Εργαστηριο Ηλεκτρονικης
Τμημα Ηλεκτρολογων Μηχανικων και Μηχανικων Υπολογιστων
Πολυτεχνικη Σχολη
Δημοκριτειο Πανεπιστημιο Θρακης

Αναπτυξη και υλοποιηση νεων τεχνικων ορασης μηχανων για εφαρμογες αναγνωρισης σε πραγματικο χρονο

Διδακτορικη Διατριβη

Λεωνιδας Κωτουλας

Τριμελης συμβουλευτικη επιτροπη:
Καθ. Ιωαννης Ανδρεαδης (Τμημα Η.Μ.&Μ.Υ., Δ.Π.Θ.), Προεδρος
Καθ. Νικολαος Παπαμαρκος (Τμημα Η.Μ.&Μ.Υ., Δ.Π.Θ.)
Αν. Καθ. Ιωαννης Θεοχαρης (Τμημα Η.Μ.&Μ.Υ., Α.Π.Θ.)

Ξανθη
Ιουνιος 2007

Περίληψη - Συμβολή στην Επιστήμη

Οι ροπές εικόνων αποτελούν συχνά χρησιμοποιούμενους περιγραφείς σχήματος σε πλήθος εφαρμογών, σε όλο το εύρος του ερευνητικού πεδίου της επεξεργασίας και ανάλυσης εικόνας. Ουσιαστικά, αποτελούν ένα χαμηλού επιπέδου εργαλείο αναγνώρισης προτύπων, το οποίο βασίζεται στη στατιστική κατανομή της εικόνας. Σε αυτή τη διατριβή, παρουσιάζονται οι βασικοί τύποι τους, και αναλύονται τα χαρακτηριστικά κάθε ενός.

Βασικός στόχος της διερεύνησης που πραγματοποιήθηκε στο μεγαλύτερο μέρος της Διδακτορικής Διατριβής ήταν η βελτίωση των χαρακτηριστικών των ροπών για την πιο αποδοτική χρήση τους. Πιο συγκεκριμένα, αρχικά παρουσιάζονται τεχνικές για την ταχεία υλοποίησή τους τόσο σε υλικό όσο και σε λογισμικό. Αναπτύχθηκαν υλοποιήσεις που στοχεύουν στις ροπές *Zernike*, *Legendre*, *Chebyshev* καθώς και στις γεωμετρικές ροπές. Η ταχύτητα των προτεινόμενων τεχνικών είναι η μεγαλύτερη που έχει αναφερθεί μέχρι σήμερα στη βιβλιογραφία, και επιτρέπει επεξεργασία πραγματικού χρόνου για εικόνες μεγάλου μεγέθους. Σε σχέση με τις υπάρχουσες τεχνικές, οι προτεινόμενες παρουσιάζουν βελτίωση στην ταχύτητα η οποία υπερβαίνει την μία τάξη μεγέθους.

Επιπλέον, περιγράφεται μία νέα τεχνική για τον υπολογισμό των ροπών αυτών σε δυαδικές εικόνες, η οποία στηρίζεται στις ιδιότητες των παραγοντικών πολυωνύμων στο διακριτό χώρο, και στην περιγραφή των δυαδικών εικόνων μέσω του περιγράμματός τους. Εκτός των συνήθων τύπων μελετήθηκε και ο πρόσφατα προταθείς Γωνιακός Ακτινικός Μετασχηματισμός, ο οποίος και υλοποιήθηκε σε μία αποδοτική αρχιτεκτονική. Στην περίπτωση των δυαδικών εικόνων, πέρα από την παραπλήσια βελτίωση στην ταχύτητα, επιτυγχάνεται και σημαντική μείωση των απαιτήσεων σε μνήμη ($\approx 50\%$), οι οποίες παίζουν καθοριστικό ρόλο σε ενσωματωμένα συστήματα.

Πέρα από τον ταχύ υπολογισμό των ροπών, παρουσιάζεται και μεθοδολογία για την μείωση των σφαλμάτων που εισάγονται κατά τον υπολογισμό τους. Λόγω της διακριτής φύσης των ψηφιακών εικόνων, ο υπολογισμός των ροπών επηρεάζεται από σφάλματα διακριτοποίησης. Πιο συγκεκριμένα, εμφανίζονται σφάλματα προσέγγισης τόσο στον υπολογισμό των ολοκληρωμάτων, όσο και στην προσέγγιση του μοναδιαίου δίσκου από καρτεσιανό πλέγμα. Η τεχνική που αναπτύχθηκε αντιμετωπίζει την πρώτη πηγή σφαλμάτων, με τη χρήση πολυωνυμικής παρεμβολής στο χώρο των ροπών. Τα πειραματικά αποτελέσματα έδειξαν ότι με τη χρήση της προτεινόμενης μεθοδολογίας, όχι μόνο είναι πιο ακριβής ο υπολογισμός των ροπών, αλλά βελτιώνονται και τα χαρακτηριστικά αμεταβλητότητάς τους, τα οποία είναι ίσως το πιο σημαντικό πλεονέκτημα στη χρήση των ροπών για την αναγνώριση προτύπων. Το σφάλμα εξαρτάται από τον τύπο των χρησιμοποιούμενων ροπών, αλλά στη γενική περίπτωση είναι υποδιπλάσιο άλλων τεχνικών.

Οι πρόσφατα προταθείσες ροπές διακριτής ορθογώνιας βάσης, παρόλο που δεν παρουσιάζουν εγγενή αμεταβλητότητα, δεν εμφανίζουν σφάλματα κατά τον υπολογισμό τους, αφού είναι ορισμένες στο διακριτό πεδίο της εικόνας. Οι συναρτήσεις βάσης για αυτές τις ροπές είναι πολυώνυμα ορθογώνια στο διακριτό χώρο της εικόνας. Τέτοια πολυώνυμα μπορούν να προκύψουν εύκολα μέσω της τεχνικής $Gram-Schmidt$, με δεδομένο ένα διάνυσμα βάρους w. Στην αναζήτηση για ένα τέτοιο διάνυσμα, το οποίο θα βελτιστοποιεί τα χαρακτηριστικά των παραγόμενων ροπών χρησιμοποιούνται εξελικτικές τεχνικές. Πιο συγκεκριμένα, αντιμετωπίσθηκε η αναζήτηση αυτή ως πρόβλημα ελαχιστοποίησης, για το οποίο χρησιμοποιήθηκαν εξελικτικές στρατηγικές. Οι παραγόμενες ροπές έδωσαν σημαντικά βελτιωμένα χαρακτηριστικά ως προς την ανακατασκευή και την ανάκτηση σχημάτων.

Παρόλα τα μειονεκτήματα των ροπών διακριτής βάσης, παρουσιάζουν και ένα σύνολο πολύ χρήσιμων ιδιοτήτων, λόγω της ευκολίας στην ανακατασκευή του σήματος. Προς εκμετάλλευση αυτών των ιδιοτήτων, παρουσιάζονται τεχνικές για τον υπολογισμό της επίδρασης γραμμικών και μη μετασχηματισμών στο χώρο των ροπών. Με βάση τα αποτελέσματα αυτής της διερεύνησης, αναπτύσσονται δύο τεχνικές για τον υπολογισμό των ροπών αντικειμένου το οποίο αποκρύπτεται μερικώς, καθώς και για τον υπολογισμό ροπών εικόνας από μονοδιάστατες προβολές της. Η πρώτη τεχνική δίνει τη δυνατότητα μερικής αμεταβλητότητας στην απόκρυψη, η οποία επηρεάζει τις τιμές των ροπών σε πολύ μεγάλο βαθμό. Η επίδραση της αμεταβλητότητας στο χώρο των ροπών δεν είχε μελετηθεί μέχρι σήμερα, λόγω του μεγάλου αριθμού των μεταβλητών οι οποίες εμπλέκονται. Με την προτεινόμενη τεχνική, όμως, επιτυγχάνεται μείωση των διαστάσεων του προβλήματος ελαχιστοποίησης σε πολύ λίγες. Η δεύτερη αποτελεί μία πολύ αποδοτική μέθοδο για το χαρακτηρισμό μίας εικόνας, όταν είναι γνωστές μόνο οι προβολές της.

Ένα βασικό πρόβλημα που απαντάται στις περισσότερες εφαρμογές υπολογιστικής όρασης, και ιδιαίτερα σε αυτές της ρομποτικής όρασης, είναι η ανάκτηση πληροφορίας βάθους. Ίσως ο πιο συχνά χρησιμοποιούμενος τρόπος για την ανάκτηση της πληροφορίας αυτής είναι η χρήση ζεύγους συσκευών απεικόνισης. Η εύρεση των αντιστοιχιών σημείων μεταξύ των δύο εικόνων αποτελεί το βασικό πρόβλημα σε τέτοιου είδους τεχνικές. Εντούτοις, το πρόβλημα συχνά διαχωρίζεται σε δύο μέρη: Στην εύρεση ενός σχετικά μικρού αριθμού σημείων αντιστοιχιών στις αρχικές, μη βαθμονομημένες εικόνες, προκειμένου να πραγματοποιηθεί η βαθμονόμηση, και στην αντιστοίχιση της πλειοψηφίας των σημείων των βαθμονομημένων εικόνων. Το πρώτο μέρος αντιμετωπίζεται με μία αρχιτεκτονική υλικού, η οποία χρησιμοποιεί τις ροπές *Zernike* ως τοπικούς περιγραφείς. Η τεχνική που προτάθηκε για τον ταχύ υπολογισμό τους, υλοποιείται για την ανάπτυξη συστήματος πραγματικού χρόνου. Για την αντιμετώπιση του δεύτερου μέρους του προβλήματος περιγράφεται μία αρχιτεκτονική υλικού, η οποία παρουσιάζει τη μεγαλύτερη ταχύτητα στη βιβλιογραφία. Παρόλο που οι παραγόμενοι χάρτες βάθους δεν παρουσιάζουν μεγάλη ακρίβεια, η εξαιρετικά υψηλή ταχύτητα επεξεργασίας καθιστά τη μέθοδο χρήσιμη σε πολλές εφαρμογές. Πιο συγκεκριμένα, η προτεινόμενη αρχιτεκτονική μπορεί να επεξεργαστεί σε πραγματικό χρόνο εικόνες μεγέθους μέχρι και 2.000×2.000 στοιχείων, κάτι που ως τώρα ήταν δυνατό μόνο για εικόνες μεγέθους μέχρι περίπου 640×480. Τα συστήματα πυκνής και αραιής αντιστοίχησης σημείων χρησιμοποιήθηκαν με επιτυχία στην ανάπτυξη του μηχανοτρονικού συστήματος *RESCUER*.

Επιγραμματικά, στην παρούσα Διδακτορική Διατριβή παρουσιάζονται τεχνικές για τον ταχύ και ακριβή υπολογισμό ροπών εικόνων, για τη βελτιστοποίηση των χαρακτηριστικών τους, και για την εφαρμογή γραμμικών και μη μετασχηματισμών. Με βάση αυτές τις τεχνικές, υλοποιούνται μέθοδοι για την αντιμετώπιση προβλημάτων που σχετίζονται με την υπολογιστική όραση, όπως η ανάκτηση πληροφορίας βάθους, και η ανακατασκευή από προβολές.

Summary - Contribution to the State of the Art

Image moments are shape descriptors, commonly used in many applications, throughout the fields of research of computer and machine vision. In this PhD Thesis, the most common types of image moments are presented, along with their respective properties.

One of the main goals of this research was the improvement of the characteristics of image moments, for their more effective use. More specifically, initially techniques are proposed for the fast hardware and software implementation of geometric, Zernike, Legendre and Chebyshev moments. The speed of the proposed architectures is the fastest reported in the literature, and allows real time processing even for large sized images. Compared to other techniques, the increase in performance is over an order of magnitude.

Furthermore, a new technique for the calculation of moments of binary images is proposed. It is based on the properties of factorial polynomials in the discrete space, and the description of binary images through their contour. Except for the common moment types, the recently proposed Angular Radial Transform was also studied, and was implemented by a new, efficient architecture. In addition to the speed increase, which is similar to the one indicated for gray-scale images, there is a considerable reduction of the memory requirements ($\approx 50\%$), a feature most important when embedded systems are considered.

Apart from the fast computation of moments, a new methodology for the reduction of errors introduced in their calculation is proposed. Due to the discrete nature of digital images, integral approximation errors, as well as unit disk discretization errors affect the resulting moment values. The proposed technique reduces the former, by introducing polynomial interpolation in the

moment space. Experimental results have shown that not only the extracted moments are more accurate, but also, their invariance properties are enhanced, which are probably their most useful characteristics of image moments in pattern recognition applications.

The recently proposed discrete orthogonal moments, although not presenting native invariance, are free of discrete approximation errors throughout their calculation, since they are defined in the discrete image space. Their basis functions are polynomials, orthogonal in a discrete subset. Such polynomials can be easily constructed using the Gram-Schmidt procedure, given a weight vector w. Evolutionary Strategies are employed, in the search for an optimal vector w, which will produce moments with desired characteristics. The derived moments presented improved characteristics regarding image reconstruction and shape retrieval.

Despite the invariance problems of discrete basis moments, they present some very useful properties, due to the straightforward and accurate image reconstruction from these moments. In order to make use of these properties, techniques suitable for the calculation of the effect of linear and non linear transformations are derived. Using these as a tool, two techniques are proposed, for the estimation of moments of a partially occluded object, as well as for the calculation of moments of an object, when only projections are available. The former can be used for implementing a retrieval system that behaves well under occlusion, which generally affects greatly moment values. The latter provides a very efficient method for shape characterization, when only projections are known.

A fundamental problem in computer vision, and especially in machine vision, is the extraction of depth information. Perhaps the most commonly used methods for obtaining such information are based on the use of a stereo camera pair. The basic problem of these techniques is the matching of points between the two images. Nevertheless, this problem is most commonly divided into two parts: The matching of a few points between the two initial, non-calibrated images, and the matching of the remaining points in the rectified stereo pair. In this Thesis, the first part is addressed through a hardware architecture, that utilizes Zernike moments as local feature descriptors. The technique that was proposed for their fast calculation is implemented in order to obtain a real-time system. For the addressing of the second part, a real-time hardware architecture is described, which presents the highest speed that can be found in the literature. Although the derived depth maps do not present very high accuracy, the very high processing speed of the implementation proves this method useful for many practical applications. The systems for dense and sparse disparity map extraction were implemented successfully in the development of the *RESCUER* mechatronic system.

Summarizing, in the present Thesis, techniques are presented for the fast and accurate calculation of image moments, for the enhancement of their properties and the effect of linear and non-linear transformations. Based on these techniques, methods for addressing fundamental problems of machine vision, such as depth information extraction and projection reconstruction, are proposed.

Πρόλογος

Η παρούσα εργασία, στο μεγαλύτερο μέρος της, παρουσιάζει τεχνικές που σχετίζονται με τις ροπές εικόνων. Αυτές αποτελούν ευρέως χρησιμοποιούμενους τοπικούς περιγραφείς εικόνων. Παραδείγματα εφαρμογών τους βρίσκονται στην πλοήγηση κινητών ρομπότ, στην κανονικοποίηση εικόνων, στην αναγνώριση προτύπων και στην υδατογράφηση δεδομένων.

Μία επισκόπηση των διαφόρων τύπων ροπών, μαζί με τους ορισμούς και τις βασικές τους ιδιότητες παρουσιάζεται στο κεφάλαιο 1. Μελετώνται τα πλεονεκτήματα και τα μειονεκτήματα κάθε τύπου, καθώς και οι ανάγκες που οδήγησαν στην ανάπτυξή τους.

Ίσως το μεγαλύτερο πρόβλημα της χρήσης των ροπών είναι οι μεγάλες υπολογιστικές τους απαιτήσεις. Η υπολογιστική τους πολυπλοκότητα καθιστά την απ' ευθείας υλοποίησή τους μη πρακτική, ιδιαίτερα σε συστήματα με περιορισμένους πόρους. Στο κεφάλαιο 2 παρουσιάζονται αρχιτεκτονικές υλικού και λογισμικού για τον αποδοτικό υπολογισμό τους. Με τη χρήση των τεχνικών αυτών, καθιστάται δυνατή η εξαγωγή τους σε πραγματικό χρόνο. Στο κεφάλαιο 6.1 παρουσιάζεται ένα σύστημα στερεοσκοπικής όρασης, το οποίο υλοποιείται με τη βοήθεια ροπών *Zernike*.

Πέρα από τον θόρυβο, ο οποίος είναι εγγενής σε κάθε σύστημα όρασης, ο υπολογισμός των ροπών υπόκειται σε υπολογιστικά σφάλματα που οφείλονται κυρίως στην διακριτή απεικόνιση των δεδομένων. Στο κεφάλαιο 3 παρουσιάζεται μία μεθοδολογία για την ελαχιστοποίηση αυτών των σφαλμάτων. Η τεχνική αυτή βελτιώνει τα χαρακτηριστικά όλων των τύπων ροπών συνεχούς βάσης.

Στο κεφάλαιο 4 παρουσιάζεται μία μεθοδολογία για τη βελτιστοποίηση των ιδιοτήτων των ροπών διακριτής βάσης με τη χρήση εξελικτικών αλγορίθμων. Πιο συγκεκριμένα, χρησιμοποιώντας εξελικτικές στρατηγικές, επιλέγεται ένα βέλτιστο αρχικό διάνυσμα βάσης, στο οποίο στηρίζεται η εξαγωγή της οικογένειας συναρτήσεων πυρήνα των εξαγόμενων ροπών. Το τελικό αποτέλεσμα αποδεικνύεται πλεονεκτικό ως προς την απεικόνιση της πληροφορίας, σε σχέση με τις τυπικά χρησιμοποιούμενες ροπές. Κάποιες χρήσιμες ιδιότητες των ροπών διακριτής βάσης παρουσιάζονται στο κεφάλαιο 5, μαζί με αρχιτεκτονικές για την ταχεία υλοποίησή τους. Βάσει αυτών των ιδιοτήτων, προτείνεται μία τεχνική για την αντιμετώπιση της απόκρυψης αντικειμένων, καθώς και μία μέθοδος για την εξαγωγή των ροπών σήματος, όταν είναι γνωστές μόνο οι προβολές του.

Ένα άλλο σημαντικό πρόβλημα της υπολογιστικής όρασης είναι η εξαγωγή τρισδιάστατης πληροφορίας για μία σκηνή, με τη χρήση περισσότερων αισθητηρίων λήψης εικόνων. Στο κεφάλαιο 6 προτείνεται μία αρχιτεκτονική υλικού η οποία έχει τη δυνατότητα πολύ γρήγορης ανάκτησης της πληροφορίας βάθους από ένα στερεοσκοπικό σύστημα.

Σε αυτό το σημείο θα ήθελα να ευχαριστήσω όλους όσους με βοήθησαν κατά τη διάρκεια της παρούσας Διδακτορικής Διατριβής, κατά τα τελευταία χρόνια. Ιδιαίτερα θα ήθελα να ευχαριστήσω τον τον πρόεδρο της τριμελούς συμβουλευτικής επιτροπής, κ. Ιωάννη Ανδρεάδη για την πολύτιμη βοήθεια και υποστήριξή του, τόσο σε επιστημονικό, όσο και σε προσωπικό επίπεδο. Η συνεισφορά του στη διατριβή υπήρξε αποφασιστική, ενώ τα σχόλια και η εποικοδομητική κριτική του στο κείμενο της διατριβής βοήθησαν καθοριστικά στην ποιότητά της. Πάνω απ' όλα όμως τον ευχαριστώ για την ελευθερία που μου άφησε στην ερευνητική κατεύθυνση της διατριβής και στην ανάπτυξη και αξιοποίηση νέων ιδεών.

Περιεχόμενα

Κατάλογος Συμβόλων

I_{xn}, I_{yn}	n-οστή μερική παράγωγος σήματος ως προς x και y, αντίστοιχα 138

Κεφάλαιο 1

Εισαγωγή

Ανάλυση εικόνας ονομάζεται μια διεργασία εξαγωγής ενός συνόλου περιγραφέων ή χαρακτηριστικών στα οποία βασίζεται μία απόφαση σχετικά με αντικείμενα.

The process of generating a set of descriptors or features on which a decision about objects in an image is based. [Online, 2007]

1.1 Ανάλυση εικόνων

Η υπολογιστική όραση, ως επιστημονικό πεδίο, ασχολείται με τη θεωρία και την τεχνολογία ανάπτυξης τεχνητών συστημάτων τα οποία εξάγουν πληροφορία μέσω εικόνων ή δεδομένων πολλών διαστάσεων. Η πληροφορία, όπως ορίστηκε από τον *Shannon*, είναι αυτό που επιτρέπει απόφαση. Από τη στιγμή που η αντίληψη μπορεί να θεωρηθεί ως η εξαγωγή πληροφορίας από αισθητήρια σήματα, η υπολογιστική όραση μπορεί να θεωρηθεί ως η επιστημονική διερεύνηση τεχνητών συστημάτων για την αντίληψη μέσω εικόνων.

Ως τεχνολογικό πεδίο, η υπολογιστική όραση επιδιώκει την εφαρμογή θεωριών και προτύπων για την κατασκευή συστημάτων υπολογιστικής όρασης. Παραδείγματα τέτοιων εφαρμογών είναι:

- Έλεγχος διεργασιών
- Ανίχνευση συμβάντων

- Οργάνωση πληροφορίας
- Μοντελοποίηση αντικειμένων ή περιβάλλοντος

Στην επιστήμη των υπολογιστών, η επεξεργασία πραγματικού χρόνου είναι η μελέτη συστημάτων υλικού και λογισμικού τα οποία υπόκεινται σε έναν περιορισμό 'πραγματικού χρόνου'. Αντίθετα, ένα σύστημα μη πραγματικού χρόνου είναι αυτό το οποίο δεν έχει δεδομένο χρονικό περιορισμό, ακόμα και αν η υψηλή απόδοση και η ταχεία απόκριση είναι επιθυμητά. Όταν ένα τέτοιο σύστημα σχετίζεται με την υπολογιστική όραση, πολλές φορές ο περιορισμός αυτός σημαίνει την επεξεργασία μίας εικόνας σε χρονικό πλαίσιο μικρότερο από το χρόνο μεταξύ δύο αλλεπάλληλων πλαισίων.

Ένα πολύ απλό παράδειγμα είναι η συμπίεση των πλαισίων σε μία εικονοσειρά, η αποστολή τους σε κάποιον παραλήπτη και η μετέπειτα αποσυμπίεση. Στην περίπτωση που η συμπίεση αυτή απαιτεί μεγαλύτερο χρόνο από το ρυθμό πλαισίων, το σύστημα θα καταρρεύσει.

Στο μεγαλύτερο μέρος της εργασίας αυτής μελετώνται τεχνικές που σχετίζονται με τις ροπές εικόνων. Αυτές αποτελούν ευρέως χρησιμοποιούμενους τοπικούς περιγραφείς εικόνων. Παραδείγματα εφαρμογών τους βρίσκονται στην πλοήγηση κινητών ρομπότ, στην κανονικοποίηση εικόνων, στην αναγνώριση προτύπων και στην υδατογράφηση δεδομένων.

Μία επισκόπηση των διαφόρων τύπων ροπών, μαζί με τους ορισμούς και τις βασικές τους ιδιότητες παρουσιάζεται στο κεφάλαιο αυτό. Μελετώνται τα πλεονεκτήματα και τα μειονεκτήματα κάθε τύπου, καθώς και οι ανάγκες που οδήγησαν στην ανάπτυξή τους.

1.2 Ροπές εικόνων

Γενικά, κάθε ποσότητα που προκύπτει από μία σχέση της μορφής:

$$M_p = \int_{\alpha}^{b} P_p(x) f(x) dx \tag{1.1}$$

μπορεί να θεωρηθεί ως ροπή τάξης p του σήματος f, ορισμένου στο διάστημα $[\alpha, b]$. Για την πιο απλή περίπτωση, για $P_p(x) = x^p$ (γεωμετρικές ροπές), και για $p = 1$ (ροπή πρώτου βαθμού), παρουσιάζονται στα Σχήματα 1.1, 1.2 και 1.3 οι γραφικές παραστάσεις για ένα παράδειγμα σήματος μήκους 128 δειγμάτων, την συνάρτηση βάσης x και τον ολοκληρωτέο όρο $f(x)x$. Η ροπή θα είναι, προφανώς, το ολοκλήρωμα της τελευταίας ποσότητας.

Η οικογένεια συναρτήσεων $P_i(x)$ αποτελεί και τον πυρήνα (*kernel*) της ροπής. Για σήματα δύο διαστάσεων, η παραπάνω σχέση γίνεται:

$$M_{p,q} = \int_{\alpha}^{b}\int_{c}^{d} P_{p,q}(x,y)f(x,y)dxdy \tag{1.2}$$

Οι διάφοροι χρησιμοποιούμενοι τύποι ροπών διαφοροποιούνται ως προς τον πυρήνα τους. Στον Πίνακα 1.1 παρουσιάζονται μερικοί από τους βασικούς.

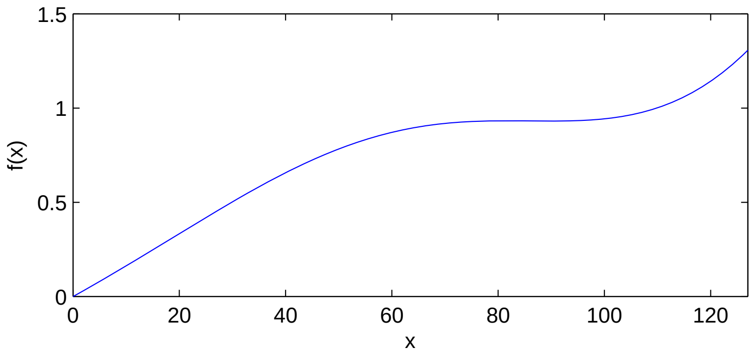
1.3 Γεωμετρικές ροπές και σταθερές του *Hu*

Χαρακτηριστικά ροπών ονομάζουμε στατιστικές εικόνων οι οποίες είναι ανεξάρτητες της περιστροφής, της μετακίνησης και της κλιμάκωσης. Είναι αμφίδρομα ορισμένες από τις εικόνες. Αυτές οι ιδιότητες δίνουν την δυνατότητα της οπτικής ανάκτησης πληροφορίας ανεξαρτήτως μεγέθους, θέσης και προσανατολισμού. Εφαρμογές των γεωμετρικών ροπών, καθώς και των σταθερών του Hu, εμφανίζονται στη βιβλιογραφία τα τελευταία 30 χρόνια, και συμπεριλαμβάνουν ταυτοποίηση σκηνών, αναγνώριση χαρακτήρων, αναγνώριση υφής, ανάκτηση εικόνων κ.α. [Dudani et al., 1977, Gonzalez and Woods, 1987, Prokop and Reeves, 1992]. Τα χαρακτηριστικά που ορίστηκαν από τον Hu [Hu, 1962] εξάγονται από τους ορισμούς των ροπών, των κεντρικών ροπών και των κανονικοποιημένων κεντρικών ροπών. Οι ορισμοί αυτοί δίνονται παρακάτω.

