαλές περιοχές, ενώ ο Mandrill είναι μια εικόνα με αρκετά έντονη υφή. Τα πειράματα που έγιναν αφορούν δύο περιπτώσεις υδατογράφησης, την περίπτωση που το περιεχόμενο της εικόνας δε λαμβάνεται υπόψη (δε χρησιμοποιείται δηλαδή κάποια μάσκα) και την περίπτωση που χρησιμοποιείται μια αντιληπτική μάσκα κατά τη διαδικασία της ένθεσης.

Στη δεύτερη περίπτωση η προτεινόμενη μάσκα, η M_e , συγκρίθηκε με τις μάσκες M_{NVF} και M_{HVS} . Η M_{HVS} περιέχει παραμέτρους, οι οποίες καθορίζονται από υποκειμενικά πειράματα (*subjective experiments*). Αυτό όμως δεν είναι καθόλου πρακτικό για αυτόματες τεχνικές υδατογράφησης. Στα πειράματά μας, για αυτές τις παραμέτρους, χρησιμοποιήσαμε τις τιμές που παρέχονται στην εργασία [78], οι οποίες καθορίζονται με βάση την αποτίμηση τριών ατόμων. Περισσότερες πληροφορίες σχετικά με αυτή τη μάσκα, την M_{HVS} , μπορεί να βρει κανείς στα [144, 153]. Πρέπει να σημειώσουμε επίσης ότι χρησιμοποιήσαμε μια κανονικοποιημένη έκδοση της M_{HVS} (η M_{HVS} διαιρείται με τη μέγιστη απόλυτη τιμή της), προκειμένου να μπορεί να συγκριθεί με τις άλλες μάσκες όσο αφορά την ενέργεια του υδατογραφήματος. Αυτή η κανονικοποίηση δεν έχει κάποια επίδραση στην απόδοση της μάσκας. Για τον υπολογισμό της μάσκας M_{NVF} και της μάσκας M_e , χρησιμοποιήσαμε μη αιτιατές γειτονιές, μεγέθους 5×5 . Οι μάσκες για τις δύο εξεταζόμενες εικόνες φαίνονται στο Σχήμα 6.3. Η M_{NVF} φαίνεται να είναι υπερφορτωμένη στις περιοχές που έχουν έντονη υφή και γύρω από ακμές, ειδικά για την εικόνα Mandrill, σε αντίθεση με την M_{HVS} και την προτεινόμενη μάσκα, την M_e , όπου οι μεγάλες τιμές είναι πιο συγκεντρωμένες σε αυτές τις περιοχές.

Προκειμένου να έχουμε ίδιες τιμές για το $PSNR$ (*Peak Signal to Noise Ratio*), το οποίο

Σχήμα 6.3: Οι μάσκες M_{NVF} , (α) και (δ), M_{HVS} , (β) και (ε), και η προτεινόμενη M_e , (γ) και (ζ), για τις εικόνες Lenna και Mandrill. Οι μάσκες M_{HVS} και M_e έχουν ενισχυθεί κατά παράγοντες 2 και 20, αντίστοιχα, προκειμένου να φαίνονται καλύτερα.

ορίζεται σαν

$$PSNR = 10 \log_{10} \frac{\max(x)^2}{\|y - x\|^2} \quad (6.2.14)$$

όπου $\|\cdot\|$ είναι η Ευκλείδεια νόρμα, ρυθμίζουμε κατάλληλα την ενέργεια (δηλαδή την τυπική απόκλιση) του υδατογραφήματος για τις συγκρινόμενες μεθόδους ένθεσης. Όπως μπορούμε να δούμε στον Πίνακα 6.1 για τη Lenna και στον Πίνακα 6.2 για το Mandrill, η νέα μάσκα επιτρέπει την ένθεση υδατογραφημάτων με σημαντικά αυξημένη ενέργεια. Αυτό έχει ως αποτέλεσμα τα υδατογραφήματα αυτά να είναι πολύ ανθεκτικά, πράγμα το οποίο επιβεβαιώνεται και από τα πειράματα που αφορούν τη διαδικασία της ανίχνευσης και τα οποία θα δούμε αργότερα.

Τι συμβαίνει όμως με την ορατότητα του υδατογραφήματος; Στην πράξη υπάρχουν δύο τρόποι για να μετράμε την ποιότητα των δεδομένων, ο υποκειμενικός και ο αντικειμενικός. Η υποκειμενική μέθοδος περιλαμβάνει μια ομάδα από παρατηρητές, στους οποίους παρουσιάζονται οι αρχικές, οι υδατογραφημένες και οι λαμβανόμενες εικόνες για διαχωρισμό. Αλλά καθώς

Πίνακας 6.1: LENNA-ΕΝΕΡΓΕΙΑ ΚΑΙ ΟΡΑΤΟΤΗΤΑ ΥΔΑΤΟΓΡΑΦΗΜΑΤΟΣ ΓΙΑ ΔΙΑΦΟΡΕΤΙΚΕΣ ΜΕΘΟΔΟΥΣ ΕΝΘΕΣΗΣ

Ένθεση	Μετρικές	PSNR:45dB	PSNR:40dB	PSNR:35dB	PSNR:30dB	PSNR:25dB
<i>Χωρίς Μάσκα</i>	<i>Ενέργεια</i>	1.435	2.552	4.538	8.069	14.350
	<i>wPSNR</i>	46.067	41.067	36.067	31.067	26.067
	<i>TPE</i>	0.0281	0.0499	0.0887	0.1578	0.2806
	<i>NLPE2</i>	0	0	0	256	256
<i>M_{NVF}</i>	<i>Ενέργεια</i>	1.671	2.971	5.284	9.397	16.711
	<i>wPSNR</i>	46.388	41.388	36.388	31.388	26.388
	<i>TPE</i>	0.0274	0.0488	0.0867	0.1542	0.2742
	<i>NLPE2</i>	0	0	3	225	255
<i>M_{HVS}</i>	<i>Ενέργεια</i>	8.515	15.142	26.927	47.884	85.151
	<i>wPSNR</i>	47.275	42.275	37.275	32.275	27.275
	<i>TPE</i>	0.0259	0.0461	0.0819	0.1457	0.2590
	<i>NLPE2</i>	0	0	10	193	255
<i>M_e</i>	<i>Ενέργεια</i>	31.977	56.854	101.120	179.820	319.770
	<i>wPSNR</i>	48.697	43.697	38.697	33.697	28.697
	<i>TPE</i>	0.0233	0.0414	0.0737	0.1311	0.2331
	<i>NLPE2</i>	0	2	41	96	219

Πίνακας 6.2: MANDRILL-ΕΝΕΡΓΕΙΑ ΚΑΙ ΟΡΑΤΟΤΗΤΑ ΥΔΑΤΟΓΡΑΦΗΜΑΤΟΣ ΓΙΑ ΔΙΑΦΟΡΕΤΙΚΕΣ ΜΕΘΟΔΟΥΣ ΕΝΘΕΣΗΣ

Ένθεση	Μετρικές	PSNR:45dB	PSNR:40dB	PSNR:35dB	PSNR:30dB	PSNR:25dB
<i>Χωρίς Μάσκα</i>	<i>Ενέργεια</i>	1.435	2.553	4.539	8.072	14.354
	<i>wPSNR</i>	49.624	44.624	39.624	34.625	29.624
	<i>TPE</i>	0.0277	0.0493	0.0878	0.1559	0.2772
	<i>NLPE2</i>	0	0	2	227	240
<i>M_{NVF}</i>	<i>Ενέργεια</i>	1.487	2.645	4.703	8.363	14.872
	<i>wPSNR</i>	49.967	44.967	39.967	34.967	29.967
	<i>TPE</i>	0.0276	0.0490	0.0872	0.1549	0.2756
	<i>NLPE2</i>	0	0	2	225	240
<i>M_{HVS}</i>	<i>Ενέργεια</i>	5.087	9.046	16.086	28.606	50.869
	<i>wPSNR</i>	51.781	46.781	41.781	36.781	31.781
	<i>TPE</i>	0.0262	0.0466	0.0830	0.1475	0.2623
	<i>NLPE2</i>	0	0	1	163	237
<i>M_e</i>	<i>Ενέργεια</i>	17.531	31.174	55.437	98.582	175.307
	<i>wPSNR</i>	54.845	49.845	44.845	39.845	34.845
	<i>TPE</i>	0.0233	0.0414	0.0737	0.1310	0.2329
	<i>NLPE2</i>	0	1	28	114	183

οι μέθοδοι υποκειμενικής αποτίμησης της ποιότητας περιλαμβάνουν αλληλεπίδραση με τον άνθρωπο, δεν είναι κατάλληλες για αυτόματο benchmarking. Για αυτόματο benchmarking της ποιότητας των δεδομένων, πρέπει να βασιστούμε σε αριθμητικές μετρήσεις [88]. Για αντικειμενική αποτίμηση της ποιότητας, πρέπει να χρησιμοποιηθούν εξελιγμένες μετρικές ποιότητας που να λαμβάνουν υπόψη τους βασικούς παραμέτρους της όρασης. Ιδανικά, ένα αντιληπτικό μοντέλο σχεδιασμένο για αυτόματο έλεγχο πιστότητας πρέπει να δίνει τα ίδια αποτελέσματα με αυτά που προκύπτουν από την ανθρώπινη παρατήρηση. Ωστόσο, για τη σύγκριση διαφορετικών αλγορίθμων υδατογράφησης αρκεί το μοντέλο να παρέχει τιμές που να έχουν μονότονη σχέση με τα αποτελέσματα που προκύπτουν από τους ελέγχους που γίνονται από ανθρώπους. Δηλαδή, δεν είναι απαραίτητο για το μοντέλο να ταιριάζει ακριβώς με τα αποτελέσματα των ελέγχων που διεξάγονται από παρατηρητές, αλλά απλά να προβλέπει τη σχετική απόδοση των διαφορετικών αλγορίθμων σε αυτούς τους ελέγχους [28].

Η διεξαγωγή υποκειμενικών ελέγχων για την αποτίμηση της ποιότητας των εικόνων είναι μια χρήσιμη αλλά και πολύπλοκη διαδικασία, καθώς αυτά τα πειράματα είναι στατιστικά. Διαφορετικοί παρατηρητές θα συμπεριφερθούν διαφορετικά. Έτσι, χρειάζεται ένας μεγάλος αριθμός παρατηρητών και ένας μεγάλος αριθμός προσπαθειών. Μια τέτοια διαδικασία, αν δε γίνει με το σωστό τρόπο, μπορεί να οδηγήσει σε λάθος συμπεράσματα. Έτσι, προκειμένου να αποτιμήσουμε την ποιότητα των υδατογραφημένων εικόνων, χρησιμοποιήσαμε μερικές μετρικές ποιότητας από το Checkmark [101, 149], το οποίο είναι ένα benchmark που διατίθεται στο δίκτυο [60].

Η πρώτη μετρική ποιότητας που χρησιμοποιήσαμε ήταν το PSNR με βάρη (wPSNR), το οποίο είναι μια παραλλαγή του κλασικού PSNR η οποία εισάγει βάρη για τις διαφορετικές περιοχές, σε αντίθεση με το PSNR που αντιμετωπίζει όλες τις περιοχές με τον ίδιο τρόπο. Έτσι, το wPSNR είναι μια πιο αξιόπιστη μετρική ποιότητας. Στο Checkmark η εισαγωγή βαρών στον υπολογισμό του PSNR γίνεται με χρήση της NVF. Μπορούμε να δούμε στα αποτελέσματα που προέκυψαν από τα πειράματά μας ότι η προτεινόμενη μάσκα παρέχει μεγαλύτερες τιμές για το wPSNR, το οποίο σημαίνει καλύτερη απόδοση όσο αφορά τη διατήρηση της ποιότητας της εικόνας. Παρατηρούμε επίσης ότι οι τιμές του wPSNR για την εικόνα Mandrill είναι μεγαλύτερες από τις αντίστοιχες τιμές για τη Lenna, το οποίο είναι λογικό αφού ένα υδατογράφημα είναι λιγότερο ορατό σε περιοχές με έντονη υφή από ότι σε μια ομαλή περιοχή.

Προκειμένου να ερευνήσουμε περαιτέρω το πόσο ορατό είναι το υδατογράφημα, χρησιμοποιήσαμε δύο ακόμα μετρικές ποιότητας, την TPE (Total Perceptual Error), η οποία είναι μια καθολική αντιληπτική μετρική, και τη NLPE2 (Number of blocks greater than the second Local Perceptual Error threshold-Αριθμός τμημάτων με τοπικό σφάλμα μεγαλύτερο από ένα προκαθορισμένο κατώφλι), η οποία είναι τοπική μετρική. Αυτές οι μετρικές προέρχονται από

το μοντέλο Watson [151] και περιλαμβάνονται στο Checkmark. Η TPE είναι ένα συνολικό αντιληπτικό σφάλμα, ανεξάρτητο από το μέγεθος της εικόνας. Η NLPE2 αποτιμά την ποιότητα της εικόνας τοπικά, σε τμήματα της εικόνας. Εδώ, χρησιμοποιήσαμε τμήματα διαστάσεων 32×32 . Τα κατώφλια για αυτές τις μετρικές, τα οποία είναι προκαθορισμένα στο Checkmark, έχουν καθοριστεί μέσω υποκειμενικών ελέγχων για διαφορετικά είδη εικόνων και έχουν υπολογιστεί οι μέσες τιμές τους [149]. Σε όλα τα πειράματα που κάναμε, οι μικρότερες τιμές για τη μετρική TPE επιτυγχάνονται για την προτεινόμενη μέθοδο ένθεσης υδατογραφήματος. Έτσι, εκτός από το wPSNR και οι τιμές της μετρικής TPE δείχνουν ότι η απόδοση της νέας μάσκας, σε σχέση με την ποιότητα της εικόνας, είναι καλύτερη από αυτή για τις μάσκες M_{NVF} και M_{HVS} , παρά το γεγονός ότι η προτεινόμενη μάσκα επιτρέπει την ένθεση υδατογραφημάτων με σημαντικά μεγαλύτερη ενέργεια.

Ωστόσο, σε μερικές περιπτώσεις, ακόμα κι αν το συνολικό αντιληπτικό σφάλμα είναι μικρό, η εικόνα αλλοιώνεται τοπικά με ορατό τρόπο. Προκειμένου να αποτιμήσουμε την ποιότητα της εικόνας τοπικά, χρησιμοποιούμε τη μετρική NLPE2. Σύμφωνα με την NLPE2, αν η ποιότητα για ένα τμήμα της εικόνας δεν είναι αποδεκτό, ολόκληρη η εικόνα απορρίπτεται ως μη αποδεκτή. Αυτό φαίνεται να είναι πολύ αυστηρό κριτήριο, γιατί αν οι αλλοιώσεις δημιουργούνται σε περιοχές με έντονη υφή, δύσκολα γίνονται αντιληπτές. Συνήθως, μια εικόνα θεωρείται ότι είναι αποδεκτή αν το PSNR είναι τουλάχιστον $45dB$. Στους Πίνακες 6.1 και 6.2, συμπεριλαμβάνονται και πιο βαριές περιπτώσεις ένθεσης. Για PSNR ίσο με $35dB$, η μέθοδός μας φαίνεται να είναι κατώτερη όσο αφορά την ποιότητα της εικόνας τοπικά, καθώς περισσότερα τμήματα των υδατογραφημένων εικόνων εμφανίζονται να έχουν μη αποδεκτή ποιότητα σε σύγκριση με τις άλλες τρεις μεθόδους. Αυτό προφανώς συμβαίνει γιατί σε κάποιες περιοχές έχει μπει παραπάνω υδατογράφημα από όσο αυτές αντέχουν. Ωστόσο, αυτές οι περιοχές έχουν έντονη υφή και ακμές και άρα η αλλοίωση δύσκολα γίνεται αντιληπτή. Από την άλλη μεριά, αν οι εικόνες υδατογραφηθούν με κάποια από τις τρεις άλλες μεθόδους, αν και εμφανίζουν καλύτερη ποιότητα τοπικά, έχουν τέτοιες τιμές για το wPSNR και το TPE που δηλώνουν χειρότερη ποιότητα συνολικά στην εικόνα, το οποίο είναι σημαντικότερο πρόβλημα. Αυτός ο ισχυρισμός επιβεβαιώνεται από τις εικόνες που φαίνονται στο Σχήμα 6.4. Οι τρεις συγκρινόμενες μάσκες έχουν χρησιμοποιηθεί για την ένθεση υδατογραφήματος και το PSNR για όλες τις υδατογραφημένες εικόνες που προκύπτουν είναι ίσο με $35dB$. Για τις μάσκες M_{NVF} και M_{HVS} , οι εικόνες φαίνεται να έχουν περισσότερο θόρυβο (πιο λίγο για τη M_{HVS}), ενώ για την προτεινόμενη μάσκα, την M_e , δεν υπάρχουν ορατές διαφοροποιήσεις. Στους Πίνακες 6.1 και 6.2 παρέχονται αποτελέσματα και για τις περιπτώσεις που το PSNR είναι ίσο με $30dB$ και με $25dB$, όπου οι εικόνες έχουν υπερφορτωθεί με υδατογράφημα. Όπως φαίνεται, σε αυτές τις περιπτώσεις η μέθοδός μας παρέχει καλύτερη ποιότητα συνολικά αλλά

Σχήμα 6.4: Υδατογραφημένες εικόνες για τις μάσκες M_{NVF} , (α) και (δ), M_{HVS} , (β) και (ε), και την προτεινόμενη M_e , (γ) και (ζ). Το PSNR είναι 35dB.

και τοπικά στην εικόνα.

Προκειμένου να δικαιολογήσουμε την απόδοση της προτεινόμενης μάσκας, πρέπει να υπενθυμίσουμε ότι είναι μια κανονικοποιημένη εκδοχή του σφάλματος πρόβλεψης. Το σφάλμα πρόβλεψης είναι πιθανό να έχει κάποιες πολύ μεγάλες (θετικές ή αρνητικές) τιμές αν δεν είναι εφικτή μια καλή πρόβλεψη σε κάποια περιοχή της εικόνας. Αυτό έχει σαν αποτέλεσμα τέτοιου είδους (μη προβλέψιμες) περιοχές να υπερφορτώνονται με υδατογράφημα. Εντούτοις, τέτοιες τιμές σφάλματος εμφανίζονται με πολύ μικρή πιθανότητα. Ας δούμε για παράδειγμα τα ιστογράμματα των σφαλμάτων πρόβλεψης για τις εικόνες Lenna και Mandrill, στο Σχήμα 6.5. Η μέγιστη τιμή του σφάλματος πρόβλεψης της Lenna είναι 58.1253, ενώ η ελάχιστη είναι -33.1791. Οι αντίστοιχες τιμές για το Mandrill είναι 108.0902 και -114.8250. Παρατηρούμε, ωστόσο, ότι για την πλειοψηφία των εικονοστοιχείων των εικόνων οι τιμές του σφάλματος πρόβλεψης κυμαίνονται από -12 έως 12 για τη Lenna και από -50 έως 50 για το Mandrill. Άρα, μπορούμε να θέσουμε κατάλληλα κατώφλια στις τιμές του σφάλματος, προκειμένου να αντιμετωπίσουμε αυτό το φαινόμενο υπερφόρτωσης των εικόνων με υδατογράφημα.

Σχήμα 6.5: Ιστογράμματα των σφαλμάτων πρόβλεψης για τη Lenna, (α), και το Mandrill, (β).

Σχήμα 6.6: Ιστογράμματα της μάσκας M_e για τη Lenna, (α), και το Mandrill, (β).

Όπως έχουμε ήδη αναφέρει, ο σκοπός μας είναι να παρέχουμε την καλύτερη δυνατή εξισορρόπηση ανάμεσα στο πόσο ορατό και στο πόσο ανθεκτικό είναι ένα υδατογράφημα. Όπως συμπεραίνουμε από την παραπάνω ανάλυση, η προτεινόμενη μάσκα παρουσιάζει πολύ καλή συμπεριφορά. Ακόμα πιο ενδιαφέρον είναι το γεγονός ότι ταυτόχρονα παρέχει υδατογραφήματα με πολύ μεγάλη ενέργεια. Αυτό έχει ως αποτέλεσμα πολύ ανθεκτικά υδατογραφήματα, γεγονός που θα επιβεβαιωθεί και αργότερα, στην ανάλυση σχετικά με την ανίχνευση των υδατογραφημάτων. Η ικανότητα της νέας μάσκας να παρέχει υδατογραφήματα με μεγάλη ενέργεια μπορεί να εξηγηθεί από τα ιστογράμματα των μασκών για τις εξεταζόμενες εικόνες, τα οποία φαίνονται στο Σχήμα 6.6. Παρατηρούμε ότι οι μέσες τιμές είναι 0.0195 και 0.0476, για τη Lenna και για το Mandrill, αντίστοιχα. Ακόμα και στην περίπτωση που κάθε μάσκα υπολογίζεται από το κατωφλιωμένο σφάλμα πρόβλεψης, όπως περιγράφηκε παραπάνω, οι μέσες τιμές

Σχήμα 6.7: Ιστογράμματα της μάσκας M_{NVF} για τη Lenna, (α), και το Mandrill, (β).

Σχήμα 6.8: Ιστογράμματα της μάσκας M_{HVS} για τη Lenna, (α), και το Mandrill, (β).

των масκών είναι 0.0895 και 0.1087. Ας παρατηρήσουμε τώρα τα ιστογράμματα των масκών M_{NVF} , στο Σχήμα 6.7, όπου οι μέσες τιμές είναι 0.8276 και 0.9501. Οι τιμές της μάσκας M_{NVF} συγκεντρώνονται στην αντίθετη κατεύθυνση σε σχέση με την προτεινόμενη μάσκα. Αυτό εξηγεί γιατί η M_{NVF} δεν μπορεί να παρέχει μεγάλη ενέργεια υδατογραφήματος και γιατί μοιάζει υπερφορτωμένη γύρω από τις ακμές. Στο Σχήμα 6.8, φαίνονται τα ιστογράμματα του μέτρου της μάσκας M_{HVS} (για να είναι συγκρίσιμα με τα ιστογράμματα των άλλων масκών). Θυμίζουν τα ιστογράμματα της προτεινόμενης μάσκας, αλλά οι μέσες τιμές είναι 0.1435 και 0.2287, που σημαίνουν μικρότερες ενέργειες υδατογραφήματος.

Συνοψίζοντας την παραπάνω μελέτη, μπορούμε να πούμε ότι τα πλεονεκτήματα της προτεινόμενης μάσκας είναι πολύ σημαντικά. Η χρήση της παρέχει πολύ υψηλές ενέργειες στα υδατογραφήματα, ενώ ταυτόχρονα η ποιότητα της εικόνας παραμένει σε πολύ καλό επίπεδο σε σχέση με τις άλλες τρεις μεθόδους ένθεσης. Επιπλέον, τα αποτελέσματα της σύγκρισης με

τη μάσκα M_{HVS} , δείχνουν ότι η προτεινόμενη μάσκα περιγράφει πολύ καλά το σύστημα της ανθρώπινης όρασης, έχοντας το πλεονέκτημα να μη χρειάζεται για τον υπολογισμό της την εκτίμηση διάφορων παραμέτρων.

6.3 Το προτεινόμενο σχήμα ανίχνευσης

Συνήθως, η διαδικασία τυφλής ανίχνευσης του υδατογραφήματος χρησιμοποιεί μια μετρική ομοιότητας βασισμένη στη συσχέτιση ανάμεσα στο υδατογράφημα και στη λαμβανόμενη εικόνα. Πρόσφατα, αναπτύχθηκαν πιο εξειδικευμένοι ανιχνευτές βασισμένοι στο πρόβλημα απλού ελέγχου υπόθεσης (*simple hypothesis testing problem*) [69]. Ωστόσο, καθώς στο χωρικό πεδίο είναι δύσκολο να βρεθεί ένα ικανοποιητικό στατιστικό μοντέλο για την περιγραφή της εικόνας, υιοθετούμε έναν ανιχνευτή βασισμένο στη συσχέτιση και προσπαθούμε να βελτιώσουμε την απόδοσή του.