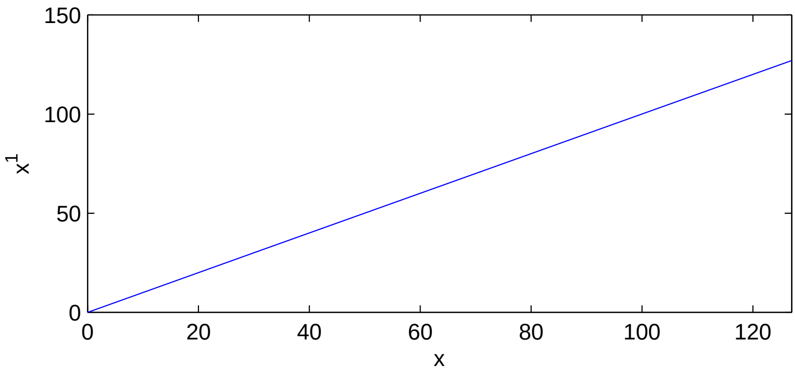
Έστω f μία συνεχής συνάρτηση ορισμένη στο χώρο R^2 . Η γεω-

Τύπος ροπής	$P_{p,q}(x,y)$
Γεωμετρικές ροπές	$x^p y^q$
Μιγαδικές ροπές	$(x+yi)^p(x-yi)^q$
Ροπές Zernike	$R_p(r)e^{-iq\theta}$
Ροπές Legendre	$L_p(x)L_q(y)$
Ροπές Fourier	$F_p(r)e^{-iq\theta}$

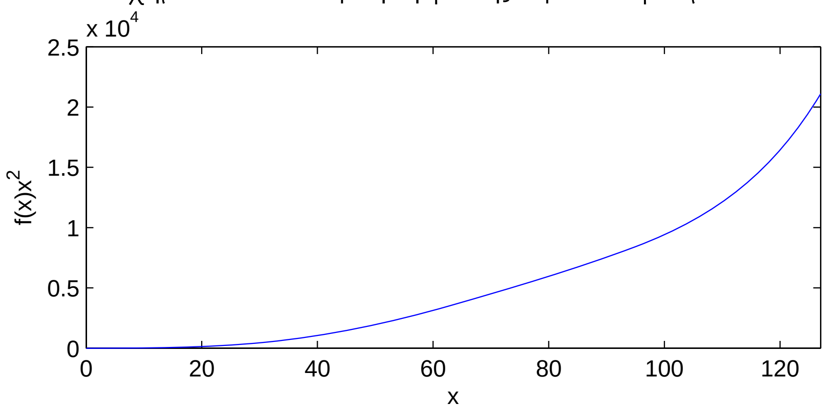
Πίνακας 1.1: Τύποι ροπών εικόνων



Σχήμα 1.1: Παράδειγμα σήματος $f(x)$



Σχήμα 1.2: Συνάρτηση βάσης πρώτου βαθμού x



Σχήμα 1.3: Ολοκληρωτέος όρος $f(x)x$

μετρική ροπή τάξης (p,q) της f ορίζεται ως:

$$m_{p,q} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p y^q f(x,y) dxdy \tag{1.3}$$

όπου $p, q \in \{0, 1, \ldots\}$.

Έχει δειχθεί ότι αν η f είναι μία σημειακά συνεχής συνάρτηση[1], τότε υπάρχουν ροπές κάθε τάξης. Οι κεντρικές ροπές ορίζονται ως εξής:

$$m_{p,q} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \bar{x})^p (y - \bar{y})^q dxdy \tag{1.4}$$

με $\bar{x} = \frac{m_{10}}{m_{00}}, \bar{y} = \frac{m_{01}}{m_{00}}$ το σημείο που ονομάζεται κέντρο (centroid) της εικόνας, και είναι αντίστοιχο του κέντρου βάρους ενός αντικειμένου.

Έστω $\alpha \in R^X$, με $X \subset Z^2$ πίνακα $m \times n$ στοιχείων. Το διακριτό ανάλογο της κεντρικής ροπής τάξης (p,q) δίνεται από την παρακάτω σχέση:

$$\mu_{p,q} = \sum_{x}^{n} \sum_{y}^{m} (x - \bar{x})^p (y - \bar{y})^q \alpha(x,y) \tag{1.5}$$

Η κανονικοποιημένη κεντρική ροπή, $\eta_{p,q}$, ορίζεται ως:

$$\eta_{p,q} = \frac{\mu_{pq}}{\mu_{00}^{\gamma}} \tag{1.6}$$

με συντελεστή κανονικοποίησης

$$\gamma = \frac{p+q}{2} + 1 \tag{1.7}$$

Ο *Hu* πρότεινε επτά ποσότητες οι οποίες, στη συνεχή περίπτωση, είναι ανεξάρτητες της περιστροφής, της μετακίνησης και του μεγέθους μίας εικόνας. Στη διακριτή περίπτωση παρουσιάζονται σφάλματα λόγω της ψηφιοποίησης.

$$\begin{aligned}
\varphi_1 &= \mu_{20} + \mu_{02} \\
\varphi_2 &= (\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2 \\
\varphi_3 &= (\mu_{30} - 3\mu_{12})^2 + (3\mu_{12} - \mu_{03})^2
\end{aligned}$$

[1] Μία συνάρτηση λέγεται σημειακά συνεχής, όταν είναι συνεχής σε ξένα σύνολα του πεδίου ορισμού της. π.χ. η βηματική συνάρτηση, που παρουσιάζει ασυνέχεια για $x = 0$, είναι σημειακά συνεχής.

$$
\begin{aligned}
\varphi_4 &= (\mu_{30}+\mu_{12})^2+(\mu_{21}-\mu_{03})^2 \\
\varphi_5 &= (\mu_{30}-3\mu_{12})(\mu_{30}+\mu_{12})\left((\mu_{30}+\mu_{12})^2-3(\mu_{21}+\mu_{03})^2\right)+ \\
&\quad (3\mu_{21}-\mu_{03})(\mu_{21}+\mu_{03})\left(3(\mu_{30}+\mu_{12})^2-(\mu_{21}+\mu_{03})^2\right) \\
\varphi_6 &= (\mu_{20}-\mu_{02})\left((\mu_{30}+\mu_{12})^2-(\mu_{21}+\mu_{03})^2\right)+ \\
&\quad 4\mu_{11}(\mu_{30}+\mu_{12})(\mu_{21}+\mu_{03}) \\
\varphi_7 &= (3\mu_{21}-\mu_{03})(\mu_{30}+\mu_{12})\left((\mu_{30}+\mu_{12})^2-3(\mu_{21}+\mu_{03})^2\right)- \\
&\quad (\mu_{30}-3\mu_{12})(\mu_{21}+\mu_{03})\left(3(\mu_{30}+\mu_{12})^2-(\mu_{21}+\mu_{03})^2\right) \qquad (1.8)
\end{aligned}
$$

1.3.1 Γεωμετρικά χαρακτηριστικά από τις γεωμετρικές ροπές

Από τις πρώτες γεωμετρικές ροπές μπορούν να προκύψουν μερικά βασικά γεωμετρικά χαρακτηριστικά για την υπό επεξεργασία εικόνα. Αυτή η δυνατότητα τις έχει κάνει πολύ δημοφιλείς σε διάφορες εφαρμογές, οι οποίες δεν εξαντλούνται στην υπολογιστική όραση. Παρακάτω δίνονται τα βασικά από αυτά, σε σχέση με τις συνεχείς ροπές. Το διακριτό ανάλογο προκύπτει με την αντικατάσταση των ολοκληρωμάτων από αθροίσματα.

1.3.2 Εμβαδόν σήματος

Είναι προφανές ότι για ένα σήμα δύο διαστάσεων $f : R^2 \to R$, το εμβαδό του δίνεται από την γεωμετρική ροπή μηδενικού βαθμού:

$$
\begin{aligned}
V &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)dxdy \\
&= m_{0,0} \qquad (1.9)
\end{aligned}
$$

Η μηδενική γεωμετρική ροπή θα δώσει και το μήκος ενός μονοδιάστατου σήματος καθώς και τον όγκο ενός τρισδιάστατου σήματος.

Κέντρο βάρους

Έστω μονοδιάστατο σήμα $f : R \to R$. Το κέντρο βάρους του δίνεται από τη σχέση:

$$\bar{x} = \frac{\int_{-\infty}^{\infty} f(x)xdx}{\int f(x)dx} \tag{1.10}$$

$$= \frac{m_1}{m_0} \tag{1.11}$$

Ομοίως, για σήμα δύο διαστάσεων $f : R^2 \to R$, ισχύει:

$$\bar{x} = \frac{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)xdxdy}{\int\int f(x,y)dxdy}$$

$$= \frac{m_{0,1}}{m_0}$$

$$\bar{y} = \frac{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)ydxdy}{\int\int f(x,y)dxdy}$$

$$= \frac{m_{0,1}}{m_0} \tag{1.12}$$

Γενικεύοντας τα παραπάνω, για N-διάστατο σήμα $f : R^N \to R$, με κέντρο βάρους $\bar{X} = [\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n]^T$ ισχύει:

$$\bar{x}_i = \frac{\int_{R^N} f(X) x_i \prod_{k=0}^{N} dx_i}{\int_{R^N} f(X) \prod_{k=0}^{N} dx_i}$$

$$= \frac{m_{0,\ldots,1,\ldots,0}}{m_{0,\ldots,0}}$$

$$\bar{X} = \begin{bmatrix} \frac{m_{1,\ldots,0}}{m_{0,\ldots,0}} \\ \frac{m_{0,1,\ldots,0}}{m_{0,\ldots,0}} \\ \vdots \\ \frac{m_{0,\ldots,1}}{m_{0,\ldots,0}} \end{bmatrix} \tag{1.13}$$

Κύριος άξονας

Η γωνία του κύριου άξονα ενός σήματος δύο διαστάσεων μπορεί να εξαχθεί βάσει των κεντρικών ροπών:

$$\theta = \frac{1}{2}\arctan\frac{\mu_{02} - \mu_{20}}{2\mu_{11}} + \frac{\pi}{4} sign\mu_{11} + \pi n, n = 0, 1 \tag{1.14}$$

1.3.3 Θόρυβος και γεωμετρικές ροπές

Οι γεωμετρικές ροπές επηρεάζονται ιδιαίτερα από την παρουσία θορύβου, λόγω του μεγάλου βάρους των ακραίων σημείων του σήματος. Παρακάτω, δίνεται ένα παράδειγμα για μονοδιάστατο σήμα.

Παράδειγμα

Έστω σήμα $f(x) = sin(\pi x/16)$ ορισμένο στα σημεία $\{0, \ldots, 16\}$ (Σχήμα 1.4). Η ροπή 4ης τάξης δίνεται από τη σχέση:

$$M_4 = \sum_{x=0}^{16} sin(\pi x/16)x^4 \tag{1.15}$$

Πραγματοποιώντας την άθροιση προκύπτει η τιμή 91.355.

Έστω ότι στο σήμα εμφανίζεται κρουστικός θόρυβος, στη θέση 16, μεταβάλλοντας την τιμή στη θέση αυτή σε 1. Η ροπή 4ης τάξης θα γίνει τώρα:

$$M_4 = \sum_{x=0}^{16} sin(\pi x/16)x^4 + 1 \times 16^4 \tag{1.16}$$

Δηλαδή, 156.890, που αντιστοιχεί σε αύξηση κατά 71%.

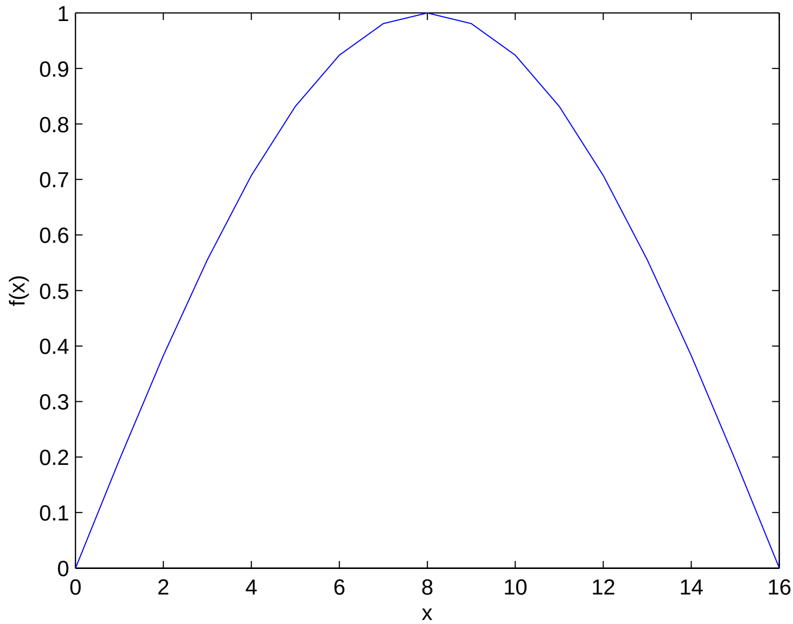
Η επίδραση του σφάλματος στις τιμές ροπών αυξάνει εκθετικά με την τάξη της ροπής. Επίσης, ο κρουστικός θόρυβος επηρεάζει πολύ περισσότερο τις ροπές από ότι ο τυχαίος, αφού ο τελευταίος μεταβάλλει γενικά λιγότερο τις τιμές του σήματος.

1.4 Ροπές *Legendre*

Οι ροπές *Legendre* ανήκουν στης κατηγορία των ροπών ορθογώνιων πολυωνύμων. Η γενική τους μορφή δίνεται από την παρακάτω σχέση:

$$Z_{p,q} = \int_{-1}^{1}\int_{-1}^{1} L_p(x)L_q(y)f(x,y)dxdy \tag{1.17}$$

με $L_p : R \to R$ πολυώνυμο Zernike βαθμού p.



Σχήμα 1.4: $sin(\pi x/16)$, για $x = \{0, 1, \ldots, 16\}$

1.4.1 Πολυώνυμα *Legendre*

Τα πολυώνυμα Legendre, επίσης αποκαλούμενα και συναρτήσεις Legendre πρώτου είδους, συντελεστές Legendre ή και αρμονικές ζώνης Lagrange [Legendre, 1785], αποτελούν τις λύσεις της διαφορικής εξίσωσης Legendre. Ορίζονται ως το επικαμπύλιο ολοκλήρωμα:

$$L_n(z) = \oint_C \frac{t^{-n-1}}{\sqrt{(1 - tz - z^2)}} dt \tag{1.18}$$

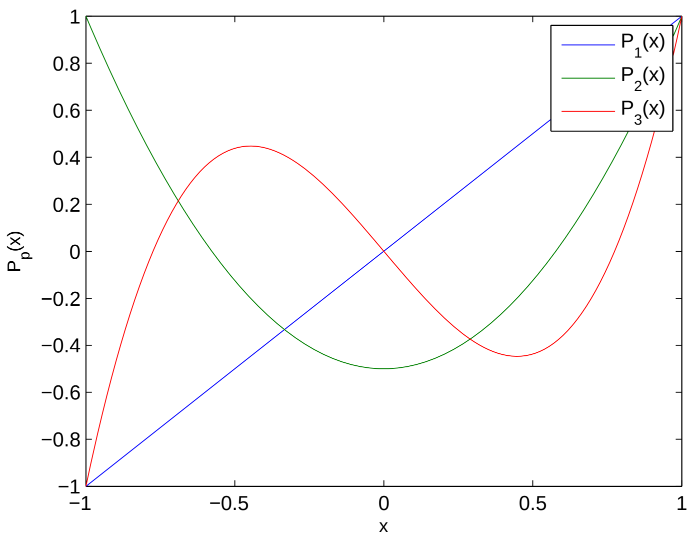
Με C καμπύλη που περικλύει την αρχή του συστήματος συντεταγμένων αριστερόστροφα.

Τα πρώτα λίγα πολυώνυμα δίνονται παρακάτω:

$$\begin{aligned} L_0(x) &= 1 \\ L_1(x) &= x \\ L_2(x) &= \frac{1}{2}(3x^2 - 1) \end{aligned}$$

$$
\begin{aligned}
L_3(x) &= \frac{1}{2}(5x^3 - 3x) \\
L_4(x) &= \frac{1}{8}(35x^4 - 30x + 3) \\
L_5(x) &= \frac{1}{8}(63x^5 - 70x^3 + 15x) \\
L_6(x) &= \frac{1}{16}(231x^6 - 315x^4 + 105x^2 - 5) \qquad (1.19)
\end{aligned}
$$

Αξίζει να σημειωθεί ότι τα πολυώνυμα Legendreείναι εύκολο να εξαχθούν με τη μέθοδο *Gram – Schmidt* για $w(x) = 1$ [Gautschi, 2004]. Στο Σχήμα 1.5 παρουσιάζονται οι γραφικές παραστάσεις για αυτά.



Σχήμα 1.5: Συναρτήσεις *Legendre*

1.5 Ροπές *Zernike*

Οι ροπές *Zernike* ανήκουν στης κατηγορία των ακτινικών ορθογώνιων μιγαδικών ροπών. Η γενική τους μορφή δίνεται από την παρακάτω σχέση:

$$Z_{p,q} = \int_0^1 \int_0^{2\pi} R_{p,q}(r) e^{-iq\theta} f(x,y) r d\theta dr \qquad (1.20)$$

με $R_{p,q} : R \rightarrow R$ πολυώνυμο Zernike βαθμού p.

1.5.1 Πολυώνυμα *Zernike*

Τα πολυώνυμα *Zernike* [Zernike, 1934] προκύπτουν από την ανάπτυξη της συνάρτησης κυματομετώπου (*wavefront*) για οπτικά συστήματα με κυκλικές κόρες [Bezdidko, 1980]. Τα άρτια και περιττά πολυώνυμα *Zernike* δίνονται ως:

$$\begin{aligned} {}^{o}U_n^m(r,\theta) &= R_n^m(r)\sin(m\theta) \\ {}^{e}U_n^m(r,\theta) &= R_n^m(r)\cos(m\theta) \end{aligned} \qquad (1.21)$$

Η ακτινική συνάρτηση R_n^m ορίζεται για ακεραίους m και n με $m \geq n \geq 0$:

$$R_n^m = \begin{cases} \sum_{l=0}^{\frac{m-n}{2}} \frac{(-1)^l (n-l)!}{l!(\frac{1}{2}(n+m)-l)!(\frac{1}{2}(n-m)-l)!} & (m-n) \text{άρτιος} \\ 0 & (m-n) \text{περιττός} \end{cases} \qquad (1.22)$$

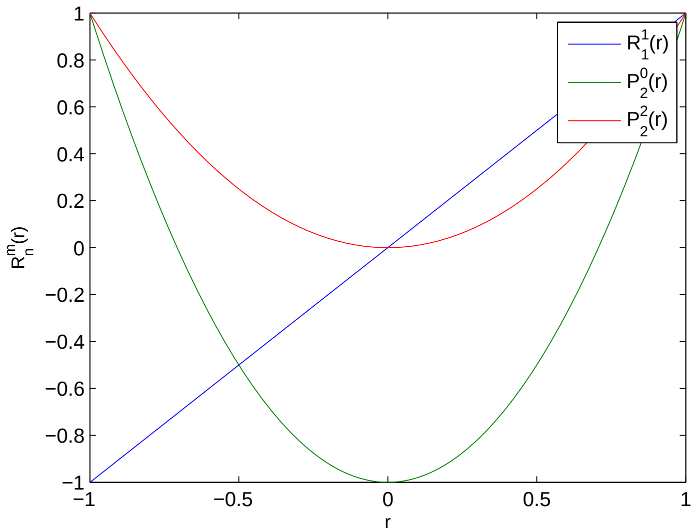
Μία άλλη κλειστή μορφή των πολυωνύμων δίνεται ως:

$$R_n^m(r) = \frac{\Gamma(n+1)\,{}_2F_1(-\frac{1}{2}(m+n), \frac{1}{2}(m-n); -n; r^{-2})}{\Gamma(\frac{1}{2}(2+n-m))\Gamma(\frac{1}{2}(2+n+m))} r^n \qquad (1.23)$$

Με $n-m$ άρτιο και $m \neq n$, $\Gamma(z)$ η συνάρτηση Γάμμα και ${}_2F_1(\alpha,b;c;z)$ υπεργεωμετρική συνάρτηση.

Τα πρώτα ακτινικά πολυώνυμα *Zernike* δίνονται παρακάτω:

$$\begin{aligned} R_0^0(r) &= 1 \\ R_1^1(r) &= r \\ R_2^0(r) &= 2r^2 - 1 \\ R_2^2(r) &= r^2 \\ R_3^1(r) &= 3r^3 - r \\ R_3^3(r) &= r^3 \\ R_4^0(r) &= 6r^4 - 6r^2 + 1 \\ R_4^2(r) &= 4r^4 - 3r^2 \\ R_4^4(r) &= r^4 \end{aligned} \qquad (1.24)$$



Σχήμα 1.6: Ακτινικές συναρτήσεις *Zernike*

Στο Σχήμα 1.6 παρουσιάζονται οι γραφικές παραστάσεις για τα τρία πρώτα.

Τα πολυώνυμα Zernike ικανοποιούν τη συνθήκη ορθογωνιότητας, δηλαδή

$$\int_0^1 R_n^m(r) R_{n'}^m(r) r dr = \delta(n' - n) c_{m,n} \tag{1.25}$$

με

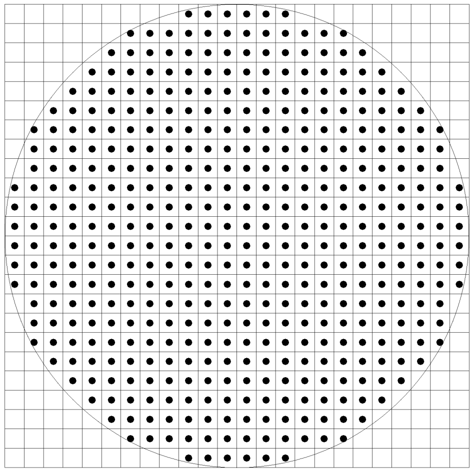
$$c_{m,n} = \frac{1}{2(n+1)} R_n^m \tag{1.26}$$

και

$$\delta(n' - n) = \begin{cases} 0 & , \quad n' \neq n \\ 1 & , \quad n' = n \end{cases} \tag{1.27}$$

Είναι εύκολο να δειχθεί ότι το μέτρο των ροπών *Zernike* δε μεταβάλλεται με την περιστροφή:

$$\left| Z_{p,q}^{\varphi} \right| \quad = \quad \left| \int_0^1 \int_0^{2\pi} R_{p,q}(r) e^{-iq(\theta+\varphi)} f(x,y) d\theta dr \right|$$



Σχήμα 1.7: Προβολή εικόνας στο μοναδιαίο δίσκο

$$
\begin{aligned}
&= \left| \int_0^1 \int_0^{2\pi} R_{p,q}(r) e^{-iq\theta} f(x,y) d\theta d r e^{-iq\varphi} \right| \\
&= \left| \int_0^1 \int_0^{2\pi} R_{p,q}(r) e^{-iq\theta} f(x,y) d\theta d r \right| \\
&= \left| Z_{p,q} \right| \qquad (1.28)
\end{aligned}
$$

Οι ροπές *Zernike* ορίζονται για συναρτήσεις στο μοναδιαίο δίσκο. Επομένως, πριν τον υπολογισμό τους, είναι αναγκαία η προβολή της εικόνας σε αυτόν, γεγονός που οδηγεί σε σφάλματα προσεγγίσεων (Σχήμα 1.7).

1.6 Ροπές *Chebychev*

Οι ροπές Chebychev προτάθηκαν σχετικά πρόσφατα ως περιγραφείς εικόνων παρουσιάζοντας ορισμένα πλεονεκτήματα σε σχέση με τις ροπές με συνεχή πυρήνα [Mukundan et al., 2001]. Τα διακριτά πολυώνυμα Chebychev είναι ορθογώνια σε διάστημα $[0, \ldots, N]$, με συνάρτηση βάρους $w(x) = 1$. Η ορθογωνιότητά τους στο πεδίο ορισμού της ψηφιακής εικόνας αποτελεί και το βασικότερο πλεονέκτημά τους, αφού τις απαλλάσσει από σφάλματα που προκύπτουν από χωρικούς μετασχηματισμούς. Η συνθήκη ορθογωνιότητας σε διακριτά πολυώνυμα ορίζεται ως:

$$\sum_{x=0}^{N-1} P_m(x)P_n(x)w(x) = \delta(m-n)\rho(n,N) \tag{1.29}$$

με

$$\rho(n,N) = \sum_{x=0}^{N-1} P_n^2(x)w(x) \tag{1.30}$$

Για κάθε φραγμένη συνάρτηση $f(x,y), 0 \leq x,y \leq N-1$, ισχύει:

$$f(x,y) = \sum_{m=0}^{N-1}\sum_{n=0}^{N-1} T_{m,n}P_m(x)P_n(y)w(x)w(y) \tag{1.31}$$

όπου

$$T_{m,n} = \frac{1}{\rho(m,N)\rho(n,N)}\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} f(x,y)P_m(x)P_n(y) \tag{1.32}$$

Η ποσότητα $T_{m,n}$ ορίζεται ως μία ροπή τάξης (m,n). Η πιο φυσική επιλογή για την οικογένεια διακριτών ορθογωνίων πολυωνύμων είναι τα πολυώνυμα Chebychev, αφού είναι ορθογώνια ως προς συνάρτηση βάρους $w(x) = 1$. Ορίζονται ως:

$$t_n(x) = (1-N)_n {}_3F_2\begin{pmatrix}-n,-x,1+n\\ 1,1-N,1\end{pmatrix} \tag{1.33}$$

με $n,x,y \in [0,N-1]$ και $(\alpha)_k$ το αύξων παραγωντικό πολυώνυμο:

$$(\alpha)_k = \alpha(\alpha+1)(\alpha+2)\ldots(\alpha+k-1) \tag{1.34}$$

Η ${}_3F_2$ είναι η γενικευμένη υπεργεωμετρική συνάρτηση:

$${}_3F_2\begin{pmatrix}\alpha_1,\alpha_2,\alpha_3\\ b_1,b_2,z\end{pmatrix} = \sum_{k=0}^{\infty}\frac{(\alpha_1)_k(\alpha_2)_k(\alpha_3)_k}{(b_1)_k(b_2)_k}\frac{z^k}{k!} \tag{1.35}$$

Βάσει των παραπάνω ορισμών, τα πολυώνυμα Chebychev μπορούν να οριστούν και ως εξής:

$$t_n(x) = n!\sum_{k=0}^{n}(-1)^{n-k}\binom{N-1-k}{n-k}\binom{n+k}{n}\binom{x}{k} \tag{1.36}$$

Το $\rho(n,N)$ της σχέσης 1.29 στην περίπτωση των πολυωνύμων Chebychev ισούται με:

$$\rho(n,N) = (2n)!\binom{N+n}{2n+1} \quad (1.37)$$

με $n = 0, 1, \ldots, N-1$. Όπως όλα τα ορθογώνια πολυώνυμα [Gautschi, 2004], τα πολυώνυμα Chebychev ικανοποιούν μία αναδρομική σχέση τριών όρων. Πιο συγκεκριμένα:

$$(n+1)t_{n+1}(x) - (2n+1)(2x-N+1)t_n(x) + n(N^2-n^2)t_{n-1}(x) = 0 \quad (1.38)$$

Αξίζει να σημειωθεί ότι στην παραπάνω μορφή τους τα πολυώνυμα αυτά δεν προσφέρονται για την εξαγωγή ροπών. Η τιμή του $t_n(x)$ είναι ανάλογη του N^n και η τιμή των ροπών, όπως ορίζονται στην εξίσωση 1.32 ανάλογη του N^{-p-q}. Για το λόγο αυτό, ορίστηκαν οι κανονικοποιημένες ροπές Chebychev, οι οποίες περιγράφονται στην επόμενη παράγραφο.

1.6.1 Κανονικοποιημένες ροπές *Chebyshev*

Τα κανονικοποιημένα πολυώνυμα *Chebyshev* ορίζονται [Mukundan et al., 2001]:

$$\tilde{t}_n(x) = \frac{t_n(x)}{\beta(n,N)} \quad (1.39)$$

με $t_n(x)$ το διακριτό πολυώνυμο *Chebychev* βαθμού n, όπως αυτό δίνεται στη σχέση 1.33 και $\beta(n,N)$ μία σταθερά κανονικοποίησης. Υπό τον παραπάνω μετασχηματισμό, η τετραγωνική νόρμα των πολυωνύμων γίνεται:

$$\tilde{\rho}(n,N) = \frac{\rho(n,N)}{\beta(n,N)^2} \quad (1.40)$$

Οι ροπές Chebychev μπορούν τώρα να οριστούν:

$$T_{p,q} = \frac{1}{\tilde{\rho}(p,N)\tilde{\rho}(q,N)} \sum_{x=0}^{N-1}\sum_{y=0}^{N-1} \tilde{t}_p(x)\tilde{t}_q(y) f(x,y) \quad (1.41)$$

με $p, q = 0, 1, 2, \ldots, N-1$.

Η παραπάνω σχέση οδηγεί και σε αντίστοιχο αντίστροφο μετασχηματισμό:

$$f(x,y) = \sum_{m=0}^{N-1}\sum_{n=0}^{N-1} T_{mn}\tilde{t}_m(x)\tilde{t}_n(y) \quad (1.42)$$

Η τελευταία σχέση περιγράφει και τη δυνατότητα ανακατασκευής μίας εικόνας από τις ροπές της.