Το προτεινόμενο σχήμα ανίχνευσης αποτελείται από δύο βήματα. Στο πρώτο βήμα, μια κατάλληλη διαδικασία προλεύκανσης εφαρμόζεται στη λαμβανόμενη εικόνα και στο υδατογράφημα με τη μάσκα, ενώ στο δεύτερο βήμα υπολογίζεται η κανονικοποιημένη συσχέτιση ανάμεσα στις ακολουθίες που έχουν προλευκανθεί. Το φίλτρο προλεύκανσης είναι το φίλτρο σφάλματος πρόβλεψης ελαχίστων τετραγώνων, το οποίο χρησιμοποιείται και για τον υπολογισμό της αντιληπτικής μάσκας. Η κανονικοποιημένη συσχέτιση που χρησιμοποιείται στο δεύτερο βήμα είναι μια από τις πιο συνηθισμένες μετρικές ανίχνευσης και προτιμάται σε σχέση με τη γραμμική συσχέτιση, εξαιτίας την ανθεκτικότητάς της σε αλλαγές του πλάτους της εικόνας. Η έξοδος του ανιχνευτή που βασίζεται στην κανονικοποιημένη συσχέτιση υποτίθεται ότι είναι κατά προσέγγιση Gaussian [28].

Υπάρχουν δύο χαρακτηριστικά του προτεινόμενου σχήματος ανίχνευσης που αξίζουν ιδιαίτερη προσοχή. Το πρώτο είναι ότι στην ανίχνευση χρησιμοποιείται το υδατογράφημα στο οποίο έχει εφαρμοστεί η μάσκα και όχι το υδατογράφημα από μόνο του. Αυτό είναι πολύ σημαντικό για την απόδοση του ανιχνευτή γιατί όπως έχουμε αποδείξει στην εργασία [62] και αναφέραμε και στο προηγούμενο κεφάλαιο, αν ισχύει ότι $P_M > \mu_M^2$, όπου P_M είναι η ισχύς της μάσκας και μ_M η μέση τιμή της, τότε είναι προτιμότερο να χρησιμοποιηθεί το υδατογράφημα με τη μάσκα κατά τη διαδικασία της ανίχνευσης. Προφανώς, η συνθήκη $P_M > \mu_M^2$ είναι πάντα αληθής, εκτός από την περίπτωση που η μάσκα είναι ίση με μια σταθερά. Υποθέτοντας ότι η λαμβανόμενη εικόνα διατηρεί (κατά προσέγγιση) τις ιδιότητες προβλεψιμότητας της αρχικής εικόνας, μπορούμε να έχουμε μια καλή εκτίμηση της μάσκας στο δέκτη. Το δεύτερο χαρακτηριστικό του προτεινόμενου σχήματος ανίχνευσης είναι ότι ενσωματώνει ένα βήμα αποδοτικής

προ-λεύκανσης, βασισμένο στο μη αιτιατό διασπαστικό φίλτρο πρόβλεψης. Στην πράξη, αποδεικνύεται ότι το προτεινόμενο σχήμα είναι μια επέκταση καλά θεμελιωμένων τεχνικών των επικοινωνιών για ανίχνευση σημάτων σε χρωματισμένο θόρυβο. Να σημειώσουμε ότι στην περίπτωση μας το σήμα που πρόκειται να ανιχνευτεί είναι το υδατογράφημα μετά την εφαρμογή της μάσκας, ενώ ο χρωματισμένος θόρυβος είναι η λαμβανόμενη εικόνα. Φυσικά, τεχνικές προ-λεύκανσης έχουν ήδη προταθεί στη βιβλιογραφία σχετικά με το υπό εξέταση πρόβλημα ([28] και αναφορές που υπάρχουν εκεί). Για παράδειγμα, η χρήση ενός φίλτρου προ-λεύκανσης προτάθηκε αρχικά στην εργασία [33], όπου ένα φίλτρο οριζόντιων διαφορών εφαρμόζεται στις γραμμές της εικόνας. Το φίλτρο σφάλματος πρόβλεψης ελαχίστων τετραγώνων που προτείνεται εδώ έχει την πολύ σημαντική ιδιότητα ότι η έξοδός του τείνει σε μια λευκή διαδικασία. Έτσι, προσαρμόζει τον εαυτό του στην είσοδο, απορροφώντας το χρώμα της. Επιπλέον, όπως θα δούμε και αργότερα, η χρήση του σε συνδυασμό με τη νέα μάσκα οδηγεί σε πολύ καλή απόδοση του συστήματος υδατογράφησης.

Στη συνέχεια, μελετάμε την απόδοση του προτεινόμενου ανιχνευτή και τη συγκρίνουμε με αυτή του κλασικού ανιχνευτή (χωρίς το βήμα της προ-λεύκανσης). Η ανάλυσή μας έγινε για την περίπτωση των πραγματικών δεδομένων. Να θυμίσουμε ότι συνήθως γίνονται κάποιες ιδανικές υποθέσεις για τη διαδικασία ανίχνευσης. Δηλαδή, κάθε υδατογράφημα θεωρείται ότι είναι μια λευκή διαδικασία εντελώς ασυσχέτιστη με άλλα υδατογραφήματα, καθώς και με την αρχική εικόνα. Εντούτοις, στην πράξη οι παραπάνω υποθέσεις ισχύουν μόνο κατά προσέγγιση. Έτσι, ποσότητες που ιδανικά θεωρούνται ίσες με το μηδέν είναι καλύτερο να λαμβάνονται υπόψη, καθώς υπάρχει περίπτωση να επηρεάζουν την απόφαση για την ύπαρξη του υδατογραφήματος, ιδιαίτερα μετά από ισχυρές επιθέσεις. Στην πράξη, όλες οι εμπλεκόμενες ποσότητες αυτο- και ετερο- συσχετίσεων πρέπει να λαμβάνονται υπόψη και μέσοι όροι πρέπει να αντικαθιστούν τους τελεστές αναμενόμενης τιμής. Επίσης, η συσχέτιση ανάμεσα στη λαμβανόμενη εικόνα (ή το σφάλμα πρόβλεψης) και ένα άλλο υδατογράφημα (διαφορετικό από αυτό που έχει εισαχθεί στην εικόνα) δε θα τίθεται πλέον ίση με το μηδέν.

Η ανάλυση που θα παρουσιάσουμε σε αυτό το κεφάλαιο αφορά τη γενική περίπτωση επίθεσης που αποτελείται από εφαρμογή γραμμικού φίλτρου και πρόσθεση θορύβου στην υδατογραφημένη εικόνα. Από τις γενικές εκφράσεις προκύπτουν πολύ εύκολα οι ειδικές περιπτώσεις της μη επίθεσης, της επίθεσης με θόρυβο και της επίθεσης με γραμμικό φίλτρο. Η ανάλυσή μας βασίστηκε σε δύο εναλλακτικές προσεγγίσεις, δηλαδή 1) την ντετερμινιστική προσέγγιση, στην οποία οι μη μηδενικοί όροι αυτο- και ετερο-συσχετίσεων θεωρούνται ότι έχουν άνω φραγμένα μέτρα, και 2) μια στατιστική προσέγγιση, στην οποία οι τιμές ανίχνευσης που υπολογίζονται μέσω της μετρικής της κανονικοποιημένης συσχέτισης θεωρούνται Gaussian τυχαίες μεταβλητές [28]. Στη δεύτερη προσέγγιση, υπολογίζονται αναλυτικά οι εκφράσεις

των μέσων τιμών και των διασπορών των εμπλεκόμενων τυχαίων μεταβλητών. Οι εκφράσεις αυτές δίνουν τη δυνατότητα να υπολογιστούν κατώφλια προσαρμοσμένα στις συνθήκες των επιθέσεων. Πρέπει να σημειώσουμε εδώ ότι ο καθορισμός των κατωφλίων γίνεται συνήθως πειραματικά.

Τρία σενάρια ανίχνευσης διερευνήθηκαν, α) η ανίχνευση ενός απλού υδατογραφήματος (στο οποίο δεν έχει εφαρμοστεί μάσκα) με χρήση του w_0 , β) η ανίχνευση ενός υδατογραφήματος με μάσκα χρησιμοποιώντας το w_0 και, γ) η ανίχνευση ενός υδατογραφήματος με μάσκα χρησιμοποιώντας το $w_0 = M \odot w_0$, όπου η μάσκα M εκτιμάται από τη λαμβανόμενη εικόνα. Παρουσιάζουμε το τρίτο σενάριο, το οποίο είναι το πιο γενικό και το πιο αποτελεσματικό.

Το σφάλμα πρόβλεψης της υδατογραφημένης εικόνας είναι $e_y(i, j) = y(i, j) - \tilde{\mathbf{a}}_y^T \tilde{\mathbf{y}}(i, j)$. Στόχος είναι να συγκρίνουμε τα παραγόμενα μέτρα συσχέτισης όταν είτε η εικόνα και το υδατογράφημα με τη μάσκα (συμβατικό σχήμα, χωρίς προ-λεύκανση), είτε το σφάλμα πρόβλεψης της εικόνας και το φιλτραρισμένο υδατογράφημα με τη μάσκα (προτεινόμενο σχήμα) χρησιμοποιούνται στη διαδικασία ανίχνευσης.

6.3.1 Ντετερμινιστική προσέγγιση

Όπως έχουμε ήδη πει, η μετρική της κανονικοποιημένης συσχέτισης υπολογισμένη για δύο δισδιάστατες ακολουθίες x_1 και x_2 μεγέθους $N_1 \times N_2$ ορίζεται από τη σχέση

$$C_{x_1, x_2} = \frac{\sum_{i=1}^{N_1} \sum_{j=1}^{N_2} x_1(i, j) x_2(i, j)}{\sqrt{\sum_{i=1}^{N_1} \sum_{j=1}^{N_2} x_1(i, j)^2} \sqrt{\sum_{i=1}^{N_1} \sum_{j=1}^{N_2} x_2(i, j)^2}} \quad (6.3.1)$$

όπου $N = N_1 \times N_2$ είναι ο αριθμός των εικονοστοιχείων της κάθε ακολουθίας. Στη ντετερμινιστική προσέγγιση θεωρούμε ότι οι συσχετίσεις ανάμεσα σε διαφορετικά υδατογραφήματα ή ανάμεσα στην εικόνα και σε ένα υδατογράφημα διαφορετικό από αυτό που έχει εισαχθεί έχουν μέτρο άνω φραγμένο από ένα μικρό θετικό αριθμό ε .

Στην περίπτωση που η επίθεση αποτελείται από γραμμικό φιλτράρισμα και πρόσθεση θορύβου η εικόνα δίνεται από την ακόλουθη σχέση

$$z(i, j) = \mathbf{h}^T \mathbf{y}(i, j) + n(i, j) \quad (6.3.2)$$

όπου το διάνυσμα $\mathbf{h}$ περιέχει τους συντελεστές ενός γραμμικού φίλτρου μεγέθους $l \times l$ κατά στήλες. Γενικά, θεωρούμε ότι $l \geq p$. Στην ανάλυσή μας, το $\mathbf{h}$ είναι το αποκομμένο κεντρικό τμήμα του φίλτρου, το οποίο έχει διαστάσεις $p \times p$, ίσες δηλαδή με τις διαστάσεις του φίλτρου σφάλματος πρόβλεψης. Η ακολουθία $n(i, j)$ είναι προσθετικός λευκός Gaussian θόρυβος, με μέση τιμή μηδέν και διασπορά σ_n^2 . Το σφάλμα πρόβλεψης για τη λαμβανόμενη εικόνα δίνεται

από τη σχέση

$$e_z(i, j) = z(i, j) - \tilde{\mathbf{a}}_z^T \tilde{\mathbf{z}}(i, j) \quad (6.3.3)$$

ενώ το φιλτραρισμένο υδατογράφημα, στο οποίο πρώτα έχει εφαρμοστεί η μάσκα, δίνεται από τη σχέση

$$e_{u_0}(i, j) = u_0(i, j) - \tilde{\mathbf{a}}_z^T \tilde{\mathbf{u}}_0(i, j). \quad (6.3.4)$$

Μετά από κατάλληλες ενέργειες (λεπτομέρειες υπάρχουν στο Παράρτημα Β) προκύπτει ότι

$$\text{αν } |h_0| P_M \sigma_{w_0}^2 L \geq \|\hat{H}\| P_M \sigma_{w_0}^2 + \gamma, \quad \text{τότε } |C_{e_z, e_{u_0}}| \geq |C_{z, u_0}| \quad (6.3.5)$$

όπου $C_{e_z, e_{u_0}}$ και C_{z, u_0} είναι οι έξοδοι των συγκρινόμενων ανιχνευτών (του προτεινόμενου και του συμβατικού, αντίστοιχα), L είναι ένας θετικός αριθμός μικρότερος από 1, γ είναι ένας μικρός βαθμωτός (στην ιδανική περίπτωση είναι ίσος με μηδέν), h_0 είναι το κεντρικό στοιχείο του γραμμικού φίλτρου και $\hat{H}$ είναι ένας πίνακας διαστάσεων $(p^2 \times p^2)$ ο οποίος ορίζεται ως

$$\hat{H} = \begin{bmatrix} 0 & h_1 & h_2 & \cdots & 0 \\ h_{-1} & 0 & h_1 & \cdots & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & \cdots & h_{-1} & 0 & h_1 \\ 0 & \cdots & h_{-2} & h_{-1} & 0 \end{bmatrix}$$

και ο οποίος είναι ένας $(p^2 + 1)/2$ -Toeplitz πίνακας με μηδενικά στην κύρια διαγώνιο. Η σχέση (6.3.5) παράγεται υποθέτοντας ότι το υδατογράφημα που εμπλέκεται στον υπολογισμό των συσχετίσεων είναι το αυθεντικό, αυτό δηλαδή που πραγματικά έχει εισαχθεί στην εικόνα, το u_0 . Στην περίπτωση ενός διαφορετικού υδατογραφήματος, u_1 , υποθέτουμε ότι οι έξοδοι και των δύο ανιχνευτών θα είναι περίπου ίδιες (στην ιδανική περίπτωση θεωρούνται ίσες με μηδέν). Αυτή η περίπτωση δεν εξετάζεται εδώ και η ανάλυσή μας αφορά μόνο την περίπτωση του αυθεντικού υδατογραφήματος.

Έτσι, αν η συνθήκη στη σχέση (6.3.5) είναι αληθής, τότε η προτεινόμενη μετρική συσχέτισης έχει καλύτερη απόδοση σε σχέση με τη συμβατική. Οι εκφράσεις για τις ειδικές περιπτώσεις της επίθεσης με γραμμικό φίλτρο, της επίθεσης με θόρυβο και της μη επίθεσης μπορούν εύκολα να παραχθούν με κατάλληλη προσαρμογή των εμπλεκόμενων παραμέτρων. Στην περίπτωση της επίθεσης με γραμμικό φίλτρο, προκύπτει ότι η ανισότητα (6.3.5) είναι αληθής αν ισχύει $|h_0| \geq \|\hat{H}\|$. Αυτή η ανισότητα είναι αληθής για φίλτρα που έχουν ενισχυμένο το κεντρικό τους τμήμα, όπως για παράδειγμα το Laplacian, το Gaussian και το Unsharp. Αντίθετα, για φίλτρα που δεν έχουν ενισχυμένο κέντρο, όπως για παράδειγμα το φίλτρο μέσου

όρου, το υδατογράφημα δεν είναι ανιχνεύσιμο. Καθώς η ανάλυση είναι ανεξάρτητη της χρησιμοποιούμενης μάσκας, το αποτέλεσμα αυτό είναι γενικό και ισχύει για οποιαδήποτε μορφή χωρικής μάσκας.

6.3.2 Στοχαστική προσέγγιση

Στη ντετερμινιστική προσέγγιση παραπάνω, η ανάλυση αφορούσε μόνο την περίπτωση του αυθεντικού υδατογραφήματος, στο οποίο έχει εφαρμοστεί η μάσκα, u_0 . Για την περίπτωση ενός υδατογραφήματος u_1 , διαφορετικού από αυτό που έχει εισαχθεί στην εικόνα, οι έξοδοι των ανιχνευτών δηλώνονται με $C_{e_z, e_{u_1}}$ και C_{z, u_1} . Στη στατιστική προσέγγιση που ακολουθεί, ασχολούμαστε και με τις δύο περιπτώσεις υδατογραφήματων (u_1 και u_0), και για τους δύο ανιχνευτές (τον κλασικό και τον προτεινόμενο), των οποίων οι έξοδοι υποθέτουμε ότι είναι τυχαίες μεταβλητές με Gaussian κατανομή και υπολογίζονται σύμφωνα με τις ακόλουθες σχέσεις

$$\begin{aligned} C_{z, u_1} &= \frac{1}{Nk_1} \sum_{i=1}^N z_i u_{1i}, & C_{e_z, e_{u_1}} &= \frac{1}{Nl_1} \sum_{i=1}^N e_{z_i} e_{u_{1i}} \\ C_{z, u_0} &= \frac{1}{Nk_0} \sum_{i=1}^N z_i u_{0i}, & C_{e_z, e_{u_0}} &= \frac{1}{Nl_0} \sum_{i=1}^N e_{z_i} e_{u_{0i}} \end{aligned}$$

όπου

$$k_1 = \sqrt{P_z \sigma_{u_1}^2}, \quad l_1 = \sqrt{\sigma_{e_z}^2 \sigma_{e_{u_1}}^2}, \quad k_0 = \sqrt{P_z \sigma_{u_0}^2} \quad \text{και} \quad l_0 = \sqrt{\sigma_{e_z}^2 \sigma_{e_{u_0}}^2}.$$

Όλες οι ποσότητες που περιέχονται στους παραπάνω υπολογισμούς είναι γνωστές ή μπορούν να εκτιμηθούν. Για παράδειγμα, το P_z είναι η ισχύς της λαμβανόμενης εικόνας z και γνωρίζοντας το κλειδί, μπορούμε να αναπαράγουμε το υδατογράφημα w_0 (ή w_1). Καθώς η ανίχνευση είναι “τυφλή” (δε γνωρίζουμε την αρχική εικόνα), το φίλτρο σφάλματος πρόβλεψης μπορεί να εκτιμηθεί από την εικόνα z . Έτσι, μπορούμε να υπολογίσουμε και τις υπόλοιπες ποσότητες, δηλαδή την ακολουθία σφάλματος πρόβλεψης e_z , το υδατογράφημα με τη μάσκα u_0 (ή u_1), το φιλτραρισμένο υδατογράφημα με τη μάσκα e_{u_0} (ή e_{u_1}) και τις αντίστοιχες διασπορές.

Έχοντας διαθέσιμες τις εξόδους των ανιχνευτών μπορούμε να υπολογίσουμε τις μέσες τιμές και τις διασπορές τους. Εδώ, παράγουμε αναλυτικές εκφράσεις για αυτές τις μέσες τιμές και τις διασπορές, για τη γενική περίπτωση της επίθεσης με γραμμικό φίλτρο και θόρυβο. Αυτό μας βοηθά να καταλάβουμε καλύτερα ποια είναι η επίδραση της επίθεσης στην ανίχνευση και μας δίνει τη δυνατότητα να υπολογίσουμε κατώφλια προσαρμόσιμα στις συνθήκες των επιθέσεων. Λεπτομέρειες για την παραγωγή αυτών των εκφράσεων, οι οποίες φαίνονται παρακάτω,

βρίσκονται στο Παράρτημα Γ.

$$\begin{aligned}\mu_{C_{z,u_1}} &= \frac{c_1}{k_1}, \quad \mu_{C_{ez,e_{u_1}}} = \frac{c_2}{l_1}, \quad \mu_{C_{z,u_0}} = \frac{1}{k_0}(h_0 P_M \sigma_{w_0}^2 + c_3) \\ \mu_{C_{ez,e_{u_0}}} &= \frac{1}{l_0}(\mathbf{a}_z^T H \mathbf{a}_z P_M \sigma_{w_0}^2 + c_4)\end{aligned}\quad (6.3.6)$$

$$\sigma_{C_{z,u_1}}^2 = \frac{1}{N k_1} \sigma_{w_1}^2 \left[P_M (\mathbf{h}^T R_x \mathbf{h} + \sigma_n^2) + h_0^2 \sigma_{w_0}^2 r_{M^2}(0) + c_5 \right] \quad (6.3.7)$$

$$\sigma_{C_{z,u_0}}^2 = \frac{1}{N k_0} \sigma_{w_0}^2 \left[P_M (\mathbf{h}^T R_x \mathbf{h} + \sigma_n^2 + c_6) + h_0^2 \sigma_{w_0}^2 (3r_{M^2}(0) - N P_M^2) + c_7 \right] \quad (6.3.8)$$

$$\begin{aligned}\sigma_{C_{ez,e_{u_1}}}^2 &= \frac{1}{N l_1} \left[P_M \sigma_{w_1}^2 \|\mathbf{a}_z\|^2 (\mathbf{g}^T R'_x \mathbf{g} + \sigma_n^2 \|\mathbf{a}_z\|^2 + c_8) \right. \\ &\quad \left. + \text{vec}(A^T)^T E \left[U_{h_{0i}}^T \otimes U_{1i} \right] \text{vec}(A) \right] \quad (6.3.9)\end{aligned}$$

$$\begin{aligned}\sigma_{C_{ez,e_{u_0}}}^2 &= \frac{1}{N l_0} P_M \sigma_{w_0}^2 \|\mathbf{a}_z\|^2 \left[\mathbf{g}^T R'_x \mathbf{g} + \sigma_n^2 \|\mathbf{a}_z\|^2 - N \frac{(\mathbf{a}_z^T H \mathbf{a}_z)^2}{\|\mathbf{a}_z\|^2} P_M \sigma_{w_0}^2 + c_9 \right] + \\ &\frac{1}{N l_0} \left[\text{vec}(A^T)^T E \left[U_{h_{0i}}^T \otimes U_{0i} \right] \text{vec}(A) + \frac{1}{N} \sum_{i \neq j} \text{vec}(A^T)^T E \left[U_{h_{0i}}^T \otimes U_{h_{0j}}' \right] \text{vec}(A) \right] \quad (6.3.10)\end{aligned}$$

όπου $\mathbf{a}_z$ είναι ένα διάνυσμα μήκους p^2 το οποίο περιέχει τους συντελεστές του φίλτρου σφάλματος πρόβλεψης κατά γραμμές, $r_{M^2}(i) = E[M^2(n)M^2(n+i)]$ και $c_1, c_2 \dots c_9$ είναι όροι τάξης ε , όπου ε είναι ένας πολύ μικρός βαθμωτός. Τέτοιες ποσότητες στην ιδανική περίπτωση θεωρούνται ίσες με το μηδέν. Εκφράσεις της μορφής $\text{vec}(\cdot)^T E[(\cdot)^T \otimes (\cdot)] \text{vec}(\cdot)$, όπου vec είναι ο τελεστής διανυσματοποίησης, $E[\cdot]$ ο τελεστής της αναμενόμενης τιμής και το $\otimes$ δηλώνει γινόμενο Kronecker, περιέχουν όρους τάξης ίσης ή μεγαλύτερης του ε , αλλά δεν αναλύονται περαιτέρω σε αυτό το σημείο. R_x είναι ο πίνακας συσχέτισης του x , $\mathbf{g}$ είναι η συνέλιξη μεταξύ του φίλτρου σφάλματος πρόβλεψης και του φίλτρου που χρησιμοποιείται στην επίθεση (με μήκος $2p^2 - 1$) και R'_x είναι ο πίνακας συσχέτισης με διαστάσεις $(2p^2 - 1) \times (2p^2 - 1)$. Έχουμε επίσης $A = \mathbf{a}_z \mathbf{a}_z^T$, $U_{0i} = \mathbf{u}_{0i} \mathbf{u}_{0i}^T$, $U_{1i} = \mathbf{u}_{1i} \mathbf{u}_{1i}^T$ και $H = h_0 I + \hat{H}$, όπου I είναι ο μοναδιαίος πίνακας. Τα $U_{h_{0i}}$ και $U_{h_{0i}}'$ δίνονται από τις σχέσεις

$$U_{h_{0i}} = \begin{bmatrix} \dots \\ \mathbf{h}^T \mathbf{u}_{0_{n+1}} \\ \mathbf{h}^T \mathbf{u}_{0_n} \\ \mathbf{h}^T \mathbf{u}_{0_{n-1}} \\ \dots \end{bmatrix}_i \begin{bmatrix} \dots & \mathbf{h}^T \mathbf{u}_{0_{n+1}} & \mathbf{h}^T \mathbf{u}_{0_n} & \mathbf{h}^T \mathbf{u}_{0_{n-1}} & \dots \end{bmatrix}_i$$

και

$$U'_{h_{0_i}} = \begin{bmatrix} \dots \\ \mathbf{h}^T \mathbf{u}_{0_{n+1}} \\ \mathbf{h}^T \mathbf{u}_{0_n} \\ \mathbf{h}^T \mathbf{u}_{0_{n-1}} \\ \dots \end{bmatrix}_i \mathbf{u}_{0_i}^T.$$

Αν και οι εκφράσεις που υπολογίστηκαν είναι κάπως πολύπλοκες, φαίνεται καθαρά ποιες είναι οι ποσότητες που επηρεάζουν τις μέσες τιμές και τις διασπορές των εξόδων των ανιχνευτών. Αυτές οι ποσότητες είναι η διασπορά του θορύβου, σ_n^2 , και κάποιες άλλες ποσότητες που εξαρτώνται από το γραμμικό φίλτρο $\mathbf{h}$, δηλαδή οι h_0 , H , $\mathbf{g}$, $U_{h_{0_i}}$ και $U'_{h_{0_i}}$. Για παράδειγμα, η διασπορά του θορύβου, σ_n^2 , εμφανίζεται στον αριθμητή των διασπορών των ανιχνευτών. Αυτό σημαίνει ότι παρουσία θορύβου, οι έξοδοι των ανιχνευτών έχουν μεγαλύτερη διασπορά, γεγονός που δυσκολεύει την ανίχνευση του υδατογραφήματος.