1.6.2 Ιδιότητες των ροπών *Chebyshev*

Συμμετρία

Η ιδιότητα της συμμετρίας μπορεί να χρησιμοποιηθεί για τη σημαντική μείωση της υπολογιστικής πολυπλοκότητας του υπολογισμού των ροπών Chebychev. Τα κανονικοποιημένα πολυώνυμα Chebychev παρουσιάζουν την ίδια συμμετρία με τα κλασσικά:

$$\tilde{t}_n(N-1-x) = (-1)^n \tilde{t}_n(x) \tag{1.43}$$

Η παραπάνω σχέση επιτρέπει τη διαίρεση του χώρου μίας εικόνας $N \times N$, (N άρτιος) σε τέσσερα ίσα μέρη, και τον υπολογισμό των πολυωνύμων μόνο στο πρώτο τεταρτημόριο $0 \leq x,y \leq (N/2-1)$. Η έκφραση 1.33 μπορεί να τροποποιηθεί, τώρα, ως εξής:

$$\begin{aligned} T_{pq} = \quad & \frac{1}{\tilde{\rho}}(p,N)\tilde{\rho}(q,N) \sum_{x=0}^{N/2-1} \sum_{y=0}^{N/2-1} \tilde{t}_p(x)\tilde{t}_q(y) \\ & \Big(f(x,y) + (-1)^p f(N-1-x,y) + (-1)^q f(x,N-1-y) \\ & +(-1)^{p+q} f(N-1-x,N-1-y) \Big) \end{aligned} \tag{1.44}$$

Εκτός από τη μείωση του υπολογιστικού κόστους στο ένα τέταρτο, η ιδιότητα της συμμετρίας είναι επίσης χρήσιμη στην ελαχιστοποίηση της απαιτούμενης μνήμης για την αποθήκευση των πολυωνύμων. Αν μία εφαρμογή απαιτεί την αποθήκευση πολυωνύμων βαθμού μέχρι M, τότε ένας πίνακας διαστάσεων $M \times (N/2)$ είναι αρκετός. Εντούτοις, σε αυτή την περίπτωση, η εφαρμογή πρέπει να υλοποιήσει αυτήν την ιδιότητα, για τον υπολογισμό των $\tilde{t}_n(x)$ όταν $x \geq (N/2)$.

Ανάπτυξη σε πολυώνυμο

Το κανονικοποιημένο πολυώνυμο $\tilde{t}_n(x)$ μπορεί να εκφραστεί ως πολυώνυμο του x. Χρησιμοποιώντας τις σχέσεις 1.36 και 1.39, προκύπτει:

$$\tilde{t}_n(x) = \frac{1}{\beta(n,N)} \sum_{k=0}^{n} C_k(n,N) \sum_{i=0}^{k} s_k^{(i)} x^i \tag{1.45}$$

με

$$C_k(n,N) = (-1)^{n-k} \frac{n!}{k!} \binom{N-1-k}{n-k} \binom{n+k}{n} \tag{1.46}$$

και $s_k^{(i)}$ τους αριθμούς *Stirling* πρώτου είδους [Scheid, 1968], οι οποίοι ικανοποιούν τη σχέση

$$\frac{x!}{(x-k)!} = \sum_{i=0}^{k} s_k^i x^i \tag{1.47}$$

Αναπαράσταση μέσω Γεωμετρικών Ροπών

Η παραπάνω ανάπτυξη είναι χρήσιμη για την έκφραση των ροπών Chebychev με όρους γεωμετρικών ροπών. Αν οι γεωμετρτικές ροπές μίας εικόνας $f(x,y)$ εκφρασθούν χρησιμοποιώντας μία διακριτή προσέγγιση ως:

$$m_{p,q} = \sum_{x=0}^{N-1}\sum_{y=0}^{N-1} x^p y^q f(x,y) \tag{1.48}$$

τότε με τη χρήση της σχέσης 1.45 οι ροπές *Chebychev* εκφράζονται:

$$T_{p,q} = A_p A_q \sum_{k=0}^{p} C_k(p,N) \sum_{l=0}^{q} C_l(q,N) \sum_{i=0}^{k}\sum_{j=0}^{l} s_k^{(i)} s_l^{(j)} m)i,j \tag{1.49}$$

με

$$A_p = \beta(p,N)\tilde{\rho}(p,N) \tag{1.50}$$

Από τη σχέση 1.49 γίνεται αντιληπτό πως οι ροπές *Chebychev* εξαρτόνται από τις γεωμετρικές ίσου και μικρότερου βαθμού. Οι εκφράσεις για τις πρώτες ροπές *Chebychev* (για $\beta(n,N) = N^n$) δίνονται παρακάτω:

$$\begin{aligned}
T_{00} &= \frac{m_{00}}{N^2} \\
T_{10} &= \frac{6m_{10} + 3(1-N)m_{00}}{N(N^2-1)} \\
T_{01} &= \frac{6m_{01} + 3(1-N)m_{00}}{N(N^2-1)} \\
T_{20} &= \frac{30m_{20} + 30(1-N)m_{10} + 5(1-N)(2-N)m_{00}}{(N^2-1)(N^2-2^2)} \\
T_{02} &= \frac{30m_{02} + 30(1-N)m_{01} + 5(1-N)(2-N)m_{00}}{(N^2-1)(N^2-2^2)} \\
T_{11} &= \frac{36m_{11} + 18(1-N)(m_{10}+m_{01}) + 9(1-N)^2 m_{00}}{(N^2-1)^2}
\end{aligned} \tag{1.51}$$

Αναδρομική Σχέση ως προς x

Τα πολυώνυμα $\tilde{t}_n(x)$ ικανοποιούν την παρακάτω αναδρομική σχέση ως προς την μεταβλητή x:

$$\begin{aligned} x(N-x)\tilde{t}_n(x) &= (-n(n+1)-(2x-1)(x-N-1)-x)\tilde{t}_n(x-1)+ \\ &\quad ((x-1)(x-N-1))\tilde{t}_n(x-2) \end{aligned} \tag{1.52}$$

Οι αρχικές τιμές της παραπάνω αναδρομής είναι:

$$\begin{aligned} \tilde{t}_n(0) &= \frac{(1-N)_n}{\beta(n,N)} \\ \tilde{t}_n(1) &= \tilde{t}_n(0)\left(1+\frac{n(n+1)}{1-N}\right) \end{aligned} \tag{1.53}$$

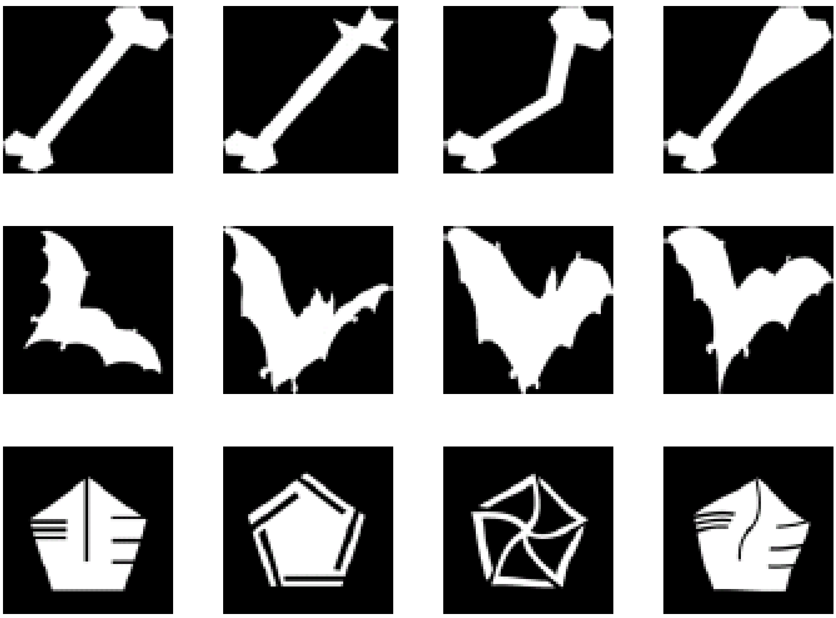
Η παραπάνω αναδρομική σχέση έχει εφαρμοστεί με επιτυχία στην μείωση του υπολογιστικού κόστους του υπολογισμού των τιμών των πολυωνύμων *Chebychev*.

1.7 Περιγραφέας *ART*

Ο Γωνιακός Ακτινικός Μετασχηματισμός (Angular Radial Transform (ART)) έχει καθοριστεί ως ο περιγραφέας σχήματος περιοχής στο πρότυπο $MPEG-7$ (Moving Picture Experts Group). Αποτελεί έναν συμπαγή περιγραφέα ο οποίος παρουσιάζει αμεταβλητότητα σε απλούς μετασχηματισμούς, παρουσιάζοντας ταυτόχρονα μικρό πλεονασμό πληροφορίας. Παραδείγματα εικόνων οι οποίες χρησιμοποιήθηκαν κατά το πρότυπο αυτό για την αποτίμηση περιγραφέων φαίνονται στο σχήμα 1.8

1.7.1 Εισαγωγή

Ο περιγραφέας *ART* μπορεί να θεωρηθεί ως ένα σύνολο ροπών με μιγαδικές, ορθογώνιες συναρτήσεις βάσης, ορισμένες στο μοναδιαίο κύκλο. Όπως και με ροπές άλλου είδους του τύπου αυτού [Teague, 1980] επιτρέπει την περιγραφή σχήματος με ελάχιστη ποσότητα πλεονάζουσας πληροφορίας. Επιπλέον, οι συντελεστές του *ART* είναι αμετάβλητοι στην περιστροφή, την κλιμάκωση και τη μετακίνηση. Ως ο περιγραφέας σχήματος του προτύπου $MPEG-7$, ορίζεται για δυαδικές εικόνες και αναπαριστάται από 140 δυαδικά ψηφία.



Σχήμα 1.8: Παραδείγματα σχημάτων στο πρότυπο $MPEG-7$

1.7.2 Περιγραφή

Ο γωνιακός ακτινικός μετασχηματισμός τάξης m,n μίας εικόνας $f(x,y)$ ορίζεται ως [Kim and Kim, 2001]:

$$F_{m,n} = \int\int V^{*}_{m,n}(x,y)f(x,y)dxdy \tag{1.54}$$

Η συνάρτηση βάσης $V_{m,n}(x,y)$ μπορεί να εκφραστεί σε πολικές συντεταγμένες ως:

$$V_{m,n}(r,\theta) = R_n(r)A_m(\theta) \tag{1.55}$$

με

$$A_m(\theta) = \frac{1}{2\pi}e^{jm\theta} \tag{1.56}$$

και

$$R_n(r) = \begin{cases} 1, & (n=0) \\ 2\cos(\pi nr), & (n>0) \end{cases} \tag{1.57}$$

Για την επίτευξη αμεταβλητότητας ως προς τη μετακίνηση, το κέντρο του πολικού συστήματος συντεταγμένων ορίζεται να είναι και το κέντρο βάρους του αντικειμένου, το οποίο μπορεί εύκολα να ανακτηθεί μέσω γεωμετρικών ροπών:

$$\begin{aligned} \bar{x} &= \frac{\int\int f(x,y)xdxdy}{\int\int f(x,y)dxdy} \\ \bar{y} &= \frac{\int\int f(x,y)ydxdy}{\int\int f(x,y)dxdy} \end{aligned} \tag{1.58}$$

Η αμεταβλητότητα ως προς την περιστροφή επιτυγχάνεται με τη χρήση του μέτρου των μιγαδικών συντελεστών *ART*.

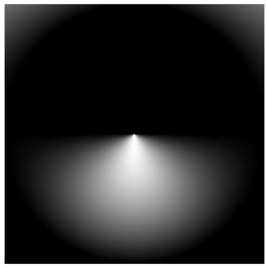
Κατά το $MPEG-7$, ο περιγραφέας *ART* εκφράζεται με 140 δυαδικά ψηφία, τα οποία αποτελούνται από 35 συντελεστές των τεσσάρων ψηφίων [Bober, 2001]. Επομένως, οι υπολογιζόμενοι συντελεστές είναι τάξης $(n<3, m<12)$, κανονικοποιημένοι με F_{00}.

1.8 Συγκριτική ανάλυση τύπων ροπών

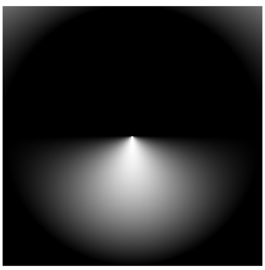
Στις προηγούμενες ενότητες παρουσιάστηκαν οι σημαντικότεροι τύποι ροπών:

- Γεωμετρικές
- *Legendre*

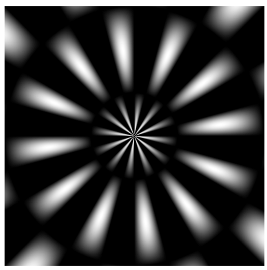
(α′) 1,0



(β′) 1,1



(γ′) 2,2



(δ′) 3,12

Σχήμα 1.9: Παραδείγματα συναρτήσεων βάσης του *ART*

- *Zernike*
- *Chebyshev*
- *ART*

Στην ενότητα αυτή συνοψίζονται οι ιδιότητές του και οι εφαρμογές στις οποίες γενικά απευθύνονται.

Γεωμετρικές Αποτελούν τον πρώτο τύπο ροπών που εμφανίστηκε στη βιβλιογραφία. Οι ροπές χαμηλής τάξης δίνουν τα βασικά χαρακτηριστικά ενός σχήματος, όπως το εμβαδόν του, το κέντρο μάζας του, τους βασικούς άξονες κ.α. Η υλοποίησή τους είναι πολύ απλή, με σχετικά μικρό υπολογιστικό κόστος.

Εντούτοις, είναι πολύ ευαίσθητες στο θόρυβο, παρουσιάζουν πολύ μεγάλο εύρος τιμών, γεγονός που συχνά οδηγεί σε αριθμητικές αστάθειες και δεν παρουσιάζουν αμεταβλητότητα σε απλούς μετασχηματισμούς. Επίσης, η ανακατασκευή της εικόνας από τις γεωμετρικές ροπές της δεν είναι εύκολη.

Legendre Αποτελούν την πιο απλή μορφή ροπών, με ορθογώνια βάση. Έτσι, είναι σχετικά απλές στην υλοποίηση, με μέτριο υπολογιστικό κόστος. Η ανοχή τους στο θόρυβο είναι πολύ καλύτερη από αυτή των γεωμετρικών ροπών, όπως και το εύρος τιμών τους, ενώ η ανακατασκευή της εικόνας από αυτές είναι εφικτή, αν και όχι ιδιαίτερα ακριβής.

Το βασικό τους μειονέκτημα είναι η μεταβολή τους μετά από απλούς γεωμετρικούς μετασχηματισμούς, η οποία μειώνει ουσιαστικά τη χρησιμότητά τους.

Zernike Αποτελούν τον πιο συχνά χρησιμοποιούμενο τύπο ροπών με πολικά ορθογώνια μιγαδικά πολυώνυμα ως βάση. Έτσι, παρουσιάζουν μικρό πλεονασμό πληροφορίας, είναι αμετάβλητες σε γραμμικούς μετασχηματισμούς και καθιστούν δυνατή την ανακατασκευή της εικόνας. Επιπλέον, συμπεριφέρονται ικανοποιητικά ως προς το θόρυβο και παρουσιάζουν μικρό εύρος τιμών.

Το μεγαλύτερο πρόβλημα αποτελεί η πολύπλοκη υλοποίησή τους και η μεγάλη υπολογιστική τους πολυπλοκότητα. Επίσης, λόγω του μετασχηματισμού συστήματος συντεταγμένων (καρτεσιανό$\leftarrow$πολικό), εισάγονται σφάλματα, πέρα από αυτά των υπολοίπων τύπων ροπών.

Τύπος ροπής	*Θ*	*Α*	*Π*	*Ι*	*Ε*
Γεωμετρικές ροπές	Μικρή	Δύσκολη	Μικρή	Όχι	Μεγάλο
Σταθερές *Hu*	Μικρή	Δύσκολη	Μικρή	Ναι	Μεγάλο
Ροπές *Legendre*	Καλή	Δυνατή	Μεσαία	Όχι	Μικρό
Ροπές *Zernike*	Καλή	Δυνατή	Μεγάλη	Ναι	Μικρό
Ροπές *Chebyshev*	Καλή	Δυνατή	Μεσαία	Όχι	Μικρό
ART	Μεσαία	Περιορισμένη	Μικρή	Ναι	Μικρό

Θ	Ανοχή στο θόρυβο
Α	Δυνατότητα ανακατασκευής
Π	Υπολογιστική πολυπλοκότητα
Ι	Αμεταβλητότητα
Ε	Εύρος τιμών

Πίνακας 1.2: Σύνοψη τύπων ροπών

Chebyshev Αποτελούν την πιο απλή μορφή ροπών με πολυώνυμα ορθογώνια στο διακριτό πεδίο της εικόνας. Επομένως, εμφανίζονται πολύ μικρά σφάλματα κατά την εξαγωγή τους. Επιπλέον, δίνουν πολύ καλά αποτελέσματα ανακατασκευής, και παρουσιάζουν μικρό εύρος τιμών.

Εντούτοις, δεν παρουσιάζουν αμεταβλητότητα στην περιστροφή, γεγονός που μειώνει τη χρησιμότητά τους, ή επιβάλλει την κανονικοποίηση της εικόνας πριν την εξαγωγή των ροπών.

ART Ο μετασχηματισμός *ART* δεν κατατάσσεται επίσημα ως ροπή, αλλά παρουσιάζει όλες τις ιδιότητες των ροπών, όπως αμεταβλητότητα σε απλούς μετασχηματισμούς και βάσεις ορθογωνίων συναρτήσεων. Είναι απλός στην υλοποίηση, με σχετικά μικρή υπολογιστική πολυπλοκότητα, και αμετάβλητος σε γραμμικούς μετασχηματισμούς.

Εντούτοις, ορίζεται μόνο σε δυαδικές εικόνες. Επίσης, χρησιμοποιούνται μόνο οι 36 πρώτοι περιγραφείς, όπως έχει οριστεί στο πρότυπο $MPEG - 7$.

1.8.1 Κριτήρια επιλογής

Προφανώς, κανένας από τους παραπάνω τύπους ροπών δεν είναι πλεονεκτικότερος για κάθε εφαρμογή. Παρόλο που ορισμένοι παρουσιάζουν εμφανή, πολύ σημαντικά μειονεκτήματα, ανάλογα με τις εκάστοτε απαιτήσεις θα πρέπει να γίνει κάποια διερεύνηση.

Γεωμετρικές Στην περίπτωση που επιθυμητά είναι μόνο τα βασικά γεωμετρικά χαρακτηριστικά ενός σχήματος (εμβαδόν, κέντρο βάρους, κύριος άξονας). Προφανές πλεονέκτημα είναι η μικρή υπολογιστική πολυπλοκότητα, μαζί με την απλή υλοποίηση σε λογισμικό. Η πρότερη επεξεργασία είναι απαραίτητη για την ελαχιστοποίηση του θορύβου.

Zernike Σε συστήματα τα οποία απαιτούν αμεταβλητότητα στην περιστροφή, τη μετακίνηση ή την κλιμάκωση. Μειονέκτημα είναι η σχετικά μεγάλη υπολογιστική πολυπλοκότητα, καθώς και η πολυπλοκότητα υλοποίησης.

Chebyshev Σε συστήματα στα οποία δεν απαιτείται αμεταβλητότητα, αλλά ακριβής ανακατασκευή.

Legendre Σε συστήματα στα οποία απαιτείται αμεταβλητότητα κατά την κλιμάκωση, εύκολη υλοποίηση και δυνατότητα ανακατασκευής.

ART Σε συστήματα που απαιτείται σχετικά εύκολη υλοποίηση και αμεταβλητότητα στην περιστροφή. Λόγω του συγκεκριμένου μήκους του περιγραφέα, δεν είναι δυνατή η αύξηση της ακρίβειας με τη χρήση περισσότερων μεταβλητών.

1.9 Αριθμητική σταθερότητα των ροπών

Μία σημαντική παρατήρηση σχετικά με την αξιοπιστία της χρήσης των διάφορων τύπων ροπών στην ανάλυση εικόνας είναι τα σφάλματα που παρουσιάζονται στον υπολογισμό τους. Αυτά οφείλονται στους παρακάτω παράγοντες:

1. Διακριτοποίηση του χώρου
2. Παρουσία θορύβου
3. Αριθμητικά σφάλματα λόγω περιορισμένης ακρίβειας αναπαράστασης

Στο κεφάλαιο 3 παρουσιάζεται αναλυτικά ο 1ος παράγοντας, μαζί με μία νέα τεχνική ελαχιστοποίησής του. Στο Α′ δίνονται αναλυτικά αριθμητικά στοιχεία για τους άλλους δύο παράγοντες.

Τύπος ροπής	Αριθμητική σταθερότητα	Εύρος τιμών
Γεωμετρικές ροπές	Ασταθείς	Πολύ μεγάλο
Σταθερές *Hu*	Ασταθείς	Μεγάλο
Ροπές *Legendre*	Σταθερές	Μικρό
Ροπές *Zernike*	Σταθερές	Μικρό
Ροπές *Chebyshev*	Σταθερές	Μικρό
ART	Σταθερές	Μικρό

Πίνακας 1.3: Αριθμητική σταθερότητα ροπών

1.10 Επίδραση του θορύβου στις ροπές

Σε προηγούμενη ενότητα δόθηκε μία γενική άποψη για το βαθμό της επίδρασης των ροπών. Για μία πιο εκτενή διερεύνηση απαιτείται μία ακριβής μαθηματική περιγραφή. Έστω μονοδιάστατο σήμα $f \in [0,1]$, ορισμένο στα σημεία $\{0, 1, \ldots, N-1\}$.

Κρουστικός θόρυβος Ο κρουστικός θόρυβος μπορεί να περιγραφεί ως εξής: Έστω h, διάνυσμα το οποίο δίνει τη θέση των σημείων που έχουν αλλοιωθεί. π.χ. $\{0, 0, 1, \ldots, 0, 1\}$. Επίσης, έστω F η εικόνα με θόρυβο. Θα ισχύει:

$$F(x) = \begin{cases} 1 & ,h(x) = 1 \\ f(x) & ,h(x) = 0 \end{cases} \tag{1.59}$$

Για τη μονοδιάστατη ροπή τάξης n:

$$\begin{aligned} M_n^F &= \sum_{x=0}^{N-1} P_n(x)F(x) \\ &= \sum_{h(x)=0} P_n(x)f(x) + \sum_{h(x)=1} P_n(x) \\ &= \sum_{h(x)=0} P_n(x)f(x) + \sum_{h(x)=1} P_n(x)(1 + f(x) - f(x)) \\ &= \sum_{x=0}^{N-1} P_n(x)f(x) + \sum_{h(x)=1} P_n(x)(1 - f(x)) \\ &= M_n^f + \sum_{h(x)=1} P_n(x)(1 - f(x)) \end{aligned} \tag{1.60}$$

Το σφάλμα που εισάγεται κατά τον υπολογισμό των ροπών είναι ίσο με $\sum_{h(x)=1} P_n(x)(1 - f(x))$. Από την παραπάνω σχέση μπορούν να προκύψουν τα εξής συμπεράσματα:

- Με δεδομένο ότι οι περισσότερες τιμές του διανύσματος h θα είναι μηδενικές, για την ελαχιστοποίηση του θορύβου απαιτείται να είναι όσο το δυνατόν μικρότερη η τιμή $P_n(x_0))/\sum|P_n(x)|$, για κάθε x_0. Πρακτικά αυτό σημαίνει ότι το εύρος τιμών της συνάρτησης βάσης $P_n(x)$ θα πρέπει να είναι όσο το δυνατόν μικρότερο[2].
- Προφανώς, όσα λιγότερα σημεία επηρεάζονται από κρουστικό θόρυβο, τόσο μικρότερη θα είναι και η μεταβολή των τιμών των ροπών.
- Οι τύποι ροπών με συνάρτηση βάσης που παίρνει τιμές θετικές και αρνητικές στατιστικά επηρεάζονται λιγότερο, λόγω της πιθανότητας αλληλοαναίρεσης των τιμών θορύβου.

Βάσει των παραπάνω, είναι προφανές ότι χειρότερα θα συμπεριφέρονται οι γεωμετρικές ροπές, που παρουσιάζουν πολύ μεγάλο εύρος τιμών και θετικά βάρη σε όλα τα σημεία.

Τυχαίος, προσθετικός θόρυβος Ο τυχαίος, προσθετικός θόρυβος μπορεί να μοντελοποιηθεί ως διάνυσμα $r : \{0, 1, \ldots, N-1\} \to [-1, 1]$. Το τελικό σήμα F θα δίνεται:

$$F(x) = \begin{cases} f(x) + r(x) & , f(x) + r(x) \leq 1, f(x) + r(x) \geq 0 \\ 1 & , f(x) + r(x) > 1 \\ 0 & , f(x) + r(x) < 0 \end{cases} \tag{1.61}$$

Για την n-οστή ροπή θα ισχύει:

$$\begin{aligned} M_n^F &= \sum_{x=0}^{N-1} P_n(x) F(x) \\ &= \sum_{x=0}^{N-1} P_n(x) \max(\min(f(x) + r(x), 1), 0) \end{aligned} \tag{1.62}$$

Για την ακριβέστερη εκτίμηση της επίδρασης του θορύβου απαιτείται εκτενέστερη διερεύνηση:

$$M_n^F = \sum_{x=0}^{N-1} P_n(x) \max(\min(f(x) + r(x), 1), 0)$$

[2]Φυσικά, αυτό έχει αρνητικό αντίκτυπο στις υπόλοιπες, επιθυμητές ιδιότητες των ροπών

$$= \sum_{\substack{f(x)+r(x)<1 \\ f(x)+r(x)>0}}^{N-1} P_n(x)(f(x)+r(x)) + \sum_{f(x)+r(x)>1} P_n(x)$$

$$= \sum_{\substack{f(x)+r(x)<1 \\ f(x)+r(x)>0}}^{N-1} P_n(x)f(x) + \sum_{f(x)+r(x)>1} P_n(x) + \sum_{\substack{f(x)+r(x)<1 \\ f(x)+r(x)>0}}^{N-1} P_n(x)r(x) \quad (1.63)$$

Από την παραπάνω σχέση προκύπτει αρχικά το προφανές συμπέρασμα ότι μικρότερο επίπεδο θορύβου επηρεάζει λιγότερο τις τιμές των ροπών. Ένα εξίσου σημαντικό συμπέρασμα, όμως, μπορεί να εξαχθεί αν λάβουμε υπ'οψιν ότι γενικά ο τυχαίος θόρυβος έχει μηδενική μέση τιμή[3]. Θεωρώντας ότι είναι ελάχιστες οι τιμές για τις οποίες η $f(x)+r(x) \notin [0,1]$, η τιμή της ροπής του τελικού σήματος είναι ίση με το άθροισμα των τιμών των ροπών του αρχικού σήματος και του σήματος θορύβου. Επομένως, το σφάλμα θα είναι ίσο με την τιμή της ροπής του θορύβου. Αυτό σημαίνει ότι και πάλι, συνάρτηση βάσης με $\sum P_n(x) = 0$, και $P_n(x_0)) << \sum |P_n(x)|$, για κάθε x_0, θα δίνει ροπές με καλύτερα χαρακτηριστικά ως προς το θόρυβο.

Παράδειγμα
Έστω σήμα $f(x) = sin(\pi x/16)$ ορισμένο στα σημεία $\{0,\ldots,16\}$ (Σχήμα 1.4). Η ροπή 4ης τάξης δίνεται από τη σχέση:

$$M_4 = \sum_{x=0}^{16} sin(\pi x/16)P_4(x) \quad (1.64)$$

Έστω ότι στο σήμα εισέρχεται τυχαίος θόρυβος χαμηλού επιπέδου $r(x)$. Μία προσέγγιση της επίδρασης του θορύβου είναι η:

$$M_4^n = \sum_{x=0}^{16} r(x)P_4(x) \quad (1.65)$$

Για γεωμετρική ροπή, η παραπάνω σχέση γίνεται:

$$M_4^n = \sum_{x=0}^{16} r(x)x^4 \quad (1.66)$$

Για θόρυβο με τυπική απόκλιση 0,1, μία τυπική τιμή είναι η 10.000. Η τιμή της αντίστοιχης ροπής του σήματος είναι

[3]Σε αντίθετη περίπτωση το συστηματικό σφάλμα αφαιρείται εύκολα πριν την επεξεργασία του σήματος.

98.000. Από τα παραπάνω, είναι προφανές ότι η τιμή της γεωμετρικής τιμής μεταβάλλεται κατά 10%.