Οι εκφράσεις για τις ειδικές περιπτώσεις (μη επίθεση, επίθεση με θόρυβο, επίθεση με γραμμικό φίλτρο) μπορούν να παραχθούν όπως και στη ντετερμινιστική προσέγγιση, παραλείποντας το θόρυβο ($\sigma_n^2 = 0$), και/ή θέτοντας το h_0 ίσο με ένα και τα h_i (για $i \neq 0$) ίσα με μηδέν. Για παράδειγμα, για την επίθεση του θορύβου, η έκφραση (6.3.7) μετατρέπεται ως εξής:

$$\sigma_{C_{z,u_1}}^2 = \frac{1}{Nk_1} \sigma_{w_1}^2 [P_M (P_x + \sigma_n^2) + \sigma_{w_0}^2 r_{M^2}(0) + c'_5]$$

όπου $c'_5 = P_M r_{xn}(0)$, το οποίο είναι τάξης ε λόγω του $r_{xn}(0)$.

Η απόδοση της ανίχνευσης υδατογραφήματος μπορεί να μετρηθεί σε σχέση με την πιθανότητα εσφαλμένης ανίχνευσης (P_F) και την πιθανότητα ανίχνευσης (P_D) για δεδομένη εικόνα. P_F είναι το ποσοστό των κλειδιών για τα οποία αποφασίζουμε ότι η εικόνα περιέχει ένα συγκεκριμένο υδατογράφημα, ενώ η εικόνα δεν είναι υδατογραφημένη ή περιέχει ένα διαφορετικό υδατογράφημα. P_D είναι το ποσοστό των κλειδιών για τα οποία σωστά αποφασίζουμε ότι η εικόνα περιέχει ένα συγκεκριμένο υδατογράφημα.

Ένας τρόπος απεικόνισης της απόδοσης του ανιχνευτή είναι να σχεδιαστεί η καμπύλη της P_D σε σχέση με την P_F . Η πιθανότητα εσφαλμένης ανίχνευσης δίνεται από τη σχέση

$$P_F = Q\left(\frac{\eta + \mu}{\sigma}\right) \quad (6.3.11)$$

και η πιθανότητα ανίχνευσης εκφράζεται σε σχέση με την P_F σαν

$$P_D = Q\left[Q^{-1}(P_F) - \sqrt{SNR}\right] \quad (6.3.12)$$

όπου $Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^{+\infty} e^{-\left(\frac{t^2}{2}\right)} dt$, η είναι ένα κατώφλι και $SNR = \frac{\mu^2}{\sigma^2}$, με μ και σ^2 να είναι η μέση τιμή και η διασπορά της εξόδου του ανιχνευτή (με την υπόθεση ότι η εικόνα είναι υδατογραφημένη). Κάθε σημείο της καμπύλης αντιστοιχεί σε μια τιμή (P_F, P_D) για ένα δεδομένο κατώφλι η . Καθώς το η αυξάνεται, η P_F μειώνεται και το ίδιο συμβαίνει και για την P_D (και αντίστροφα). Αυτή η μορφή απεικόνισης της απόδοσης ονομάζεται *χαρακτηριστική λειτουργίας του δέκτη* (*Receiver Operating Characteristic-ROC*) [69]. Έτσι, όπως φαίνεται από τη σχέση (7.3.13), η ROC του ανιχνευτή υδατογραφήματος εξαρτάται αποκλειστικά από την τιμή του SNR . Όσο μεγαλύτερη είναι η τιμή του SNR , τόσο μεγαλύτερη είναι και η P_D σε σχέση με μια προκαθορισμένη P_F και, συνεπώς, τόσο πιο καλή θα είναι και η απόδοση του ανιχνευτή. Παρουσιάζουμε ROC καμπύλες στην ενότητα 6.5, καθώς είναι πολύ χρήσιμες για την αποτίμηση της συμπεριφοράς του ανιχνευτή σε μεθόδους υδατογράφησης [77].

Αν η συνθήκη (6.3.5) είναι αληθής, μπορεί να αποδειχτεί ότι $SNR_{C_{ez,eu_0}} > SNR_{C_{z,u_0}}$, όπου $SNR_{C_{ez,eu_0}} = \frac{\mu_{C_{ez,eu_0}}^2}{\sigma_{C_{ez,eu_0}}^2}$ και $SNR_{C_{z,u_0}} = \frac{\mu_{C_{z,u_0}}^2}{\sigma_{C_{z,u_0}}^2}$, ενώ οι εμπλεκόμενες μέσες τιμές και διασπορές ορίζονται σύμφωνα με τις εξισώσεις (6.3.6), (6.3.8) και (6.3.10). Έτσι, ο προτεινόμενος ανιχνευτής έχει καλύτερη απόδοση σε σχέση με τον κλασικό ανιχνευτή. Αυτό επιβεβαιώνεται και από τα πειραματικά αποτελέσματα που παρουσιάζονται στην ενότητα 6.5. Με βάση τις εκφράσεις που παρήχθησαν για τις μέσες τιμές και τις διασπορές των συγκρινόμενων εξόδων των ανιχνευτών, μπορούν να υπολογιστούν κατάλληλα κατώφλια, προσαρμοσμένα στις συνθήκες των επιθέσεων.

6.4 Καταλληλότητα του προτεινόμενου ανιχνευτή

Όπως έχει ήδη αναφερθεί, εξαιτίας της δυσκολίας στην εύρεση ικανοποιητικών στατιστικών μοντέλων για πραγματικές εικόνες στο χωρικό πεδίο [55, 96], συνήθως γίνεται η υπόθεση ότι τα δεδομένα ακολουθούν Gaussian κατανομή και χρησιμοποιούνται ανιχνευτές με βάση τη συσχέτιση. Σε άλλα πεδία (FFT, DCT, wavelets), τα δεδομένα της εικόνας μπορεί να περιγράφονται καλύτερα από κάποια άλλη *συνάρτηση πυκνότητας πιθανότητας* (*probability density function-pdf*) και βέλτιστοι ανιχνευτές μπορούν να κατασκευαστούν.

Σε αυτή την ενότητα θα προσπαθήσουμε να επιβεβαιώσουμε την καταλληλότητα του προτεινόμενου ανιχνευτή, ο οποίος βασίζεται στη συσχέτιση, για το πρόβλημά μας. Ας υποθέσουμε ότι οι εικόνες στο χωρικό πεδίο μπορούν να προσεγγιστούν από μια γενικευμένη Gaussian κατανομή. Στην ανάλυσή μας χρησιμοποιούμε το στάσιμο γενικευμένο Gaussian

μοντέλο, εξαιτίας της ευρείας χρήσης του στην αποκατάσταση εικόνας και στην αποθορυβοποίηση [17, 94]. Μια γενικευμένη Gaussian pdf μέσης τιμής μηδέν δίνεται από την έκφραση

$$f_x(x) = Ae^{-|\beta x|^c} \quad (6.4.1)$$

όπου

$$A = \frac{\beta c}{2\Gamma(1/c)} \quad \text{και} \quad \beta = \frac{1}{\sigma} \left[\frac{\Gamma(3/c)}{\Gamma(1/c)} \right]^{1/2}. \quad (6.4.2)$$

Οι Gaussian και Laplacian κατανομές είναι ειδικές περιπτώσεις αυτής της pdf, για $c = 2$ και $c = 1$, αντίστοιχα. Ο βέλτιστος ανιχνευτής για το γενικευμένο Gaussian μοντέλο [54] δίνεται από τη σχέση

$$l(z) = \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} \beta^c [|z(i, j)|^c - |z(i, j) - u_0(i, j)|^c] \quad (6.4.3)$$

όπου $l(z)$ είναι η *log-likelihood* συνάρτηση, z είναι η εικόνα που ελέγχουμε και u_0 είναι το υδατογράφημα στο οποίο έχει εφαρμοστεί η μάσκα και το οποίο υποθέτουμε ότι είναι γνωστό. Τα β και c είναι παράμετροι της pdf, οι οποίες μπορούν να εκτιμηθούν από την υπό έλεγχο εικόνα. Αυτό είναι ένα μειονέκτημα του συγκεκριμένου ανιχνευτή (6.4.3), καθώς πρέπει να προστεθεί ένα στάδιο εκτίμησης παραμέτρων πριν από το στάδιο της ανίχνευσης. Για την εκτίμηση αυτών των παραμέτρων, μπορεί να χρησιμοποιηθεί μια μέθοδος *moment matching* [91].

Αφού, λοιπόν, ο ανιχνευτής που ορίζεται στην (6.4.3) είναι παραμετρικός, η συμπεριφορά του μπορεί να μελετηθεί για διαφορετικές κατανομές. Αν το $l(z)$ έχει καλή απόδοση για τιμές του c κοντά στο 2 (που αντιστοιχεί στη Gaussian κατανομή), αυτό σημαίνει ότι ένας ανιχνευτής βασισμένος στη συσχέτιση θα είναι κατάλληλος. Έτσι, πραγματοποιήσαμε τέτοια πειράματα για τιμές του c μεταξύ του 0.1 και του 2, με βήμα 0.1. Επιπλέον, μελετήσαμε την επίδραση της προτεινόμενης προ-λεύκανσης πάνω στον ανιχνευτή που ορίζεται στην (6.4.3). Η αντίστοιχη έκφραση είναι

$$l(e_z) = \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} \beta^c [|e_z(i, j)|^c - |e_z(i, j) - e_{u_0}(i, j)|^c] \quad (6.4.4)$$

όπου e_z είναι το σφάλμα πρόβλεψης της λαμβανόμενης εικόνας, z , και e_{u_0} είναι το φιλτραρισμένο (με το φίλτρο σφάλματος πρόβλεψης) υδατογράφημα με τη μάσκα. Οι παράμετροι β και c εκτιμώνται σε αυτή την περίπτωση από την ακολουθία σφάλματος πρόβλεψης, e_z .

Στα πειράματά μας, μειώσαμε τις διαστάσεις των εικόνων Lenna και Mandrill (128×128 και 119×125 , αντίστοιχα), προκειμένου να ελαττωθεί ο χρόνος των υπολογισμών. Οι διαδικασίες ένθεσης και ανίχνευσης υδατογραφήματος έγιναν για 1000 υδατογραφήματα, τα οποία

Σχήμα 6.9: Το SNR σαν συνάρτηση του c όταν δεν υπάρχει επίθεση, για τη Lenna, (α), και το Mandrill, (β).

παρήχθησαν από 1000 διαφορετικά κλειδιά, και το SNR (έχει οριστεί στην ενότητα 6.3.2) υπολογίστηκε σαν συνάρτηση της παραμέτρου c . Η ένθεση ήταν τέτοια ώστε το PSNR ανάμεσα στην αρχική και στην υδατογραφημένη εικόνα να είναι ίσο με $40dB$. Αρχικά, εξετάζεται η περίπτωση που δεν υπάρχει επίθεση και τα αποτελέσματα φαίνονται στο Σχήμα 6.9. Παρατηρούμε ότι ο ανιχνευτής αποδίδει καλύτερα για $c = 2$. Επιπλέον, η διαδικασία προλεύκανσης βελτιώνει την απόδοση του ανιχνευτή, ιδιαίτερα για την εικόνα της Lenna.

Στη συνέχεια, εξετάζουμε την περίπτωση της επίθεσης της θεωρητικής μας ανάλυσης (δηλαδή την εφαρμογή γραμμικού φίλτρου και την πρόσθεση θορύβου). Η επίθεση έγινε με προσαρμοστικό φίλτρο Wiener (το οποίο χρησιμοποιείται στο Checkmark για αποθορυβοποίηση) μεγέθους 5×5 και πρόσθεση λευκού Gaussian θορύβου ($15dB$). Μετά την πρόσθεση θορύβου, είναι λογικό η κατανομή της εξεταζόμενης εικόνας να μοιάζει περισσότερο με Gaussian. Αυτό επιβεβαιώνεται από το Σχήμα 6.10, όπου μπορούμε να δούμε ότι ο ανιχνευτής έχει καλύτερη απόδοση για $c = 2$. Παρατηρούμε επιπλέον ότι η προλεύκανση έχει ευεργετική επίδραση στον ανιχνευτή.

Τα παραπάνω πειράματα δικαιολογούν την επιλογή ενός ανιχνευτή που βασίζεται στη συσχέτιση για το συγκεκριμένο πρόβλημα, ο οποίος είναι επίσης ένας απλός ανιχνευτής που δεν απαιτεί εκτίμηση παραμέτρων. Επιπλέον, συμπεραίνουμε ότι η προτεινόμενη προλεύκανση είναι χρήσιμη και για έναν ανιχνευτή διαφορετικό από το συσχετιστή (*correlator*), αφού βελτιώνει την έξοδο του σχεδόν για κάθε τιμή του c .

Σχήμα 6.10: Το SNR σαν συνάρτηση του c μετά από εφαρμογή προσαρμοστικού φίλτρου Wiener και πρόσθεση θορύβου (15dB) για τη Lenna, (α), και το Mandrill, (β).

6.5 Πειραματικά αποτελέσματα

Προκειμένου να ελέγξουμε την ανθεκτικότητα του προτεινόμενου σχήματος ανίχνευσης σε διαφορετικά είδη επιθέσεων, πραγματοποιήθηκε μεγάλος αριθμός πειραμάτων. Επιπλέον, ελέγχθηκε η απόδοση του προτεινόμενου ανιχνευτή σε σχέση με τη νέα μάσκα. Οι διαστάσεις των εικόνων Lenna και Mandrill μετασχηματίστηκαν σε 128×128 και 119×125 , αντίστοιχα, προκειμένου να μειωθεί ο χρόνος των υπολογισμών. Σε όλες τις περιπτώσεις, οι διαδικασίες ένθεσης και ανίχνευσης εφαρμόστηκαν για 1000 κλειδιά, επιλεγμένα με τυχαίο τρόπο. Η ένθεση ήταν τέτοια, ώστε το PSNR ανάμεσα στην αρχική και στην υδατογραφημένη εικόνα να είναι 40dB . Για κάθε περίπτωση που εξετάσαμε, υπολογίσαμε το SNR (που ορίζεται στην ενότητα 6.3.2) και σχεδιάσαμε τις ROC καμπύλες για προκαθορισμένες τιμές της πιθανότητας εσφαλμένης ανίχνευσης, P_F .

Μελετήσαμε τις ακόλουθες περιπτώσεις ανίχνευσης. Πρώτα, υπολογίζεται η κανονικοποιημένη συσχέτιση ανάμεσα στη λαμβανόμενη εικόνα και το απλό υδατογράφημα ($\langle z, w \rangle$). Στη συνέχεια, υπολογίζεται η κανονικοποιημένη συσχέτιση ανάμεσα στη λαμβανόμενη εικόνα και το υδατογράφημα με τη μάσκα, για κάθε μία από τις συγκρινόμενες μάσκες, δηλαδή την M_{NVF} , την M_{HVS} και την προτεινόμενη M_e ($\langle z, u \rangle$ με NVF , $\langle z, u \rangle$ με HVS , και $\langle z, u \rangle$ με M_e). Τέλος, ελέγχεται το προτεινόμενο σχήμα ανίχνευσης για τις τρεις περιπτώσεις μασκών ($\langle e_z, e_u \rangle$ με NVF , $\langle e_z, e_u \rangle$ με HVS , και $\langle e_z, e_u \rangle$ με M_e).

Σχήμα 6.11: ROC μετά την εφαρμογή προσαρμοστικού φίλτρου Wiener και πρόσθεση θορύβου για τη Lenna, (α), και το Mandrill, (β).

Για τις παραπάνω περιπτώσεις, ελέγξαμε πρώτα την ανθεκτικότητα σε γραμμικό φιλτράρισμα και πρόσθεση θορύβου, που είναι η περίπτωση της θεωρητικής μας ανάλυσης. Οι εξεταζόμενες εικόνες πέρασαν από προσαρμοστικό φίλτρο Wiener διαστάσεων 5×5 , το οποίο είναι ένα φίλτρο αφαίρεσης θορύβου. Στη συνέχεια, προστέθηκε λευκός Gaussian θόρυβος ($15dB$). Οι ROC καμπύλες για τις δύο εξεταζόμενες εικόνες φαίνονται στο Σχήμα 6.11, όπου μπορούμε να δούμε ότι ο προτεινόμενος ανιχνευτής σε συνδυασμό με τη νέα μάσκα έχει καλύτερη απόδοση σε σχέση με τις υπόλοιπες περιπτώσεις. Αυτό επιβεβαιώνει τη θεωρητική ανάλυση και τον ισχυρισμό μας ότι η χρήση της προτεινόμενης μάσκας παρέχει περαιτέρω βελτίωση στον προτεινόμενο ανιχνευτή. Πρέπει, επίσης, να σημειώσουμε ότι η χρήση της προτεινόμενης μάσκας βελτιώνει θεαματικά και τον κλασικό ανιχνευτή (χωρίς το στάδιο της προλεύκανσης). Ειδικότερα, για τη Lenna, ο κλασικός ανιχνευτής σε συνδυασμό με την προτεινόμενη μάσκα δίνει παρόμοια αποτελέσματα με τον προτεινόμενο ανιχνευτή συνδυασμένο με τη M_{NMF} . Αυτό είναι λογικό, αφού η προτεινόμενη μάσκα κάνει δυνατή την ένθεση υδατογραφημάτων με πολύ μεγάλη ενέργεια και συνεπώς πολύ ανθεκτικών.

Ακολούθως, εξετάσαμε την περίπτωση εκτίμησης και αφαίρεσης του υδατογραφήματος. Χρησιμοποιήθηκε η επίθεση *template removal* του Chechmark, όπου το υδατογράφημα εκτιμάται από ένα προσαρμοστικό φίλτρο Wiener. Στα πειράματά μας χρησιμοποιήσαμε ένα φίλτρο διαστάσεων 5×5 . Τα αποτελέσματα φαίνονται στο Σχήμα 6.12. Η πιθανότητα εσφαλμένης ανίχνευσης παίρνει τιμές από 10^{-4} έως 0.1 με βήμα $5 \cdot 10^{-3}$, προκειμένου να δούμε τις διαφορές ανάμεσα στις συγκρινόμενες περιπτώσεις ανίχνευσης. Διαφορετικά, καθώς οι ROC καμπύλες

Σχήμα 6.12: ROC καμπύλες μετά από επίθεση *template removal* (από το *Checkmark*) για τη Lenna, (α), και για το Mandrill, (β).

είναι σχεδόν ευθείες γραμμές για την προτεινόμενη ανίχνευση, που σημαίνει ότι η πιθανότητα ανίχνευσης είναι ουσιαστικά ίση με 1, δε θα μπορούσαμε να διακρίνουμε κάποια διαφορά ανάμεσα στις συγκρινόμενες μάσκες. Όπως φαίνεται στο Σχήμα 6.12, η προτεινόμενη μάσκα βελτιώνει και για αυτή την επίθεση την απόδοση και των δύο ανιχνευτών (με και χωρίς τη προλεύκανση).

Μια ακόμα επίθεση από το *Checkmark* χρησιμοποιήθηκε για τον έλεγχο της απόδοσης των συγκρινόμενων τεχνικών, η DPR (Denoising followed by Perceptual Remodulation), όπου το μέγεθος του φίλτρου αποθορυβοποίησης ήταν 5×5 . Η μάσκα μας παραμένει η καλύτερη, όπως φαίνεται στο Σχήμα 6.13, όπου η πιθανότητα εσφαλμένης ανίχνευσης παίρνει ίδιες τιμές με το Σχήμα 6.12. Πρέπει να σημειώσουμε ότι για τη Lenna ο προτεινόμενος ανιχνευτής αποδίδει πολύ καλύτερα για τη μάσκα M_{NVF} από ότι για την μάσκα M_{HVS} . Για το Mandrill οι καμπύλες με την καλύτερη απόδοση αντιστοιχούν στην προτεινόμενη μάσκα, M_e .

Τέλος, εξετάσαμε την περίπτωση της συμπίεσης JPEG. Αυτή η περίπτωση είναι ιδιαίτερης σημασίας, καθώς οι εικόνες συνήθως αποθηκεύονται και διανέμονται σε αρχεία τύπου .jpg. Η συμπίεση μπορεί να προσεγγιστεί από ένα συνδυασμό φιλτραρίσματος (που επιλέγει τις κατάλληλες συχνότητες) και θορύβου (χβαντισμός). Έτσι, περιμένουμε ότι ο προτεινόμενος ανιχνευτής θα είναι ανθεκτικός σε συμπίεση JPEG. Αυτό επιβεβαιώνεται στα Σχήματα 6.14 και 6.15. Στο Σχήμα 6.14, φαίνεται η συνηθισμένη περίπτωση συμπίεσης (όπου ο παράγοντας ποιότητας είναι ίσος με το 75). Η πιθανότητα εσφαλμένης ανίχνευσης παίρνει τιμές από 10^{-8} έως 10^{-3} με βήμα $5 \cdot 10^{-5}$ για την Lenna και από 10^{-13} έως 10^{-5} με βήμα $5 \cdot 10^{-7}$ για

Σχήμα 6.13: ROC μετά από την επίθεση DPR του Checkmark για τη Lenna, (α), και το Mandrill, (β).

Σχήμα 6.14: ROC καμπύλες μετά από συμπίεση JPEG με παράγοντα ποιότητας (Quality Factor) 75, για τη Lenna, (α), και το Mandrill, (β).

Σχήμα 6.15: ROC καμπύλες για συμπίεση JPEG με παράγοντα ποιότητας 30, για τη Lenna, (α), και το Mandrill, (β).

το Mandrill. Όπως μπορούμε να δούμε στο Σχήμα 6.14, η χρήση της προτεινόμενης μάσκας βελτιώνει την απόδοση και των δύο ανιχνευτών. Πρέπει να σημειωθεί ότι για τον προτεινόμενο ανιχνευτή η M_{NVF} και η M_{HVS} έχουν παρόμοια απόδοση για τη Lenna, ενώ η M_{NVF} είναι καλύτερη για το Mandrill. Στο Σχήμα 6.15, εξετάζεται μια περίπτωση μεγαλύτερης συμπίεσης (με παράγοντα ποιότητας ίσο με 30), η οποία έχει παρόμοια αποτελέσματα για την εικόνα της Lenna. Για το Mandrill, ο προτεινόμενος ανιχνευτής σε συνδυασμό με την M_{HVS} είναι ελαφρώς καλύτερος.

Τα αποτελέσματα για όλες τις παραπάνω επιθέσεις φαίνονται επίσης στο Σχήμα 6.16. Όπως έχει ήδη αναφερθεί, οι ROC καμπύλες εξαρτώνται αποκλειστικά από την τιμή του SNR. Έτσι, στο Σχήμα 6.16, φαίνεται το SNR για κάθε ανιχνευτή, για όλες τις εξεταζόμενες επιθέσεις. Η νέα μάσκα σε συνδυασμό με τον προτεινόμενο ανιχνευτή έχει την καλύτερη απόδοση για πολλές ακόμα επιθέσεις, εκτός από αυτές που παρουσιάζονται εδώ. Ένας μεγάλος αριθμός πειραμάτων, για ένα μεγάλο σύνολο εικόνων, επιβεβαιώνουν τα παραπάνω αποτελέσματα. Επιπλέον, επιβεβαιώνεται η ανωτερότητα της προτεινόμενης μάσκας M_e σε σχέση με τις άλλες δύο.