Στις ροπές ορθογώνιας βάσης, η αντίστοιχες τιμές είναι πολύ μικρότερες. Για παράδειγμα, η ροπή Legendre, 4ης τάξης έχει τιμή -9.13, και η μεταβολή λόγω σφάλματος είναι -0.3, που αντιστοιχεί σε 3%.

1.11 Συμπεράσματα

Οι ροπές εικόνων αποτελούν ένα πολύ χρήσιμο εργαλείο στην ανάλυσή τους. Προσφέρουν ισχυρούς περιγραφείς, οι οποίοι μπορεί να είναι αμετάβλητοι σε γραμμικούς μετασχηματισμούς. Βασικό τους πλεονέκτημα είναι και το μικρό μέγεθος των περιγραφέων αυτών, γεγονός που αποτελεί βασικό προτέρημα στην ταξινόμηση εικόνων.

Τα βασικά προβλήματα των ροπών είναι το σχετικά μεγάλο υπολογιστικό τους κόστος, που σχετίζεται με τον υπολογισμό σχετικά πολύπλοκων συναρτήσεων για τον χωρικό μετασχηματισμό της εικόνας εισόδου, καθώς και τα εισερχόμενα σφάλματα στον υπολογισμό τους. Στα επόμενα κεφάλαια αναπτύσσονται τεχνικές για την αντιμετώπιση των δυσκολιών αυτών, με στόχο την ταχύτερη και πιο ακριβή αναπαράσταση εικόνων.

Κεφάλαιο 2

Ταχύς υπολογισμός ροπών

Οι ροπές εικόνων αποτελούν ισχυρούς περιγραφείς σχήματος, με πλήθος εφαρμογών. Εντούτοις, το υπολογιστικό κόστος για τον υπολογισμό τους, περιορίζει τη χρήση τους σε πρακτικές εφαρμογές. Παρακάτω παρουσιάζονται τεχνικές υλικού και λογισμικού για τον αποδοτικό υπολογισμό τους. Αυτές είναι οι ταχύτερες που αναφέρονται στη βιβλιογραφία μέχρι σήμερα.

Οπως περιγράφηκε στο προηγούμενο κεφάλαιο, οι ροπές εικόνων αποτελούν ισχυρούς περιγραφείς σχήματος. Οι ροπές *Zernike* χρησιμοποιούνται εκτενώς λόγω των επιθυμητών τους χαρακτηριστικών ως προς την αμεταβλητότητα, την ανοχή στο θόρυβο και τη δυνατότητα ανακατασκευής του αρχικού σήματος. Εντούτοις, ακόμα και σήμερα, οι πιο κοινά χρησιμοποιούμενες ροπές είναι οι γεωμετρικές. Αυτό διότι είναι απλές στην υλοποίηση και δίνουν απευθείας τα βασικά γεωμετρικά χαρακτηριστικά ενός σχήματος. Τέλος, οι ροπές διακριτής βάσης καλούνται να συνδυάσουν την ορθογωνιότητα των ροπών *Zernike* με την ταχύτητα των γεωμετρικών. Σε αυτό το κεφάλαιο παρατίθενται αρχιτεκτονικές υλοποίησης γεωμετρικών ροπών, καθώς και ροπών *Zernike*, *Legendre* και *Chebyshev*. Η ανάγκη υλοποίησης τέτοιων τεχνικών είναι προφανής, αν ληφθεί υπ'όψιν ότι ακόμα και στην περίπτωση που ολόκληρο το σύνολο των συναρτήσεων βάσης των ροπών είναι αποθηκευμένο στη μνήμη, ο υπολογισμός μίας ροπής απαιτεί N^2 προσθέσεις και N^2 πολλαπλασιασμούς για εικόνα μεγέθους $N \times N$. Βεβαίως, η αποθήκευση όλων των τιμών κάθε συνάρτησης βάσης είναι μη πρακτική και το πραγματικό υπολογιστικό κόστος είναι κατά πολύ μεγαλύτερο. Οι αρχιτεκτονικές αυτές δίνουν την δυνατότητα υπολογισμού των ροπών αυτών σε πραγματικό χρόνο, ακόμα και για εικόνες μεγάλου μεγέθους.

2.1 Γεωμετρικές ροπές

Τέσσερις αρχιτεκτονικές υλικού πραγματικού χρόνου για τον υπολογισμό των γεωμετρικών ροπών αναλύονται εκτενώς στην εργασία [Kotoulas, 2004] καθώς και στην [Kotoulas and Andreadis, 2004]. Οι αρχιτεκτονικές αυτές κλήθηκαν να αντιμετωπίσουν το πρόβλημα του σχετικά υψηλού υπολογιστικού κόστους των γεωμετρικών ροπών.

Χάριν πληρότητας, παρατίθενται τα βασικά σημεία τους.

2.1.1 Υπολογισμός των ροπών με τη χρήση φίλτρων πόλων

Το φίλτρο ενός πόλου $Y_1(z) = 1/(z-1)$ ισοδυναμεί με έναν συσσωρευτή [Scheid, 1968] . Η απόκρισή του στο πεδίο του χρόνου σε μία είσοδο $f[n]$ είναι:

$$y_1 = \sum_{n=0}^{N-1} f(n) \tag{2.1}$$

Ένα φίλτρο πόλων $Y_1(z) = 1/(z-1)^r$ μπορεί να υλοποιηθεί με πολλαπλούς συσσωρευτές συνδεδεμένους σε σειρά. Για μία είσοδο $f[n] \rightarrow F(z)$ η απόκρισή του συστήματος είναι:

$$y_r = \frac{1}{r!} \sum_{n=0}^{N-1} (N-n)(N-n+1)\ldots(N-n+r)f(n) \tag{2.2}$$

όπου r ο αριθμός των συσσωρευτών. Ο παράγοντας $(N-n)(N-n+1)\ldots(N-n+r)$ είναι ένα παραγοντικό πολυώνυμο το οποίο μπορεί να εκφραστεί σε κανονική μορφή πολυωνύμου χρησιμοποιώντας τους αριθμούς του Stirling :

$$\frac{1}{r!}(N-n)(N-n+1)\ldots(N-n+r) = \sum_{t=0}^{r} s_t^r (N-n)^t \tag{2.3}$$

Μία αναλυτική περιγραφή των αριθμών του Stirling μπορεί να βρεθεί στην βιβλιογραφία [Scheid, 1968]. Περισσότερες πληροφορίες για τα φίλτρα πόλων παρατίθενται στο Παράρτημα Β΄

Η σχέση 2.2 μπορεί, τώρα, να πάρει την εξής μορφή:

$$y_r = \sum_{t=0}^{r} s_t^r \sum_{n=0}^{N-1} (N-n)^t f(n) \tag{2.4}$$

Το τελευταίο άθροισμα ισούται με την μονοδιάστατη διακριτή γεωμετρική ροπή τάξης t ως προς το τελευταίο σημείο της συνάρτησης εισόδου:

$$y_r = \sum_{t=0}^{r} s_t^r \mu_t \tag{2.5}$$

ή

$$\begin{bmatrix} y_1 \\ y_2 \\ \dots \\ y_r \end{bmatrix} = S \begin{bmatrix} \mu_1 \\ \mu_2 \\ \dots \\ \mu_r \end{bmatrix} \tag{2.6}$$

Αντιστρέφοντας τον πίνακα S, η παραπάνω σχέση μπορεί να λυθεί ως προς μ_t:

$$\mu_r = \sum_{t=0}^{r} \dot{s}_t^r y_t \tag{2.7}$$

, όπου $S\dot{S} = I$. Ο S είναι τριγωνικός, αντιστρέψιμος πίνακας [Scheid, 1968]. Για τη γενίκευση του παραπάνω για ροπές ως προς οποιοδήποτε σημείο μπορεί να χρησιμοποιηθεί το διωνυμικό θεώρημα. Πραγματοποιώντας κατάλληλους μετασχηματισμούς προκύπτει τελικά:

$$\mu_r^X = \sum_{t=0}^{r} \frac{r!}{t!(t-r)!} (X-N)^{r-t} \sum_{i=0}^{t} \dot{s}_i^t y_i \tag{2.8}$$

Χρησιμοποιώντας την παραπάνω σχέση η μονοδιάστατη γεωμετρική ροπή οποιασδήποτε τάξης ως προς οποιοδήποτε σημείο X μπορεί να υπολογιστεί από μία συστοιχία συσσωρευτών. Τα αποτελέσματα αυτά μπορούν να χρησιμοποιηθούν για την εξαγωγή παρόμοιων σχέσεων για γεωμετρικές ροπές δύο διαστάσεων. Τώρα απαιτείται ένα φίλτρο δύο διαστάσεων, το οποίο υλοποιείται από ένα πλέγμα συσσωρευτών. Η τελική σχέση ανάμεσα σε μία γεωμετρική ροπή δύο διαστάσεων και τις εξόδους ενός πλέγματος συσσωρευτών που προκύπτει μετά από

Μέγεθος εικόνας	Υλικό	Λογισμικό	
		Προσθέσεις	Πολλαπλασιασμοί
(α)			
128×128	$409ns$	16.384	327.680
256×256	$1,6ms$	65.536	1.310.720
1.024×1.024	$26ms$	1.048.576	20.971.520
2.048×2.048	$105ms$	4.194.304	83.886.080
(b)			
128×128	$409ns$	16.384	1.146.880
256×256	$1,6ms$	65.536	4.587.520
1.024×1.024	$26ms$	1.048.576	73.400.320
2.048×2.048	$105ms$	4.194.304	293.601.280

Πίνακας 2.1: Υπολογιστικό κόστος υλοποίησης γεωμετρικών ροπών για τις πρώτες (α) 10 και (b) 21 ροπές

παρόμοιες αντικαταστάσεις είναι:

$$\mu_{r,s}^{X,U} = \sum_{t=0}^{r} \frac{r!}{t!(t-r)!}(X-N_x)^{r-t} \sum_{\xi=0}^{\varphi} \frac{\varphi!}{\xi!(\varphi-\xi)!}(Y-N_y)^{\varphi-\xi} \sum_{e=0}^{t} \dot{s}_t^t \sum_{i=0}^{\xi} \dot{s}_i^{\xi} y_{t,i} \quad (2.9)$$

Αυτή η σχέση μπορεί να χρησιμοποιηθεί για τον υπολογισμό γεωμετρικών ροπών δύο διαστάσεων, ως προς σημείο (X,Y) από τις εξόδους του πλέγματος συσσωρευτών. Εφόσον οι τιμές των $y_{t,i}$ έχουν υπολογιστεί από το υλικό, οι απαιτούμενοι υπολογισμοί μεταβλητής υποδιαστολής για τις πρώτες 10 γεωμετρικές ροπές είναι 274 πολλαπλασιασμοί και 74 αθροίσεις, όπως προκύπτει από καταμέτρηση των πράξεων της 2.9. Αξίζει να σημειωθεί ότι ο άμεσος υπολογισμός των πρώτων n ροπών απαιτεί $O(n^2N^2)$ πολλαπλασιασμούς και $O(nN^2)$ αθροίσεις για μία εικόνα χωρικής ανάλυσης NxN στοιχείων. Η σχέση μεταξύ μ και y είναι γραμμική και μπορεί να επεκταθεί και για τον υπολογισμό γεωμετρικών ροπών σε περισσότερες διαστάσεις. Στον Πίνακα 2.1 παρατίθενται αριθμητικά στοιχεία για το υπολογιστικό κόστος του προτεινόμενου συστήματος, σε σχέση με τον υπολογισμό σε λογισμικό. Οι χρόνοι υλικού αναφέρονται σε συσκευή *FPGA* τυπικής συχνότητας λειτουργίας $100MHz$.

2.2 Ροπές *Zernike*

Η αποδοτική εξαγωγή των ροπών Zernikeμε τη χρήση φίλτρων πόλων μελετήθηκε επίσης στην εργασία [Kotoulas, 2004] και την [Kotoulas and Andreadis, 2005]. Παρακάτω παρατίθενται τα βασικά σημεία της υλοποίησης αυτής. Σε αντίθεση με τις γεωμετρικές ροπές, στην περίπτωση των ροπών Zernike, είναι προτιμότερο αυτές να υπολογιστούν από τις γεωμετρικές, και όχι άμεσα από τις εξόδους ενός φίλτρου πόλων.

2.2.1 Σχέση ροπών *Zernike* με γεωμετρικές ροπές

Έστω συνεχές σήμα $f(x,y) : R^2 \to R$. Ισχύει:

$$
\begin{aligned}
Z_{p,q} &= \frac{p+1}{\pi}\int_0^1\int_0^{2\pi} rR_{pq}(r)e^{-iq\theta}f(r,\theta)d\theta dr \\
&= \frac{p+1}{\pi}\int_0^1\int_0^{2\pi} r\sum_{\substack{k=q \\ p-|k| \\ even}}^{p} B_{pqk}r^k e^{-iq\theta}f(r,\theta)d\theta dr \\
&= \frac{p+1}{\pi}\int_0^1\int_0^{1} r\sum_{\substack{k=q \\ p-|k| \\ even}}^{p} B_{pqk}(x^2+y^2)^{\frac{k}{2}}\frac{(x-yi)^q}{(x^2+y^2)^{\frac{q}{2}}}f(x,y)dxdy \\
&= \frac{p+1}{\pi}\int_0^1\int_0^{1} r\sum_{\substack{k=q \\ p-|k| \\ even}}^{p} B_{pqk}(x^2+y^2)^{\frac{k-q}{2}}\sum_{m=0}^{q}\binom{q}{m}x^m(-i)^{q-m}y^{q-m}f(x,y)dxdy \\
&= \frac{p+1}{\pi}\int_0^1\int_0^{1} r\sum_{\substack{k=q \\ p-|k| \\ even}}^{p} B_{pqk}\sum_{n=0}^{\frac{k-q}{2}}\binom{\frac{k-q}{2}}{n}x^{2n}y^{k-q-2n} \\
&\quad \sum_{m=0}^{q}\binom{q}{m}x^m(-i)^{q-m}y^{q-m}f(x,y)dxdy \\
&= \frac{p+1}{\pi}\sum_{\substack{k=q \\ p-|k| \\ even}}^{p} B_{pqk}\sum_{n=0}^{\frac{k-q}{2}}\binom{\frac{k-q}{2}}{n}\sum_{m=0}^{q}\binom{q}{m}\int_0^1\int_0^1 x^{m+2n}y^{k-2n-m}f(x,y)dxdy
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{p+1}{\pi}\sum_{\substack{k=q\\ p-|k|\\ even}}^{p} B_{pqk}\sum_{n=0}^{\frac{k-q}{2}}\binom{\frac{k-q}{2}}{n}\sum_{m=0}^{q}\binom{q}{m}\int_0^N\int_0^N N^{-k-2}u^{m+2n}w^{k-2n-m}f(w,u)dudw \\
Z_{p,q} &= \frac{p+1}{\pi}\sum_{\substack{k=q\\ p-|k|\\ even}}^{p} B_{pqk}\sum_{n=0}^{\frac{k-q}{2}}\binom{\frac{k-q}{2}}{n}\sum_{m=0}^{q}\binom{q}{m} N^{-k-2}M_{m+2n,k-2n-m} \qquad (2.10)
\end{aligned}
$$

2.2.2 Ταχύς υπολογισμός ροπών *Zernike*

Με βάση την παραπάνω σχέση, καθίσταται δυνατός ο αποδοτικός υπολογισμός των ροπών *Zernike*, χρησιμοποιώντας τις σχέσεις που περιγράφηκαν στην προηγούμενη ενότητα. Πιο συγκεκριμένα, μετά την εξαγωγή των γεωμετρικών ροπών, υπολογίζονται οι ροπές *Zernike* με χρήση της σχέσης 2.10. Η τελευταία παρουσιάζει ασήμαντη υπολογιστική πολυπλοκότητα σε σχέση με την απ' ευθείας εξαγωγή των ροπών από την εικόνα.

2.2.3 Αποτίμηση τεχνικής

Στη βιβλιογραφία μπορούν να βρεθούν πολλές μέθοδοι για την ταχεία υλοποίηση των ροπών *Zernike*. Οι περισσότερες εκμεταλλεύονται τα χαρακτηριστικά συμμετρίας των πολυωνύμων, καθώς και διάφορες αναδρομικές σχέσεις. Πιο συγκεκριμένα, οι *Mukundan* και *Ramakrishnan* [Mukundan and Ramakrishnan, 1995], με μία προσέγγιση των πολικών συντεταγμένων έδωσαν μη ακριβείς περιγραφείς, οι οποίοι υπολογίζονται με σημαντική αύξηση στην ταχύτητα. Ο *Belkasim* [B. Belkasim et al., 1996] καθώς και ο *Gu* [Gu et al., 2002], εκμεταλλευόμενοι τις αναδρομικές ιδιότητες, έδωσαν ακριβείς εκφράσεις οι οποίες μειώνουν σημαντικά την υπολογιστική πολυπλοκότητα της εξαγωγής των ροπών,

Μία σύγκριση της προτεινόμενης τεχνικής με τις αναφερόμενες στη βιβλιογραφία παρουσιάζεται στον Πίνακα 2.2. Και σε αυτήν την περίπτωση, οι χρόνοι υλικού αναφέρονται σε συσκευή τυπικής συχνότητας 100*MHz*.

2.3 Ροπές *Chebyshev*

Οι ροπές Chebychev προτάθηκαν πρόσφατα για την αντιμετώπιση προβλημάτων των ροπών με συνεχείς βάσεις. Είναι ροπές διακριτής βάσης,

N=1.024, M=40	Αριθμός Αθροίσεων	Αριθμός Πολ/σμών
Mukundan and Ramakrishnan (1995)	$8,42\,10^8$	$1,64\,10^9$
Belkasim(1996)	41.983.920	$8,38\,10^8$
Gu (2002)	$5,09\,10^8$	84.705.280
Προτεινόμενη τεχνική	450.877	598.969

Πίνακας 2.2: Κόστος Υπολογισμού Ροπών *Zernike*

ορθογώνιες στο πεδίο της εικόνας, ένα χαρακτηριστικό που παρουσιάζει έναν αριθμό πλεονεκτημάτων σε σχέση με τις συνεχείς ορθογώνιες ροπές. Μέχρι τώρα δεν έχουν γίνει προσπάθειες για τον αποδοτικό υπολογισμό τους. Σε αυτό το κεφάλαιο παρουσιάζεται μία αρχιτεκτονική υλικού η οποία επιτρέπει την ανάλυση εικόνων μεγέθους μέχρι 4MPixels σε πραγματικό χρόνο. Η αρχιτεκτονική αυτή υλοποιήθηκε σε ένα FPGA με τυπική συχνότητα λειτουργίας 100MHz [Kotoulas and Andreadis, 2006].

2.3.1 Υλοποίηση ροπών *Chebyshev* με τη χρήση φίλτρων πόλων

Μονοδιάστατες ροπές *Chebyshev*

Μία μονοδιάστατη ροπή Chebychev τάξης n ενός διακριτού σήματος f δίνεται από την παρακάτω σχέση [Mukundan et al., 2001]:

$$T_n = \frac{1}{\tilde{\rho}(n,N)} \sum_{x=0}^{N-1} \tilde{t}_p(x) f(x) \tag{2.11}$$

με $\tilde{t}_n(x)$ το ορθογώνιο πολυώνυμο Chebychev τάξης n:

$$\begin{aligned} \tilde{t}_n(x) &= \frac{t_n(x)}{\beta(n,N)} \\ &= \frac{n!}{\beta(n,N)} \sum_{k=0}^{n} (-1)^{n-k} \binom{N-1-k}{n-k} \binom{n+k}{n} \binom{x}{k} \end{aligned} \tag{2.12}$$

όπου $\beta(n,N)$ ένας παράγοντας κανονικοποίησης. Διάφορες εκφράσεις του μπορούν να βρεθούν στην [Mukundan et al., 2001] . Η απλούστερη επιλογή είναι η N^n. Αναπτύσσοντας τους συνδυαστικούς όρους η παραπάνω σχέση γίνεται:

$$\tilde{t}_n(x) = \frac{n!}{\beta(n,N)} \sum_{k=0} n(-1)^{n-k} \frac{(N-1-k)!}{(n-k)!(N-n-1)!} \frac{(N+k)!x!}{n!k!k!(x-k)!} \quad (2.13)$$

ή

$$\tilde{t}_n(x) = \frac{1}{\beta(n,N)(N-n-1)!} \sum_{k=0} n(-1)^{n-k} \frac{(N-1-k)!}{(n-k)!} \frac{(N+k)!x!}{n!(k!)^2(x-k)!} \quad (2.14)$$

Η εξίσωση 2.11 τώρα γίνεται:

$$T_n = \frac{1}{\tilde{\rho}(n,N)\beta(n,N)(N-n-1)!} \sum_{x=0}^{N-1} \sum_{k=0}^{n} (-1)^{n-k} \frac{(n+k)!(N-1-k)!x!}{(n-k)!(k!)^2(x-k)!} f[x] \quad (2.15)$$

ή

$$\begin{aligned} T_n &= \frac{1}{\tilde{\rho}(n,N)\beta(n,N)(N-n-1)!} \\ & N \sum_{k=0}^{n} (-1)^{n-k} \frac{(n+k)!(N-1-k)!}{(n-k)!(k!)} \sum_{x=0}^{N-1} \frac{x!}{(x-k)!k!} f[x] \quad (2.16) \end{aligned}$$

Το τελευταίο άθροισμα της παραπάνω σχέσης ισοδυναμεί με την έξοδο ενός φίλτρου k πόλων. Επομένως, οι μονοδιάστατες ροπές *Chebychev* μπορούν να υπολογισθούν από τις εξόδους φίλτρων πόλων σύμφωνα με την παρακάτω σχέση:

$$T_n = \frac{1}{\tilde{\rho}(n,N)\beta(n,N)(N-n-1)!} \sum_{k=0}^{n} (-1)^{n-k} \frac{(n+k)!(N-1-k)!}{(n-k)!(k!)} y_k[N] \quad (2.17)$$

Η υλοποίηση της παραπάνω σχέσης απαιτεί έναν πολύ μικρό αριθμό πράξεων, από τη στιγμή που οι όροι $\tilde{\rho}$ και β μπορούν να υπολογισθούν εκ των προτέρων. Πιο συγκεκριμένα, αποθηκεύοντας σε πίνακα μεγέθους $(n+1) \times (n+1)$ τις τιμές της παράστασης:

$$(-1)^{n-k} \frac{(n+k)!(N-1-k)!}{(n-k)!(k!)} \quad (2.18)$$

οι απαιτούμενες πράξεις είναι n προσθέσεις και $n+2$ πολλαπλασιασμοί για ροπή τάξης n. Λαμβάνοντας υπ'οψη ότι ο απ' ευθείας υπολογισμός θα απαιτούσε περίπου n προσθέσεις και n πολλαπλασιασμούς για κάθε σημείο της εικόνας, είναι προφανής η βελτίωση στην ταχύτητα του συστήματος.

Ροπές *Chebychev* δύο διαστάσεων

Μία ροπή *Chebychec* τάξης p,q μπορεί να οριστεί βάσει της παρακάτω σχέσης [Mukundan et al., 2001]:

$$T_{p,q} = \frac{1}{\tilde{\rho}(p,N)\tilde{\rho}(q,N)} \sum_{y=0}^{N-1}\sum_{x=0}^{N-1} \tilde{t}_p(x)\tilde{t}_q(y) f[x,y] \tag{2.19}$$

Χρησιμοποιώντας τα αποτελέσματα της προηγούμενης διερεύνησης, μπορούν να υπολογισθούν οι ροπές δύο διαστάσεων:

$$T_{p,q} = \frac{1}{\tilde{\rho}(p,N)\tilde{\rho}(q,N)} \sum_{y=0}^{N-1} \tilde{t}_q(y) \sum_{x=0}^{N-1} \tilde{t}_p(x) f[x,y] \tag{2.20}$$

Από τη σχέση 1.33 προκύπτει:

$$\begin{aligned} T_{p,q} &= \frac{1}{\tilde{\rho}(q,N)} \sum_{y=0}^{N-1} \tilde{t}_q(y) \frac{1}{\tilde{\rho}(p,N)} \sum_{x=0}^{N-1} \tilde{t}_p(x) f(x,y) \\ &= \frac{1}{\tilde{\rho}(q,N)} \sum_{y=0}^{N-1} \tilde{t}_q(y) T_p(y) \end{aligned} \tag{2.21}$$

Εκφράζοντας τον τελευταίο όρο σύμφωνα με τα προηγούμενα, προκύπτει:

$$\begin{aligned} T_{p,q} = {} & \frac{q!}{\tilde{\rho}(q,N)\tilde{\rho}(p,N)\beta(q,N)\beta(p,N)(N-p-1)!} \times \\ & \sum_{y=0}^{N-1}\sum_{t=0}^{q} (-1)^{q-t} \binom{N-t-1}{q-t} \binom{q+t}{q} \binom{y}{t} \times \\ & \sum_{k=0}^{p} (-1)^{p-k} \frac{(p+k)!(N-k-1)!}{(p-k)!k!} y_t \end{aligned} \tag{2.22}$$

ή

$$T_{p,q} = \frac{q!}{\tilde{\rho}(q,N)\tilde{\rho}(p,N)\beta(q,N)\beta(p,N)(N-p-1)!} \times$$

$$\sum_{k=0}^{p} \frac{(p+k)!(N-k-1)!}{(p-k)!k!} \times$$
$$\sum_{y=0}^{N-1} \sum_{t=0}^{q} (-1)^{q-t} \frac{(N-1-t)!(q+t)!y!}{(q-t)!t!t!(y-t)!} y_t \tag{2.23}$$

και

$$\begin{aligned} T_{p,q} &= \frac{q!}{\tilde{\rho}(q,N)\tilde{\rho}(p,N)\beta(q,N)\beta(p,N)(N-p-1)!} \times \\ &\times \sum_{k=0}^{p} (-1)^{p-k} \frac{(p+k)!(N-k-q)!}{(p-k)!k!} \sum_{t=0}^{p} (-1)^{q-t} \frac{(q+t)!(N-t-q)!}{(q-t)!t!} \\ &\times \sum_{y=0}^{N-1} \frac{y!}{t!(y-t)!} y_t \end{aligned} \tag{2.24}$$

Τελικά, οι ροπές *Chebychev* συσχετίζονται με τις εξόδους φίλτρων πόλων σύμφωνα με την παρακάτω σχέση:

$$\begin{aligned} T_{p,q} &= \frac{q!}{\tilde{\rho}(q,N)\tilde{\rho}(p,N)\beta(q,N)\beta(p,N)(N-p-1)!} \times \\ &\sum_{k=0}^{p} (-1)^{p-k} \frac{(p+k)!(N-k-q)!}{(p-k)!k!} \times \\ &\sum_{t=0}^{p} (-1)^{q-t} \frac{(q+t)!(N-t-q)!}{(q-t)!t!} y_{t,k} \end{aligned} \tag{2.25}$$

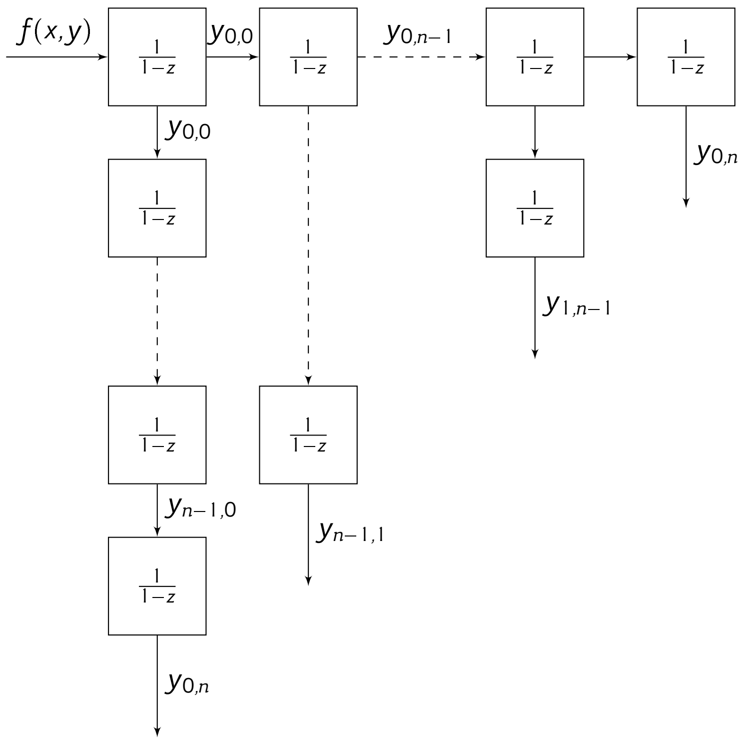
Ορίζοντας:

$$A(p,q,N) = \frac{q!}{\tilde{\rho}(q,N)\tilde{\rho}(p,N)\beta(q,N)\beta(p,N)(N-p-1)!} \tag{2.26}$$

τότε

$$\begin{aligned} T_{p,q} &= A(p,q,N) \sum_{k=0}^{p} (-1)^{p-k} \frac{(p+k)!(N-k-q)!}{(p-k)!k!} \times \\ &\sum_{t=0}^{p} (-1)^{q-t} \frac{(q+t)!(N-t-q)!}{(q-t)!t!} y_{t,k} \end{aligned} \tag{2.27}$$

με $y_{t,k}$ την έξοδο ενός φίλτρου τάξης (t,k).



Σχήμα 2.1: Πλέγμα συσσωρευτών

Χρησιμοποιώντας τη σχέση 2.27, η αποδοτική εξαγωγή των ροπών *Chebychev* από απλά φίλτρα πόλων είναι εφικτή. Ο όρος $A(p,q,N)$ μπορεί να υπολογισθεί εκ των προτέρων, ή και ταυτόχρονα με την επεξεργασία υλικού, μειώνοντας την υπολογιστική πολυπλοκότητα σε μερικούς απλούς υπολογισμούς. Αξίζει να σημειωθεί ότι οι παραπάνω σχέσεις προέκυψαν από τον αρχικό ορισμό των ροπών *Chebychev*, χωρίς προσεγγίσεις. Επομένως, τα αποτελέσματα της μεθόδου είναι αυτά του αρχικού ορισμού.

2.3.2 Αρχιτεκτονική υλικού

Η υλοποίηση ενός φίλτρου πόλων δύο διαστάσεων αποτελείται από ένα πλέγμα συσσωρευτών, όπως αυτό φαίνεται στο σχήμα 2.1.

Στο σχήμα 2.2 παρουσιάζεται ένας απλός συσσωρευτής.

Αποτελείται από έναν αθροιστή, έναν καταχωρητή και έναν μετρη-

Σχήμα 2.2: Συσσωρευτής

τή. Ο μετρητής υλοποιήθηκε προκειμένου να μειωθεί η απαιτούμενη επιφάνεια πυριτίου, αφού η απευθείας υλοποίηση θα σήμαινε έναν μεγαλύτερο αθροιστή. Στο Σχήμα 2.3 παρουσιάζονται οι βηματικές αποκρίσεις των τριών πρώτων οριζόντιων συσσωρευτών.

Το σύστημα μπορεί να βελτιστοποιηθέι λαμβάνοντας υπ'οψιν ότι μόνο οι συσσωρευτές γραμμής απαιτούν μεταβολή της κατάστασής τους σε κάθε κύκλο ρολογιού. Οι συσσωρευτές στήλης μπορούν να υλοποιηθούν από απλούς σειριακούς αθροιστές.

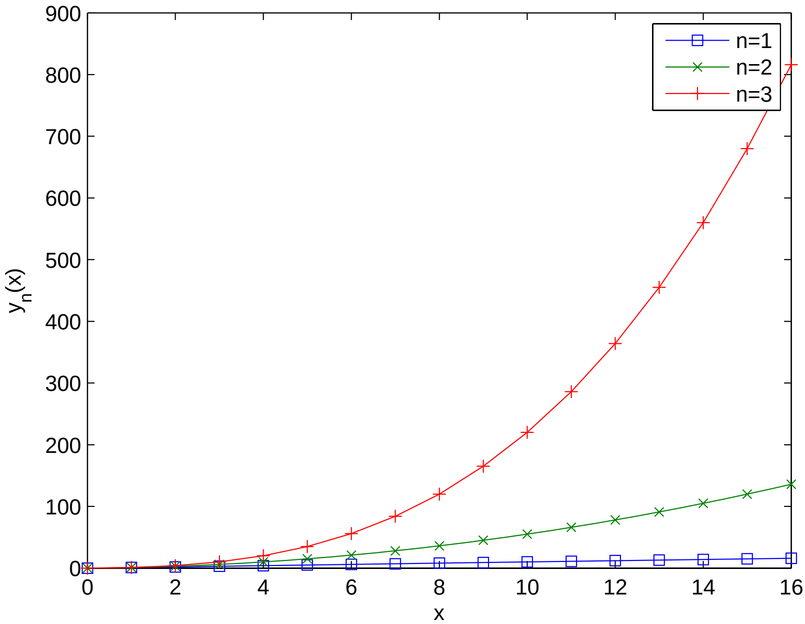
Ο αριθμός των ψηφίων συσσώρευσης που απαιτούνται για την υλοποίηση N ροπών για μία εικόνα χωρικής ανάλυσης $A \times B$ στοιχείων και χρωματικής ανάλυσης n στοιχείων δίνεται από την παρακάτω έκφραση:

$$\begin{aligned} N_{adder-bits} &= \frac{1}{6}(2\sqrt{N}-1)(4\sqrt{N}-1)(2\sqrt{N})(\log_2 A + \\ & \log_2 B) + \frac{(2\sqrt{N}-2)(2\sqrt{N}-1)}{2}n \end{aligned} \tag{2.28}$$

Η παραπάνω σχέση προκύπτει ως εξής:

Ο πρώτος συσσωρευτής πρέπει να έχει εύρος $n(\log_2 A + \log_2 B)$, προκειμένου να μπορεί να εξυπηρετήσει το άθροισμα των στοιχείων μίας εικόνας, ακόμα και στην περίπτωση που όλα έχουν τη μέγιστη φωτεινότητα. Για κάθε διαδοχικό συσσωρευτή, το εύρος αυξάνει κατά έναν παράγοντα $(\log_2 A + \log_2 B)$. Αυτό σημαίνει ότι για να υπολογιστούν ροπές μέχρι βαθμού K απαιτούνται:

$$\begin{aligned} N_{ac-bits} &= \sum_{k=0}^{K}\Big((k+1)^2(\log_2 A + \log_2 B) + kn\Big) \\ &= \frac{1}{6}(K+1)(K=2)(2K+3)(\log_2 A + \log_2 B) + \frac{K(K-1)}{2}n \end{aligned} \tag{2.29}$$



Σχήμα 2.3: Βηματικές αποκρίσεις για φίλτρα πόλων

Το πλέγμα των συσσωρευτών μπορεί να υλοποιηθεί με άμεσο τρόπο. Η ταχύτητα είναι κρίσιμη μόνο για τους συσσωρευτές της πρώτης γραμμής, αφού τα δεδομένα εισόδου εισέρχονται σε αυτούς σε κάθε κύκλο του ρολογιού. Οι συσσωρευτές στήλης αλλάζουν την κατάστασή τους σε κάθε αλλαγή γραμμής. Οι πρώτοι υλοποιήθηκαν με αθροιστές πρόβλεψης κρατουμένου, ενώ οι δεύτεροι με απλούς αθροιστές διάδοσης κρατουμένου. Οι συσσωρευτές στήλης μπορούν να υλοποιηθούν και με σειριακούς αθροιστές, με την προϋπόθεση ότι επεξεργάζονται μόνο εικόνες σχετικά μεγάλου μεγέθους. Για μικρές εικόνες (μικρότερες από 10,000 στοιχεία), η ταχύτητα των σειριακών αθροιστών δεν επαρκεί, και επομένως, τοπικές γεωμετρικές ροπές δεν μπορούν να υπολογιστούν.

2.3.3 Υπολογιστική πολυπλοκότητα των ροπών Chebyshev

Ο γρήγορος υπολογισμός των ροπών *Chebychev* μπορεί να επιτευχθεί με τον υπολογισμό όλων τον όρων πλην των εξόδων των φίλτρων $y_{t,k}$ και την αποθήκευσή τους σε πίνακες έρευνας (*LUTs*).

Ο όρος $A(p,q,N)$ απαιτείται να υπολογισθεί το πολύ μία φορά ανά ροπή. Για την εξαγωγή ενός αριθμού ροπών, αρκεί ένας υπολογισμός ανά δύο ροπές, αφού $A(p,q,N) = A(q,p,N)$.

Θεωρώντας $\beta(n,N) = \rho(n,N)$ έχουμε:

$$A(p,q,N) = \frac{1}{(N-q-1)!(N-p-1)!} \tag{2.30}$$

απλοποιώντας επιπλέον τον υπολογισμό.

Θέτοντας:

$$C_N(q,t) = (-1)^{q-t}\frac{(N-t-1)!(q+t)!}{t!(q-t)!} \tag{2.31}$$

Και αποθηκεύοντας τα αποτελέσματα σε πίνακα έρευνας, η σχέση 2.27 γίνεται:

$$T_{p,q} = A(p,q,N)\sum_{k=0}^{p} C_B(p,k)\sum_{t=0}^{q} C_N(q,t)y_{t,k} \tag{2.32}$$

απαιτώντας μόνο $(q+1)(P+1)+1$ πολλαπλασιασμούς και $(p+q)$ αθροίσεις.

Τα μεγέθη των απαιτούμενων *LUTs* είναι T^2 για $C_N(q,t)$ και $T(T+3)/4$ για τον $A(p,q,N)$.

Ο άμεσος υπολογισμός μίας ροπής *Chebychev* δύο διαστάσεων από τον ορισμό τους απαιτεί $N^2+N+2N(p+q-2)$ πολλαπλασιασμούς και $N^2+N(p+q-2)-1$ αθροίσεις.

Η υλοποίηση ροπών μέχρι τάξης T απαιτεί $1/24(T+1)(T+2)(T^2+7T+24)$ πολλαπλασιασμούς και $(1/3)T^3+T^2+(2/3)T$ προσθέσεις, σε αντιδιαστολή με τον άμεσο υπολογισμό, ο οποίος θα απαιτούσε $(1/6)N(T+2)(T+1)(4T+3N-9)$ πολλαπλασιασμούς και $(1/6)(T+2)(T+1)(2NT+3N^2-6N-3)$ αθροίσεις. Οι Wang και Wang [Wang and Wang, 2006] πρότειναν πρόσφατα μία αναδρομική μέθοδο η οποία στόχευε κυρίως στο ανάστροφο πρόβλημα. Σε σχέση με τον υπολογισμό ροπών, η μέθοδός τους μειώνει το υπολογιστικό κόστος σε $N^2/2+3N/2+1$ πολλαπλασιασμούς και $5N^2/4+N$ προσθέσεις ανά υπολογιζόμενη ροπή. Στον πίνακα 2.3 παρουσιάζονται συγκριτικά αποτελέσματα της προτεινόμενης μεθόδου σε σχέση με την άμεση και την αναδρομική μέθοδο.

$N = 128,\ T = 16$	Προσθέσεις	Πολ/σμοί
Άμεσος υπολογισμός	2.676.327	2.865.792
Αναδρομική μέθοδος	1.006.200	2.472.960
Προτεινόμενη μέθοδος	1.632	4.998&16.384 κύκλοι ρολογιού

Πίνακας 2.3: Συγκριτικά αποτελέσματα για τον υπολογισμό ροπών *Chebychev* μέχρι βαθμού 16, για εικόνα 128×128 στοιχείων

$N = 128,\ T = 16$	Χρόνος προσθέσεων	Χρόνος Πολ/σμών
Άμεσος υπολογισμός	$1,3ms$	$2,6ms$
Αναδρομική μέθοδος	$0,5ms$	$2,5ms$
Προτεινόμενη μέθοδος	$0,8\mu s$	$5\mu s$

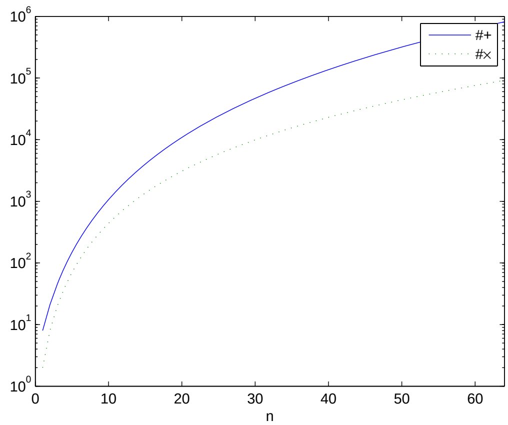
Πίνακας 2.4: Απαιτούμενος χρόνος για τον υπολογισμό ροπών *Chebychev* μέχρι βαθμού 16, για εικόνα 128×128 στοιχείων

Στον πίνακα 2.4 παρουσιάζεται αντίστοιχα ο απαιτούμενος χρόνος, υποθέτοντας μία συσκευή *FPGA* με τυπική συχνότητα λειτουργίας 100*MHz* και έναν προσωπικό υπολογιστή 2*GHz*, ικανό για μία πρόσθεση ανά κύκλο και έναν πολλαπλασιασμό ανά δύο κύκλους.

Τα παραπάνω ισχύουν για επεξεργασία εικόνων με γνωστό μέγεθος $N \times N$. Αν το μέγεθος της εικόνας δεν είναι γνωστό εκ των προτέρων, τότε οι όροι $C_N(q,T)$ και $A(p,q,N)$ πρέπει να υπολογιστούν, απαιτώντας επιπλέον $N+p+q$ πολλαπλασιασμούς. Εντούτοις, αυτοί οι πολλαπλασιασμοί μπορούν να εκτελεσθούν κατά τη διάρκεια λειτουργίας του υλικού. Στα σχήματα 2.4, 2.5 και 2.6 παρουσιάζονται κάποιες γραφικές αναπαραστάσεις της απόδοσης της προτεινόμενης αρχιτεκτονικής, συγκρινόμενης με των άλλων δύο.

2.3.4 Συμπεράσματα

Σε αυτό το μέρος της διατριβής παρουσιάστηκε μία νέα αρχιτεκτονική για τον υπολογισμό των ροπών *Chebychev*. Με τη χρήση μίας μονάδας υλικού, ο υπολογισμός αυτός παρουσιάζει ασήμαντο υπολογιστικό κόστος για την μονάδα λογισμικού, ενώ το υλικό, εκμεταλλευόμενο αρχιτεκτονική διοχέτευσης, επεξεργάζεται ένα στοιχείο ανά κύκλο. Η τεχνική είναι επεκτάσιμη για οποιοδήποτε μέγεθος εικόνας και οποιονδήποτε βαθμό ροπών. Μπορεί να χρησιμοποιηθεί για την εξαγωγή ροπών *Chebychev* σε πραγματικό χρόνο μέχρι και 16ου βαθμού για εικόνες μεγέθους μέχρι 4 εκατομμυρίων στοιχείων.



Σχήμα 2.4: Αριθμός απαιτούμενων πράξεων για εικόνα $N \times N$

2.4 Ροπές *Legendre*

Μία ροπή *Legendre* τάξης $(m+n)$ ορίζεται ως [Teague, 1980]:

$$\lambda_{m,n} = \frac{1}{4}(2m+1)(2n+1)\int_{-1}^{1}\int_{-1}^{1} P_m(x)P_n(y)f(x,y)dxdy \tag{2.33}$$

με $m,n = 0,1,\ldots,\infty$, P_m και P_n τα πολυώνυμα *Legendre* και $f(x,y)$ τη συνεχή εικόνα. Τα πολυώνυμα *Legendre* αποτελούν μία πλήρη ορθογώνια βάση, ορισμένη στο διάστημα $[-1,1]$. Για να ισχύει η ορθογωνιότητα στις ροπές, η συνάρτηση της εικόνας, $f(x,y)$ ορίζεται στο ίδιο διάστημα με της βάσης. Το πολυώνυμο *Legendre* βαθμού n ορίζεται ως:

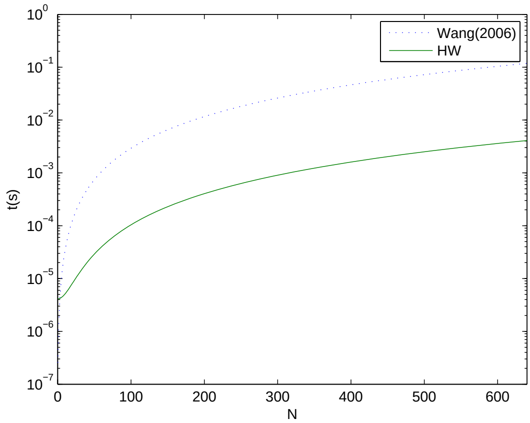
$$P_n(x) = \sum_{j=0}^{n} \alpha_{n,j} x^j \tag{2.34}$$

και $\alpha_{n,j}$ είναι οι συντελεστές *Legendre*:

$$\alpha_{n,j} = (-1)^{(n-j)/2} 2^{-n} \frac{(n+j)!}{((n-j)/2)!((n+j)/2)!j!} \tag{2.35}$$

με $(n-j)$ άρτιο. Στη διακριτή περίπτωση, η σχέση 2.33 γίνεται:

$$\lambda_{m,n} = \frac{1}{4}(2m+1)(2n+1)\sum_{x}^{N}\sum_{y}^{N} P_m(x)P_n(y)f(x,y) \tag{2.36}$$



Σχήμα 2.5: Απαιτούμενος χρόνος (t) για την εξαγωγή ροπών μέχρι 16 τάξης από εικόνα μεγέθους $N \times N$

με $x, y \in [-1, 1]$. Αντιστοιχίζοντας το άθροισμα στο σύστημα συντεταγμένων της εικόνας, η σχέση (2.36) γίνεται:

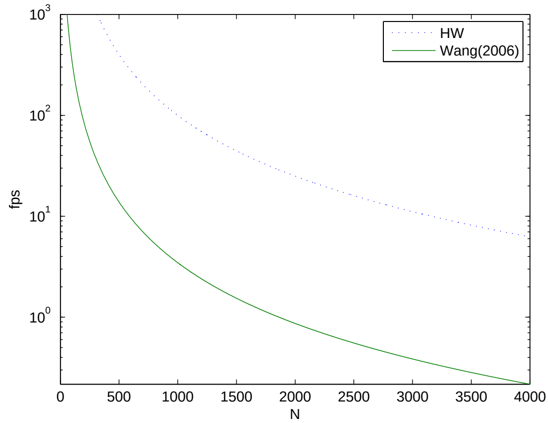
$$\begin{aligned}\lambda_{m,n} &= \frac{1}{4}(2m+1)(2n+1) \\ &\sum_{x}^{N}\sum_{y}^{N} P_m(2x/N-1)P_n(2y/N-1)f(x,y)\end{aligned} \tag{2.37}$$

Από τη σχέση (2.34) προκύπτει:

$$\begin{aligned}P_n(2x/N-1) &= \sum_{j=0}^{n}\alpha_{n,j}(2x/N-1)^j \\ &= \sum_{j=0}^{n}\alpha_{n,j}\sum_{t=0}^{j}\binom{j}{t}\frac{2^t}{N^t}(-1)^{j-t}x^t \\ &= \sum_{t=0}^{n}\frac{2^t}{N^t}x^t\sum_{j=t}^{n}\binom{j}{t}\alpha_{n,j}(-1)^{j-t}\end{aligned} \tag{2.38}$$

Η Σχέση (2.37) μπορεί τώρα να εκφραστεί ως εξής:

$$\lambda_{m,n} = \frac{1}{4}(2m+1)(2n+1)$$

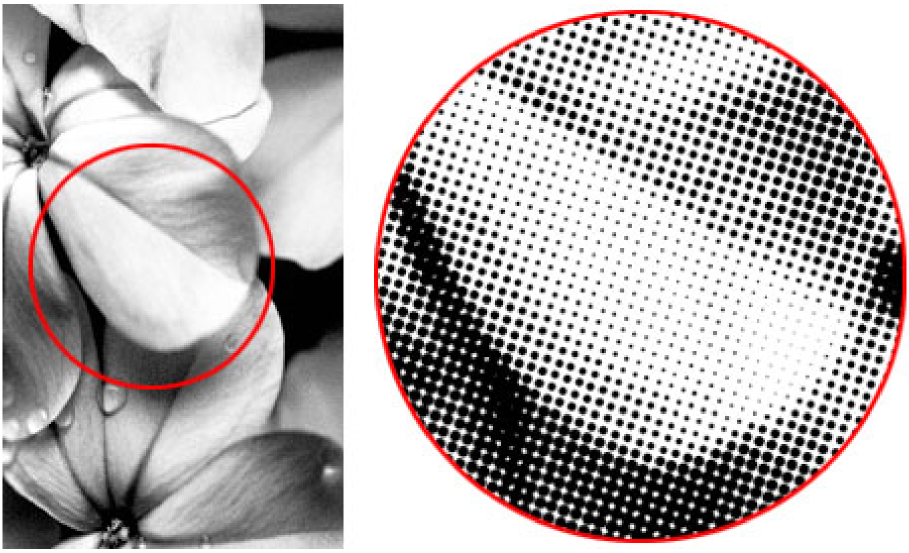


Σχήμα 2.6: Αριθμός επεξεργαζόμενων ανά δευτερόλεπτο εικόνων (*fps*) μεγέθους $N \times N$ για ροπές τάξης μέχρι 16

$$
\begin{aligned}
& \sum_{x}^{N}\sum_{y}^{N}\sum_{t=0}^{n}\frac{2^t}{N^t}x^t\sum_{j=t}^{n}\binom{j}{t}\alpha_{n,j}(-1)^{j-t} \\
& \sum_{s=0}^{m}\frac{2^s}{N^s}y^s\sum_{i=s}^{m}\binom{i}{s}\alpha_{m,i}(-1)^{i-s}f(x,y) \\
= \; & \frac{1}{4}(2m+1)(2n+1) \\
& \sum_{t=0}^{n}\frac{2^t}{N^t}\sum_{j=t}^{n}\binom{j}{t}\alpha_{n,j}(-1)^{j-t} \\
& \sum_{s=0}^{m}\frac{2^s}{N^s}\sum_{i=s}^{m}\binom{i}{s}\alpha_{m,i}(-1)^{i-s}\sum_{y}^{N}\sum_{x}^{N}x^t y^s f(x,y)
\end{aligned}
\tag{2.39}
$$

Εκτός από τον τελευταίο παράγοντα, ο οποίος είναι η γεμετρική ροπή τάξης t,s, όλα τα άλλα μέλη της εξίσωσης (2.39) είναι ανεξάρτητα της υπό ανάλυση εικόνας. Επομένως, χρειάζεται να υπολογιστούν μόνο μία φορά. Τέλος, χρησιμοποιώντας τις σχέσεις που παρατίθενται στο παράρτημα Β΄, μπορεί να εξαχθεί μία σχέση της μορφής:

$$\lambda_{m,n} = \sum_{t=0}^{m}\sum_{s=0}^{n} A_t^m A_s^n M_{t,s} \tag{2.40}$$



Σχήμα 2.7: Παράδειγμα δυαδικής εικόνας η οποία δεν εμφανίζει απλό περίγραμμα

με $M_{t,s}$ τη ροπή συσσώρευσης τάξης t,s. Πιο συγκεκριμένα, εκφράζοντας τα πολυώνυμα *Legendre* σε παραγοντική μορφή, οι συντελεστές τους θα είναι οι A_t^m.

2.5 Εξαγωγή Ροπών Δυαδικών Εικόνων

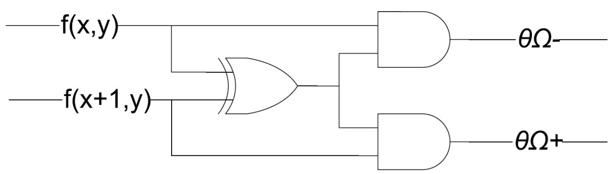
Οι παραπάνω τεχνικές μπορούν να εφαρμοστούν τόσο σε εικόνες αποχρώσεων του γκρι, όσο και σε δυαδικές. Εντούτοις, για δυαδικές εικόνες οι οποίες εμφανίζουν ένα απλό περίγραμμα, προτείνεται παρακάτω μία τεχνική κατάλληλη για πολύ γρήγορη επεξεργασία, η οποία απαιτεί και ελάχιστο αριθμό υπολογιστικών πόρων. Με τον όρο απλό περίγραμμα, εννοούνται εικόνες οι οποίες παρουσιάζουν ένα σχετικά μικρό αριθμό μεταβάσεων από λογικό 1 σε λογικό 0. Αυτές συμπεριλαμβάνουν σχήματα όπως σύμβολα, γράμματα κλπ., αλλά όχι εικόνες που έχουν προκύψει από τεχνικές *halftoning*. Οι εικόνες του Σχήματος 1.8 παρουσιάζουν απλό περίγραμμα, σε αντίθεση με αυτή του Σχήματος 2.7.

2.5.1 Περίγραμμα κατά *Phillips*

Στην εργασία [Phillips, 1993] προτάθηκε ο παρακάτω ορισμός για σημεία περιγράμματος μίας δυαδικής εικόνας. Για ένα πεπερασμένο σύνολο Ω:

$$\begin{aligned} \theta\Omega- &= \{(x,y)|(x,y) \notin \Omega, (x+1,y) \in \Omega\} \\ \theta\Omega+ &= \{(x,y)|(x,y) \in \Omega, (x+1,y) \notin \Omega\} \end{aligned} \quad (2.41)$$

Ο ορισμός αυτός δίνει τη δυνατότητα για πολύ γρήγορη επεξεργασία. Σε αντίθεση με ένα τελεστή διαφορών, τα σημεία $\theta\Omega-$ μπορεί να μην ανήκουν στο πεδίο της εικόνας. Είναι προφανές ότι σε κάθε γραμμή, το πλήθος των σημείων που ανήκουν στο $\theta\Omega-$ θα είναι ίδιο με αυτών που ανήκουν στο $\theta\Omega+$. Η υλοποίηση ενός φίλτρου που θα δίνει αυτά τα σημεία φαίνεται στο σχήμα 2.8



Σχήμα 2.8: Φίλτρο εξαγωγής σημείων περιγράμματος δυαδικής εικόνας

2.5.2 Εξαγωγή συσσωρευτικών ροπών από το περίγραμμα

Περιγραφή διαδικασίας

Η συσσωρευτική ροπή τάξης m, n, για μία δυαδική εικόνα $f(x,y) \in \{0,1\}$ δίνεται ως (Β΄.7):

$$\mu_{m,n} = \sum_{x}^{N} \sum_{y}^{N} f(x,y) x^{\overline{m}} y^{n} \quad (2.42)$$

ή

$$\mu_{m,n} = \sum_{y}^{N} y^{(n)} \sum_{x}^{N} x^{(\overline{m})} f(x,y) \quad (2.43)$$

Σε κάθε γραμμή y υπάρχουν n_y σημεία που ανήκουν στο $\theta\Omega-$ και n_y σημεία που ανήκουν στο $\theta\Omega+$. Για δύο διαδοχικά τέτοια σημεία

$(x^- \in \theta\Omega-,\psi)$, $(x^+ \in \theta\Omega+,y)$, η συσσωρευτική ροπή τάξης m,n του οριζόμενου τμήματος είναι:

$$\mu_{m,n} = y^n \sum_{x=x^-}^{x^+} x^m \tag{2.44}$$

Λόγω της:

$$\sum_{k=0}^{n} k^m = \frac{n^{m+1}}{m+1} \tag{2.45}$$

η (2.44) μπορεί να γραφεί ως:

$$\mu_{mn} = y^n \frac{(x^-)^{m+1} - (x^+)^{m+1}}{m+1} \tag{2.46}$$

Αθροίζοντας για k_y ευθύγραμμα τμήματα για τη γραμμή y:

$$\mu_{mn}(y) = y^n \sum_{i=0}^{k_y} \frac{(x_i^-)^{m+1} - (x_i^+)^{m+1}}{m+1} \tag{2.47}$$

Η τελική έκφραση προκύπτει αθροίζοντας την παραπάνω για κάθε γραμμή:

$$M_{mn} = \sum_{y=0}^{N} y^n \sum_{i=0}^{k_y} \frac{(x_i^-)^{m+1} - (x_i^+)^{m+1}}{m+1} \tag{2.48}$$

ή

$$M_{mn} = (m+1)^{-1} \sum_{y=0}^{N} y^n \sum_{i=0}^{k_y} ((x_i^-)^{m+1} - (x_i^+)^{m+1}) \tag{2.49}$$

Υπολογιστικό κόστος

Η (2.49) μπορεί να χρησιμοποιηθεί για τον γρήγορο υπολογισμό των ροπών συσσώρευσης. Για μία εικόνα N οριζόντιων γραμμών σάρωσης, με n ακραία σημεία ανά γραμμή, κατά μέση τιμή, οι απαιτούμενες πράξεις είναι $3N$ πολλαπλασιασμοί και $2nN+N$ προσθέσεις. Το παραπάνω προκύπτει ως εξής:

- Για κάθε x και y μέχρι τη διάσταση της εικόνας N, απαιτείται ένας πολλαπλασιασμός για τον υπολογισμό του x^m από το x^{m-1}.