6.6 Συμπεράσματα, μελλοντικές ερευνητικές κατευθύνσεις

Δύο βασικές συνεισφορές της διδακτορικής διατριβής σχετικά με το πρόβλημα της υδατογράφησης εικόνων στο χωρικό πεδίο παρουσιάστηκαν σε αυτό το κεφάλαιο. Πρώτα, παρουσιάσαμε

Σχήμα 6.16: Το SNR για διαφορετικές επιθέσεις, για τη Lenna, (α), και το Mandrill, (β).

μια νέα χωρική αντιληπτική μάσκα, η οποία περιγράφει πολύ καλά τις ιδιότητες της ανθρώπινης όρασης. Βασίζεται στο σφάλμα πρόβλεψης της αρχικής εικόνας και αυξάνει πάρα πολύ την ενέργεια των υδατογραφημάτων, ενώ ταυτόχρονα η ποιότητα της εικόνας διατηρείται σε πολύ καλό επίπεδο.

Στη συνέχεια, παρουσιάσαμε ένα βελτιωμένο “τυφλό” ανιχνευτή, ο οποίος βασίζεται σε κατάλληλη προ-λεύκανση. Η βελτιωμένη απόδοσή του αποδείχτηκε θεωρητικά για τις περιπτώσεις που δεν υπάρχει επίθεση, που γίνεται επίθεση με γραμμικό φίλτρο, καθώς και επίθεση με θόρυβο. Η θεωρητική ανάλυση είναι ανεξάρτητη της νέας μάσκας και τα αποτελέσματα ισχύουν για οποιαδήποτε τεχνική υδατογράφησης βασίζεται σε χωρική μάσκα. Αυτό επιβεβαιώνεται και από τα πειραματικά αποτελέσματα, όπου φαίνεται ότι η προτεινόμενη ανίχνευση έχει καλή απόδοση ακόμα και για μάσκες διαφορετικές από την M_e . Ωστόσο, όταν χρησιμοποιείται η M_e , η απόδοση του προτεινόμενου ανιχνευτή βελτιώνεται ακόμη περισσότερο.

Τέλος, επιπλέον εκφράσεις έχουν υπολογιστεί (μέσες τιμές και διασπορές των μετρικών συσχέτισης) για τις περιπτώσεις της θεωρητικής μας ανάλυσης, οι οποίες δίνουν τη δυνατότητα υπολογισμού κατωφλίων ανίχνευσης προσαρμοσίμων στις συνθήκες των επιθέσεων.

Όπως και στο κεφάλαιο 5, μια πιθανή ερευνητική κατεύθυνση θα μπορούσε να είναι η θεωρητική μελέτη της απόδοσης του προτεινόμενου σχήματος υδατογράφησης για μη γραμμικές επιθέσεις. Επιπλέον, το προτεινόμενο σχήμα, το οποίο ανήκει στην κατηγορία των τεχνικών “πληροφορημένης” ένθεσης, θα μπορούσε ίσως να συνδυαστεί με μια κατάλληλη τεχνική “πληροφορημένης” κωδικοποίησης, καταλήγοντας σε ένα ακόμη πιο ισχυρό σχήμα υδατογράφησης εικόνων.

Κεφάλαιο 7

Βελτιωμένη ένθεση πολλαπλασιαστικών υδατογραφημάτων μέσω χωρο-χρονικής κωδικοποίησης μπλοκ

Στο κεφάλαιο αυτό παρουσιάζουμε ένα νέο σχήμα για υδατογράφηση εικόνων στο χωρικό πεδίο, το οποίο βασίζεται σε χωρο-χρονική κωδικοποίηση μπλοκ (*STBC-Space-Time Block Coding*) [66]. Ειδικότερα, χρησιμοποιούμε μία 4×4 πραγματική, ορθογώνια διάταξη για την ένθεση ενός πολλαπλασιαστικού υδατογραφήματος στην εικόνα. Η εικόνα χωρίζεται σε τέσσερα τμήματα, τα οποία μετά από κάποιες πράξεις μπορούν να θεωρηθούν σαν τέσσερα διαφορετικά κανάλια με επίπεδη εξασθένιση (*flat fading channels*). Στο δέκτη γίνεται αποκωδικοποίηση μέγιστης πιθανοφάνειας (*maximum likelihood*), η οποία βασίζεται μόνο σε γραμμική επεξεργασία και έχει χαμηλό κόστος. Αυτό το σχήμα αποδεικνύεται πιο αποδοτικό από την επαναληπτική υδατογράφηση και εκμεταλλεύεται τα πλεονεκτήματα της χωρο-χρονικής κωδικοποίησης μπλοκ, δηλαδή επιτυγχάνει πλήρη διαφορικότητα (*diversity*) ενώ ταυτόχρονα προσφέρει το μέγιστο δυνατό ρυθμό μετάδοσης (*transmission rate*) για πραγματικές διατάξεις συμβόλων (*constellations*).

7.1 Εισαγωγή

Όπως έχουμε ήδη αναφέρει, η προστασία των πνευματικών δικαιωμάτων ψηφιακών δεδομένων και η πιστοποίηση της αυθεντικότητας τους μέσω της υδατογράφησης αποτελεί αντικείμενο

εντατικής έρευνας τα τελευταία χρόνια. Στην υδατογράφιση, όμως, υπάρχουν κάποιες αντιφατικές απαιτήσεις που πρέπει να ικανοποιούνται ταυτόχρονα. Ένα υδατογράφημα πρέπει να μην είναι ορατό, αλλά ταυτόχρονα πρέπει να είναι ανθεκτικό και να φέρει αρκετή πληροφορία. Είναι γνωστό ότι η επανάληψη του υδατογραφήματος σαν μια διαδικασία κωδικοποίησης, μαζί με τη χρήση κάποιας αντιληπτικής μάσκας κατά την ένθεση αυξάνουν σημαντικά την ανθεκτικότητα. Ωστόσο, η επανάληψη του υδατογραφήματος έχει κάποια μειονεκτήματα. Καταρχήν, μπορεί να θεωρηθεί σαν ένας απλός κώδικας επανάληψης, ο οποίος δεν είναι κατάλληλος για κωδικοποίηση καναλιού. Επίσης, η εξασθένιση εξαιτίας κάποιας προσαρμοστικής αντιληπτικής μάσκας ή κάποιας πιθανής επίθεσης μπορεί να μειώσει την απόδοση του τελικού συστήματος υδατογράφησης. Επιπλέον, πρέπει να γίνει ένας συμβιβασμός ανάμεσα στην ανθεκτικότητα του υδατογραφήματος στην *επίθεση της αποκοπής (cropping attack)* και στο ρυθμό του κώδικα (*code rate*). Αυτό όμως απαιτεί την εξάπλωση του υδατογραφήματος σε ολόκληρη την εικόνα. Έτσι, πρέπει να βρεθεί ο κατάλληλος συμβιβασμός ανάμεσα στον πλεονασμό του υδατογραφήματος και στην αποδοτικότητα της κωδικοποίησης. Κάποιες πιθανές λύσεις, προκειμένου να επιτευχθεί αυτός ο συμβιβασμός, είναι η χρήση τεχνικών διαφορικότητας ή σύνθετων σχημάτων κωδικοποίησης, όπως η *κωδική διαμόρφωση με χρήση trellis (trellis-coded modulation)* [147].

Στην εργασία [75] χρησιμοποιείται διαφορικότητα μέσω της επανάληψης του υδατογραφήματος. Εισάγεται ένα μη στάσιμο, παράλληλο, δυαδικό και συμμετρικό κανάλι και η εκτίμησή του επιτυγχάνεται με τη βοήθεια ενός υδατογραφήματος αναφοράς. Οι συγγραφείς της εργασίας [147] θεωρούν διαφορικότητα με μια *βέλτιστη κωδικοποίηση σήματος ανά περιόδους (interperiod optimal signal encoding)* βασισμένη σε επαναληπτικούς κώδικες, όπως οι κώδικες turbo και οι *κώδικες ισότητας χαμηλής πυκνότητας (low-density parity check codes)*, με τον αποκωδικοποιητή να κάνει εκτίμηση της κατάστασης του καναλιού υδατογράφησης με τη βοήθεια ενός πιλοτικού υδατογραφήματος αναφοράς. Στην ένθεση χρησιμοποιείται *διαμόρφωση BPSK (Binary Phase Shift Keying modulation)*, ενώ ακόμα θεωρείται ένα γενικευμένο κανάλι υδατογράφησης το οποίο περιλαμβάνει γεωμετρικές επιθέσεις, εξασθένιση και πρόσθεση προσθετικού μη-Gaussian θορύβου.

Στην εργασία [82], παρουσιάζεται μια τεχνική υδατογράφησης στο χωρικό πεδίο βασισμένη σε μια ειδική μορφή διαφορικότητας στη μετάδοση. Ειδικότερα, οι RGB συνιστώσες των έγχρωμων εικόνων θεωρούνται σαν τρία ανεξάρτητα κανάλια με αργή εξασθένιση. Το ίδιο υδατογράφημα με διαφορετική μορφή, διαμορφωμένο από διαφορετικές ψευδοτυχαίες ακολουθίες και *interleavers*, τοποθετείται στα τρία χρωματικά κανάλια ταυτόχρονα. Η εξαγωγή του υδατογραφήματος στηρίζεται στη σύγκριση των ενεργειών των τριών καναλιών ή στη συσώρευση της ενέργειάς τους. Στην εργασία [83], παρουσιάζεται η αντίστοιχη διαδικασία στο

πεδίο του μετασχηματισμού DCT.

Σε αυτό το κεφάλαιο παρουσιάζουμε μια καινούρια τεχνική υδατογράφησης, εμπνευσμένη από τη χωρο-χρονική κωδικοποίηση μπλοκ, η οποία έχει προταθεί σχετικά πρόσφατα στις ασύρματες επικοινωνίες σαν μια τεχνική μετάδοσης με διαφορικότητα [132, 133]. Η χωρο-χρονική κωδικοποίηση μπλοκ αποτελεί γενίκευση του σχήματος μετάδοσης που προτάθηκε από τον Alamouti [1] για αυθαίρετο αριθμό κεραιών μετάδοσης. Στο σχήμα του Alamouti χρησιμοποιούνται δύο κεραίες μετάδοσης και μια κεραία λήψης. Επίσης, το σχήμα αυτό μπορεί να γενικευτεί σε μια διάταξη με δύο κεραίες μετάδοσης και M κεραίες λήψης, παρέχοντας διαφορικότητα τάξης $2M$. Η χωρο-χρονική κωδικοποίηση μπλοκ έχει κάποια σημαντικά χαρακτηριστικά. Καταρχήν, επιτυγχάνει τη μέγιστη διαφορικότητα που μπορεί να παρέχει μια διάταξη κεραιών μετάδοσης και λήψης και είναι λιγότερο σύνθετη από τη χωρο-χρονική κωδικοποίηση trellis. Δεν απαιτεί τη μετάδοση κάποιας πληροφορίας για την κατάσταση του καναλιού από το δέκτη στον πομπό και, επιπλέον, ένας πολύ απλός αλγόριθμος αποκωδικοποίησης μέγιστης πιθανοφάνειας χρησιμοποιείται στο δέκτη, ο οποίος βασίζεται μόνο σε γραμμική επεξεργασία. Τέλος, για διατάξεις πραγματικών συμβόλων (όπως η PAM-Pulse Amplitude Modulation κωδικοποίηση), παρέχει το μέγιστο δυνατό ρυθμό μετάδοσης που επιτρέπει η σχετική θεωρία [134].

Θεωρώντας την υδατογράφηση σαν ένα πρόβλημα επικοινωνιών, προσπαθούμε να βελτιώσουμε την ανίχνευση του υδατογραφήματος εφαρμόζοντας διαφορικότητα βασισμένη σε ορθογώνιους κώδικες κατά τη μετάδοση του. Μελετάμε την περίπτωση τεσσάρων πομπών και ενός δέκτη, αν και επεκτάσεις και σε άλλες διατάξεις είναι δυνατές ακολουθώντας παρόμοια διαδικασία. Η εικόνα χωρίζεται σε τέσσερα τμήματα, τα οποία μετά από μερικές απλές πράξεις στο χωρικό πεδίο μπορούν να θεωρηθούν σαν τέσσερα διαφορετικά κανάλια επίπεδης εξασθένισης. Το υδατογράφημα εισάγεται στην εικόνα με πολλαπλασιαστικό τρόπο σύμφωνα με μια πραγματική ορθογώνια διάταξη, διαστάσεων 4×4 [133]. Στο δέκτη η αποκωδικοποίηση μέγιστης πιθανοφάνειας επιτυγχάνεται με απλό τρόπο. Ένα κατάλληλο φίλτρο (*matched filter*) εφαρμόζεται στο λαμβανόμενο σήμα, το οποίο αποσυνδέει τα μεταδιδόμενα σήματα, ενώ για την ανίχνευση του υδατογραφήματος χρησιμοποιείται η γραμμική συσχέτιση. Αποδεικνύεται ότι το προτεινόμενο σχήμα υδατογράφησης έχει καλύτερη απόδοση σε σχέση με την επαναληπτική υδατογράφηση και είναι ανθεκτικό σε αρκετές διαφορετικές επιθέσεις, όπως η πρόσθεση προσθετικού λευκού Gaussian θορύβου, το φιλτράρισμα, η JPEG συμπίεση, και άλλες.

7.2 Διατύπωση του προβλήματος και μοντέλο καναλιού

Η εικόνα που πρόκειται να υδατογραφηθεί χωρίζεται σε τέσσερα τμήματα και μετά από μερικές πράξεις κάθε ένα από αυτά τα τμήματα αντιστοιχεί σε ένα διαφορετικό κανάλι επίπεδης

εξασθένησης. Η ένθεση είναι πολλαπλασιαστική και η υδατογραφημένη εικόνα δίνεται από τη σχέση

$$y_i = x_i + x_i w_i a = x_i(1 + w_i a) \quad (7.2.1)$$

όπου το x είναι η αρχική εικόνα, το w είναι το υδατογράφημα, a είναι ένας πραγματικός αριθμός που καθορίζει την ενέργεια του υδατογραφήματος και το $i = 1, 2, 3, 4$ είναι ένας δείκτης που δηλώνει το τμήμα της εικόνας. Αν χρησιμοποιούσαμε επαναληπτική υδατογράφιση, τότε τα τέσσερα υδατογραφήματα w_i που θα αντιστοιχούσαν στα τέσσερα τμήματα της εικόνας θα ήταν ίσα με ένα μοναδικό υδατογράφημα. Υποθέτουμε ότι το υδατογράφημα αυτό έχει μέση τιμή μηδέν και προκύπτει από μια γεννήτρια που παράγει τους αριθμούς $\{-1, 1\}$ με ψευδοτυχαίο τρόπο, με βάση ένα μυστικό κλειδί. Στο προτεινόμενο σχήμα τα τέσσερα υδατογραφήματα είναι διαφορετικά, αλλά σχετίζονται κατάλληλα υπό την έννοια της χωρο-χρονικής κωδικοποίησης μπλοκ.

Όπως έχει ήδη αναφερθεί, θέλουμε να εκφράσουμε τα τμήματα της εικόνας σαν τέσσερα διαφορετικά κανάλια με επίπεδη εξασθένιση. Μπορούμε λοιπόν να γράψουμε την αρχική εικόνα x σαν άθροισμα της μέσης τιμής της και του αντίστοιχου σφάλματος, όπως φαίνεται στην ακόλουθη σχέση

$$x_i = m_i + e_i \quad (7.2.2)$$

όπου e_i είναι μια ακολουθία μέσης τιμής μηδέν. Χρησιμοποιώντας τη σχέση (7.2.2) μέσα στη σχέση (7.2.1), προκύπτει ότι η υδατογραφημένη εικόνα είναι ίση με

$$y_i = m_i(1 + w_i a) + e_i(1 + w_i a). \quad (7.2.3)$$

Η χρήση πολλαπλασιαστικής ένθεσης βοηθά στο να εκφράσουμε τα τμήματα της εικόνας σαν κανάλια επίπεδης εξασθένησης. Αν θεωρήσουμε ότι το m_i (η μέση τιμή) είναι μια σταθερά για κάθε μπλοκ, τότε αυτό μπορεί να θεωρηθεί σαν την παράμετρο εξασθένησης του αντίστοιχου καναλιού επίπεδης εξασθένησης. Ο όρος που απομένει στη σχέση (7.2.3), η ακολουθία σφάλματος e_i πολλαπλασιασμένη σημείο προς σημείο με τη συνάρτηση του υδατογραφήματος $1 + w_i a$, μπορεί να θεωρηθεί θόρυβος. Υποθέτουμε, επίσης, ότι στην υδατογραφημένη εικόνα προστίθεται λευκός Gaussian θόρυβος n , με μηδενική μέση τιμή, κατά τη διάρκεια της μετάδοσής της.

7.3 Προτεινόμενος αλγόριθμος

Η έννοια της διαφορικότητας έχει μελετηθεί εκτενώς στις ασύρματες επικοινωνίες και έχει αποδειχτεί πολύ ισχυρό εργαλείο, το οποίο βελτιώνει την απόδοση των ασύρματων συστημάτων.

Σχήμα 7.1: Προτεινόμενη ένθεση υδατογραφήματος

Καθώς η υδατογράφηση μπορεί να θεωρηθεί σαν ένα πρόβλημα επικοινωνιών, προσπαθούμε να βρούμε έναν τρόπο να βελτιώσουμε την ανίχνευση του υδατογραφήματος, εφαρμόζοντας διαφορικότητα στη μετάδοση βασισμένη σε ορθογώνιους κώδικες. Για λόγους απλότητας, χρησιμοποιούμε μια διάταξη με τέσσερεις πομπούς και ένα δέκτη. Έτσι, η ακόλουθη πραγματική ορθογώνια διάταξη [133], διαστάσεων 4×4 , χρησιμοποιείται κατά την ένθεση με κατάλληλο τρόπο

$$\begin{pmatrix} w(4k+1) & w(4k+2) & w(4k+3) & w(4k+4) \\ -w(4k+2) & w(4k+1) & -w(4k+4) & w(4k+3) \\ -w(4k+3) & w(4k+4) & w(4k+1) & -w(4k+2) \\ -w(4k+4) & -w(4k+3) & w(4k+2) & w(4k+1) \end{pmatrix} \quad (7.3.1)$$

όπου για δεδομένο k η i -οστή στήλη περιέχει τέσσερα διαφορετικά σύμβολα, τα οποία μεταδίδονται από τον αντίστοιχο i -οστό πομπό. Πρέπει να σημειώσουμε ότι τα ίδια σύμβολα, με διαφορετική διάταξη και πρόσημο, μεταδίδονται από όλους τους πομπούς.

7.3.1 Ένθεση υδατογραφήματος

Χρησιμοποιώντας την ορθογώνια διάταξη (7.3.1), το αρχικό υδατογράφημα μετασχηματίζεται σε τέσσερα διαφορετικά υδατογραφήματα, τα οποία εισάγονται στα τέσσερα τμήματα της εικόνας, όπως φαίνεται στο Σχήμα 7.1. Τα δεδομένα που πρόκειται να μεταδοθούν μέσω

των τεσσάρων διαφορετικών καναλιών (τμημάτων της εικόνας) είναι $w(4k+1)$, $w(4k+2)$, $w(4k+3)$ και $w(4k+4)$, όπου $k = 0, 1, \dots, \frac{N^2}{16} - 1$ και N^2 είναι ο αριθμός των εικονοστοιχείων της εικόνας. Κάθε τμήμα της εικόνας περιέχει $N^2/4$ εικονοστοιχεία, των οποίων η ανάγνωση γίνεται κατά στήλες. Προφανώς, ο αριθμός των bits του υδατογραφήματος που θα εισαχθεί στην εικόνα είναι ίσος με $N^2/4$. Στο σημείο $(1, 1)$ του κάθε μπλοκ της εικόνας υπάρχει ένα διαφορετικό στοιχείο του υδατογραφήματος και όλα μαζί αντιστοιχούν στην πρώτη γραμμή της διάταξης (7.3.1). Στα σημεία $(2, 1)$ των τεσσάρων τμημάτων της εικόνας εισάγονται τα στοιχεία του υδατογραφήματος που αντιστοιχούν στην δεύτερη γραμμή της διάταξης (7.3.1), ενώ η ίδια λογική ακολουθείται και στη συνέχεια, κατά στήλες, μέχρι να εισαχθούν όλα τα στοιχεία του υδατογραφήματος στην εικόνα. Έτσι, τα τέσσερα διαφορετικά υδατογραφήματα που εισάγονται στα τέσσερα τμήματα της εικόνας σχετίζονται κατάλληλα μεταξύ τους, υπό την έννοια της χωρο-χρονικής κωδικοποίησης μπλοκ. Προσθέτοντας τα τέσσερα τμήματα της εικόνας στο δέκτη, τα λαμβανόμενα σήματα που αντιστοιχούν στους δείκτες $4k+1$, $4k+2$, $4k+3$ και $4k+4$ είναι

$$\begin{aligned} r(4k+1) = & a[m_1w(4k+1) + m_2w(4k+2) + m_3w(4k+3) + m_4w(4k+4)] \\ & + \epsilon(4k+1) + (m_1 + m_2 + m_3 + m_4) + n(4k+1) \end{aligned} \quad (7.3.2)$$

όπου $\epsilon(4k+1) = e(4k+1)[1+w(4k+1)a] + e(4k+2)[1+w(4k+2)a] + e(4k+3)[1+w(4k+3)a] + e(4k+4)[1+w(4k+4)a]$,

$$\begin{aligned} r(4k+2) = & a[m_2w(4k+1) - m_1w(4k+2) + m_4w(4k+3) - m_3w(4k+4)] \\ & + \epsilon(4k+2) + (m_1 + m_2 + m_3 + m_4) + n(4k+2) \end{aligned} \quad (7.3.3)$$

όπου $\epsilon(4k+2) = e(4k+2)[1+w(4k+1)a] + e(4k+1)[1-w(4k+2)a] + e(4k+4)[1+w(4k+3)a] + e(4k+3)[1-w(4k+4)a]$,

$$\begin{aligned} r(4k+3) = & a[m_3w(4k+1) - m_4w(4k+2) - m_1w(4k+3) + m_2w(4k+4)] \\ & + \epsilon(4k+3) + (m_1 + m_2 + m_3 + m_4) + n(4k+3) \end{aligned} \quad (7.3.4)$$

όπου $\epsilon(4k+3) = e(4k+3)[1+w(4k+1)a] + e(4k+4)[1-w(4k+2)a] + e(4k+1)[1-w(4k+3)a] + e(4k+2)[1+w(4k+4)a]$, και

$$\begin{aligned} r(4k+4) = & a[m_4w(4k+1) + m_3w(4k+2) - m_2w(4k+3) - m_1w(4k+4)] \\ & + \epsilon(4k+4) + (m_1 + m_2 + m_3 + m_4) + n(4k+4) \end{aligned} \quad (7.3.5)$$

όπου $\epsilon(4k+4) = e(4k+4)[1+w(4k+1)a] + e(4k+3)[1+w(4k+2)a] + e(4k+2)[1-w(4k+3)a] + e(4k+1)[1-w(4k+4)a]$.

Αν ορίσουμε τα διανύσματα

$$\mathbf{r} = \begin{bmatrix} r(4k+1) \\ r(4k+2) \\ r(4k+3) \\ r(4k+4) \end{bmatrix}, \quad \mathbf{w} = \begin{bmatrix} w(4k+1) \\ w(4k+2) \\ w(4k+3) \\ w(4k+4) \end{bmatrix}, \quad \boldsymbol{\epsilon} = \begin{bmatrix} \epsilon(4k+1) \\ \epsilon(4k+2) \\ \epsilon(4k+3) \\ \epsilon(4k+4) \end{bmatrix}, \quad \mathbf{n} = \begin{bmatrix} n(4k+1) \\ n(4k+2) \\ n(4k+3) \\ n(4k+4) \end{bmatrix}$$

και $\mathbf{1}_{4 \times 1} = \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}^T$ και τον πίνακα

$$H = \begin{bmatrix} m_1 & m_2 & m_3 & m_4 \\ m_2 & -m_1 & m_4 & -m_3 \\ m_3 & -m_4 & -m_1 & m_2 \\ m_4 & m_3 & -m_2 & -m_1 \end{bmatrix}$$

οι λαμβανόμενες ακολουθίες μπορούν να εκφραστούν μέσω της σχέσης

$$\mathbf{r} = H\mathbf{w}a + \boldsymbol{\epsilon} + \mathbf{n} + (m_1 + m_2 + m_3 + m_4)\mathbf{1}_{4 \times 1}. \quad (7.3.6)$$

Να σημειωθεί ότι ο πίνακας που εκφράζει το κανάλι, H , είναι ορθογώνιος και ισχύει

$$H^T H = (m_1^2 + m_2^2 + m_3^2 + m_4^2)I. \quad (7.3.7)$$

7.3.2 Ανίχνευση υδατογραφήματος

Υποθέτοντας ότι οι μέσες τιμές των τμημάτων των εικόνων παραμένουν κατά προσέγγιση οι ίδιες μετά την υδατογράφηση, ο πίνακας H μπορεί να εκτιμηθεί από την υδατογραφημένη εικόνα. Αυτό σημαίνει ότι η ανίχνευση είναι “τυφλή”, ότι δηλαδή δε χρειάζεται γνώση της αρχικής εικόνας στον ανιχνευτή. Ο πίνακας H^T μπορεί να χρησιμοποιηθεί σαν το matched φίλτρο στο δέκτη, το οποίο θα αποσυνδέσει τα μεταδιδόμενα σήματα. Πολλαπλασιάζοντας το λαμβανόμενο σήμα $\mathbf{r}$ με τον πίνακα H^T , έχουμε

$$\begin{aligned} \mathbf{z} = H^T \mathbf{r} &= H^T H \mathbf{w}a + H^T \boldsymbol{\epsilon} + H^T \mathbf{n} + (m_1 + m_2 + m_3 + m_4)H^T \mathbf{1}_{4 \times 1} \\ &= a(m_1^2 + m_2^2 + m_3^2 + m_4^2)\mathbf{w} + \mathbf{v} + \mathbf{c} \end{aligned} \quad (7.3.8)$$

όπου $\mathbf{v} = H^T \boldsymbol{\epsilon} + H^T \mathbf{n}$ και $\mathbf{c} = (m_1 + m_2 + m_3 + m_4)H^T \mathbf{1}_{4 \times 1}$. Το διάνυσμα $\mathbf{c}$ αποτελείται από σταθερά στοιχεία. Ειδικότερα,

$$\mathbf{c} = \begin{bmatrix} (m_1 + m_2 + m_3 + m_4)^2 \\ (m_2 + m_3)^2 - (m_1 + m_4)^2 \\ (m_3 + m_4)^2 - (m_1 + m_2)^2 \\ (m_2 + m_4)^2 - (m_1 + m_3)^2 \end{bmatrix}$$

Σχήμα 7.2: Επαναληπτική ένθεση υδατογραφήματος

και μπορεί να εκτιμηθεί και αυτό από την υδατογραφημένη εικόνα. Αφαιρώντας το $\mathbf{c}$ από την έξοδο του matched φίλτρου $\mathbf{z}$, παίρνουμε

$$\mathbf{z}' = a(m_1^2 + m_2^2 + m_3^2 + m_4^2)\mathbf{w} + H^T\boldsymbol{\epsilon} + H^T\mathbf{n} \quad (7.3.9)$$

και αυτή είναι η ακολουθία που θα χρησιμοποιηθεί κατά τη διαδικασία ανίχνευσης του υδατογραφήματος. Να σημειωθεί ότι οι ακολουθίες $H^T\boldsymbol{\epsilon} + H^T\mathbf{n}$ περιέχουν μεταβλητές σφάλματος και θόρυβο. Αν αυτοί οι όροι ήταν γνωστοί, θα μπορούσαμε να εκτιμήσουμε το $\mathbf{w}$ από το $\mathbf{z}'$, με κατάλληλες πράξεις.

Στην περίπτωση μας, οι παραπάνω όροι δεν είναι γνωστοί και έτσι ακολουθήσαμε μια κλασική στρατηγική για την ανίχνευση του υδατογραφήματος, χρησιμοποιώντας έναν ανιχνευτή βασισμένο στη γραμμική συσχέτιση. Υπολογίζουμε τη γραμμική συσχέτιση ανάμεσα στο $\mathbf{z}'$ και στο $\mathbf{w}$ και καθώς το διάνυσμα $\mathbf{w}$ (υδατογράφημα) είναι ασυσχέτιστο με τους όρους $H^T\boldsymbol{\epsilon} + H^T\mathbf{n}$ (σφάλμα και θόρυβος), ισχύει ότι

$$E[\mathbf{z}'\mathbf{w}] = a(m_1^2 + m_2^2 + m_3^2 + m_4^2)\sigma_w^2 \quad (7.3.10)$$

όπου το E δηλώνει την αναμενόμενη τιμή (*expectation*).

Για λόγους σύγκρισης θα εξετάσουμε στη συνέχεια την περίπτωση της επαναληπτικής ένθεσης υδατογραφήματος, η οποία απεικονίζεται στο Σχήμα 7.2. Σε αυτή την περίπτωση, το υδατογράφημα εισάγεται και στα τέσσερα τμήματα της εικόνας και η αναμενόμενη τιμή της γραμμικής συσχέτισης ανάμεσα στο λαμβανόμενο σήμα y' και στο υδατογράφημα w είναι

$$E[y'w] = a(m_1 + m_2 + m_3 + m_4)\sigma_w^2 \quad (7.3.11)$$

όπου $y' = y_1 + y_2 + y_3 + y_4 + n$. Είναι προφανές ότι το $E[y'w]$ έχει μικρότερη τιμή από το $E[\mathbf{z}'\mathbf{w}]$.

Η απόδοση του ανιχνευτή μπορεί να μετρηθεί σε σχέση με την πιθανότητα εσφαλμένης ανίχνευσης (P_F) και την πιθανότητα ανίχνευσης (P_D). Υπενθυμίζουμε ότι P_F είναι το ποσοστό των κλειδιών για τα οποία αποφασίζουμε ότι το επιθυμητό υδατογράφημα είναι παρόν, ενώ δεν είναι, και P_D είναι το ποσοστό των κλειδιών για τα οποία σωστά αποφασίζουμε ότι το επιθυμητό υδατογράφημα είναι παρόν.

Ένας τρόπος απεικόνισης της απόδοσης του ανιχνευτή είναι η γραφική παράσταση της πιθανότητας P_D σε σχέση με την P_F . Η πιθανότητα της εσφαλμένης ανίχνευσης ορίζεται σαν

$$P_F = Q\left(\frac{\eta + \mu}{\sigma}\right) \quad (7.3.12)$$

και η πιθανότητα ανίχνευσης ορίζεται σε σχέση με την P_F σαν:

$$P_D = Q\left[Q^{-1}(P_F) - \sqrt{SNR}\right] \quad (7.3.13)$$

όπου $Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^{+\infty} e^{-\left(\frac{t^2}{2}\right)} dt$, η είναι ένα κατώφλι και $SNR = \frac{m^2}{\sigma^2}$, με m και σ^2 να είναι η μέση τιμή και η διασπορά της εξόδου του ανιχνευτή (με την υπόθεση ότι το υδατογράφημα είναι παρόν). Κάθε σημείο της γραφικής παράστασης αντιστοιχεί σε μια τιμή (P_F, P_D) , για ένα δεδομένο κατώφλι η . Καθώς αυξάνεται το η , μειώνεται το P_F και το ίδιο συμβαίνει και για το P_D (και αντίστροφα). Όπως είπαμε και στο κεφάλαιο 6, αυτή η μορφή απεικόνισης ονομάζεται χαρακτηριστική λειτουργίας του δέκτη (ROC-Receiver Operating Characteristic) [69]. Από τη σχέση (7.3.13) φαίνεται ότι η ROC καμπύλη του ανιχνευτή υδατογραφήματος εξαρτάται αποκλειστικά από την τιμή του SNR . Όσο μεγαλύτερο είναι το SNR , τόσο μεγαλύτερη είναι η πιθανότητα ανίχνευσης P_D , η οποία σχετίζεται με κάποια συγκεκριμένη P_F , και κατά συνέπεια τόσο καλύτερη απόδοση έχει ο ανιχνευτής.

Εδώ, μπορεί να αποδειχτεί ότι $SNR_{E[\mathbf{z}'\mathbf{w}]} > SNR_{E[\mathbf{y}'\mathbf{w}]}$, όπου $SNR_{E[\mathbf{z}'\mathbf{w}]} = \frac{m_{E[\mathbf{z}'\mathbf{w}]}^2}{\sigma_{E[\mathbf{z}'\mathbf{w}]}^2}$ και $SNR_{E[\mathbf{y}'\mathbf{w}]} = \frac{m_{E[\mathbf{y}'\mathbf{w}]}^2}{\sigma_{E[\mathbf{y}'\mathbf{w}]}^2}$. Αυτό επιβεβαιώνεται επίσης από μεγάλο αριθμό πειραματικών αποτελεσμάτων. Συμπεραίνουμε λοιπόν ότι το προτεινόμενο σχήμα έχει πολύ καλύτερη απόδοση σε σχέση με την επαναληπτική υδατογράφηση.

7.4 Πειραματικά αποτελέσματα

Προκειμένου να ελεγχθεί η απόδοση του ανιχνευτή του προτεινόμενου σχήματος, σε σχέση με την επαναληπτική υδατογράφηση, πραγματοποιήσαμε μεγάλο αριθμό πειραμάτων. Παρουσιάζουμε κάποια ενδεικτικά αποτελέσματα για την εικόνα της Lenna, διαστάσεων (256×256) . Σε όλες τις περιπτώσεις που θα παρουσιάσουμε, οι διαδικασίες ένθεσης και ανίχνευσης του

(α)

(β)

Σχήμα 7.3: (α) Αρχική εικόνα, (β) Υδατογραφημένη εικόνα

υδατογραφήματος εφαρμόστηκαν για 1000 διαφορετικά κλειδιά, επιλεγμένα με τυχαίο τρόπο. Η παράμετρος a που καθορίζει την ενέργεια του υδατογραφήματος ήταν τέτοια, ώστε το PSNR (Peak Signal to Noise Ratio) ανάμεσα στην αρχική και στην υδατογραφημένη εικόνα να είναι $40dB$. Η αρχική και η υδατογραφημένη εικόνα φαίνονται στο Σχήμα 7.3. Για κάθε περίπτωση που εξετάσαμε, υπολογίσαμε το SNR και σχεδιάσαμε τις ROC καμπύλες για συγκεκριμένες τιμές της πιθανότητας εσφαλμένης ανίχνευσης.

Εξετάσαμε επιθέσεις όπως ο θόρυβος, η συμπίεση JPEG, το φιλτράρισμα (γραμμικό και μη-γραμμικό) και η υπο-δειγματοληψία της εικόνας στα $3/4$ των αρχικών της διαστάσεων και ακολούθως η υπερ-δειγματοληψία της στις αρχικές της διαστάσεις. Υπολογίσαμε το $SNR_{E[z'w]}$ και το $SNR_{E[y'w]}$ για τις δύο συγκρινόμενες τεχνικές, καθώς και τη διαφορά

Πίνακας 7.1: ΤΙΜΕΣ ΤΟΥ SNR(dB) ΓΙΑ ΔΙΑΦΟΡΕΤΙΚΕΣ ΕΠΙΘΕΣΕΙΣ

Επιθέσεις	$SNR_{E[z'w]}$	$SNR_{E[y'w]}$	SNR_{gain}
Καμιά επίθεση	15.5012	7.7878	7.7133
<i>AWGN</i>			
15dB	11.1565	7.7547	3.4017
10dB	10.2439	7.5988	2.6451
5dB	8.2833	7.3628	0.9205
<i>Συμπίεση JPEG</i>			
QF:75	7.4756	0.5278	6.9478
QF:50	3.8729	-3.7032	7.5761
Wiener (3×3)	1.0322	-6.4998	7.5320
Median (3×3)	2.2647	-5.4501	7.7148
Sample down-up	9.2423	2.6375	6.6048

Σχήμα 7.4: ROC καμπύλες για (α) καμία επίθεση και (β) *sampling down-up*

τους $SNR_{gain} = SNR_{E[z'w]} - SNR_{E[y'w]}$. Τα αποτελέσματα φαίνονται στον Πίνακα 7.1. Είναι προφανές ότι η απόδοση της ανίχνευσης είναι σημαντικά βελτιωμένη για το προτεινόμενο σχήμα, ακόμα και στις περιπτώσεις σοβαρών επιθέσεων. Στη συνέχεια, φαίνονται κάποιες χαρακτηριστικές ROC καμπύλες, οι οποίες επιβεβαιώνουν τα αποτελέσματα του Πίνακα 7.1. Στο Σχήμα 7.4, μπορούμε να δούμε τις καμπύλες που αφορούν την περίπτωση που δεν υπάρχει επίθεση και την περίπτωση που έχει γίνει υπό- και υπερ-δειγματοληψία της εικόνας. Οι καμπύλες για την επίθεση του θορύβου (10dB) και της JPEG συμπίεσης ($QF : 75$) και οι καμπύλες για τα φίλτρα Wiener και median φαίνονται στα Σχήματα 7.5 και 7.6, αντίστοιχα.

7.5 Συμπεράσματα, μελλοντικές ερευνητικές κατευθύνσεις

Στο κεφάλαιο αυτό προτάθηκε ένα νέο σχήμα για υδατογράφιση εικόνων, το οποίο βασίζεται σε χωρο-χρονική κωδικοποίηση μπλοκ. Πιο συγκεκριμένα, χρησιμοποιήθηκε μια πραγματική ορθογώνια διάταξη, διαστάσεων 4×4 . Ένα τέτοιο σχήμα κωδικοποίησης παρέχει πλήρη διαφορικότητα, το μεγαλύτερο δυνατό ρυθμό μετάδοσης και χαμηλή πολυπλοκότητα. Η απόδοση του προτεινόμενου σχήματος (όσο αφορά την ανίχνευση) είναι πολύ καλύτερη σε σχέση με την κλασική μέθοδο επαναληπτικής υδατογράφησης.

Ακολουθώντας αντίστοιχη διαδικασία, το προτεινόμενο σχήμα μπορεί να γενικευτεί για M πομπούς. Επιπλέον, μπορεί να συνδυαστεί κατάλληλα με ένα σχήμα επαναληπτικής ένθεσης, προκειμένου να είναι ανθεκτικό σε περισσότερες επιθέσεις, όπως για παράδειγμα η αποκοπή

(*cropping*). Τέλος, μπορεί να χρησιμοποιηθεί σε συνδυασμό με μια αντιληπτική μάσκα που αυξάνει σημαντικά την ενέργεια του υδατογραφήματος [65], παρέχοντας ακόμα πιο βελτιωμένη απόδοση.

Σχήμα 7.5: ROC καμπύλες για (α) AWGN και (β) συμπίεση JPEG

Σχήμα 7.6: ROC καμπύλες για φίλτρο (α) Wiener και (β) median

Κεφάλαιο 8

Γενικά συμπεράσματα. Μελλοντικές ερευνητικές κατευθύνσεις

Σε αυτή τη διδακτορική διατριβή, ασχοληθήκαμε με δύο πολύ σημαντικά προβλήματα της ψηφιακής επεξεργασίας εικόνων, την αντιστοίχιση και την ψηφιακή υδατογράφηση εικόνων. Η αντιστοίχιση εικόνων έχει σαν σκοπό την εύρεση γεωμετρικών και άλλων διαφορών ανάμεσα σε δύο ή περισσότερες εικόνες μιας σκηνής και η ψηφιακή υδατογράφηση εικόνων έρχεται να προσφέρει κατοχύρωση των πνευματικών δικαιωμάτων των ψηφιακών δεδομένων. Το πρόβλημα της αντιστοίχισης μπορεί να αποτελέσει τμήμα της ψηφιακής υδατογράφησης, στη φάση της ανίχνευσης του υδατογραφήματος (συγχρονισμός). Επιπλέον, για την ανίχνευση του υδατογραφήματος χρησιμοποιούνται παρόμοιες ή και ίδιες μετρικές με αυτές που χρησιμοποιούνται στην αντιστοίχιση εικόνων (συσχέτιση, κανονικοποιημένη συσχέτιση, κ.τ.λ.). Έτσι, τα δύο προβλήματα δεν είναι ξένα μεταξύ τους και οποιαδήποτε βελτίωση αφορά την αντιστοίχιση εικόνων μπορεί να έχει θετικές επιδράσεις και στην ψηφιακή υδατογράφηση.

Η έρευνα που έγινε στα πλαίσια της διδακτορικής διατριβής σε σχέση με το πρόβλημα της αντιστοίχισης εικόνων αφορά τις τεχνικές που βασίζονται στη συσχέτιση των εικόνων. Ένα από τα μειονεκτήματα της χωρικής συσχέτισης είναι ότι η περιοχή γύρω από τη μέγιστη τιμή της μπορεί να έχει μεγάλο εύρος και να επηρεάζει την ακρίβεια εκτίμησης του σωστού μετασχηματισμού που πρέπει να εφαρμοστεί σε μια εικόνα προκειμένου να αντιστοιχηθεί με μια άλλη. Για την αντιμετώπιση αυτού του προβλήματος, προτάθηκε στο κεφάλαιο 2 μια διαδικασία προ-λεύκανσης των εικόνων, βασισμένη στο φίλτρο σφάλματος πρόβλεψης, καθώς και ένας επαναληπτικός αλγόριθμος αντιστοίχισης, ο οποίος υπολογίζει μετατοπίσεις και περιστροφές [12]. Ο αλγόριθμος αυτός εφαρμόζεται στη συνέχεια σε ακολουθίες ιατρικών εικόνων από

τον τράχηλο της μήτρας με σκοπό τη διάγνωση δυσπλασιών και κακοηθειών [64].

Η διάγνωση βασίζεται στο γεγονός ότι το χρώμα του προβληματικού ιστού αλλάζει μετά την τοπική εφαρμογή διαλύματος οξικού οξέος στην εξεταζόμενη περιοχή. Αυτό εισάγει άλλη μια παράμετρο στο πρόβλημα της αντιστοίχισης, εκτός από αυτές που αφορούν τις γεωμετρικές παραμορφώσεις. Στα πειράματά μας θεωρήσαμε ότι η αλλαγή αυτή είναι γραμμική και άρα μπορούμε να αντιμετωπίσουμε το παραπάνω πρόβλημα χρησιμοποιώντας σαν μετρική αντιστοίχισης την κανονικοποιημένη συσχέτιση ανάμεσα σε ακολουθίες οι οποίες έχουν υποστεί προ-λεύκανση. Μελλοντικά, η παραπάνω υπόθεση να μπορούσε να γίνει για μικρότερα παράθυρα των εικόνων, δηλαδή θα ήταν ενδιαφέρον να εξεταστεί το πρόβλημα τοπικά. Επίσης, θα μπορούσε να γίνει έλεγχος για περισσότερα είδη γεωμετρικών παραμορφώσεων, αν και εξετάζοντας το πρόβλημα τοπικά, η μετατόπιση και η περιστροφή συνήθως αρκούν για να αποκαταστήσουν τις γεωμετρικές παραμορφώσεις.

Ένα δεύτερο μειονέκτημα των τεχνικών που βασίζονται στη χωρική συσχέτιση είναι το μεγάλο υπολογιστικό τους κόστος. Έτσι, στο κεφάλαιο 3, προτείνουμε ένα γρήγορο σχήμα υπολογισμού της συσχέτισης, το οποίο βασίζεται σε κατάλληλη τμηματοποίηση της εικόνας και στη χρήση του μετασχηματισμού Fourier [12, 64]. Επιπλέον, το πιο απαιτητικό κομμάτι στη διαδικασία της αντιστοίχισης είναι ο υπολογισμός της χρησιμοποιούμενης μετρικής σαν συνάρτηση της σχετικής θέσης των εικόνων. Στο ίδιο κεφάλαιο, για την περίπτωση που υπάρχουν μόνο μετατοπίσεις ανάμεσα στις συγκρινόμενες εικόνες, προτείνουμε ένα αποδοτικό επαναληπτικό σχήμα για τον υπολογισμό του συντελεστή συσχέτισης στο χωρικό πεδίο, το οποίο μειώνει σημαντικά τον αριθμό των ψαξμάτων που απαιτούνται για την εύρεση του μεγίστου και παρέχει ακρίβεια εικονοστοιχείου [68, 67].

Και οι δύο αλγόριθμοι αφορούν την περίπτωση της μετατόπισης (*translation*). Μια πιθανή κατεύθυνση για μελλοντική έρευνα θα μπορούσε να είναι η προσπάθεια εύρεσης αποδοτικών αλγορίθμων, για την περίπτωση διαφορετικών γεωμετρικών παραμορφώσεων, όπως για παράδειγμα η περιστροφή και η κλιμάκωση.

Πολλές εφαρμογές απαιτούν αντιστοίχιση με ακρίβεια υποδιαίρεσης εικονοστοιχείου. Οι πιο συχνά χρησιμοποιούμενες τεχνικές που παρέχουν τέτοια ακρίβεια βασίζονται σε παρεμβολή. Στο κεφάλαιο 4, προτείνουμε μια νέα τεχνική αντιστοίχισης με ακρίβεια υποδιαίρεσης εικονοστοιχείου, η οποία βασίζεται στη μεγιστοποίηση του συντελεστή συσχέτισης. Η τεχνική αυτή δεν απαιτεί την ανακατασκευή των τιμών της έντασης και παρέχει μια λύση κλειστού τύπου για την εκτίμηση της μετατόπισης [67, 68]. Επιπλέον, έχει καλύτερη απόδοση σε σχέση με τις υπάρχουσες τεχνικές.

Κάποιες κατευθύνσεις για μελλοντική έρευνα σχετικά με την προτεινόμενη τεχνική θα

μπορούσαν να είναι η μελέτη της απόδοσή της για την περίπτωση που ανάμεσα στις συγκρινόμενες εικόνες υπάρχουν μετατοπίσεις οι οποίες διαφοροποιούνται τοπικά, καθώς και η διερεύνηση πιθανής επέκτασής της στην περίπτωση άλλων γεωμετρικών παραμορφώσεων.

Όσο αφορά τώρα την έρευνα που έγινε σχετικά με το πρόβλημα της υδατογράφησης εικόνων, είχε σαν στόχο την ένθεση ισχυρών υδατογραφήματων στο χωρικό πεδίο, χωρίς να επηρεάζεται η ποιότητα της εικόνας. Επιπλέον, προτάθηκαν κατάλληλα σχήματα “τυφλής” ανίχνευσης με εξαιρετικά καλή απόδοση.

Καταρχήν, στο κεφάλαιο 5, προτείναμε τη χρήση μιας χωρική μάσκας για την ένθεση υδατογραφήματων σε εικόνες, η οποία κατασκευάζεται με βάση την τοπική διασπορά του σφάλματος πρόβλεψης της εικόνας. Επιπλέον, αναπτύξαμε ένα βελτιωμένο σχήμα ανίχνευσης. Το σχήμα αυτό είναι “τυφλό”, δεν απαιτεί δηλαδή τη γνώση της αρχικής εικόνας, και χρησιμοποιεί σαν μετρική ανίχνευσης την κανονικοποιημένη συσχέτιση ανάμεσα στο υδατογράφημα και το σφάλμα πρόβλεψης της εικόνας που φτάνει στο δέκτη. Η βελτίωση στην απόδοση του ανιχνευτή έχει αποδειχτεί θεωρητικά, αρχικά για την περίπτωση που υδατογράφημα και εικόνα, αλλά και διαφορετικά υδατογραφήματα μεταξύ τους, θεωρούνται εντελώς ασυσχέτιστα [62]. Στη συνέχεια, μελετήθηκε η περίπτωση που αυτή η ιδανική υπόθεση δεν ισχύει, όπως άλλωστε συμβαίνει στην πράξη, και αποδείχτηκε θεωρητικά ότι ο προτεινόμενος ανιχνευτής έχει βελτιωμένη απόδοση, ακόμα και όταν η υδατογραφήμενη εικόνα έχει υποστεί επίθεση με γραμμικό φίλτρο και θόρυβο. Επίσης, υπολογίστηκαν εκφράσεις που επιτρέπουν τον υπολογισμό κατάλληλων κατωφλίων ανίχνευσης, προσαρμόσιμων στις συνθήκες των επιθέσεων [63].