- Εφ'οσον έχουν υπολογιστεί τα παραπάνω, απαιτούνται $2n_yN$ προσθέσεις για τον υπολογισμό του αθροίσματος $\sum_{i=1}^{i=n_y}\left((x_i^-)^{\underline{m+1}} - (x_i^+)^{\underline{m+1}}\right)$ για κάθε $y \leq N$.
- Τέλος, $2N$ πολλαπλασιασμοί και N προσθέσεις απαιτούνται για την τελική τιμή κάθε ροπής συσσώρευσης.

Για μία τυπική εικόνα, με $N = 128$ και $n = 4$, τα παραπάνω μεταφράζονται σε 384 πολλαπλασιασμούς και 1.152 προσθέσεις ανά υπολογιζόμενη ροπή.

Σύμφωνα με το θεώρημα 1, στο Παράρτημα Β′, οι γεωμετρικές ροπές μπορούν να εκφραστούν μέσω συσσωρευτικών ροπών. Όπως και στις προηγούμενες παραγράφους του κεφαλαίου, από τις ροπές συσσώρευσης, οι οποίες αποτελούν και τις εξόδους των φίλτρων πόλων, μπορούν να προκύψουν ροπές *Zernike*, *Chebyshev* και *Legendre*.

Παράδειγμα
Έστω μία γραμμή εικόνας, μήκους 64 δειγμάτων, με τιμές $\{0,0,0,1,1,1,1,\ldots,1,1,0\}$. Η σχέση που δίνει την αναλυτική εξαγωγή της συσσωρευτικής ροπής τάξης m είναι:

$$M_m = \sum_{x=0}^{64} f(x)x^{\overline{m}} \tag{2.50}$$

Ακόμα και αν θεωρηθεί ότι έχουν υπολογιστεί εκ των προτέρων όλες οι τιμές $x^{\overline{m}}$, για κάθε $x \in [0,64]$, απαιτούνται 60 προσθέσεις. Στην περίπτωση που δεν ισχύει κάτι τέτοιο, οι απαιτούμενες πράξεις είναι $60+60m$ προσθέσεις και $60+60m$ πολλαπλασιασμοί.

Αν έχουν εξαχθεί τα δύο σημεία περιγράμματος (θέσεις 3 και 63) οι απαιτούμενες πράξεις είναι μία αφαίρεση και μία διαίρεση, με αποθηκευμένα τα $x^{\overline{m}}$, ή $2m+2$ πολλαπλασιασμοί και $2m+3$ προσθέσεις για εξάρχής υπολογισμό.

2.6 Περιγραφέας *ART*

Παρά την εκτεταμένη χρήση του, έχει καταβληθεί λίγη προσπάθεια για τον ταχύ υπολογισμό του περιγραφέα *ART*, ο οποίος είναι ιδιαίτερα σημαντικός σε συστήματα με μειωμένους υπολογιστικούς πόρους. Στην ενότητα αυτή παρουσιάζεται μία αποδοτική υλοποίηση του *ART*, με τη σύνδεσή του με ροπές συσσώρευσης. Οι χρησιμοποιούμενες προσεγγίσεις δεν επηρεάζουν τον τελικό περιγραφέα των 140 ψηφίων, και η αρχιτεκτονική είναι

Ο περιγραφέας *ART* παρουσιάζει την ιδιομορφία της μη πολυωνυμικής συνάρτησης βάσης. Επομένως, δεν είναι δυνατόν να εκφραστεί ως γραμμική συνάρτηση ροπών συσσώρευσης. Για τον ταχύ υπολογισμό του, προτείνεται μία τεχνική στηριζόμενη στην πολυωνυμική προσέγγιση των συναρτήσεων βάσης. Η αρχιτεκτονική που προκύπτει παρουσιάζει μικρότερες απαιτήσεις σε υπολογιστικούς πόρους τόσο σε σχέση με την άμεση μέθοδο καθώς και με τη μέθοδο που προτείνεται στην [Hwang and Kim, 2006].

2.6.1 Προτεινόμενη τεχνική

Το ακτινικό μέρος της συνάρτησης $V_{mn}(r,\theta)$ μπορεί να γραφεί ως εξής:

$$R_n(r,\theta) = 2T_n(\cos(\pi r)) - \delta(n) \tag{2.51}$$

Έστω πολυώνυμο $\tilde{T}$ το οποίο ορίζεται ως εξής:

$$\tilde{T}_n(x) = 2T_n(\cos(\pi r)) - \delta(n) \tag{2.52}$$

με $T_n(x)$ το πολυώνυμο *Chebyshev* πρώτου είδους και τάξης n. Τα πολυώνυμα *Chebyshev* του πρώτου και δεύτερου είδους μπορούν να εκφραστούν ως εξής [Mason, 2003]:

$$\begin{aligned} T_n(\cos(x)) &= cos(nx) \\ U_n(\cos(x)) &= \frac{\sin((n+1)x)}{\sin(x)} \end{aligned} \tag{2.53}$$

Με παρόμοιο τρόπο, το γωνιακό μέρος της $V_{mn}(r,\theta)$ μπορεί να εκφραστεί ως εξής:

$$\begin{aligned} A_m(\theta) &= \frac{\cos(m\theta) + j\sin(m\theta)}{2\pi} \\ &= \frac{T_n(\cos(\theta)) + jU_{n-1}(\cos(\theta))sin(\theta)}{2\pi} \end{aligned} \tag{2.54}$$

Από τις σχέσεις (2.51) και (2.54) φαίνεται ότι η $V_{mn}(r,\theta)$ εκφράσθηκε ως συνάρτηση των $\cos(\pi r)$, $\sin(\theta)$ και $\cos(\theta)$.

Για την εξαγωγή των τιμών των παραπάνω μεταβλητών, μπορεί να χρησιμοποιηθεί πολυωνυμική προσσέγγιση δύο διαστάσεων:

$$\cos(\pi r(x,y)) \approx P_k^{(1)}(x,y) \tag{2.55}$$

με $P_k^{(1)}(x,y)$ ένα πολυώνυμο δύο διαστάσεων βαθμού k.

Ομοίως,

$$\cos(\theta(x,y)) \approx P_k^{(2)}(x,y) \tag{2.56}$$

και

$$\sin(\theta(x,y)) \approx P_k^{(3)}(x,y) \tag{2.57}$$

Η $V_{mn}(r,\theta)$ εκφρασμένη ως πολυώνυμο των (x,y) παίρνει την παρακάτω μορφή:

$$V_{mn}^*(r,\theta) = Q_{mn}(x,y) \tag{2.58}$$

όπου

$$\begin{aligned} Q_{mn}(x,y) \quad = \quad & \frac{1}{2\pi} T_m(P_k^{(1)}(x,y)) \Big(T_n(P_k^{(2)}(x,y)) - \\ & jU_{n-1}(P_k^{(2)}(x,y)) P_k^{(3)}(x,y) \Big) \end{aligned} \tag{2.59}$$

Είναι προφανές ότι το Q είναι βαθμού $(m+n)k$.

Όπως αναλύεται στο Παράρτημα Β΄, κάθε πολυώνυμο μπορεί να εκφραστεί ως πολυώνυμο παραγωντικών δυνάμεων. Επομένως, εκφράζοντας το Q με τέτοιο τρόπο, είναι δυνατή η εξαγωγή των συντελεστών *ART* από ροπές συσσώρευσης. Οι τελευταίες εξάγονται όπως περιγράφεται στην Ενότητα 2.5, αφού ο περιγραφέας ορίζεται μόνο για δυαδικές εικόνες.

Από τον ορισμό του *ART* και τη σχέση (2.58) προκύπτει:

$$F_{m,n} = \int_0^{N-1} \int_0^{N-1} Q_{mn}(x,y) f(x,y) dxdy \tag{2.60}$$

ή

$$F_{m,n} = \sum_{t=0}^{2mnk} \sum_{s=0}^{2mnk} \alpha_{ts}^{mn} \int_0^{N-1} \int_0^{N-1} x^{\underline{m}} y^{\underline{n}} f(x,y) dxdy \tag{2.61}$$

με α_{ts}^{mn} τους συντελεστές του $P_{mn}(x,y)$. Τελικά

$$F_{m,n} = \sum_{t=0}^{2mnk} \sum_{s=0}^{2mnk} \alpha_{ts}^{mn} \mu_{ts} \tag{2.62}$$

2.6.2 Ακρίβεια

Το σφάλμα που εισάγεται στην περιγραφόμενη τεχνική, οφείλεται στην πολυωνυμική προσέγγιση των $\cos(\pi r)$, $\sin(\theta)$ και $\cos(\theta)$. Οι αναλυτικές εκφράσεις των μεταβλητών αυτών είναι:

$$\cos(\pi r) = \cos\left(\pi \frac{\sqrt{(x-\bar{x})^2+(y-\bar{y})^2}}{\max\left[\sqrt{(x-\bar{x})^2+(y-\bar{y})^2}\right]}\right) \tag{2.63}$$

$$\cos(\theta) = \frac{(x-\bar{x})}{\sqrt{(x-\bar{x})^2+(y-\bar{y})^2}} \tag{2.64}$$

$$\sin(\theta) = \frac{(y-\bar{y})}{\sqrt{(x-\bar{x})^2+(y-\bar{y})^2}} \tag{2.65}$$

Προφανώς, ο βαθμός του πολυωνύμου προσέγγισης ορίζει και την ακρίβεια των εξαγόμενων συντελεστών. Για την επίτευξη υψηλών ταχυτήτων επεξεργασίας, οι συντελεστές των πολυωνύμων είναι υπολογισμένοι εκ των προτέρων, και αποθηκευμένοι σε πίνακες αναζήτησης (*LUTs*).

Σχετικά με τον βαθμό προσέγγισης, με τη χρήση πολυωνύμων 3ου βαθμού, το μέσο σφάλμα των συντελεστών ΑΡΤ βρέθηκε στην περιοχή του 10^{-3} για εικόνες μεγέθους 256×256 . Από τη στιγμή που οι συντελεστές *ART* διακριτοποιούνται με ακρίβεια 4 δυαδικών ψηφίων, ο υπολογισμός είναι ακριβής. Για μικρότερες εικόνες, πολυώνυμα παρεμβολής 2ου βαθμού βρέθηκαν έδωσαν ακριβή αποτελέσματα, στην απαιτούμενη ακρίβεια.

Για πιο αποδοτική προσέγγιση, είναι πλεονεκτικός ο χωρισμός της εικόνας σε τεταρτημόρια. Επομένως, οι συντελεστές *ART* υπολογίζονται μέσω 8 αθροισμάτων. 4 για το πραγματικό και 4 για το φανταστικό μέρος τους:

$$\Re(F_{mn}) = \int\int_C \Re(V^*_{mn}(r,x,y))f(x,y)dxdy \tag{2.66}$$

$$\Im(F_{mn}) = \int\int_C \Im(V^*_{mn}(r,x,y))f(x,y)dxdy \tag{2.67}$$

με $C = \{0\ldots\bar{x}, 0\ldots\bar{y}\}$ ή $C = \{0\ldots\bar{x}, \bar{y}\ldots N\}$ ή $C = \{\bar{x}\ldots N, 0\ldots\bar{y}\}$ ή $C = \{\bar{x}\ldots N, \bar{y}\ldots N\}$

2.6.3 Αποτίμηση επιδόσεων

Παρακάτω δίνεται ένα περίγραμμα της προτεινόμενης τεχνικής:

1. Εξαγωγή των πολυωνύμων δύο διαστάσεων για την προσέγγιση των $\cos(\pi r)$, $\sin(\theta)$ και $\cos(\theta)$, στα 4 τεταρτημόρια.

2. Εξαγωγή των συντελεστών των 2 πολυωνύμων (πραγματικών και φανταστικών) της σχέσης (2.59). Χρησιμοποιώντας παρεμβολή 2ου βαθμού, αυτοί είναι 729 συντελεστές. Για παρεμβολή 3ου βαθμού, το πλήθος τους αυξάνεται σε 1.600.

3. Υπολογισμός των τριών πρώτων γεωμετρικών ροπών τις εικόνας.

4. Υπολογισμός του βαρύκεντρου της εικόνας.

5. Υπολογισμός των ροπών συσσώρευσης, όπως περιγράφεται στην Ενότητα 2.5, για τα 4 τεταρτημόρια της εικόνας .

6. Εξαγωγή των συντελεστών *ART* σύμφωνα με την (2.62).

7. Κανονικοποίηση με την F_{00} και διακριτοποίηση σε 4 δυαδικά ψηφία βάσει πίνακα διακριτοποίησης.

Τα βήματα 1 και 2 μπορούν να υπολογισθούν ανεξάρτητα της εικόνας και επομένως, θεωρούνται προϋπολογισμένα. Το βασικό πλεονέκτημα της τεχνικής σε σχέση με τη συμβατική, καθώς και με αυτή της [Hwang and Kim, 2006] είναι οι πολύ μικρές απαιτήσεις σε μνήμη. Για έναν αριθμό από 128 πιθανών μεγίστων ακτίνων, απαιτείται η αποθήκευση 186.624 συντελεστών. Εντούτοις, αν θεωρήσουμε τη μέγιστη αυτή τιμή φραγμένη σε μικρότερη ακτίνα, αυτές οι απαιτήσεις μειώνονται σημαντικά. Για παράδειγμα, επιτρέποντας 16 τιμές ακτίνων, το μέγεθος των πινάκων αναζήτησης γίνεται 11.664.

Η συμβατική μέθοδος θα απαιτούσε 35 πίνακες αναζήτησης ίδιου μεγέθους με την εικόνα, ενώ η τεχνική της [Hwang and Kim, 2006] μειώνει το μέγεθος αυτό στο 25%. Επομένως, για μία εικόνα μεγέθους 256×256, οι απαιτούμενοι συντελεστές είναι 2.293.760 και 573.440 αντίστοιχα.

Τα βήματα 3 και 4 απαιτούν ένα μικρό αριθμό πράξεων Πιο συγκεκριμένα, για τον υπολογισμό των τριών πρώτων ροπών απαιτούνται $6nN+6N+3$ πολλαπλασιασμοί και $6nN+3N$ προσθέσεις, ακολουθούμενοι από 2 διαιρέσεις.

Όπως περιγράφηκε στην Ενότητα 2.5, για εικόνα N οριζόντιων γραμμών σάρωσης, με κατά μέσο όρο n σημεία περιγράμματος ανά

Πίνακας 2.5: Υπολογιστική πολυπλοκότητα των μεθόδων εξαγωγής των συντελεστών *ART* για εικόνα μεγέθους $N \times N$

Μέθοδος	# Προσθέσεων	#Πολλαπλασιασμών
Συμβατική	$72N^2$	$72N^2$
[Hwang and Kim, 2006]	$72N^2$	$16N^2$
Προτεινόμενη (προς. 3ου βαθμού)	$4,400N + 182,270$	$1,600N + 182,160$
Προτεινόμενη (προς. 2ου βαθμού)	$3,520N + 81,070$	$1,280 + 81,070$

Πίνακας 2.6: Απαιτήσεις μνήμης για τον υπολογισμό των συντελεστών *ART* για εικόνα μεγέθους $N \times N$

Μέθοδος	Αριθμός τιμών
Συμβατική	$72N^2$
Hwang & Kim [Hwang and Kim, 2006]	$16N^2$
Προτεινόμενη(προς. 3ου βαθμού)	$3,200 radius\_range$
Προτεινόμενη(προς. 2ου βαθμού)	$1,458 radius\_range$

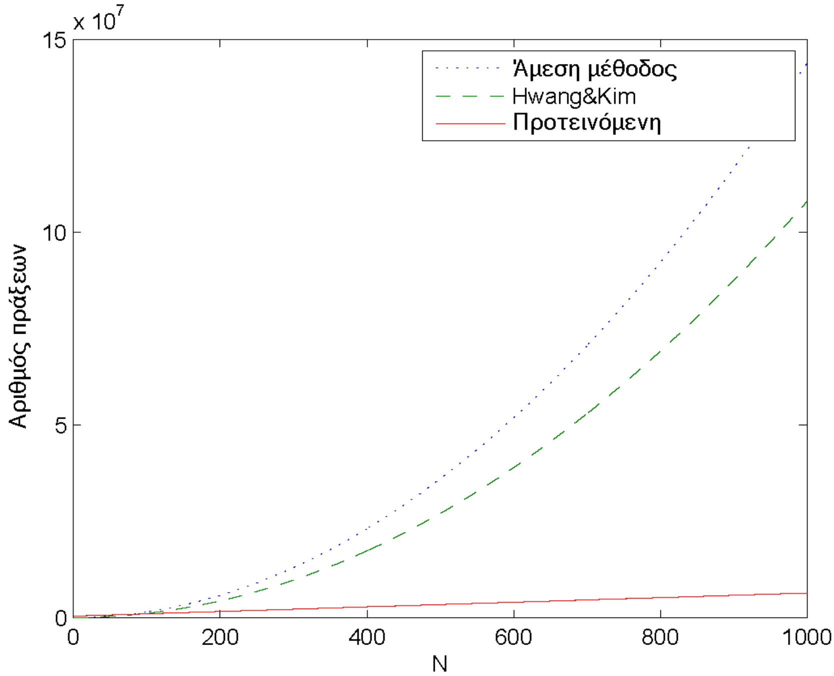
γραμμή, οι απαιτούμενες πράξεις είναι $2nN + 2N + 1$ πολλαπλασιασμοί και $2nN + N$ προσθέσεις για το βήμα 5. Από τη στιγμή που οι ροπές αυτές υπολογίζονται και για τα 4 τεταρτημόρια, οι συνολικές πράξεις είναι το πολύ $(4nN + 4N + 4)M$ πολλαπλασιασμοί και $(4nN + 4N)M$ προσθέσεις. Επειδή οι πράξεις αυτές εκτελούνται σε μιγαδικούς αριθμούς, το πλήθος των πράξεων από υπολογιστική άποψη είναι διπλάσιο από το προαναφερθέν.

Η σχέση (2.62) απαιτεί $2(2mnk)^2$ πολλαπλασιασμούς και $2(2mnk)^2$ αθροίσεις ανά συντελεστή. Επομένως, για τον υπολογισμό και των 36 συντελεστών *ART* απαιτούνται 80.960 προσθέσεις και πολλαπλασιασμοί, για προσέγγιση 2ου βαθμού ή 182.160 για 3ου. Η άμεση μέθοδος θα απαιτούσε $2 \times 36N^2$. Στην μέθοδο της [Hwang and Kim, 2006], το πλήθος των πολλαπλασιασμών υποτετραπλασιάζεται, ενώ αυτό των αθροίσεων μένει σταθερό. Στον Πίνακα 2.5 παρουσιάζεται μία επισκόπηση της υπολογιστικής πολυπλοκότητας των τριών μεθόδων, ενώ στον 2.6 των απαιτήσεων σε μνήμη.

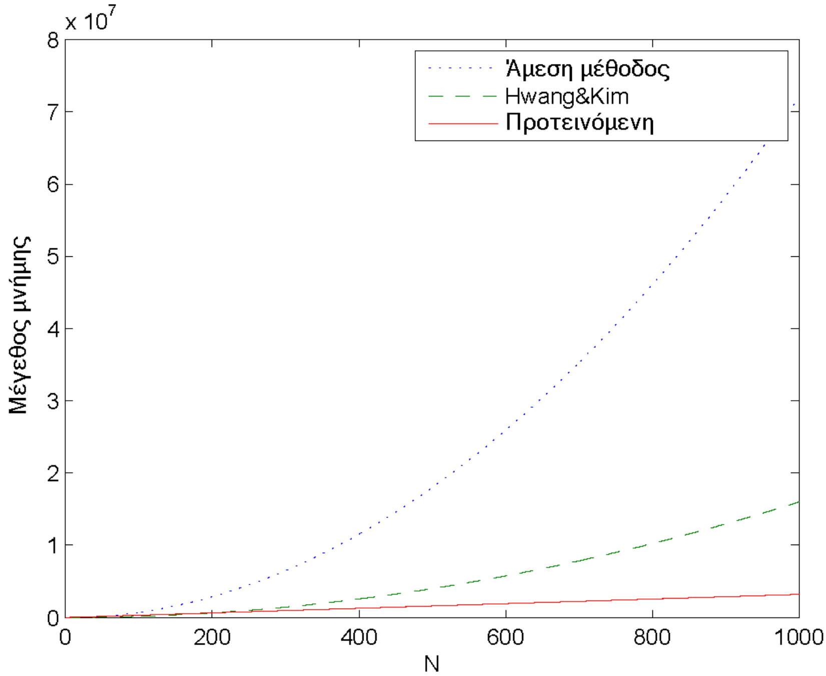
Στα σχήματα 2.9 και 2.10 παρουσιάζονται γραφικές παραστάσεις της υπολογιστικής πολυπλοκότητας και των απαιτήσεων σε μνήμη για τις τρεις τεχνικές. Φαίνεται ότι η προτεινόμενη τεχνική παρουσιάζει μεγάλο πλεονέκτημα σε σχέση με τις προηγούμενες, όταν οι εικόνες προς ανάλυση είναι μεγάλες. Πιο συγκεκριμένα, για εικόνες μεγέθους μέχρι περίπου 40×40, η τεχνική των Hwang & Kim [Hwang and Kim, 2006] είναι οριακά γρηγορότερη. Εντούτοις, για εικόνες μεγέθους 128×128 απαιτεί 40% λιγότερους πόρους, ενώ για 256×256 η βελτίωση είναι 70%. Για εικόνες μεγέθους *VGA*, η βελτίωση σε πόρους και ταχύτητα είναι 89%, η οποία αντιστοιχεί σε περίπου μία τάξη μεγέθους.

2.6.4 Συμπεράσματα

Στην ενότητα αυτή, παρουσιάστηκε μία νέα τεχνική για τον γρήγορο και αποδοτικό, σε σχέση με τη μνήμη, υπολογισμό του μετασχηματισμού *ART*. Προσεγγίζοντας τις συναρτήσεις βάσης του *ART* μέσω πολυωνύμων, καθιστάται δυνατή η έκφραση των συντελεστών μέσω ροπών συσσώρευσης. Με τη χρήση της τεχνικής που περιγράφηκε σε προηγούμενη ενότητα για την εξαγωγή των ροπών αυτών για δυαδικές εικόνες, το τελικό σύστημα παρουσιάζει βελτίωση τόσο στον τομέα της ταχύτητας, όσο και στις απαιτήσεις μνήμης. Η τελική βελτίωση είναι περίπου μία τάξη μεγέθους για εικόνες μεγέθους *VGA*.



Σχήμα 2.9: Πλήθος απαιτούμενων πράξεων για εικόνα μεγέθους $N \times N$



Σχήμα 2.10: Μέγεθος απαιτούμενης μνήμης για εικόνα μεγέθους $N \times N$

2.7 Σύνοψη κεφαλαίου-Συμπεράσματα

Στο κεφάλαιο αυτό παρουσιάστηκαν τεχνικές για την αποδοτική εξαγωγή χαρακτηριστικών ροπών εικόνων. Οι τεχνικές αυτές μπορούν να υλοποιηθούν σε υλικό ή λογισμικό και επιτυγχάνουν πολύ σημαντική βελτίωση στους απαιτούμενους υπολογιστικούς πόρους. Πιο συγκεκριμένα, οι τεχνικές αυτές στοχεύουν στην ταχεία εξαγωγή των παρακάτω χαρακτηριστικών σχήματος:

- Γεωμετρικές ροπές
- Ροπές *Zernike*
- Ροπές *Chebyshev*
- Ροπές *Legendre*
- Περιγραφέας *ART*

Η βασική αρχή των παραπάνω αρχιτεκτονικών είναι η χρήση των συσσωρευτικών ροπών οι οποίες προκύπτουν από απλά φίλτρα πόλων. Σε κάθε περίπτωση, οι προτεινόμενες τεχνικές είναι οι πιο ταχείς που αναφέρονται στη βιβλιογραφία, ακόμα και σε υλοποίησή τους σε λογισμικό. Οι αρχιτεκτονικές υλικού παρουσιάζουν βελτιωμένη ταχύτητα κατά περίπου 1 τάξη μεγέθους, σε σχέση με άλλες ταχείς μεθόδους, ενώ η υλοποίηση σε λογισμικό παρουσιάζει βελτίωση κατά περίπου 50%. Με τη χρήση τους, καθιστάται δυνατός ο υπολογισμός των περιγραφέων αυτών σε πραγματικό χρόνο ακόμα και για εικόνες μεγάλου μεγέθους.

Κεφάλαιο 3

Ακριβής υπολογισμός ροπών

Σε όλα τα ψηφιακά συστήματα που σχετίζονται με την επεξεργασία και την ανάλυση εικόνας απαντώνται σφάλματα που σχετίζονται με την διακριτοποίηση των πραγματικών σκηνών. Οι ροπές εικόνων επηρεάζονται ιδιαίτερα από αυτά, αφού ουσιαστικά αποτελούν ολοκληρωτικούς μετασχηματισμούς, με συναρτήσεις βάσης σε όλο το εύρος των συχνοτήτων. Στο κεφάλαιο αυτό παρουσιάζεται μία μεθοδολογία για την ελαχιστοποίηση των σφαλμάτων αυτών, με τη χρήση πολυωνυμικής παρεμβολής.

Οι ορισμοί για διάφορους τύπους ροπών, οι οποίοι χρησιμοποιούνται εκτενώς σε ένα μεγάλο αριθμό εφαρμογών, δόθηκαν στα προηγούμενα Κεφάλαια.Επίσης, προτάθηκαν τεχνικές για την ταχεία εξαγωγή των χαρακτηριστικών αυτών, που επιτρέπουν τη χρήση τους σε πραγματικό χρόνο. Πέρα από το θόρυβο, ο οποίος είναι εγγενής σε κάθε εικόνα, ο υπολογισμός των ροπών υπόκειται σε σφάλματα που σχετίζονται με τη διακριτή προσέγγιση του συνεχούς πεδίου ορισμού τους, την διακριτοποίηση των τιμών τους κ.λ.π. Στο κεφάλαιο αυτό, προτείνεται μία τεχνική για τη μείωση των σφαλμάτων αυτών, με τη χρήση πολυωνυμικής παρεμβολής. Πέρα από την προφανή βελτίωση της ακρίβειας των τιμών των εξαγόμενων περιγραφέων, βελτιώνονται και χαρακτηριστικά όπως η αμεταβλητότητά τους στην περιστροφή ή την κλιμάκωση. Η τεχνική απαιτεί πολύ μικρό υπολογιστικό κόστος, σε σχέση με τον ίδιο τον υπολογισμό των ροπών. Επίσης, δίνονται περιγραφές για γεωμετρικές ροπές, καθώς και για ροπές *Zernike*.

3.1 Εισαγωγή

Μία γεωμετρική ροπή τάξης p, q ενός σήματος $f : C \subset R^2 \to R$ δίνεται από την παρακάτω σχέση:

$$m_{p,q} = \int_C \int x^p y^q f(x,y) dxdy \qquad (3.1)$$

Στην περίπτωση, όμως, των ψηφιακών εικόνων, η παραπάνω σχέση προσεγγίζεται από το διακριτό ανάλογό της:

$$m_{p,q} = \sum_{0}^{N-1} \sum_{0}^{M-1} x^p y^q f(x,y) dxdy \qquad (3.2)$$

Το αποτέλεσμα αυτής της προσέγγισης είναι η εισαγωγή σφαλμάτων διακριτοπίησης, τα οποία εμφανίζονται ιδιαίτερα σε εικόνες μικρού μεγέθους ή σε ροπές υψηλής τάξης. Το σφάλμα αυτό δημιουργεί σημαντικά προβλήματα στις ιδιότητες αμεταβλητότητας των ροπών.