Μια πιθανή επέκταση της παραπάνω δουλειάς θα μπορούσε να αφορά τη θεωρητική μελέτη της απόδοσης του προτεινόμενου σχήματος υδατογράφησης στην περίπτωση μη γραμμικών επιθέσεων. Το πρόβλημα αυτό είναι πολύ δύσκολο καθώς δεν υπάρχει ένα γενικό μοντέλο μη γραμμικής επίθεσης και κάθε είδος τέτοιας επίθεσης πρέπει να μελετηθεί ξεχωριστά.

Στο κεφάλαιο 6, προτείναμε τη χρήση μιας νέας χωρικής μάσκας, η οποία επιτρέπει την ένθεση υδατογραφήματων με εξαιρετικά μεγάλη ενέργεια, τα οποία κατά συνέπεια είναι πολύ ανθεκτικά. Η μάσκα αυτή βασίζεται στο σφάλμα πρόβλεψης της αρχικής εικόνας (και όχι στη διασπορά του σφάλματος πρόβλεψης, όπως στην προηγούμενη μάσκα που προτείναμε). Η απόδοσή της συγκρίθηκε με πολύ γνωστές και ευρέως χρησιμοποιούμενες μάσκες, και προέκυψε σημαντικά καλύτερη. Επιπλέον, στη φάση της ανίχνευσης του υδατογραφήματος προτάθηκε η χρήση μιας κατάλληλης διαδικασίας προ-λεύκανσης τόσο της εικόνας που φτάνει στο δέκτη, όσο και του υδατογραφήματος μετά την εφαρμογή της μάσκας, η οποία εκτιμάται στο δέκτη από τη λαμβανόμενη εικόνα. Η ανίχνευση παρουσιάζει σημαντικά βελτιωμένη απόδοση, η οποία αποδεικνύεται θεωρητικά. Η θεωρητική ανάλυση είναι ανεξάρτητη της μάσκας και τα

αποτελέσματά της μπορούν να χρησιμοποιηθούν σε κάθε τεχνική υδατογράφησης βασισμένη σε χωρικές μάσκες. Είναι σημαντικό όμως να τονίσουμε ότι τα πειραματικά αποτελέσματα έδειξαν ότι ο συνδυασμός του προτεινόμενου ανιχνευτή με την προτεινόμενη μάσκα δίνει την καλύτερη απόδοση [65]. Τέλος, υπολογίστηκαν κάποιες επιπλέον εκφράσεις (μέσες τιμές και διασπορές των μετρικών συσχέτισης) για τις περιπτώσεις της θεωρητικής μας μελέτης, οι οποίες δίνουν τη δυνατότητα υπολογισμού κατωφλίων ανίχνευσης προσαρμοσμένων στις συνθήκες των επιθέσεων.

Μελλοντικά, η απόδοση του προτεινόμενου σχήματος υδατογράφησης θα μπορούσε να μελετηθεί θεωρητικά για μη γραμμικές επιθέσεις. Επιπλέον, το προτεινόμενο σχήμα, το οποίο ανήκει στην κατηγορία των τεχνικών “πληροφορημένης” ένθεσης, θα μπορούσε ίσως να συνδυαστεί με μια κατάλληλη τεχνική “πληροφορημένης” κωδικοποίησης, καταλήγοντας σε ένα ακόμη πιο ισχυρό σχήμα.

Η χρήση μιας αντιληπτικής μάσκας σε συνδυασμό με την επαναληπτική ένθεση του υδατογραφήματος σαν μια διαδικασία κωδικοποίησης αυξάνουν σημαντικά την ανθεκτικότητα του υδατογραφήματος. Ωστόσο, υπάρχουν κάποια μειονεκτήματα από πλευράς αποδοτικότητας της κωδικοποίησης και πρέπει να βρεθεί ο κατάλληλος συμβιβασμός ανάμεσα σε αυτή και στις επαναλήψεις του υδατογραφήματος. Πιθανές λύσεις σε αυτό το πρόβλημα είναι η χρήση τεχνικών *διαφορικότητας* (*diversity*) ή σύνθετων σχημάτων κωδικοποίησης. Εμείς προτείνουμε ένα σχήμα υδατογράφησης στο χωρικό πεδίο, βασισμένο σε χωρο-χρονική κωδικοποίηση μπλοκ [66], το οποίο παρουσιάστηκε στο κεφάλαιο 7. Ειδικότερα, χρησιμοποιούμε μια 4×4 πραγματική, ορθογώνια διάταξη, για την ένθεση ενός πολλαπλασιαστικού υδατογραφήματος στην εικόνα. Η εικόνα χωρίζεται σε τέσσερα τμήματα τα οποία με κατάλληλες πράξεις μπορούν να θεωρηθούν σαν τέσσερα διαφορετικά κανάλια με *επίπεδη εξασθένιση* (*flat fading*). Ένα τέτοιο σχήμα κωδικοποίησης παρέχει πλήρη διαφορικότητα, το μεγαλύτερο δυνατό ρυθμό μετάδοσης και χαμηλή πολυποκότητα. Αποδείχθηκε ότι το προτεινόμενο σχήμα έχει πολύ καλύτερη απόδοση σε σχέση με την επαναληπτική υδατογράφηση.

Ακολουθώντας αντίστοιχη διαδικασία, το σχήμα αυτό θα μπορούσε να γενικευτεί για M πομπούς. Επιπλέον, μπορεί να συνδυαστεί κατάλληλα με ένα σχήμα επαναληπτικής ένθεσης, προκειμένου να είναι ανθεκτικό σε περισσότερες επιθέσεις, όπως για παράδειγμα η *αποκοπή* (*cropping*). Τέλος, μπορεί να χρησιμοποιηθεί σε συνδυασμό με μια αντιληπτική μάσκα που αυξάνει σημαντικά την ενέργεια του υδατογραφήματος, παρέχοντας ακόμα πιο βελτιωμένη απόδοση.

Παράρτημα Α

Απόδειξη Θεωρήματος 1 (Κεφάλαιο 4)

Υπολογίζουμε τις μερικές παραγώγους του συντελεστή συσχέτισης που ορίζεται στην εξίσωση (4.2.10) σε σχέση με τις παραμέτρους μετατόπισης t_x και t_y , και τις εξισώνουμε με το μηδέν. Έτσι, παίρνουμε τις ακόλουθες εξισώσεις

$$2a_1(b_0 + b_1t_x + b_2t_y + b_3t_xt_y + b_4t_x^2 + b_5t_y^2) = (a_0 + a_1t_x + a_2t_y)(b_1 + b_3t_y + 2b_4t_x) \quad (\text{A-1})$$

και

$$2a_2(b_0 + b_1t_x + b_2t_y + b_3t_xt_y + b_4t_x^2 + b_5t_y^2) = (a_0 + a_1t_x + a_2t_y)(b_2 + b_3t_x + 2b_5t_y) \quad (\text{A-2})$$

οι οποίες είναι μη γραμμικές σε σχέση με τις παραμέτρους μετατόπισης t_x και t_y . Ωστόσο, διαιρώντας την (A-1) με την (A-2), λαμβάνουμε την ακόλουθη γραμμική σχέση ανάμεσα στο t_x και στο t_y

$$t_x = \frac{a_2b_3 - 2a_1b_5}{a_1b_3 - 2a_2b_4}t_y + \frac{a_2b_1 - a_1b_2}{a_1b_3 - 2a_2b_4}. \quad (\text{A-3})$$

Αντικαθιστώντας την εξίσωση (A-3) στην (A-1), προκύπτει η ακόλουθη σχέση

$$\begin{aligned} & [(2a_1b_0 - a_0b_1)(a_1b_3 - 2a_2b_4) + (a_1b_1 - 2a_0b_4)(a_2b_1 - a_1b_2)] + \\ & [(a_1b_1 - 2a_0b_4)(a_2b_3 - 2a_1b_5) + (a_1b_3 - 2a_2b_4)(a_1b_2 - a_0b_3)]t_y + \\ & [(a_1b_3 - 2a_2b_4)(a_2b_3 - 2a_1b_5) + (2a_1b_5 - a_2b_3)(a_1b_3 - 2a_2b_4)]t_y^2 = 0 \end{aligned} \quad (\text{A-4})$$

η οποία μαζί με την (A-3) αποτελούν ένα σύστημα εξισώσεων, ισοδύναμο με αυτό που αποτελείται από τις σχέσεις (A-1) και (A-2). Μπορούμε να δούμε στη (A-4) ότι ο συντελεστής του t_y^2 είναι ίσος με το μηδέν. Έτσι, μπορούν εύκολα να προκύψουν οι επιθυμητές σχέσεις (4.2.12) και (4.2.13).

Είναι γνωστό, [115], ότι μια ικανή συνθήκη για να είναι το σημείο (t'_x, t'_y) μέγιστο είναι ο πίνακας των δεύτερων μερικών παραγώγων (ο πίνακας Hessian), $C(x + t_x, y + t_y)$, υπολογισμένος στο σημείο (t'_x, t'_y) , να είναι αρνητικά ορισμένος. Υπολογίζοντας τις δεύτερες μερικές παραγώγους στο σημείο (t'_x, t'_y) , το οποίο ορίζεται στις σχέσεις (4.2.12) και (4.2.13), προκύπτουν οι ακόλουθες σχέσεις

$$\left. \frac{\partial^2 C(x + t_x, x + t_y)}{\partial t_x^2} \right|_{(t_x, t_y) = (t'_x, t'_y)} = (a_1 b_1 - 2a_0 b_4) + (a_1 b_3 - 2a_2 b_4) t'_y \quad (\text{A-5})$$

$$\left. \frac{\partial^2 C(x + t_x, x + t_y)}{\partial t_y^2} \right|_{(t_x, t_y) = (t'_x, t'_y)} = (a_2 b_2 - 2a_0 b_5) - (2a_1 b_5 - a_2 b_3) t'_x \quad (\text{A-6})$$

$$\begin{aligned} \left. \frac{\partial^2 C(x + t_x, x + t_y)}{\partial t_x \partial t_y} \right|_{(t_x, t_y) = (t'_x, t'_y)} &= (2a_1 b_2 - a_0 b_3 - a_2 b_1) + (a_1 b_3 - 2a_2 b_4) t'_x \\ &\quad + 2(2a_1 b_5 - a_2 b_3) t'_y \end{aligned} \quad (\text{A-7})$$

οι οποίες μπορούν να χρησιμοποιηθούν, προκειμένου να ελεγχθεί αν ο Hessian πίνακας είναι αρνητικά ορισμένος.

Παράρτημα Β

Απόδειξη της σχέσης (6.3.5)

Ο σκοπός είναι να βρούμε κάτω από ποιες συνθήκες είναι αληθής η ανισότητα $|C_{z,u_0}| \leq |C_{e_z,e_{u_0}}|$. Πρώτα, ορίζονται οι ποσότητες που περιλαμβάνονται σε αυτή την ανισότητα. Σημειώνουμε ότι το $\hat{E}[\cdot]$ που χρησιμοποιείται παρακάτω δηλώνει μέσο όρο δειγμάτων και δεν πρέπει να συγχέεται με τον τελεστή της αναμενόμενης τιμής, $E[\cdot]$. Έχουμε, λοιπόν,

$$C_{z,u_0} = \frac{\hat{E}[zu_0]}{\sqrt{\hat{E}[z^2]}\sqrt{\hat{E}[u_0^2]}} \quad , \quad C_{e_z,e_{u_0}} = \frac{\hat{E}[e_z e_{u_0}]}{\sqrt{\hat{E}[e_z^2]}\sqrt{\hat{E}[e_{u_0}^2]}} \quad \text{και}$$

$$\begin{aligned} \hat{E}[zu_0] &= \mu_M \mathbf{h}^T \mathbf{r}_{xw_0}(0) + \mathbf{h}^T \text{diag}[\mathbf{r}_M(0)] \mathbf{r}_{w_0}(0) + \mu_M r_{nw_0}(0), & \hat{E}[z^2] &= P_z \\ \hat{E}[u_0^2] &= P_M \sigma_{w_0}^2, & \hat{E}[e_z e_{u_0}] &= \hat{E}[zu_0] \|\mathbf{a}_z\|^2 + S_0, & \hat{E}[e_z^2] &= \sigma_{e_z}^2, & \hat{E}[e_{u_0}^2] &= \sigma_{e_{u_0}}^2. \end{aligned}$$

Το $\mathbf{r}_{xw_0}(i) = [\cdots \ r_{xw_0}(i+1) \ r_{xw_0}(i) \ r_{xw_0}(i-1) \ \cdots]^T$, ενώ τα $\mathbf{r}_M(i)$ και $\mathbf{r}_{w_0}(i)$ ορίζονται με παρόμοιο τρόπο. Το S_0 υπολογίζεται από τη σχέση $S_0 = \mu_M \mathbf{a}_z^T \hat{H}_{xw_0} \mathbf{a}_z + \mathbf{a}_z^T \hat{H}_{M \odot w_0} \mathbf{a}_z + \mu_M \mathbf{a}_z^T \hat{R}_{nw_0} \mathbf{a}_z$ και για τις υπόλοιπες ποσότητες έχουμε:

$$\hat{H}_{xw_0} = \begin{bmatrix} 0 & \mathbf{h}^T \mathbf{r}_{xw_0}(-1) & \cdots & \mathbf{h}^T \mathbf{r}_{xw_0}(-p^2+1) \\ \mathbf{h}^T \mathbf{r}_{xw_0}(1) & 0 & \cdots & \mathbf{h}^T \mathbf{r}_{xw_0}(-p^2+2) \\ \vdots & \ddots & \ddots & \vdots \\ \mathbf{h}^T \mathbf{r}_{xw_0}(p^2-1) & \cdots & \mathbf{h}^T \mathbf{r}_{xw_0}(1) & 0 \end{bmatrix}$$

όπου οι πίνακες $\hat{H}_{M \odot w_0}$ και $\hat{R}_{nw_0}$ ορίζονται αντίστοιχα με το $\hat{H}_{xw_0}$. Τα μη μηδενικά στοιχεία τους είναι $\mathbf{h}^T \mathbf{r}_{M \odot w_0}(i)$ και $r_{nw_0}(i)$ αντίστοιχα, όπου

$$\mathbf{r}_{M \odot w_0}(i) = [\cdots \ r_M(i)r_{w_0}(i+1) \ r_M(i)r_{w_0}(i) \ r_M(i-1)r_{w_0}(i-1) \ \cdots]^T = \text{diag}[\mathbf{r}_M(i)] \mathbf{r}_{w_0}(i).$$

Για τις εμπλεκόμενες αυτο- και ετερο-συσχετίσεις, οι οποίες ιδανικά θεωρούνται ότι είναι ίσες με το μηδέν, υποθέτουμε ότι έχουν μέτρο άνω φραγμένο από ένα μικρό θετικό βαθμωτό ε :

$$|r_{xw_0}(i)|, |r_{nw_0}(i)| \leq \varepsilon, \quad \forall i, \quad \varepsilon > 0 \quad \text{και} \quad |r_M(i)|, |r_{w_0}(i)| \leq \varepsilon, \quad i \neq 0, \quad \varepsilon > 0.$$

Στη συνέχεια, αντικαθιστώντας κάποιες ποσότητες στην ανισότητα $|C_{z,u_0}| \leq |C_{e_z,e_{u_0}}|$ έχουμε ότι

$$|\hat{E}[zu_0]| \sqrt{\frac{\sigma_{e_z}^2 \sigma_{e_{u_0}}^2}{P_z P_M \sigma_{w_0}^2}} \leq |\hat{E}[zu_0]| \|\mathbf{a}_z^2\| + |S_0|. \quad (\text{B-1})$$

$$\text{Αν } |\hat{E}[zu_0]| \cdot \|\mathbf{a}_z^2\| \geq |S_0|, \quad \text{τότε } |\hat{E}[zu_0]| \cdot \|\mathbf{a}_z^2\| - |S_0| \leq |\hat{E}[zu_0]| \|\mathbf{a}_z^2\| + |S_0| \quad (\text{B-2})$$

και συνεπώς η σχέση (B-1) μπορεί να γραφεί ως εξής

$$|\hat{E}[zu_0]| \left(\|\mathbf{a}_z^2\| - \sqrt{\frac{\sigma_{e_z}^2 \sigma_{e_{u_0}}^2}{P_z P_M \sigma_{w_0}^2}} \right) \geq |S_0|. \quad (\text{B-3})$$

Έτσι, προκειμένου να βρούμε κάτω από ποιους περιορισμούς είναι αληθής η σχέση (B-3), πρέπει να καθορίσουμε πότε ισχύει η συνθήκη $|\hat{E}[zu_0]| \cdot \|\mathbf{a}_z^2\| \geq |S_0|$ της σχέσης (B-2). Για το λόγο αυτό, θα βρούμε ένα κάτω όριο για το $|\hat{E}[zu_0]|$ και ένα άνω όριο για το $|S_0|$.

Κάτω όριο για το $|\hat{E}[zu_0]|$

$$\text{Αν } |S_1| \geq |S_2|, \quad \text{τότε } |\hat{E}[zu_0]| \geq |S_1| - |S_2| \quad (\text{B-4})$$

όπου

$$\begin{aligned} S_1 &= \mu_M h_0 r_{xw_0}(0) + h_0 P_M \sigma_{w_0}^2 + \mu_M r_{nw_0}(0), \\ S_2 &= \mu_M \tilde{\mathbf{h}}^T \tilde{\mathbf{r}}_{xw_0}(0) + \tilde{\mathbf{h}}^T \text{diag}[\tilde{\mathbf{r}}_M(0)] \tilde{\mathbf{r}}_{w_0}(0) + \mu_M \tilde{\mathbf{r}}_{nw_0}(0). \end{aligned}$$

Με το σύμβολο “ $\sim$ ” σημειώνουμε τα διανύσματα τα οποία δεν περιέχουν το κεντρικό στοιχείο των αντίστοιχων διανυσμάτων χωρίς το “ $\sim$ ”. Θα βρούμε στη συνέχεια ένα κάτω όριο για το $|S_1|$ και ένα άνω όριο για το $|S_2|$.

- Κάτω όριο για το $|S_1|$

$$\begin{aligned} \text{Αν } |h_0| P_M \sigma_{w_0}^2 &\geq \mu_M |h_0 r_{xw_0}(0) + r_{nw_0}(0)|, \\ \text{τότε } |S_1| &\geq |h_0| P_M \sigma_{w_0}^2 - \mu_M |h_0 r_{xw_0}(0) + r_{nw_0}(0)|. \end{aligned} \quad (\text{B-5})$$

Η συνθήκη της σχέσης (B-5) μπορεί να γραφεί σαν $|h_0| P_M \sigma_{w_0}^2 \geq \mu_M \varepsilon (|h_0| + 1)$ και είναι πάντα αληθής, καθώς το ε είναι ένας πολύ μικρός βαθμωτός κοντά στο μηδέν (στην ιδανική

περίπτωση θεωρείται μηδέν). Έτσι,

$$|S_1| \geq |h_0|P_M\sigma_{w_0}^2 - \mu_M\varepsilon(|h_0| + 1). \quad (\text{B-6})$$

- Άνω όριο για το $|S_2|$

$$\begin{aligned} |S_2| &\leq \|\tilde{\mathbf{h}}\| (\mu_M \|\tilde{\mathbf{r}}_{xw_0}(0)\| + \|\tilde{\mathbf{r}}_M(0) \odot \tilde{\mathbf{r}}_{w_0}(0)\|) + \mu_M \|\tilde{\mathbf{r}}_{nw_0}(0)\| \\ &\leq \varepsilon\sqrt{P-1} \left[(\mu_M + \varepsilon) \|\tilde{\mathbf{h}}\| + \mu_M \right] \end{aligned} \quad (\text{B-7})$$

Έτσι, η συνθήκη της σχέσης (B-4) μπορεί να γραφεί σαν

$$|h_0|P_M\sigma_{w_0}^2 - \mu_M\varepsilon(|h_0| + 1) \geq \varepsilon\sqrt{P-1} \left[(\mu_M + \varepsilon) \|\tilde{\mathbf{h}}\| + \mu_M \right]$$

η οποία είναι πάντα αληθής για μικρό ε . Με βάση τις σχέσεις (B-6) και (B-7), το κάτω όριο για το $|\hat{E}[zu_0]|$ δίνεται από τη σχέση

$$|\hat{E}[zu_0]| \geq |h_0|P_M\sigma_{w_0}^2 - S_3 \quad (\text{B-8})$$

$$\text{όπου } S_3 = \mu_M\varepsilon(|h_0| + 1) + \varepsilon\sqrt{P-1} \left[(\mu_M + \varepsilon) \|\tilde{\mathbf{h}}\| + \mu_M \right].$$

Άνω όριο για το $|S_0|$

$$|S_0| \leq \|\mathbf{a}_z\|^2 \left(\mu_M \|\hat{H}_{xw_0}\| + \|\hat{H}_{M \odot w_0}\| + \mu_M \|\hat{R}_{nw_0}\| \right) \leq \|\mathbf{a}_z\|^2 \left[\|\hat{H}\|P_M\sigma_{w_0}^2 + S_4 \right] \quad (\text{B-9})$$

όπου

$$\begin{aligned} S_4 &= S_5 + \mu_M\varepsilon\sqrt{P(P-1)} \left[\|\mathbf{h}\|\sqrt{P} + 1 \right], \\ S_5 &= \frac{1}{2}\sqrt{P-1} \left[\sqrt{S_6(3P-1)} + \|\mathbf{h}\|\varepsilon^2\sqrt{P(P+1)} \right], \\ S_6 &= 2\varepsilon^2|h_i|P_M\sigma_{w_0}^2 \left(\sum_{j \neq i} |h_j| \right) + \varepsilon^4 \left(\sum_{j \neq i} |h_j| \right)^2. \end{aligned} \quad (\text{B-10})$$

Με βάση τις σχέσεις (B-8) και (B-9), η συνθήκη της (B-2) μπορεί να γραφεί σαν

$$|h_0|P_M\sigma_{w_0}^2 \geq \|\hat{H}\|P_M\sigma_{w_0}^2 + S_3 + S_4. \quad (\text{B-11})$$

Καθώς το S_3+S_4 είναι ένας μικρός βαθμωτός (λόγω του ε), η παραπάνω ανισότητα (B-11) είναι αληθής αν $|h_0| \geq \|\hat{H}\|$. Θεωρώντας ότι αυτός ο περιορισμός ικανοποιείται, αντικαθιστούμε

στην ανισότητα (B-3) τα $|\hat{E}[zu_0]|$ και $|S_0|$ με τα κάτω και άνω όριά τους, αντίστοιχα. Το αποτέλεσμα είναι

$$|h_0|P_M\sigma_{w_0}^2 L \geq \|\hat{H}\|P_M\sigma_{w_0}^2 + \gamma \quad (\text{B-12})$$

όπου $L = 1 - \sqrt{\frac{\sigma_{e_z}^2 \sigma_{e_{u_0}}^2}{P_z P_M \sigma_{w_0}^2 \|\mathbf{a}_z\|^4}}$ και $\gamma = S_3 L + S_4$. Έτσι, συμπεραίνουμε ότι, προκειμένου να ισχύει ότι $|C_{e_z, e_{u_0}}| \geq |C_{z, u_0}|$, πρέπει να ικανοποιείται η σχέση (B-12).

Παράρτημα Γ

Παραγωγή μέσων τιμών και διασπορών των συσχετίσεων (Κεφάλαιο 6)

Οι μέσες τιμές και οι διασπορές θα υπολογιστούν για τους αριθμητές των μετρικών κανονικοποιημένης συσχέτισης. Οι παρανομαστές χρησιμοποιούνται σαν σταθερές κανονικοποίησης και άρα τα k_0 , k_1 , l_0 και l_1 που ορίζονται στην ενότητα 6.3.2 δεν περιλαμβάνονται στους υπολογισμούς μας. Οι συσχετίσεις που μας ενδιαφέρουν είναι οι:

$$C_{z,u_1} = \frac{1}{N} \sum_{i=1}^N \mathbf{h}^T \mathbf{x}_i u_{1_i} + \frac{1}{N} \sum_{i=1}^N \mathbf{h}^T \mathbf{u}_{0_i} u_{1_i} + \frac{1}{N} \sum_{i=1}^N n_i u_{1_i} \quad (\Gamma-1)$$

$$C_{e_z, e_{u_1}} = \frac{1}{N} \sum_{i=1}^N \mathbf{a}_z^T \begin{bmatrix} \dots \\ \mathbf{h}^T \mathbf{x}_{n+1} \\ \mathbf{h}^T \mathbf{x}_n \\ \mathbf{h}^T \mathbf{x}_{n-1} \\ \dots \end{bmatrix}_i \mathbf{u}_{1_i}^T \mathbf{a}_z + \frac{1}{N} \sum_{i=1}^N \mathbf{a}_z^T \begin{bmatrix} \dots \\ \mathbf{h}^T \mathbf{u}_{0_{n+1}} \\ \mathbf{h}^T \mathbf{u}_{0_n} \\ \mathbf{h}^T \mathbf{u}_{0_{n-1}} \\ \dots \end{bmatrix}_i \mathbf{u}_{1_i}^T \mathbf{a}_z + \frac{1}{N} \sum_{i=1}^N \mathbf{a}_z^T \mathbf{n}_i \mathbf{u}_{1_i}^T \mathbf{a}_z \quad (\Gamma-2)$$

και C_{z,u_0} , $C_{e_z, e_{u_0}}$, οι οποίες υπολογίζονται αντίστοιχα με τα C_{z,u_1} , $C_{e_z, e_{u_1}}$, αντικαθιστώντας τα u_{1_i} και $\mathbf{u}_{1_i}$ από τα u_{0_i} και $\mathbf{u}_{0_i}$, αντίστοιχα.

Πρώτα, υπολογίζουμε τις μέσες τιμές για τις παραπάνω συσχετίσεις. Έχουμε:

$$\mu_{C_{z,u_1}} = \mu_M \mathbf{h}^T \mathbf{r}_{xw_1}(0) + h_0 P_{Mr_{w_0w_1}}(0) + \mu_M r_{nw_1}(0) = c_1 \quad (\Gamma-3)$$

$$\mu_{C_{ez}, e_{u_1}} = \mu_M \mathbf{a}_z^T H_{xw_1} \mathbf{a}_z + \mathbf{a}_z^T H \mathbf{a}_z P_M r_{w_0 w_1}(0) + \mu_M \mathbf{a}_z^T R_{nw_1} \mathbf{a}_z = c_2 \quad (\Gamma-4)$$

$$\mu_{C_{z, u_0}} = \mu_M \mathbf{h}^T \mathbf{r}_{xw_0}(0) + h_0 P_M \sigma_{w_0}^2 + \mu_M r_{nw_0}(0) = h_0 P_M \sigma_{w_0}^2 + c_3 \quad (\Gamma-5)$$

$$\begin{aligned} \mu_{C_{ez}, e_{u_0}} &= \mu_M \mathbf{a}_z^T H_{xw_0} \mathbf{a}_z + \mathbf{a}_z^T H \mathbf{a}_z P_M \sigma_{w_0}^2 + \mu_M \mathbf{a}_z^T R_{nw_0} \mathbf{a}_z \\ &= \mathbf{a}_z^T H \mathbf{a}_z P_M \sigma_{w_0}^2 + c_4. \end{aligned} \quad (\Gamma-6)$$

$H_{xw_0} = \mathbf{h}^T \mathbf{r}_{xw_0}(0)I + \hat{H}_{xw_0}$ και $R_{nw_0} = r_{nw_0}(0)I + \hat{R}_{nw_0}$, όπου τα $\hat{H}_{xw_0}$ και $\hat{R}_{nw_0}$ έχουν οριστεί στο Παράρτημα Β. Οι πίνακες H_{xw_1} και R_{nw_1} ορίζονται με παρόμοιο τρόπο (τα στοιχεία τους είναι $\mathbf{h}^T \mathbf{r}_{xw_1}(i)$ και $r_{nw_1}(i)$, αντίστοιχα). Οι ποσότητες c_1 , c_2 , c_3 και c_4 περιέχουν συσχετίσεις, οι οποίες στην ιδανική περίπτωση θεωρούνται ίσες με το μηδέν. Εδώ, υποθέτουμε ότι αυτές οι συσχετίσεις είναι τάξης ε , όπου το ε είναι ένας πολύ μικρός βαθμωτός. Συνεπώς, τα c_1 , c_2 , c_3 και c_4 είναι μικροί βαθμωτοί. Να σημειώσουμε ότι οι όροι τάξης ίσης ή μεγαλύτερης από ε^2 αγνοούνται στον υπολογισμό των μέσων τιμών και των διασπορών.

Στη συνέχεια, ακολουθεί ο υπολογισμός των διασπορών.

$$\begin{aligned} \sigma_{C_{z, u_1}}^2 &= E[C_{z, u_1}^2] = \frac{1}{N^2} \sum_{i=1}^N E[(\mathbf{h}^T \mathbf{x}_i u_{1i})^2] + \frac{1}{N^2} \sum_{i=1}^N E[(\mathbf{h}^T \mathbf{u}_{0i} u_{1i})^2] \\ &\quad + \frac{1}{N^2} \sum_{i=1}^N E[n_i^2 u_{1i}^2] + \frac{2}{N^2} \sum_{i=1}^N E[\mathbf{h}^T \mathbf{x}_i u_{1i} n_i u_{1i}] \\ &= \dots = \frac{1}{N} \sigma_{w_1}^2 \left[P_M (\mathbf{h}^T R_x \mathbf{h} + \sigma_n^2) + h_0^2 \sigma_{w_0}^2 r_{M^2}(0) + c_5 \right] \end{aligned} \quad (\Gamma-7)$$

όπου το $c_5 = 2P_M \mathbf{h}^T \mathbf{r}_{xn}(0) + \sigma_{w_0}^2 \sum_{i \neq 0} h^2(i) r_{M^2}(i)$ είναι τάξης ε εξαιτίας του $\mathbf{r}_{xn}(0)$ και του $r_{M^2}(i)$.

$$\begin{aligned} \sigma_{C_{z, u_0}}^2 &= E[C_{z, u_0}^2] - E^2[C_{z, u_0}] = \frac{1}{N^2} \sum_{i=1}^N E[(\mathbf{h}^T \mathbf{x}_i u_{0i})^2] + \frac{1}{N^2} \sum_{i=1}^N E[(\mathbf{h}^T \mathbf{u}_{0i} u_{0i})^2] \\ &\quad + \frac{1}{N^2} \sum_{i \neq j} E[\mathbf{h}^T \mathbf{u}_{0i} u_{0i} \mathbf{h}^T \mathbf{u}_{0j} u_{0j}] + \frac{1}{N^2} \sum_{i=1}^N E[n_i^2 u_{0i}^2] \\ &\quad + \frac{2}{N^2} \sum_{i=1}^N E[\mathbf{h}^T \mathbf{x}_i u_{1i} n_i u_{1i}] - h_0^2 P_M^2 \sigma_{w_0}^4 - 2c_3 h_0 P_M \sigma_{w_0}^2 \\ &= \dots = \frac{1}{N} \sigma_{w_0}^2 \left[P_M (\mathbf{h}^T R_x \mathbf{h} + \sigma_n^2 + c_6) + h_0^2 \sigma_{w_0}^2 (3r_{M^2}(0) - N P_M^2) + c_7 \right] \end{aligned} \quad (\Gamma-8)$$

όπου επίσης τα $c_6 = 2\mathbf{h}^T \mathbf{r}_{xn}(0) - 2Nc_3h_0$ και $c_7 = \sigma_{w_0}^2 \left[\sum_{i \neq 0} h_i^2 r_{M^2}(i) + \frac{1}{N} h_0^2 \sum_{i \neq 0} r_{M^2}(i) \right]$ είναι τάξης ε .

$$\begin{aligned} \sigma_{C_{ez, eu_1}}^2 &= E \left[C_{ez, eu_1}^2 \right] = \frac{1}{N^2} \sum_{i=1}^N E \left[\left(\mathbf{a}_z^T \begin{bmatrix} \dots \\ \mathbf{h}^T \mathbf{x}_{n+1} \\ \mathbf{h}^T \mathbf{x}_n \\ \mathbf{h}^T \mathbf{x}_{n-1} \\ \dots \end{bmatrix} \mathbf{u}_{1_i}^T \mathbf{a}_z \right)^2 \right] \\ &+ \frac{1}{N^2} \sum_{i=1}^N E \left[\left(\mathbf{a}_z^T \begin{bmatrix} \dots \\ \mathbf{h}^T \mathbf{u}_{0_{n+1}} \\ \mathbf{h}^T \mathbf{u}_{0_n} \\ \mathbf{h}^T \mathbf{u}_{0_{n-1}} \\ \dots \end{bmatrix} \mathbf{u}_{1_i}^T \mathbf{a}_z \right)^2 \right] + \frac{1}{N^2} \sum_{i=1}^N E \left[(\mathbf{a}_z^T \mathbf{n}_i)^2 (\mathbf{a}_z^T \mathbf{u}_{1_i})^2 \right] \\ &+ \frac{2}{N^2} \sum_{i=1}^N E \left[\mathbf{a}_z^T \begin{bmatrix} \dots \\ \mathbf{h}^T \mathbf{x}_{n+1} \\ \mathbf{h}^T \mathbf{x}_n \\ \mathbf{h}^T \mathbf{x}_{n-1} \\ \dots \end{bmatrix} \mathbf{u}_{1_i}^T \mathbf{a}_z \mathbf{a}_z^T \mathbf{n}_i \mathbf{u}_{1_i}^T \mathbf{a}_z \right] \quad (\Gamma-9) \end{aligned}$$

Ο όρος μέσα στο πρώτο άθροισμα της σχέσης (Γ-9) μπορεί να γραφεί σαν

$$\begin{aligned} \mathbf{a}_z^T \begin{bmatrix} \mathbf{h}^T R_x(0) \mathbf{h} & \mathbf{h}^T R_x(-1) \mathbf{h} & \dots & \mathbf{h}^T R_x(-p^2 + 1) \mathbf{h} \\ \mathbf{h}^T R_x(1) \mathbf{h} & \mathbf{h}^T R_x(0) \mathbf{h} & \dots & \mathbf{h}^T R_x(-p^2 + 2) \mathbf{h} \\ \vdots & & \ddots & \vdots \\ \mathbf{h}^T R_x(p^2 - 1) \mathbf{h} & \dots & \mathbf{h}^T R_x(-1) \mathbf{h} & \mathbf{h}^T R_x(0) \mathbf{h} \end{bmatrix}_i \mathbf{a}_z \\ \cdot \mathbf{a}_z^T E [\mathbf{u}_{1_i} \mathbf{u}_{1_i}^T] \mathbf{a}_z. \quad (\Gamma-10) \end{aligned}$$

Μετά από μερικές πράξεις ο πρώτος παράγοντας του γινομένου της σχέσης (Γ-10) μπορεί να γραφεί σαν $\mathbf{g}^T R'_x \mathbf{g}$, όπου το $\mathbf{g}$ είναι η συνέλιξη ανάμεσα στο φίλτρο σφάλματος πρόβλεψης $\mathbf{a}_z$ και στο φίλτρο $\mathbf{h}$ που χρησιμοποιείται για την επίθεση. Το διάνυσμα $\mathbf{g}$ έχει μήκος $(2p^2 - 1)$ και ο πίνακας συσχέτισης R'_x έχει διαστάσεις $(2p^2 - 1) \times (2p^2 - 1)$. Ο δεύτερος όρος του γινομένου μπορεί εύκολα να υπολογιστεί και είναι ίσος με $P_M \sigma_{w_1}^2 \|\mathbf{a}_z\|^2$.

Το περιεχόμενο του δεύτερου όρου άθροισης της σχέσης (Γ-9) μπορεί να γραφεί σαν $E [\mathbf{a}_z^T U_{h_{0_i}} A U_{1_i} \mathbf{a}_z]$, όπου τα $U_{h_{0_i}}$, A και U_{1_i} έχουν οριστεί στην ενότητα 6.3.2. Έτσι, έχουμε ότι

$$\begin{aligned} E [\mathbf{a}_z^T U_{h_{0_i}} A U_{1_i} \mathbf{a}_z] &= E \left[\text{tr} \left(U_{h_{0_i}} A U_{1_i} \mathbf{a}_z \mathbf{a}_z^T \right) \right] = E \left[\text{tr} \left(U_{h_{0_i}} A U_{1_i} A \right) \right] \\ &= \text{vec}(A^T)^T E [U_{h_{0_i}} \otimes U_{1_i}] \text{vec}(A). \quad (\Gamma-11) \end{aligned}$$

Οι υπόλοιποι όροι άθροισης της σχέσης (Γ-9) μπορούν εύκολα να υπολογιστούν. Αντικαθιστώντας τις υπολογισμένες ποσότητες στη σχέση (Γ-9) έχουμε ότι

$$\sigma_{C_{ez,eu_1}}^2 = \frac{1}{N} \left[P_M \sigma_{w_1}^2 \|\mathbf{a}_z\|^2 (\mathbf{g}^T R'_x \mathbf{g} + \sigma_n^2 \|\mathbf{a}_z\|^2 + c_8) + \text{vec}(A^T)^T E \left[U_{h_{0_i}}^T \otimes U_{1_i} \right] \text{vec}(A) \right] \quad (\Gamma-12)$$

όπου το $c_8 = \mathbf{a}_z^T \hat{R}_n \mathbf{a}_z + 2\mathbf{a}_z^T H_{xn} \mathbf{a}_z$ είναι τάξης ε , ενώ τα $\hat{R}_n$ και H_{xn} ορίζονται αντίστοιχα με τα $\hat{R}_{nw_0}$ και H_1 , με τα στοιχεία τους να είναι $r_n(i)$ ($i \neq 0$) και $\mathbf{h}^T \mathbf{r}_{xn}(i)$, αντίστοιχα.

Τέλος, για τη $\sigma_{C_{ez,eu_0}}^2$ έχουμε

$$\begin{aligned} \sigma_{C_{ez,eu_0}}^2 &= E \left[C_{ez,eu_0}^2 \right] - E^2 \left[C_{ez,eu_0} \right] = \frac{1}{N^2} \sum_{i=1}^N E \left[\left(\mathbf{a}_z^T \begin{bmatrix} \dots \\ \mathbf{h}^T \mathbf{x}_{n+1} \\ \mathbf{h}^T \mathbf{x}_n \\ \mathbf{h}^T \mathbf{x}_{n-1} \\ \dots \end{bmatrix}_i \mathbf{u}_{0_i}^T \mathbf{a}_z \right)^2 \right] \\ &+ \frac{1}{N^2} \sum_{i=1}^N E \left[\left(\mathbf{a}_z^T \begin{bmatrix} \dots \\ \mathbf{h}^T \mathbf{u}_{0_{n+1}} \\ \mathbf{h}^T \mathbf{u}_{0_n} \\ \mathbf{h}^T \mathbf{u}_{0_{n-1}} \\ \dots \end{bmatrix}_i \mathbf{u}_{0_i}^T \mathbf{a}_z \right)^2 \right] + \frac{1}{N^2} \sum_{i \neq j} E \left[\mathbf{a}_z^T U'_{h_{0_i}} \mathbf{a}_z \mathbf{a}_z^T U'_{h_{0_j}} \mathbf{a}_z \right] \\ &+ \frac{1}{N^2} \sum_{i=1}^N E \left[(\mathbf{a}_z^T \mathbf{n}_i)^2 (\mathbf{a}_z^T \mathbf{u}_{0_i})^2 \right] + \frac{2}{N^2} \sum_{i=1}^N E \left[\mathbf{a}_z^T \begin{bmatrix} \dots \\ \mathbf{h}^T \mathbf{x}_{n+1} \\ \mathbf{h}^T \mathbf{x}_n \\ \mathbf{h}^T \mathbf{x}_{n-1} \\ \dots \end{bmatrix}_i \mathbf{u}_{0_i}^T \mathbf{a}_z \mathbf{a}_z^T \mathbf{n}_i \mathbf{u}_{0_i} \mathbf{a}_z \right] \\ &- (\mathbf{a}_z^T H \mathbf{a}_z)^2 P_M^2 \sigma_{w_0}^4 - 2c_4 \mathbf{a}_z^T H \mathbf{a}_z P_M \sigma_{w_0}^2 \end{aligned} \quad (\Gamma-13)$$

όπου το $U'_{h_{0_i}}$ έχει οριστεί στην ενότητα 6.3.2. Ακολουθώντας παρόμοια βήματα με την περίπτωση της $\sigma_{C_{ez,eu_1}}^2$, έχουμε

$$\begin{aligned} \sigma_{C_{ez,eu_0}}^2 &= \frac{1}{N} P_M \sigma_{w_0}^2 \|\mathbf{a}_z\|^2 \left[\mathbf{g}^T R'_x \mathbf{g} + \sigma_n^2 \|\mathbf{a}_z\|^2 - N \frac{(\mathbf{a}_z^T H \mathbf{a}_z)^2}{\|\mathbf{a}_z\|^2} P_M \sigma_{w_0}^2 + c_9 \right] + \\ &\frac{1}{N} \left[\text{vec}(A^T)^T E \left[U_{h_{0_i}}^T \otimes U_{0_i} \right] \text{vec}(A) + \frac{1}{N} \sum_{i \neq j} \text{vec}(A^T)^T E \left[U_{h_{0_i}}'^T \otimes U_{h_{0_j}}' \right] \text{vec}(A) \right] \end{aligned} \quad (\Gamma-14)$$

όπου το $c_9 = \mathbf{a}_z^T \hat{R}_n \mathbf{a}_z - 2Nc_4 \frac{\mathbf{a}_z^T H \mathbf{a}_z}{\|\mathbf{a}_z\|^2}$ είναι τάξης ε .

Οι παραπάνω μέσες τιμές και διασπορές διαιρούνται στη συνέχεια με τις κατάλληλες σταθερές κανονικοποίησης και έτσι προκύπτουν οι τελικές μέσες τιμές και διασπορές των εξεταζόμενων κανονικοποιημένων συσχετίσεων που παρουσιάστηκαν στην ενότητα 6.3.2.

Βιβλιογραφία

- [1] S. M. Alamouti, *A simple transmit diversity technique for wireless communications*, IEEE Journal on Selected Areas in Communications **16** (1998), no. 8, 1451–1458.
- [2] R. J. Althof, M. G. J. Wind, and J. T. Dobbins, *A rapid and automatic image registration algorithm with subpixel accuracy*, IEEE Transactions on Medical Imaging **16** (1996), no. 3, 308–316.
- [3] R. J. Anderson and F. A. P. Petitcolas, *On the limits of steganography*, IEEE Journal on Selected Areas of Communications **16** (1998), 474–481.
- [4] P. E. Anuta, *Spatial registration of multispectral and multitemporal digital imagery using Fast Fourier Transform*, IEEE Transactions on Geoscience Electronics **8** (1970), 353–368.
- [5] R. Bajcsy and S. Kovacic, *Multiresolution elastic matching*, Computer Vision, Graphics and Image Processing **46** (1989), 1–21.
- [6] C. Balas, *A novel optical imaging method for the early detection, quantitative grading and mapping of cancerous and precancerous lesions of cervix*, IEEE Transactions on Biomedical Engineering **48** (2001), 96–104.
- [7] D. I. Barnea and H. F. Silverman, *A class of algorithms for fast digital image registration*, IEEE Transactions on Computers **C-21** (1972), 179–186.
- [8] M. Barni, F. Bartolini, V. Capellini, and A. Piva, *A DCT-domain system for robust image watermarking*, Signal Processing **66** (1998), 357–372.
- [9] M. Barni, F. Bartolini, V. Cappellini, A. Lippi, and A. Piva, *A DWT-based technique for spatio-frequency masking of digital signatures*, Proc. of SPIE Int. Conf. Security and Watermarking of Multimedia Contents, vol. 3657, Jan. 1999, San Jose, CA, pp. 31–39.

- [10] M. Barni, F. Bartolini, and A. Piva, *Improved wavelet-based watermarking through pixel-wise masking*, IEEE Transactions Image Processing **10** (2001), no. 5, 783–791.
- [11] M. Barni, F. Bartolini, A. De Rosa, and A. Piva, *A new decoder for the optimum recovery of nonadditive watermarks*, IEEE Transactions Image Processing **10** (2001), no. 5, 755–766.
- [12] K. Berberidis and I. Karybali, *A new efficient cross-correlation based image registration technique with improved performance*, Proc. of 11th European Signal Processing Conference (EUSIPCO 2002), vol. I, Sep. 2002, Toulouse, France, pp. 13–16.
- [13] K. Berberidis and S. Theodoridis, *A new fast block adaptive algorithm*, IEEE Transactions on Signal Processing **47** (1999), no. 1, 75–87.
- [14] K. A. Birney and T. R. Fischer, *On the modeling of the DCT and subband image data for compression*, IEEE Transactions on Image Processing **4** (1995), 186–193.
- [15] F. L. Bookstein, *Principal warps: Thin-plate splines and the decomposition of deformations*, IEEE Transactions on Pattern Analysis and Machine Intelligence **11** (1989), 567–585.
- [16] G. Borgefors, *Hierarchical chamfer matching: A parametric edge matching algorithm*, IEEE Transactions on Pattern Analysis and Machine Intelligence **10** (1988), 849–865.
- [17] C. Bouman and K. Sauer, *A generalized gaussian image model for edge-preserving map estimation*, IEEE Transactions on Image Processing **2** (1993), no. 3, 296–310.
- [18] L. Brown, *A survey of image registration techniques*, ACM Computing Surveys **24** (1992), no. 4, 325–376.
- [19] B. Chen and G. W. Wornell, *An information-theoretic approach to the design of robust digital watermarking systems*, Proc. of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 1999), vol. 4, 1999, pp. 2061–2064.
- [20] ———, *Quantization index modulation: A class of provably good methods for digital watermarking and information embedding*, IEEE Transactions on Information Theory **47** (2001), no. 4, 1423–1443.

- [21] Q. Chen, M. Defrise, and F. Deconinck, *Symmetric phase-only matched filtering of Fourier-Mellin transform for image registration and recognition*, IEEE Transactions on Pattern Recognition and Machine Intelligence **16** (1994), 1156–1168.
- [22] Q. Cheng and T. S. Huang, *An additive approach to transform-domain information hiding and optimum detection structure*, IEEE Transactions on Multimedia **3** (2001), no. 3, 273–284.
- [23] ———, *Robust optimum detection of transform domain multiplicative watermarks*, IEEE Transactions on Signal Processing **51** (2003), no. 4, 906–924.
- [24] R. J. Clarke, *Transform coding of images*, New York: Academic, 1985.
- [25] M. Costa, *Writing on dirty paper*, IEEE Transactions on Information Theory **29** (1983), 439–441.
- [26] I. J. Cox, J. Killian, F. T. Leighton, and T. Shamoan, *Secure spread spectrum watermarking for multimedia*, IEEE Transactions on Image Processing **6** (1997), 1673–1687.
- [27] I. J. Cox, J. Killian, T. Leighton, and T. Shamoan, *Secure spread spectrum watermarking for images, audio and video*, Proc. of IEEE International Conference on Image Processing (ICIP 1996), 1996, Lausanne, Switzerland, pp. 243–246.
- [28] I. J. Cox, M. L. Miller, and J. A. Bloom, *Digital watermarking*, Morgan Kaufman Publishers, 2002.
- [29] I. J. Cox, M. L. Miller, and A. L. McKellips, *Watermarking as communications with side information*, Proceedings of the IEEE **87** (1999), no. 7, 1127–1141.
- [30] X. Dai and S. Khorram, *A feature-based image registration algorithm using improved chain-code representation combined with invariant moments*, IEEE Transactions on Geoscience and Remote Sensing **37** (1999), 2351–2362.
- [31] M. H. Davis, A. Khotanzad, D. P. Flaming, and S. E. Harms, *A physics-based coordinate transformation for 3D image matching*, IEEE Transactions on Medical Imaging **16** (1996), no. 3, 317–328.
- [32] J. F. Delaigle, C. De Vleeschouwer, and B. Macq, *Watermarking algorithm based on a human visual model*, Signal Processing **66** (1998), no. 3, 319–335.