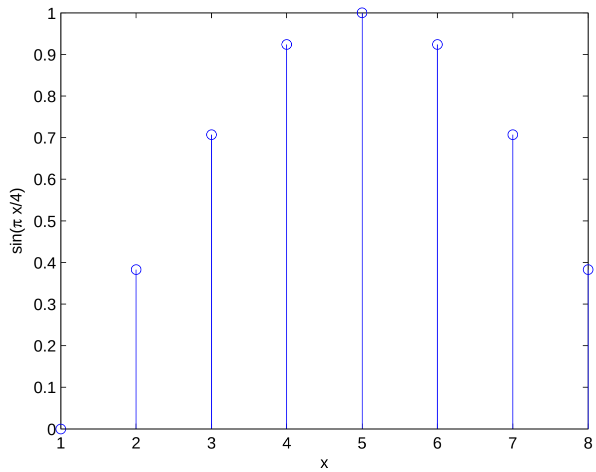
3.2 Παρεμβολή

3.2.1 Περιγραφή - Βιβλιογραφική έρευνα

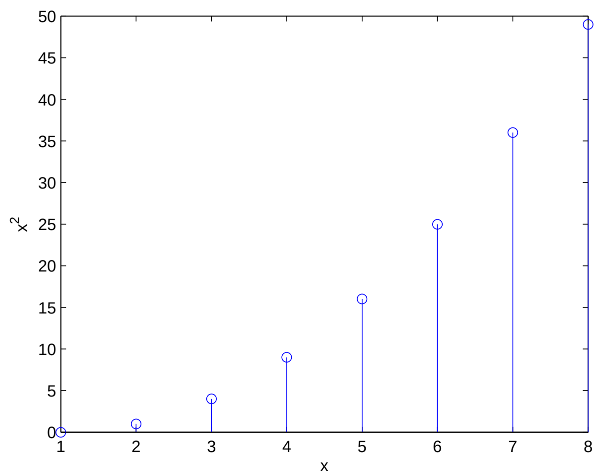
Η λογική πίσω στην οποία στηρίζεται η βελτίωση της ακρίβειας βάσει παρεμβολής δίνεται στα παρακάτω σχήματα.

Στο Σχήμα 3.1 παρουσιάζεται η συνάρτηση $sin(\pi x/16$, για 9 δείγματα. Στο Σχήμα 3.2 παρουσιάζεται η συνάρτηση βάσης x^2 (Γεωμετρική ροπή 2ης τάξης), ενώ στο 3.3 ο όρος $f(x)x^2$, ο οποίος μετά από προσεγγιστική ολοκλήρωση θα δώσει την αντίστοιχη ροπή.

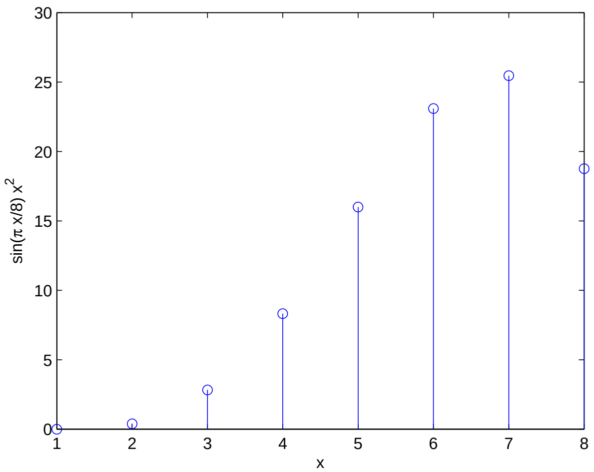
Η προσέγγιση που προτείνεται στο Κεφάλαιο αυτό, στηρίζεται στην παρεμβολή του τρίτου σήματος, στο χώρο των ροπών. Στη βιβλιογραφία περιγράφεται ένας αριθμός άλλων τεχνικών για τον ακριβή υπολογισμό των ροπών κατά τις οποίες εφαρμόζεται πολυωνυμική παρεμβολή στις συναρτήσεις βάσης.



Σχήμα 3.1: Σήμα 9 δειγμάτων



Σχήμα 3.2: Η συνάρτηση βάσης x^2



Σχήμα 3.3: Ο ολοκληρωτέος όρος, $f(x)x^2$

Στην [Flusser, 1998] προτείνεται μία τεχνική για το γρήγορο και ακριβή υπολογισμό των ροπών δυαδικών εικόνων, βάσει του περιγράμματος τους, με τρόπο παρόμοιο με αυτόν που περιγράφηκε στην Ενότητα 2.5. Επιλύοντας αναλυτικά το ολοκλήρωμα των συναρτήσεων βάσης των γεωμετρικών ροπών για τιμές στο διάστημα $[-1,1]$ και ολοκληρώνοντας εκ των προτέρων για διάφορα ευθύγραμμα τμήματα, κατέστη δυνατός ο ταχύς υπολογισμός των ροπών, με αυξημένη ακρίβεια.

Στην [Liao and Pawlak, 1996] δόθηκαν αναλυτικές περιγραφές για τις πηγές σφαλμάτων στον υπολογισμό των ροπών, καθώς και μία τεχνική για την μείωσή τους. Τα αποτελέσματα επεκτάθηκαν αργότερα στην [Liao and Pawlak, 1998], στην οποία γίνεται παρόμοια διερεύνηση για τις ροπές *Zernike*.

Η βασική διαφορά της προτεινόμενης τεχνικής με τις παραπάνω, είναι ότι στις παραπάνω το σήμα εισόδου θεωρείται σημειακά σταθερό. Πια αναλυτικά, για κάθε σημείο του μονοδιάστατου ή δισδιάστατου σήματος, θεωρείται $f(x,y) = f(\lfloor x, \lfloor y)$. Η γραμμική ή και πολυωνυμική παρεμβολή αναφέρεται μόνο στη συνάρτηση βάσης. Όπως φαίνεται και στα πειραματικά αποτελέσματα, σε επόμενη ενότητα, η προτεινόμενη τεχνική υπερτερεί έναντι των παραπάνω τόσο σε ακρίβεια, όσο και σε ταχύτητα, αφού ήδη υπάρχουσες τεχνικές εξαγωγής ροπών μπορούν να χρησιμοποιηθούν.

3.2.2 Γραμμική παρεμβολή

1Δ περίπτωση

Μία συνεχής συνάρτηση μπορεί να προσεγγισθεί γραμμικά χρησιμοποιώντας την παρακάτω σχέση:

$$f(x) = f(\lfloor x \rfloor + \alpha) = (1-\alpha)f(\lfloor x \rfloor) + \alpha f(\lfloor x \rfloor + 1) \tag{3.3}$$

Η ροπή τάξης p γίνεται:

$$M'_p = \sum_{x=0}^{N-1} \int_0^1 \left((1-\alpha)f(\lfloor x \rfloor) + \alpha f(\lfloor x \rfloor + 1)\right)(\lfloor x \rfloor + \alpha)^p d\alpha \tag{3.4}$$

Για το πρώτο άθροισμα στο ολοκλήρωμα ισχύει:

$$(1-\alpha)\sum_{x=0}^{N-1} f(\lfloor x \rfloor)(\lfloor x \rfloor + \alpha)^p = (1-\alpha)\sum_{x=0}^{N-1} f(\lfloor x \rfloor)\sum_{m=0}^{p}\binom{p}{m}\lfloor x \rfloor^m \alpha^{p-m} \tag{3.5}$$

ή

$$(1-\alpha)\sum_{x=0}^{N-1} f(\lfloor x \rfloor)(\lfloor x \rfloor + \alpha)^p = (1-\alpha)\sum_{m=0}^{p}\binom{p}{m}\alpha^{p-m}\sum_{x=0}^{N-1} f(\lfloor x \rfloor)\lfloor x \rfloor^m \tag{3.6}$$

και τελικά

$$(1-\alpha)\sum_{m=0}^{p}\binom{p}{m}\alpha^{p-m}M_m \tag{3.7}$$

Για το δεύτερο άθροισμα:

$$\alpha\sum_{x=0}^{N-1} f(\lfloor x \rfloor + 1)(\lfloor x \rfloor + \alpha)^p = \alpha\sum_{x=0}^{N-1} f(\lfloor x \rfloor + 1)(\lfloor x \rfloor + 1 + \alpha - 1)^p \tag{3.8}$$

και

$$\alpha\sum_{x=0}^{N-1} f(\lfloor x \rfloor + 1)(\lfloor x \rfloor + \alpha)^p = \alpha\sum_{x=0}^{N-1} f(\lfloor x \rfloor + 1)\sum_{m=0}^{p}\binom{p}{m}(\lfloor x \rfloor + 1)^m(\alpha - 1)^{p-m} \tag{3.9}$$

ή

$$\alpha\sum_{x=0}^{N-1} f(\lfloor x \rfloor + 1)(\lfloor x \rfloor + \alpha)^p = \alpha\sum_{m=0}^{p}\binom{p}{m}(\alpha - 1)^{p-m}\sum_{x=0}^{N-1} f(\lfloor x \rfloor + 1)(\lfloor x \rfloor + 1)^m \tag{3.10}$$

Θέτοντας $t = \lfloor x \rfloor + 1$:

$$\alpha \sum_{x=0}^{N-1} f(\lfloor x \rfloor + 1)(\lfloor x \rfloor + \alpha)^p = \alpha \sum_{m=0}^{p} \binom{p}{m} (\alpha - 1)^{p-m} \sum_{t=1}^{N} f(t)(t)^m \tag{3.11}$$

ή

$$\alpha \sum_{m=0}^{p} \binom{p}{m} (\alpha - 1)^{p-m} (\sum_{t=0}^{N-1} f(t)(t)^m - f(0)(0)^m + f(N)N^m) \tag{3.12}$$

, θεωρώντας $f(N) = 0$:[1] άρα

$$\alpha \sum_{m=0}^{p} \binom{p}{m} (\alpha - 1)^{p-m} \sum_{t=0}^{N-1} (f(t)t^m - f(0)(0)^m) \tag{3.13}$$

ή

$$\alpha \sum_{m=0}^{p} \binom{p}{m} (\alpha - 1)^{p-m} \sum_{t=0}^{N-1} f(t)t^m - \alpha(\alpha - 1)^p f(0) \tag{3.14}$$

και τελικά

$$\alpha \sum_{m=0}^{p} \binom{p}{m} (\alpha - 1)^{p-m} M_t + \alpha(\alpha - 1)^p f(0) \tag{3.15}$$

Από τις σχέσεις 3.7 και 3.15:

$$M'_p = \int_0^1 (1-\alpha) \sum_{m=0}^{p} \binom{p}{m} \alpha^{p-m} M_m + \alpha \sum_{m=0}^{p} \binom{p}{m} (\alpha - 1)^{p-m} M_m + \alpha(\alpha - 1)^p f(0) d\alpha \tag{3.16}$$

και μετά την ολοκλήρωση προκύπτει:

$$\begin{aligned} M'_p \quad = \quad & \sum_{m=0}^{p} \frac{1}{(p-m+1)(p-m+2)} \binom{p}{m} M_m + \\ & \sum_{m=0}^{p} \binom{p}{m} \frac{(-1)^{p-m}}{(p-m+1)(p-m+2)} M_m + \frac{(-1)^p}{(p+1)(p+2)} f(0) \end{aligned} \tag{3.17}$$

[1] Τα δείγματα του σήματος βρίσκονται στα σημεία $[0, \ldots, N-1]$. Δεν υπάρχει, δηλαδή, δείγμα στο N και επομένως η τιμή $f(N)$ μπορεί να θεωρηθεί ίση με μηδέν.

Τελικά,

$$M'_p = \sum_{m=0}^{p} \binom{p}{m} \frac{M_m(1+(-1)^{p-m})}{(p-m+1)(p-m+2)} + \frac{(-1)^p}{(p+1)(p+2)} f(0) \tag{3.18}$$

$$M'_p = \frac{1}{(p+1)(p+2)} \sum_{m=0}^{p} \binom{p+2}{m} (M_m(1+(-1)^{p-m})) + \frac{(-1)^p}{(p+1)(p+2)} f(0) \tag{3.19}$$

Ορίζοντας

$$G_p = \frac{1}{(p+1)(p+2)} \sum_{m=0}^{p} x^m \binom{p+2}{m} (1+(-1)^{p-m}) \tag{3.20}$$

προκύπτει τελικά μία καλύτερη προσέγγιση στις συνεχείς γεωμετρικές ροπές, χρησιμοποιώντας ως πυρήνα όχι τα μονώνυμα x^p αλλά τα πολυώνυμα $G_p(x)$:

$$\int x^p f(x) dx \approx \sum_{x=0}^{N-1} f(x) G_p(x) \tag{3.21}$$

με

$$\left| \int x^p f(x) dx - \sum_{x=0}^{N-1} f(x) G_p(x) \right| \le \left| \int x^p f(x) dx - \sum_{x=0}^{N-1} f(x) x^p \right| \tag{3.22}$$

Οι τελευταίες εκφράσεις ισχύουν κατά προσέγγιση, αφού θεωρείται ότι $f(0) << \sum f(x)$. Παρόλο που μία τέτοια θεώρηση ισχύει στη γενική περίπτωση, σε ορισμένα σήματα μπορεί να οδηγήσει σε σημαντικά σφάλματα. Το σημαντικό πλεονέκτημα της προσέγγισης αυτής, όμως, είναι η απλή έκφραση των γεωμετρικών ροπών αυξημένης ακρίβειας, σε σχέση με την συμβατική, διακριτή προσέγγιση.

Εκφράσεις για τα πρώτα επτά πολυώνυμα δίνονται παρακάτω:

$$\begin{aligned}
P_0(x) &= 1 \\
P_1(x) &= x \\
P_2(x) &= x^2 + \frac{1}{6} \\
P_3(x) &= x^3 + \frac{1}{2}x \\
P_4(x) &= x^4 + x^2 + \frac{1}{15}
\end{aligned}$$

$$
\begin{aligned}
P_5(x) &= x^5 + \frac{5}{3}x^3 + \frac{1}{3}x \\
P_6(x) &= x^6 + \frac{5}{2}x^4 + x^2 + \frac{1}{28}
\end{aligned} \tag{3.23}
$$

2Δ περίπτωση

Η γραμμική παρεμβολή στις δύο διαστάσεις μπορεί να εκφραστεί με την παρακάτω σχέση:

$$
\begin{aligned}
M'_{p,q} &= \int_0^N \int_0^N f(x,y)x^p y^q dxdy \\
&= \sum_{x=0}^{N-1}\sum_{y=0}^{N-1} \int_0^1 \int_0^1 f(\lfloor x \rfloor + \alpha, \lfloor y \rfloor + b)(\lfloor x \rfloor + \alpha)^p(\lfloor y \rfloor + b)^q d\alpha db
\end{aligned} \tag{3.24}
$$

και

$$
\begin{aligned}
M'_{p,q} = & \int_0^N \int_0^N f(x,y)x^p y^q dxdy \\
= & \sum_{x=0}^{N-1}\sum_{y=0}^{N-1} \int_0^1 \int_0^1 (1-\alpha)(1-b)f(\lfloor x \rfloor, \lfloor y)) \\
& +\alpha(1-b)f(\lfloor x \rfloor + 1, \lfloor y \rfloor) + (1-\alpha)bf(\lfloor x \rfloor, \lfloor y \rfloor + 1) \\
& +\alpha bf(\lfloor x \rfloor + 1, \lfloor y \rfloor + 1)(\lfloor x \rfloor + \alpha)^p(\lfloor y \rfloor + b)^q d\alpha db
\end{aligned} \tag{3.25}
$$

Για τον πρώτο όρο ισχύει

$$
A_1 = \sum_{x=0}^{N-1}\sum_{y=0}^{N-1}(1-\alpha)(1-b)f(\lfloor x \rfloor, \lfloor y))(\lfloor x \rfloor + \alpha)^p(\lfloor y \rfloor + b)^q \tag{3.26}
$$

και με τη χρήση του δυωνυμικού θεωρήματος

$$
A_1 = \sum_{x=0}^{N-1}\sum_{y=0}^{N-1}(1-\alpha)(1-b)f(\lfloor x \rfloor, \lfloor y)) \sum_{m=0}^{p} \binom{p}{m} (\lfloor x \rfloor)^m \alpha^{p-m} \sum_{n=0}^{q} \binom{q}{n} (\lfloor y \rfloor)^n b^{q-n} \tag{3.27}
$$

Ανακατατάσσοντας την παρακάτω έκφραση:

$$A_1 = (1-\alpha)(1-b)\sum_{n=0}^{q}\binom{q}{n}\sum_{m=0}^{p}\binom{p}{m}b^{q-n}\alpha^{p-m}\sum_{x=0}^{N-1}\sum_{y=0}^{N-1}f(\lfloor x\rfloor, \lfloor y)(\lfloor x\rfloor)^m(\lfloor y\rfloor)^n \tag{3.28}$$

Το τελευταίο διπλό άθροισμα είναι η γεωμετρική ροπή τάξης m,n:

$$A_1 = (1-\alpha)(1-b)\sum_{n=0}^{q}\binom{q}{n}\sum_{m=0}^{p}\binom{p}{m}b^{q-n}\alpha^{p-m}M_{m,n} \tag{3.29}$$

Για τον δεύτερο όρο:

$$A_2 = \sum_{x=0}^{N-1}\sum_{y=0}^{N-1}\alpha(1-b)f(\lfloor x\rfloor+1, \lfloor y\rfloor)(\lfloor x\rfloor+\alpha)^p(\lfloor y\rfloor+b)^q \tag{3.30}$$

ή

$$A_2 = \sum_{x=0}^{N-1}\sum_{y=0}^{N-1}\alpha(1-b)f(\lfloor x\rfloor+1, \lfloor y\rfloor)(\lfloor x\rfloor+1+\alpha-1)^p(\lfloor y\rfloor+b)^q \tag{3.31}$$

πάλι, με τη χρήση του διωνυμικού θεωρήματος

$$\begin{aligned} A_2 \;=\; & \sum_{x=0}^{N-1}\sum_{y=0}^{N-1}\alpha(1-b)f(\lfloor x\rfloor+1, \lfloor y\rfloor) \\ & \sum_{m=0}^{p}\binom{p}{m}(\lfloor x\rfloor+1)^m(\alpha-1)^{p-m}\sum_{n=0}^{q}\binom{q}{n}(\lfloor y\rfloor)^n b^{q-n} \end{aligned} \tag{3.32}$$

Θέτοντας $t = \lfloor x\rfloor + 1$:

$$A_2 = \alpha(1-b)\sum_{m=0}^{p}\binom{p}{m}\sum_{n=0}^{q}\binom{q}{n}(\alpha-1)^{p-m}b^{q-n}\sum_{t=1}^{N}\sum_{y=0}^{N-1}f(t, \lfloor y\rfloor)t^m(\lfloor y\rfloor)^n \tag{3.33}$$

αναπτύσσοντας το τελευταίο διπλό άθροισμα:

$$\sum_{t=1}^{N}\sum_{y=0}^{N-1}f(t, \lfloor y\rfloor)t^m(\lfloor y\rfloor)^n = \sum_{t=0}^{N-1}\sum_{y=0}^{N-1}f(t, \lfloor y\rfloor)t^m(\lfloor y\rfloor)^n - \sum_{y=0}^{N-1}f(0, \lfloor y\rfloor)0^m(\lfloor y\rfloor)^n \tag{3.34}$$

Ο δεύτερος όρος είναι μη μηδενικός μόνο για $m = 0$, ενώ ο πρώτος είναι η διακριτή γεωμετρική ροπή τάξης m, n:

$$\begin{aligned} A_2 &= \alpha(1-b)\sum_{m=0}^{p}\binom{p}{m}\sum_{n=0}^{q}\binom{q}{n}(\alpha-1)^{p-m}b^{q-n}\times \\ &\left(M_{m,n} - \sum_{y=0}^{N-1} f(0,\lfloor y\rfloor)0^m(\lfloor y\rfloor)^n\right) \end{aligned} \tag{3.35}$$

άρα

$$\begin{aligned} A_2 &= \alpha(1-b)\sum_{m=0}^{p}\binom{p}{m}\sum_{n=0}^{q}\binom{q}{n}(\alpha-1)^{p-m}b^{q-n}M_{m,n} \\ &-\alpha(1-b)\sum_{n=0}^{q}(\alpha-1)^{p}b^{q-n}\binom{q}{n}\sum_{y=0}^{N-1} f(0,\lfloor y\rfloor)(\lfloor y\rfloor)^n \end{aligned} \tag{3.36}$$

Για τον τρίτο όρο ισχύει:

$$A_3 = \sum_{x=0}^{N-1}\sum_{y=0}^{N-1}(1-\alpha)bf(\lfloor x\rfloor, \lfloor y\rfloor + 1)(\lfloor x\rfloor + \alpha)^p(\lfloor y\rfloor + b)^q \tag{3.37}$$

Ακολουθώντας παρόμοια διερεύνηση με αυτή του προηγούμενου όρου, προκύπτει:

$$\begin{aligned} A_3 &= b(1-\alpha)\sum_{n=0}^{q}\binom{q}{n}\sum_{m=0}^{p}\binom{p}{m}(b-1)^{q-n}\alpha^{p-m}M_{m,n} \\ &-b(1-\alpha)\sum_{m=0}^{p}\binom{p}{m}(b-1)^{q}\alpha^{p-m}\sum_{x=0}^{N-1} f(\lfloor x\rfloor, 0)(\lfloor x\rfloor)^m \end{aligned} \tag{3.38}$$

Για τον τέταρτο όρο:

$$\begin{aligned} A_4 &= \sum_{x=0}^{N-1}\sum_{y=0}^{N-1}\alpha bf(\lfloor x\rfloor + 1, \lfloor y\rfloor + 1)(\lfloor x\rfloor + \alpha)^p(\lfloor y\rfloor + b)^q \\ &= \sum_{x=0}^{N-1}\sum_{y=0}^{N-1}\alpha bf(\lfloor x\rfloor + 1, \lfloor y\rfloor + 1)(\lfloor x\rfloor + 1 + \alpha - 1)^p(\lfloor y\rfloor + 1 + b - 1)^q \end{aligned}$$

$$= \sum_{x=0}^{N-1}\sum_{y=0}^{N-1} \alpha b f(\lfloor x\rfloor+1,\lfloor y\rfloor+1)\sum_{m=0}^{p}\binom{p}{m}(\lfloor x\rfloor+1)^m(\alpha-1)^{p-m}\times \sum_{n=0}^{q}\binom{q}{n}(\lfloor y\rfloor+1)^n(b-1)^{q-n} \quad (3.39)$$

Για $t=\lfloor x\rfloor+1$ και $s=\lfloor y\rfloor+1$:

$$\begin{aligned}
A_4 &= \sum_{t=1}^{N}\sum_{s=1}^{N}\alpha b f(t,s)\sum_{m=0}^{p}\binom{p}{m}t^m(\alpha-1)^{p-m}\sum_{n=0}^{q}\binom{q}{n}s^n(b-1)^{q-n}\\
&= \alpha b\sum_{m=0}^{p}\binom{p}{m}(\alpha-1)^{p-m}\sum_{n=0}^{q}\binom{q}{n}(b-1)^{q-n}\sum_{t=1}^{N}\sum_{s=1}^{N}f(t,s)t^m s^n\\
&= \alpha b\sum_{m=0}^{p}\binom{p}{m}(\alpha-1)^{p-m}\sum_{n=0}^{q}\binom{q}{n}(b-1)^{q-n}\sum_{t=1}^{N-1}\sum_{s=1}^{N-1}f(t,s)t^m s^n\\
&= \alpha b\sum_{m=0}^{p}\binom{p}{m}(\alpha-1)^{p-m}\sum_{n=0}^{q}\binom{q}{n}(b-1)^{q-n}\times\\
&\quad \sum_{t=1}^{N-1}\left(\sum_{s=0}^{N-1}f(t,s)t^m s^n - f(t,0)t^m0^n\right)\\
&= \alpha b\sum_{m=0}^{p}\binom{p}{m}(\alpha-1)^{p-m}\sum_{n=0}^{q}\binom{q}{n}(b-1)^{q-n}\times\\
&\quad \left[\sum_{t=0}^{N-1}\left(\sum_{s=0}^{N-1}f(t,s)t^m s^n - f(t,0)t^m0^n\right) - \right.\\
&\quad \left.\left(\sum_{s=0}^{N-1}f(0,s)0^m s^n - f(0,0)0^m0^n\right)\right]
\end{aligned} \quad (3.40)$$

ή

$$\begin{aligned}
A_4 &= \alpha b\sum_{m=0}^{p}\binom{p}{m}(\alpha-1)^{p-m}\sum_{n=0}^{q}\binom{q}{n}(b-1)^{q-n}\\
&\quad \left[M_{m,n}-\sum_{t=0}^{N-1}f(t,0)t^m0^n-\sum_{s=0}^{N-1}f(0,s)0^m s^n+f(0,0)0^m0^n\right]
\end{aligned}$$

Τελικά

$$
\begin{aligned}
A_4 &= \alpha b \sum_{m=0}^{p} \binom{p}{m} (\alpha-1)^{p-m} \sum_{n=0}^{q} \binom{q}{n} (b-1)^{q-n} M_{m,n} && (3.41)\\
&\quad -\alpha b \sum_{m=0}^{p} \binom{p}{m} (\alpha-1)^{p-m} (b-1)^{q} \sum_{t=0}^{N-1} f(t,0) t^{m} && (3.42)\\
&\quad -\alpha b (\alpha-1)^{p-m} \sum_{n=0}^{q} \binom{q}{n} (b-1)^{q-n} \sum_{s=0}^{N-1} f(0,s) s^{n} && (3.43)\\
&\quad +\alpha b (\alpha-1)^{p} (b-1)^{q} f(0,0) && (3.44)
\end{aligned}
$$

$$
\begin{aligned}
&A_1 + A_2 + A_3 + A_4 = \\
&(1-\alpha)(1-b) \sum_{n=0}^{q} \binom{q}{n} \sum_{m=0}^{p} \binom{p}{m} b^{q-n} \alpha^{p-m} M_{m,n} \\
&+\alpha(1-b) \sum_{m=0}^{p} \binom{p}{m} \sum_{n=0}^{q} \binom{q}{n} (\alpha-1)^{p-m} b^{q-n} M_{m,n} \\
&-\alpha(1-b) \sum_{n=0}^{q} \binom{q}{n} (\alpha-1)^{p} b^{q-n} \sum_{y=0}^{N-1} f(0,\lfloor y \rfloor)(\lfloor y \rfloor)^{n} \\
&+b(1-\alpha) \sum_{n=0}^{q} \binom{q}{n} \sum_{m=0}^{p} \binom{p}{m} (b-1)^{q-n} \alpha^{p-m} M_{m,n} \\
&-b(1-\alpha) \sum_{m=0}^{p} \binom{p}{m} (b-1)^{q} \alpha^{p-m} \sum_{x=0}^{N-1} f(\lfloor x \rfloor, 0)(\lfloor x \rfloor)^{m} \\
&+\alpha b \sum_{m=0}^{p} \binom{p}{m} (\alpha-1)^{p-m} \sum_{n=0}^{q} \binom{q}{n} (b-1)^{q-n} M_{m,n} \\
&-\alpha b \sum_{m=0}^{p} \binom{p}{m} (\alpha-1)^{p-m} (b-1)^{q} \sum_{t=0}^{N-1} f(t,0) t^{m} \\
&-\alpha b (\alpha-1)^{p} \sum_{n=0}^{q} \binom{q}{n} (b-1)^{q-n} \sum_{s=0}^{N-1} f(0,s) s^{n} \\
&+\alpha b (\alpha-1)^{p} (b-1)^{q} f(0,0) && (3.45)
\end{aligned}
$$

Το μέρος της παραπάνω σχέσης το οποίο αναφέρεται στα οριακά δείγματα της εικόνας δίνεται από την παρακάτω έκφραση:

$$
\begin{aligned}
&R = \\
&-\alpha(1-b)\sum_{n=0}^{q}\binom{q}{n}(\alpha-1)^{p}b^{q-n}\sum_{y=0}^{N-1}f(0,\lfloor y\rfloor)(\lfloor y\rfloor)^{n} \\
&-b(1-\alpha)\sum_{m=0}^{p}\binom{p}{m}(b-1)^{q}\alpha^{p-m}\sum_{x=0}^{N-1}f(\lfloor x\rfloor,0)(\lfloor x\rfloor)^{m} \\
&-\alpha b\sum_{m=0}^{p}\binom{p}{m}(\alpha-1)^{p-m}(b-1)^{q}\sum_{t=0}^{N-1}f(t,0)t^{m} \\
&-\alpha b(\alpha-1)^{p}\sum_{n=0}^{q}\binom{q}{n}(b-1)^{q-n}\sum_{s=0}^{N-1}f(0,s)s^{n} \\
&+\alpha b(\alpha-1)^{p}(b-1)^{q}f(0,0)
\end{aligned}
\tag{3.46}
$$

Ολοκληρώνοντας, προκύπτει:

$$
\begin{aligned}
&\int_0^1\int_0^1 R d\alpha db = \\
&-\frac{(-1)^{p}}{(p+1)(p+2)}\sum_{n=0}^{q}\binom{q}{n}\frac{1}{(q-n+1)(q-n+2)}\sum_{y=0}^{N-1}f(0,\lfloor y\rfloor)(\lfloor y\rfloor)^{n} \\
&-\frac{(-1)^{q}}{(q+1)(q+2)}\sum_{m=0}^{p}\binom{p}{m}\frac{1}{(p-m+1)(p-m+2)}\sum_{x=0}^{N-1}f(\lfloor x\rfloor,0)(\lfloor x\rfloor)^{m} \\
&-\frac{(-1)^{q}}{(q+1)(q+2)}\sum_{m=0}^{p}\binom{p}{m}\frac{(-1)^{p-m}}{(p-m+1)(p-m+2)}\sum_{t=0}^{N-1}f(t,0)t^{m} \\
&-\frac{(-1)^{p}}{(p+1)(p+2)}\sum_{n=0}^{q}\binom{q}{n}\frac{(-1)^{q-n}}{(q-n+1)(q-n+2)}\sum_{s=0}^{N-1}f(0,s)s^{n} \\
&+\frac{(-1)^{p}}{(p+1)(p+2)}\frac{(-1)^{q}}{(q+1)(q+2)}f(0,0)
\end{aligned}
\tag{3.47}
$$

ή

$$
\int_0^1\int_0^1 R d\alpha db =
$$

$$
\begin{aligned}
&-\frac{(-1)^p}{(p+1)(p+2)}\sum_{n=0}^{q}\binom{q}{n}\frac{1+(-1)^{q-n}}{(q-n+1)(q-n+2)}\sum_{y=0}^{N-1}f(0,\lfloor y\rfloor)(\lfloor y\rfloor)^n\\
&-\frac{(-1)^q}{(q+1)(q+2)}\sum_{m=0}^{p}\binom{p}{m}\frac{1+(-1)^{p-m}}{(p-m+1)(p-m+2)}\sum_{x=0}^{N-1}f(\lfloor x\rfloor,0)(\lfloor x\rfloor)^m\\
&+\frac{(-1)^p}{(p+1)(p+2)}\frac{(-1)^q}{(q+1)(q+2)}f(0,0) \qquad (3.48)
\end{aligned}
$$

Οι υπόλοιποι όροι της σχέσης 3.45 μετασχηματίζονται στην παρακάτω σχέση:

$$
\begin{aligned}
&\int_0^1\int_0^1 A_1+A_2+A_3+A_4-Rd\alpha db=\\
&\sum_{n=0}^{q}\binom{q}{n}\sum_{m=0}^{p}\binom{p}{m}\frac{1}{(p-m+1)(p-m+2)}\frac{1}{(q-n+1)(q-n+2)}M_{m,n}\\
&+\sum_{m=0}^{p}\binom{p}{m}\sum_{n=0}^{q}\binom{q}{n}\frac{(-1)^{p-m}}{(p-m+1)(p-m+2)}\frac{1}{(q-n+1)(q-n+2)}M_{m,n}\\
&+\sum_{n=0}^{q}\binom{q}{n}\sum_{m=0}^{p}\binom{p}{m}\frac{1}{(p-m+1)(p-m+2)}\frac{(-1)^{q-n}}{(q-n+1)(q-n+2)}M_{m,n}\\
&+\sum_{m=0}^{p}\binom{p}{m}(\alpha-1)^{p-m}\sum_{n=0}^{q}\binom{q}{n}\times\\
&\quad\frac{(-1)^{p-m}}{(p-m+1)(p-m+2)}\frac{(-1)^{q-n}}{(q-n+1)(q-n+2)}M_{m,n} \qquad (3.49)
\end{aligned}
$$

Η τελευταία έκφραση αποτελεί μία ακριβέστερη προσέγγιση στις συνεχείς γεωμετρικές ροπές, σε σχέση με το απλό διακριτό άθροισμα. Αυτό συμβαίνει λόγω της πολυωνυμικής παρεμβολής του σήματος αντί της θεώρησης του ως σταθερού, ανάμεσα στα σημεία δειγματοληψίας.