- [33] G. Depovere, T. Kalker, and J. P. Linnartz, *Improved watermark detection using filtering before correlation*, Proc. of IEEE International Conference on Image Processing (ICIP 1998), 1998, pp. 430–434.
- [34] J. Eggers and B. Girod, *Informed watermarking*, Kluwer Academic Publishers, 2002.
- [35] J. J. Eggers, R. Bäuml, R. Tzschoppe, and B. Girod, *Scalar Costa scheme for information embedding*, IEEE Transactions on Signal Processing **51** (2003), no. 4, 1003–1019.
- [36] J. J. Eggers, J. K. Su, and B. Girod, *An blind watermarking scheme based on structured codebooks*, Proc. of IEEE Secure Images and Image Authentication, 2000, London, UK.
- [37] ———, *Performance of a practical blind watermarking scheme*, Proc. of SPIE Security and Watermarking of Multimedia Contents III, vol. 4314, 2001, San Jose, Ca.
- [38] J. Flusser, *An adaptive method for image registration*, Pattern Recognition **25** (1992), 45–54.
- [39] ———, *Object matching by means of matching likelihood coefficients*, Pattern Recognition Letters **16** (1995), 893–900.
- [40] J. Flusser and T. Suk, *A moment-based approach to registration of images with affine geometric distortion*, IEEE Transactions on Geoscience and Remote Sensing **32** (1994), 382–387.
- [41] H. Foroosh (Shekarforoush), J. B. Zerubia, and M. Berthod, *Extension of phase correlation to subpixel registration*, IEEE Transactions on Image Processing **11**, no. 3.
- [42] Signal Processing Forum, *What is the future for watermarking? (Part II)*, IEEE Signal Processing Magazine (2003), 53–57.
- [43] J. Fridrich, *Robust bit extracion from images*, Proc. of IEEE ICMCS, Jun. 1999, Florence, Italy.
- [44] R. C. Gonzalez and R. E. Woods, *Digital image processing*, Prentice Hall, 2002.
- [45] A. Goshtasby, *Image registration by local approximation methods*, Image and Vision Computing **6** (1988), 255–261.

- [46] A. Goshtasby and G. C. Stockman, *Point pattern matching using convex hull edges*, IEEE Transactions on Systems, Man and Cybernetics **15** (1985), 631–637.
- [47] A. Goshtasby, G. C. Stockman, and C. V. Page, *A region-based approach to digital image registration with subpixel accuracy*, IEEE Transactions on Geoscience and Remote Sensing **24** (1986), 390–399.
- [48] V. Govindu and C. Shekhar, *Alignment using distributions of local geometric properties*, IEEE Transactions on Pattern Analysis and Machine Intelligence **21** (1999), 1031–1043.
- [49] E. Guest, E. Berry, R. A. Baldock, M. Fidrich, and M. A. Smith, *Robust point correspondence applied to two- and three-dimensional image registration*, IEEE Transactions on Pattern Analysis and Machine Intelligence **23** (2001), 165–179.
- [50] J. V. Hanjal, D. L. G. Hill, and D. J. Hawkes, *Medical image registration*, CRC Press, 2001.
- [51] F. Hartung and M. Kutter, *Multimedia watermarking techniques*, Proceedings of the IEEE **87** (1999), no. 7, 1079–1107.
- [52] S. Haykin, *Adaptive filter theory*, fourth ed., Prentice Hall, 2002.
- [53] C. Heegard and A. El Gamal, *On the capacity of computer memory with defects*, IEEE Transactions on Information Theory **29** (1983), 731–739.
- [54] J. R. Hernández, M. Amado, and F. Pérez-González, *DCT-domain watermarking techniques for still images: Detector performance analysis and a new stutture*, IEEE Transactions on Image Processing **9** (2000), no. 1, 55–68.
- [55] J. R. Hernández and F. Pérez-González, *Statistical analysis of watermarking schemes for copyright protection of images*, Proceedings of the IEEE **87** (1999), no. 7, 1142–1166.
- [56] D. L. G. Hill, P. G. Batchelor, M. Holden, and D. J. Hawkes, *Medical image registration*, Physics in Medicine and Biology **46** (2001), no. 3, R1–R45.
- [57] Y. C. Hsieh, D. M. McKeown, and F. P. Perlant, *Performance evaluation of scene registration and stereo matching for cartographic feature extraction*, IEEE Transactions on Pattern Analysis and Machine Intelligence **14** (1992), 214–237.

- [58] http://lcavwww.epfl.ch/reproducible_research/VandewalleSV05.
- [59] http://vision.unige.ch/certimark/public/CMK_D22.pdf.
- [60] <http://watermarking.unige.ch/Checkmark/index.html>.
- [61] <http://www.petitcolas.net/fabien/watermarking/StirMark/index.html>.
- [62] I. Karybali and K. Berberidis, *Blind image-adaptive watermarking*, Proc. of 10th IEEE International Conference on Electronics Circuits and Systems (ICECS 2003), Dec. 2003, Sharjah, UAE, pp. 894–897.
- [63] ———, *A new blind image-adaptive watermarking scheme: Theoretical and experimental results*, Proc. of XII European Signal Processing Conference (EUSIPCO 2004), Sep. 2004, Vienna, Austria, pp. 1489–1492.
- [64] I. Karybali, K. Berberidis, and C. Balas, *A high-performance image registration technique and application to multispectral imaging for medical diagnosis*, Proc. of 16th International Eurasip Conference BIOSIGNAL, vol. 16, Jun. 2002, Brno, Czech Republic, pp. 255–258.
- [65] I. G. Karybali and K. Berberidis, *Efficient spatial image watermarking via new perceptual masking and blind detection schemes*, IEEE Transactions on Image Processing, Accepted with minor revisions.
- [66] ———, *Improved embedding of multiplicative watermarks via space-time block coding*, Proc. of 13th European Signal Processing Conference (EUSIPCO 2005), Sep. 2005, Antalya, Turkey.
- [67] I. G. Karybali, K. Berberidis, E. Z. Psarakis, and G. D. Evangelidis, *A spatial domain approach to subpixel image registration*, IEEE Transactions on Image Processing, Submitted.
- [68] I. G. Karybali, E. Z. Psarakis, K. Berberidis, and G. D. Evangelidis, *Efficient image registration with subpixel accuracy*, IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2006), Submitted.
- [69] S. M. Kay, *Fundamentals of statistical signal processing, Volume II, Detection theory*, Prentice-Hall, 1998.

- [70] D. Keren, S. Peleg, and R. Brada, *Image sequence enhancement using sub-pixel displacement*, Proc. of IEEE International Conference on Computer Vision and Pattern Recognition (ICCV-PR 1988), Jun. 1988, pp. 742–746.
- [71] S. P. Kim and W. Y. Su, *Subpixel accuracy image registration by spectrum cancellation*, Proc. of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 1993), vol. 5, Apr. 1993, pp. 153–156.
- [72] O. Koval, S. Voloshynovskiy, and F. Pérez-González, *Quantization-based watermarking performance improvement using host statistics: AWGN attack case*, Proc. of ACM Multimedia and Security Workshop, Sep. 2004, Magneburg, Germany.
- [73] O. Koval, S. Voloshynovskiy, F. Pérez-González, F. Deguillaume, and T. Pun, *Spread spectrum watermarking for real images: Is everything so hopeless?*, Proc. of XII European Signal Processing Conference (EUSIPCO 2004), Sep. 2004, Vienna, Austria, pp. 1477–1480.
- [74] D. Kundur and D. Hatzinakos, *A robust digital image watermarking scheme using wavelet-based fusion*, Proc. of IEEE International Conference on Image Processing (ICIP 1997), 1997, Santa Barbara, CA, pp. 544–547.
- [75] ———, *Diversity and attack characterization for improved robust watermarking*, IEEE Transactions on Signal Processing **49** (2001), no. 10, 2383–2396.
- [76] M. Kutter, F. Jordan, and F. Bossen, *Digital signature of color images using amplitude modulation*, Proc. of SPIE Electronic Imaging, Storage and Retrieval for Image and Video Databases V, Feb. 1997, San Jose, CA, pp. 518–526.
- [77] M. Kutter and F. Petitcolas, *A fair benchmark for image watermarking systems*, SPIE, vol. 3657, Jun. 1999, San Jose, pp. 226–239.
- [78] M. Kutter and S. Winkler, *A vision based masking model for spread-spectrum image watermarking*, IEEE Transactions on Image Processing **11** (2002), no. 1, 16–25.
- [79] G. C. Langelaar, I. Setyawan, and R. L. Lagendijk, *Watermarking digital image and video data, a state-of-the-art overview*, IEEE Signal Processing Magazine **17** (2000), no. 6, 20–46.
- [80] J. P. Lewis, *Fast normalized cross-correlation*, Vision Interface (1995).

- [81] H. Li, B. S. Manjunath, and S. K. Mitra, *A contour-based approach to multisensor image registration*, IEEE Transactions on Image Processing **4** (1995), 320–334.
- [82] X. Li, Z. Du, W. Li, P. Lu, and X. Xue, *Multi-channel data hiding scheme for color images*, Proc. of IEEE ITCC 2003, 2003, Las Vegas, NV, USA, pp. 569–573.
- [83] X. Li and X. Xue, *Improved robust watermarking in DCT domain for color images*, Proc. of IEEE AINA 2004, 2004, Fukuoka, Japan, pp. 53–57.
- [84] J. S. Lim, *Two-dimensional signal and image processing*, Prentice Hall, 1990.
- [85] C.-Y. Lin, M. Wu, J. A. Bloom, I. J. Cox, M. L. Miller, and Y. M. Lui, *Rotation scale and translation resilient watermarking for images*, IEEE Transactions on Image Processing **10** (2001), no. 5, 767–782.
- [86] L. Lucchese and G. M. Cortelazzo, *A noise-robust frequency domain technique for estimating planar roto-translations*, IEEE Transactions on Signal Processing **48** (2000), no. 6, 1769–1786.
- [87] A. Morse M. Anderson, J. Jordan and F. Sharp, *Intergrated colposcopy*, London: Chapman & Hall Medical, 1996.
- [88] B. Macq, J. Dittmann, and E. J. Delp, *Benchmarking of image watermarking algorithms for digital rights management*, Proceedings of the IEEE **92** (2004), no. 6, 971–984.
- [89] B. M. Macq and J. J. Quisquater, *Cryptology for digital TV broadcasting*, Proceedings of the IEEE **83** (1995), no. 6, 944–957.
- [90] J. B. A Maintz and M. A. Viergever, *A survey of medical image registration*, Medical Image Analysis **2** (1998), 1–36.
- [91] S. G. Mallat, *A theory for multiresolution signal decomposition: The wavelet representation*, IEEE Transactions on Pattern Analysis and Machine Intelligence **11** (1989), no. 7, 674–693.
- [92] B. Marcel, M. Briot, and R. Murrieta, *Calcul de translation et rotation par la transformation de Fourier*, Traitement du Signal **14** (1997), no. 2, 135–149.

- [93] M. L. Miller, G. J. Doërr, and I. J. Cox, *Applying informed coding and embedding to design a robust high-capacity watermark*, IEEE Transactions on Image Processing **13** (2004), no. 6, 792–807.
- [94] P. Moulin and J. Liu, *Analysis of multiresolution image denoising schemes using generalized Gaussian and complexity priors*, IEEE Transactions on Information Theory **45** (1999), no. 3, 909–919.
- [95] P. Moulin and J. A. Ó Sullivan, *Information-theoretic analysis of information hiding*, IEEE Transactions on Information Theory **49** (2003), no. 3, 563–593.
- [96] A. N. Netravali and B. G. Haskell, *Digital pictures. representation, compression and standards*, New York: Plenum, 1995.
- [97] M. Ohki and S. Hashiguchi, *Two-dimensional LMS adaptive filters*, IEEE Transactions on Consumer Electronics **37** (1991), 66–73.
- [98] N. R. Pal and S. K. Pal, *A review on image segmentation techniques*, Pattern Recognition **26** (1993), 1277–1294.
- [99] J. A. Parker, R. V. Kenyon, and D. E. Troxel, *Comparison of interpolating methods for image resampling*, IEEE Transactions in Medical Imaging **2** (1983), 31–39.
- [100] T. Peli, *An algorithm for recognition and localization of rotated and scaled objects*, Proceedings of the IEEE **69** (1981), 483–485.
- [101] S. Pereira, S. Voloshynovskiy, M. Madueno, S. Marchand-Maillet, and T. Pun, *Second generation benchmarking and application oriented evaluation*, Proc. of Information Hiding Workshop, Apr. 2001, Pittsburg, PA, USA.
- [102] S. Pereira, S. Voloshynovskiy, and T. Pun, *Optimal transform domain watermark embedding via linear programming*, Signal Processing **81** (2001), 1251–1260.
- [103] F. Pérez-González, F. Balado, and J. R. Hernández Martín, *Performance analysis of existing and new methods for data hiding with known-host information in additive channels*, IEEE Transactions on Signal Processing **51** (2003), no. 4, 960–980.
- [104] S. Periaswamy and H. Farid, *Elastic registration in the presence of intensity variations*, IEEE Transactions on Medical Imaging **22** (2003), no. 7, 865–874.

- [105] F. A. P. Petitcolas, R. J. Anderson, and M. G. Kuhn, *Attacks on copyright marking systems*, Proc. of Information Hiding 2nd International Workshop, 1998, pp. 219–239.
- [106] I. Pitas, *A method for signature casting on digital images*, Proc. of IEEE International Conference on Image Processing (ICIP 1996), 1996, Lausanne, Switzerland, pp. 215–218.
- [107] A. Piva, M. Barni, F. Bartolini, and V. Capellini, *DCT-based watermark recovering without resorting to uncorrupted original image*, Proc. of IEEE International Conference on Image Processing (ICIP 1997), 1997, Santa Barbara, CA, pp. 520–527.
- [108] 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** (2003), no. 8, 986–1004.
- [109] C. Podilchuk and W. Zeng, *Perceptual watermarking of still images*, Proc. of IEEE Workshop on Multimedia Signal Processing, 1997, Princeton, NJ, pp. 363–368.
- [110] C. I. Podilchuk and W. Zeng, *Image-adaptive watermarking using visual models*, IEEE Journal on Selected Areas in Communications **16** (1998), no. 4, 525–539.
- [111] W. K. Pratt, *Correlation techniques for image registration*, IEEE Transactions on Aerospace and Electronic Systems **10** (1972), 353–358.
- [112] ———, *Digital image processing*, John Wiley & Sons Inc, 2nd Ed., 1991.
- [113] E. Z. Psarakis and G. D. Evangelidis, *An ENCC based self-weighted similarity measure tailored to the stereo correspondence problem*, To be submitted to IEEE Transactions on Pattern Analysis and Machine Intelligence.
- [114] ———, *An enhanced correlation-based method for stereo correspondence with sub-pixel accuracy*, Proc. of 10th IEEE International Conference on Computer Vision (ICCV 2005), Oct. 2005, Beijing, China.
- [115] S. S. Rao, *Engineering optimization, theory and practice*, John Wiley & Sons Inc, 1996.
- [116] B. S. Reddy and B. N. Chatterji, *An FFT-based technique for translation, rotation and scale-invariant image registration*, IEEE Transactions on Image Processing **5** (1996), 1266–1271.

- [117] A. Roche, G. Malandain, X. Pennec, and N. Ayache, *The correlation ratio as a new similarity measure for multimodal image registration*, vol. 1496, 1998, Lecture Notes in Computer Science, Cambridge, USA, pp. 1115–1124.
- [118] K. Rohr, *Landmark-based image analysis: Using geometric and intensity models*, vol. 21, Kluwer Academic Publishers, 2001, Computational Imaging and Vision Series.
- [119] M. Roux, *Automatic registration of SPOT images and digitized maps*, Proc. of IEEE International Conference on Image Processing (ICIP 1996), 1996, Lausanne, Switzerland, pp. 625–628.
- [120] J. J. K. Ó Ruanaidh, W. J. Dowling, and F. M. Boland, *Phase watermarking of digital images*, Proc. of IEEE International Conference on Image Processing (ICIP 1996), 1996, Lausanne, Switzerland, pp. 239–242.
- [121] J. J. K. Ó Ruanaidh and T. Pun, *Rotation scale and translation invariant spread spectrum digital image watermarking*, Signal Processing **66** (1998), no. 3, 303–317.
- [122] A. M. Sequeira and C. W. Terrien, *A new 2-D fast RLS algorithm*, Proc. of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 1990), vol. 3, Apr. 1990, pp. 1401–1404.
- [123] C. E. Shannon, *Channels with side information at the transmitter*, IBM Journal of Research and Development (1958), 289–293.
- [124] H. Shekarforoush, M. Berthod, and J. Zerubia, *Subpixel image registration by estimating the polyphase decomposition of cross power spectrum*, Computer Vision Pattern Recognition (1994), 532–537.
- [125] R. M. Shier, *The colposcopy unit*, Obstetrics and Gynecology Clinics of N. America **40**(1) (1993), 55–58.
- [126] D. Shin, J. K. Pollard, and J. P. Muller, *Accurate geometric correction of ATRS images*, IEEE Transactions on Geoscience and Remote Sensing **35** (1997), 997–1006.
- [127] V. Solachidis, A. Tefas, N. Nikolaidis, S. Tsekeridou, and A. Nikolaidis, *A benchmarking protocol for watermarking methods*, Proc. of IEEE International Conference on Image Processing (ICIP 2001), 2001, pp. 1023–1026.

- [128] G. Stockman, S. Kopstein, and S. Benett, *Matching images to models for registration and object detection via clustering*, IEEE Transactions on Pattern Analysis and Machine Intelligence **4** (1982), 229–241.
- [129] H. S. Stone, M. T. Orchard, E-C. Chang, and S. A. Matrucci, *A fast direct Fourier-based algorithm for subpixel registration of images*, IEEE Transactions on Geoscience and Remote Sensing **39** (2001), no. 10, 2235–2243.
- [130] H. S. Stone and R. Wolpov, *Blind cross-spectral image registration using prefiltering and Fourier-based translation detection*, IEEE Transactions on Geoscience and Remote Sensing **40** (2002), no. 3, 637–650.
- [131] M. D. Swanson, M. Kobayashi, and A. H. Tewfik, *Multimedia data-embedding and watermarking technologies*, Proceedings of the IEEE **86** (1998), 1064–1087.
- [132] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, *The application of orthogonal designs to wireless communication*, Proc. of Information Theory Workshop, Killarney, Ireland.
- [133] ———, *Space-time block codes from orthogonal designs*, IEEE Transactions on Information Theory **45** (1999), no. 5, 1456–1467.
- [134] V. Tarokh, S. Seshadri, and A. R. Calderbank, *Space-time codes for high data rate wireless communication: Performance analysis and code construction*, IEEE Transactions on Information Theory **44** (1998), 744–765.
- [135] E. De Castro and C. Morandi, *Registration of translated and rotated images using finite Fourier transforms*, IEEE Transactions on Pattern Analysis and Machine Intelligence **PAMI-9** (1987), 700–703.
- [136] J. De Moigne and I. Zavorin, *Use of wavelets for image registration*, Proc. of SPIE Aerospace 2000, Wavelet Applications VIII, Apr. 2000.
- [137] S. L. Marple Jr., *Digital spectral analysis*, Prentice Hall, 1987.
- [138] P. Thévenaz, U. E. Ruttimann, and M. Unser, *A pyramidal approach to subpixel registration based on intensity*, IEEE Transactions on Image Processing **7** (1998), 27–41.
- [139] Q. Tian and M. N. Huhns, *Algorithms for subpixel registration*, Computer Vision, Graphics, and Image Processing **35** (1986), 220–233.

- [140] A. Z. Tirkel, C. F. Osbourne, and T. E. Hall, *Image and watermark registration*, Signal Processing **66** (1998), no. 3, 373–383.
- [141] J. Ton and A. K. Jain, *Registering landsat images by point matching*, IEEE Transactions on Geoscience and Remote Sensing **27** (1989), 642–651.
- [142] R. Tsai and T. Huang, *Multiframe image restoration and registration*, Advances in Computer Vision and Image Processing **1** (1984), 317–339.
- [143] P. A. van den Elsen, E. J. D. Pol, and M. A. Viergever, *Medical image matching - a review with classification*, IEEE Engineering in Medicine and Biology **12** (1993), no. 1, 26–39.
- [144] P. Vandergheynst and J. F. Gobbers, *Directional dyadic wavelet transforms: Design and algorithms*, IEEE Transactions on Image Processing **11** (2002), no. 4, 363–372.
- [145] P. Vandewalle, S. Süsstrunk, and M. Vetterli, *A frequency domain approach to registration of aliased images with application to super-resolution*, Eurasip Journal on Applied Signal Processing, To appear in SI on Super-Resolution Imaging.
- [146] P. Viola and W. M. Wells, *Alignment by maximization of mutual information*, Proc. of 5th IEEE International Conference on Computer Vision (ICCV 1995), Jun. 1995, pp. 16–23.
- [147] S. Voloshynovskiy, F. Deguillaume, S. Pereira, and T. Pun, *Optimal adaptive diversity watermarking with channel state estimation*, Proc. of SPIE Security and Watermarking of Multimedia Contents III, vol. 4314, 2001, San Jose, Ca.
- [148] S. Voloshynovskiy, A. Herrigel, N. Baumgaertner, and T. Pun, *A stochastic approach to content adaptive digital image watermarking*, Proc. of 3rd International Workshop on Information Hiding, Sep. 1999, Dresden, Germany, pp. 211–236.
- [149] S. Voloshynovskiy, S. Pereira, V. Iquise, and T. Pun, *Attack modelling: Towards a second generation benchmark*, Signal Processing **81** (2001), 1177–1214.
- [150] C. Y. Wang, H. Sun, S. Yadas, and A. Rosenfeld, *Some experiments in relaxation image matching using corner features*, Pattern Recognition **16** (1983), 167–182.
- [151] A. B. Watson, *DCT quantization matrices visually optimized for individual images*, Proc. of SPIE Conf. Human Vision, Visual Processing and Digital Display VI, 1995.

- [152] W. M. Wells and P. Viola et al., *Multi-modal volume registration by maximization of mutual information*, Medical Image Analysis **1** (1996), 35–51.
- [153] S. Winkler, *Vision models and quality metrics for image processing applications*, Ph.D. no. 2313, Lausanne, EPFL, 2000.
- [154] G. Wolberg and S. Zokai, *Robust image registration using log-polar transform*, Proc. of IEEE International Conference on Image Processing (ICIP 2000), 2000, Canada.
- [155] R. B. Wolfgang and E. J. Delp, *A watermark for digital images*, Proc. of IEEE International Conference on Image Processing (ICIP 1996), 1996, Lausanne, Switzerland, pp. 219–222.
- [156] R. B. Wolfgang, C. I. Podilchuk, and E. J. Delp, *Perceptual watermarks for digital images and video*, Proceedings of the IEEE **87** (1999), no. 7, 1108–1126.
- [157] X. G. Xia, C. G. Boncelet, and G. R. Arce, *A multiresolution watermark for digital images*, Proc. of IEEE International Conference on Image Processing (ICIP 1997), 1997, Santa Barbara, CA, pp. 548–551.
- [158] W. Zeng and B. Liu, *On resolving rightful ownership of digital images by invisible watermarks*, Proc. of IEEE International Conference on Image Processing (ICIP 1997), 1997, Santa Barbara, CA, pp. 552–555.
- [159] Q. Zheng and R. Chellapa, *A computational vision approach to image registration*, IEEE Transactions on Image Processing **2** (1993), 311–325.
- [160] Z. Zheng, H. Wang, and E. K. Teoh, *Analysis of grey level corner detection*, Pattern Recognition Letters **20** (1999), 149–162.
- [161] D. Ziou and S. Tabbone, *Edge detection techniques - An overview*, International Journal of Pattern Recognition and Image Analysis **8** (1998), 537–559.
- [162] B. Zitová and J. Flusser, *Image registration methods: A survey*, Elsevier Image and Vision Computing **21** (2003), 977–1000.
- [163] B. Zitová, J. Flusser, and F. Sroubek, *Application of image processing for the conservation of the medieval mosaic*, Proc. of IEEE International Conference on Image Processing (ICIP 2002), 2002, Rochester, pp. 993–996.