3.2.3 Πολυωνυμική Παρεμβολή

Τα παραπάνω αποτελέσματα μπορούν να γενικευθούν για κυβική παρεμβολή ή παρεμβολή μεγαλύτερης τάξης. Με δεδομένα n σημεία ελέγχου, το πολυώνυμο παρεμβολής μπορεί να περιγραφεί με τη χρήση του πίνακα Vandermonde [Hoffman and Kunze, 1971]:

$$V\begin{bmatrix}\alpha_n\\ \alpha_{n-1}\\ \vdots\\ \alpha_0\end{bmatrix}=\begin{bmatrix}y_0\\ y_1\\ \vdots\\ y_n\end{bmatrix} \tag{3.50}$$

όπου

$$V=\begin{bmatrix}x_0^n & x_0^{n-1} & \cdots & x_0^0\\ x_1^n & x_1^{n-1} & \cdots & x_1^0\\ \vdots & \vdots & \ddots & \vdots\\ x_n^n & x_n^{n-1} & \cdots & x_n^0\end{bmatrix} \tag{3.51}$$

είναι ο πίνακας Vandermonde , (x_n,y_n) τα δεδομένα σημεία και

$$P(x)=\sum_{t=0}^{n}\alpha_t x^t \tag{3.52}$$

το πολυώνυμο παρεμβολής βαθμού n. Η σημειολογία μπορεί να απλουστευθεί υποθέτοντας ότι:

$$(x_0,x_1,\ldots,x_n)=(x-n/2+1,x-n/2+2,\ldots,x+n/2) \tag{3.53}$$

ή

$$x_k=x-n/2+k+1 \tag{3.54}$$

Η εξίσωση 3.50 μπορεί τώρα να πάρει την παρακάτω μορφή:

$$V\begin{bmatrix}\alpha_n\\ \alpha_{n-1}\\ \vdots\\ \alpha_0\end{bmatrix}=\begin{bmatrix}f(x-n/2+1)\\ f(x-n/2+2)\\ \vdots\\ f(x-n/2+1)\end{bmatrix} \tag{3.55}$$

με

$$V=\begin{bmatrix}(x-n/2+1)^n & (x-n/2+1)^{n-1} & \cdots & (x-n/2+1)^0\\ (x-n/2+2)^n & (x-n/2+2)^{n-1} & \cdots & (x-n/2+2)^0\\ \vdots & \vdots & \ddots & \vdots\\ (x+(n-1)/2)^n & (x+(n-1)/2)^{n-1} & \cdots & (x+(n-1)/2)^0\end{bmatrix} \tag{3.56}$$

Λύνοντας την εξίσωση 3.55 το πολυώνυμο παρεμβολής γίνεται:

$$P(x)=\sum_{t=0}^{n}\alpha_t x^t=\sum_{t=0}^{n}\sum_{k=0}^{n}V_{n+1}^{-1}(n-t+1,k)f(x-\frac{n+3}{2}+k)x^t \tag{3.57}$$

με V_{n+1}^{-1} τον αντίστροφο του $(n+1)x(n+1)$ πίνακα Vandermonde , όπως αυτός ορίζεται στην εξίσωση 3.50. Η γεωμετρική ροπή ενός σήματος $f(x)$ για το διάστημα $[x, x+1]$ μπορεί να οριστεί ως:

$$
\begin{aligned}
M_p|_x^{x+1} &= \int_0^1 (x+\mu)^p f(x+\mu)d\mu = \int_0^1 (x+\mu)^p P(\mu)d\mu \\
&= \int_0^1 (x+\mu)^p \sum_{t=0}^{n}\sum_{k=0}^{n} V_{n+1}^{-1}(n-t+1,k) f(x - \frac{n+3}{2} + k)\mu^t d\mu
\end{aligned}
\tag{3.58}
$$

Αθροίζοντας για όλο το σήμα, παίρνουμε:

$$
\begin{aligned}
M_p &= \sum_{x=0}^{N-1} \int_0^1 (x+\mu)^p \sum_{t=0}^{n}\sum_{k=0}^{n} V_{n+1}^{-1}(n-t+1,k) f\left(x - \frac{n+3}{2} + k\right)\mu^t d\mu \\
&= \int_0^1 \sum_{t=0}^{n}\sum_{k=0}^{n} V_{n+1}^{-1}(n-t+1,k)\mu^t \sum_{x=0}^{N-1} f\left(x - \frac{n+3}{2} + k\right)(x+\mu)^p d\mu \\
&= \int_0^1 \sum_{t=0}^{n}\sum_{k=0}^{n} V_{n+1}^{-1}(n-t+1,k)\mu^t \sum_{x=0}^{N-1} f\left(x - \frac{n+3}{2} + k\right) \sum_{l=0}^{p} \binom{p}{l} x^l \mu^{p-l} d\mu \\
&= \int_0^1 \sum_{t=0}^{n}\sum_{k=0}^{n} V_{n+1}^{-1}(n-t+1,k)\mu^t \sum_{l=0}^{p} \binom{p}{l} \mu^{p-l} \sum_{x=0}^{N-1} f\left(x - \frac{n+3}{2} + k\right) x^l d\mu
\end{aligned}
\tag{3.59}
$$

Το τελευταίο άθροισμα είναι ισοδύναμο με μία γεωμετρική ροπή τάξης l ως προς το σημείο $(n+3/2)-k$:

$$
\begin{aligned}
M_p &= \int_0^1 \sum_{t=0}^{n}\sum_{k=0}^{n} V_{n+1}^{-1}(n-t+1,k)\mu^t \sum_{l=0}^{p} \binom{p}{l} \mu^{p-l} \sum_{x=0}^{N-1} f\left(x - \frac{n+3}{2} + k\right) x^l d\mu \\
&= \int_0^1 \sum_{t=0}^{n}\sum_{k=0}^{n} V_{n+1}^{-1}(n-t+1,k)\mu^t \sum_{l=0}^{p} \binom{p}{l} \mu^{p-l} M_l^{\frac{n+3}{2}-k} d\mu \\
&= \sum_{t=0}^{n}\sum_{k=0}^{n} V_{n+1}^{-1}(n-t+1,k) \sum_{l=0}^{p} \binom{p}{l} M_l^{\frac{n+3}{2}-k} \int_0^1 \mu^{p-l+t} d\mu
\end{aligned}
\tag{3.60}
$$

και τελικά:

$$M_p = \sum_{t=0}^{n} \sum_{k=0}^{n} V_{n+1}^{-1}(n-t+1,k) \sum_{l=0}^{p} \binom{p}{l} M_l^{\frac{n+3}{2}-k}(p-l+t+1)^{-1} \qquad (3.61)$$

Η τελευταία έκφραση μπορεί να χρησιμοποιηθεί για τον ακριβή υπολογισμό των γεωμετρικών ροπών ενός διακριτού σήματος. Πιο συγκεκριμένα, για κάθε σήμα ορισμένο σε διάστημα C ισχύει ότι:

$$\left| M_p - \int_C f(x)x^p \right| \leq \left| \sum_C f(x)x^p - \int_C f(x)x^p \right| \qquad (3.62)$$

Προφανώς, όσο χειρότερες είναι οι συνθήκες τις αρχικής διακριτοποίησης, τόσο μεγαλύτερη θα είναι η βελτίωση της προτεινόμενης τεχνικής σε σχέση με το απλό άθροισμα. Στην μία ακραία περίπτωση, στην οποία η δειγματοληψία είναι πολύ μεγάλης συχνότητας, και στην οποία $f(x) \approx f(x-1)$, η παρούσα μέθοδος εκφυλίζεται σε αυτήν της [Liao and Pawlak, 1996]. Στην περίπτωση, όμως, που το ψηφιακό σήμα παρουσιάζει μεγάλες συχνότητες, τα αποτελέσματα της παρούσας τεχνικής είναι σημαντικά βελτιωμένα, και μπορούν να ξεπεράσουν τη μία τάξη μεγέθους μείωσης στο σφάλμα. Οι ροπές του μετατοπισμένου σήματος, προκύπτουν εύκολα από τις αρχικά υπολογιζόμενες γεωμετρικές ροπές, όπως αναλύεται και στην ενότητα 5.1.1. Αξίζει να σημειωθεί ότι η έκφραση απλουστεύεται σημαντικά, αν αντικατασταθεί ο βαθμός του πολυωνύμου παρεμβολής. Για παράδειγμα, για ομοιογενή, σημειακή, κυβική παρεμβολή, η εξίσωση 3.61 γίνεται:

$$M_p = \sum_{m=0}^{p} \binom{p}{m} \frac{4(2^{p-m+2}-1)(1+(-1)^{p-m})}{(p+1)(p+2)(p+3)(p+4)} M_m \qquad (3.63)$$

Η παραπάνω διαδικασία μπορεί να επεκταθεί σε δύο ή και περισσότερες διαστάσεις, με παρόμοιο τρόπο.

3.3 Ακριβής υπολογισμός ροπών *Zernike*

Οι ροπές Zernike μπορούν να εκφραστούν συναρτήσει των γεωμετρικών ροπών:

$$Z_{p,q} = \frac{p+1}{\pi} \int_0^1 \int_0^{2\pi} rR_{pq}(r)e^{-iq\theta} f(r,\theta)d\theta dr$$

$$
\begin{aligned}
&= \frac{p+1}{\pi} \int_0^1 \int_0^{2\pi} r \sum_{k=q,p-|k|even}^{p} B_{pqk} r^k e^{-iq\theta} f(r,\theta) d\theta dr \\
&= \frac{p+1}{\pi} \sum_{k=q,p-|k|even}^{p} B_{pqk} \sum_{n=0}^{\frac{k-q}{2}} \binom{\frac{k-q}{2}}{n} \sum_{m=0} q \binom{q}{m} (-i)^{q-m} \\
&\quad \int_0^N \int_0^N (\sqrt{2}N)^{-k-2} u^{m+2n} w^{k-2n-m} f(u,w) du dw \qquad (3.64)
\end{aligned}
$$

και

$$
Z_{p,q} = \frac{p+1}{\pi} \sum_{k=q,p-|k|even}^{p} (\sqrt{2}N)^{-k-2} B_{pqk} \sum_{n=0}^{\frac{k-q}{2}} \binom{\frac{k-q}{2}}{n} \sum_{m=0}^{q} \binom{q}{m} (-i)^{q-m} M'_{m+2n,k-2n-m} \qquad (3.65)
$$

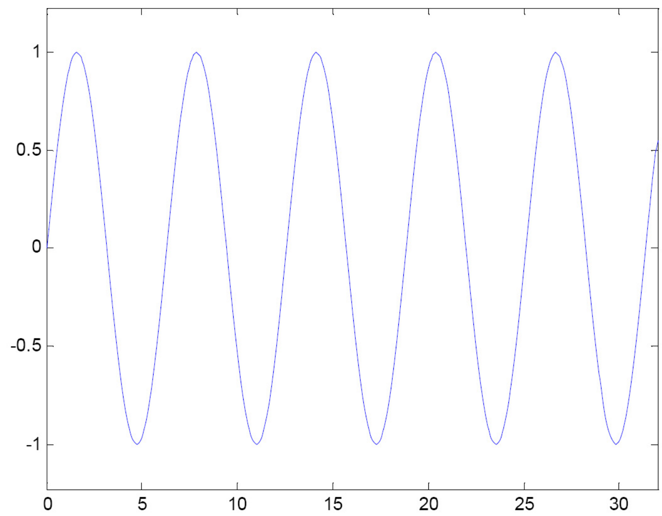
Η σχέση 3.65 συνδέει τις ροπές *Zernike* με τις συνεχείς γεωμετρικές. Η σχέση γίνεται γραμμική μόνο αν το μέγεθος του σήματος N θεωρηθεί σταθερό. Οι τελευταίες μπορούν να υπολογιστούν χρησιμοποιώντας πολυώνυμα παρεμβολής, όπως αυτά περιγράφηκαν στην παράγραφο 3.2.

3.4 Υπολογιστική Πολυπλοκότητα

Η εισαγωγή της πολυωνυμικής παρεμβολής στον υπολογισμό των ροπών προσθέτει κάποιο υπολογιστικό κόστος. Οι προτεινόμενες ροπές μπορούν να εξαχθούν από τις συνήθεις χρησιμοποιώντας την παρακάτω γενικότερη έκφραση:

$$
M_{p,q} = \sum_{m=0}^{q} \sum_{n=0}^{q} A_{p,q}(m,n) M_{p,q} \qquad (3.66)
$$

Τα $A_{p,q}(m,n)$ εξαρτώνται μόνο από τον τύπο της παρεμβολής. Έτσι, είναι δυνατόν να έχουν υπολογιστεί εκ των προτέρων, μειώνοντας την υπολογιστική πολυπλοκότητα σε pq πολλαπλασιασμούς και pq προσθέσεις, η οποία απαιτεί ασήμαντο υπολογιστικό χρόνο. Λόγω συμμετρίας, $A_{p,q}(m,n) = A_{q,p}(n,m)$. Έτσι, το μέγεθος της απαιτούμενης μνήμης μειώνεται στο μισό.



Σχήμα 3.4: Ημιτονοειδές σήμα

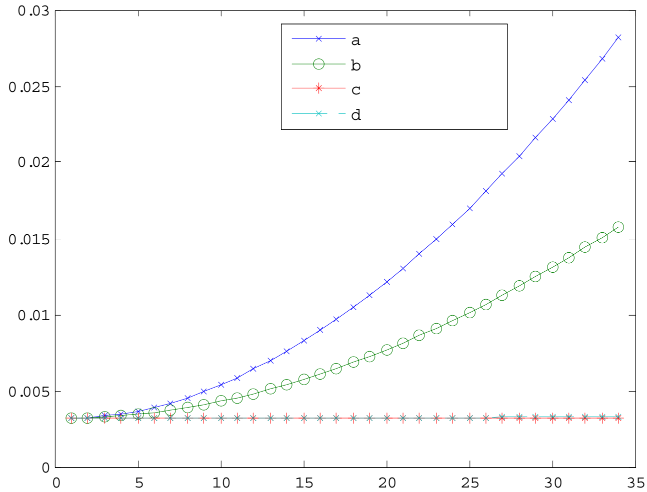
3.5 Πειραματικά αποτελέσματα

3.5.1 Μονοδιάστατα σήματα

Σε αυτήν την Ενότητα παρουσιάζουμε πειραματικά αποτελέσματα σχετικά με την ακρίβεια των γεωμετρικών ροπών σε μονοδιάστατα σήματα. Σαν ενδεικτικό παράδειγμα παρουσιάζεται η εξαγωγή των πρώτων 16 ροπών ενός απλού ημιτονοειδούς σήματος, όπως φαίνεται στο σχήμα 3.4. Στο σχήμα 3.5 παρουσιάζεται το σφάλμα που προκύπτει από τον υπολογισμό ροπών υπολογισμένων βάσει της εξίσωσης 3.2, καθώς και με τη χρήση πολυωνύμων παρεμβολής 1ου, 3ου και 5ου βαθμού.

3.5.2 Σήματα δύο διαστάσεων

Για την δοκιμή της προτεινόμενης τεχνικής σε σήματα δύο διαστάσεων, υπολογίστηκαν οι ροπές τάξης μέχρι $(3,3)$ για την εικόνα στου σχήματος 3.6. Αυτό το σήμα περιγράφεται από τη σχέση $sin(x/3)sin(y/3)+1$. Χρησιμοποιώντας την αναλυτική περιγραφή του, οι ακριβείς τιμές των ροπών υπολογίστηκαν με τη χρήση της σχέσης 3.1. Το απόλυτο σφάλμα που προέκυψε εμφανίζεται στο σχήμα 3.7. Η χρήση πολυωνύμων 5ου βαθμού μείωσε το προσεγγιστικό σφάλμα κατά περίπου 3

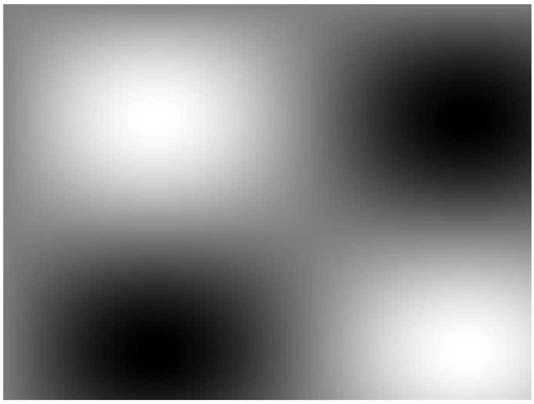


Σχήμα 3.5: Σχετικό σφάλμα για υπολογισμό ροπών μονοδιάστατου σήματος. Προσεγγίσεις: (*a*) Διακριτή (*b*) σημειακά σταθερή, (*c*) κυβική παρεμβολή, (*d*) γραμμική παρεμβολή

τάξεις μεγέθους.

3.5.3 Ροπές *Zernike*

Ίσως το πιο χρήσιμο χαρακτηριστικό των ροπών *Zernike*είναι η αμεταβλητότητά τους στην περιστροφή. Το σφάλμα, όμως, που εισάγεται από την διακριτή προσέγγιση της εικόνας επηρεάζει αυτήν την ιδιότητα, ειδικά όταν επεξεργάζονται εικόνες μικρού μεγέθους. Χρησιμοποιώντας την τεχνική που περιγράφεται στην Ενότητα 3.3, είναι δυνατό να βελτιωθεί η ιδιότητα αυτή. Στα σχήματα 3.8 και 3.9 παρουσιάζεται η επίδραση της περιστροφής στα μέτρα των ροπών Zernike βαθμών 10 και 6 αντίστοιχα, για την εικόνα του σχήματος 3.10. Παρατηρείται ότι ενώ η άμεση, διακριτή υλοποίηση οδηγεί σε σφάλματα μέχρι 50% για ροπές 10ης τάξης και 15% για ροπές 6ης τάξης, χρησιμοποιόντας κυβική προσέγγιση τα σφάλματα αυτά μειώνονται σε 10% και 1%, αντίστοιχα. Σε μεγαλύτερες εικόνες, τα σφάλματα αυτά μειώνονται σημαντικά, όπως φαίνεται στα σχήματα 3.11, 3.12 και 3.13. Σε κάθε περίπτωση, οι ροπές παρεμβολής είναι πιο ακριβείς. Η ίδια διαδικασία επαναλήφθηκε για την αποτίμηση της αμεταβλητότητας στην κλιμάκωση. Η εικόνα μεγεθύνθηκε με λόγους από 0.5 μέχρι 10, και εξάχθηκαν

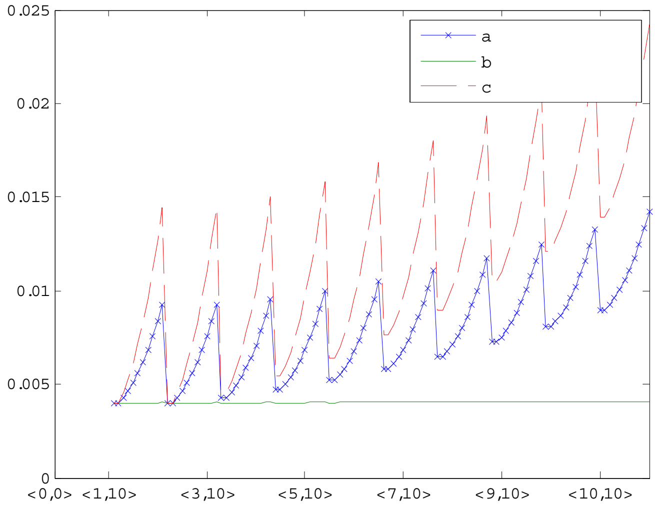


Σχήμα 3.6: Δοκιμαστική εικόνα ημιτόνων

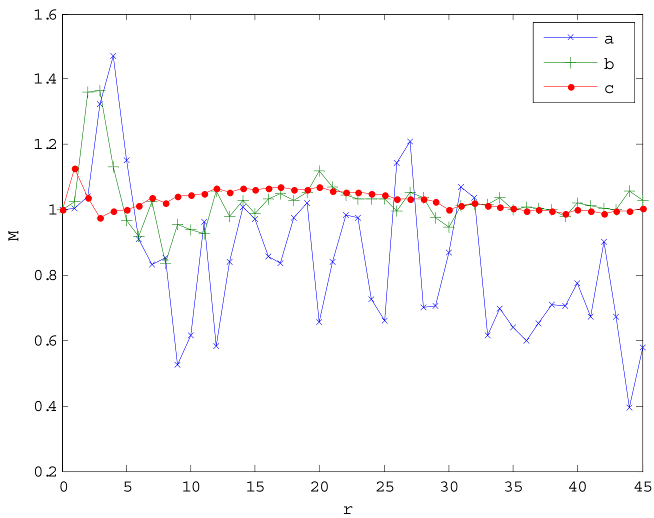
οι ροπές *Zernike*. Τα μέτρα των ροπών 4ης τάξης σε σχέση με το αρχικό παρουσιάζονται στο σχήμα 3.14. Παρατηρείται ότι ενώ για ακέραιους λόγους η διακριτή προσέγγιση δίνει ακριβή αποτελέσματα, ένα σφάλμα της τάξης του 1% εμφανίζεται για κλασματικούς, το οποίο ελαχιστοποιείται μέσω της χρήσης ροπών παρεμβολής.

3.6 Συμπεράσματα

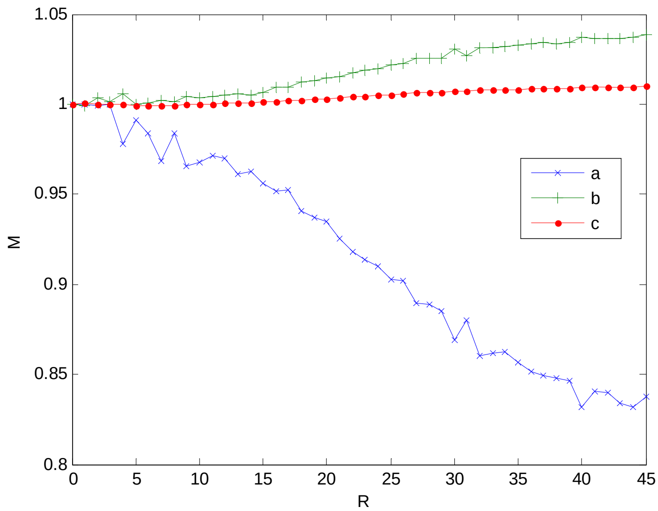
Στο κεφάλαιο αυτό έγινε ανάλυση μίας από τις σημαντικότερες πηγές σφαλμάτων κατά τον υπολογισμό ροπών εικόνων. Στη συνέχεια, χρησιμοποιώντας γραμμική, κυβική και γενικότερα πολυωνυμική παρεμβολή, κατέστη δυνατός ο πιο ακριβής υπολογισμός τόσο των γεωμετρικών ροπών, όσο και των ροπών *Zernike*. Το επιπλέον υπολογιστικό κόστος για την παρεμβολή είναι αμελητέο σε σχέση με αυτό του υπολογισμού των ροπών. Τα παραπάνω αποτελέσματα μπορούν εύκολα να επεκταθούν και σε άλλους τύπους ροπών, όπως οι μιγαδικές και οι Legendre.



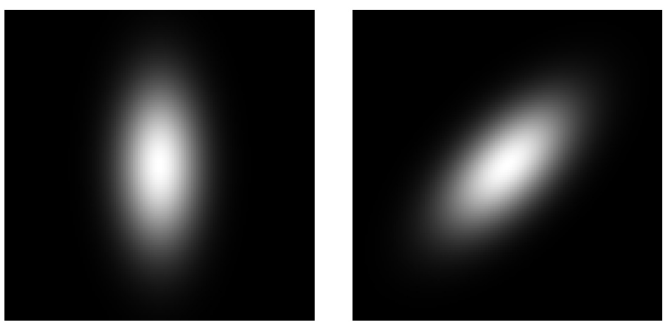
Σχήμα 3.7: Σχετικό σφάλμα προσέγγισης για: (α) σημειακά σταθερή προσέγγιση, (*b*) γραμμική προσέγγιση, (*c*) διακριτή προσέγγιση



Σχήμα 3.8: Επίδραση περιστροφής (γωνία *r*) στην κανονικοποιημένη τιμή ροπών (*M*) 10ης τάξης συνθετικής εικόνας.(α) Διακριτή προσέγγιση, (*b*) γραμμική προσέγγιση, (*c*) κυβική προσέγγιση

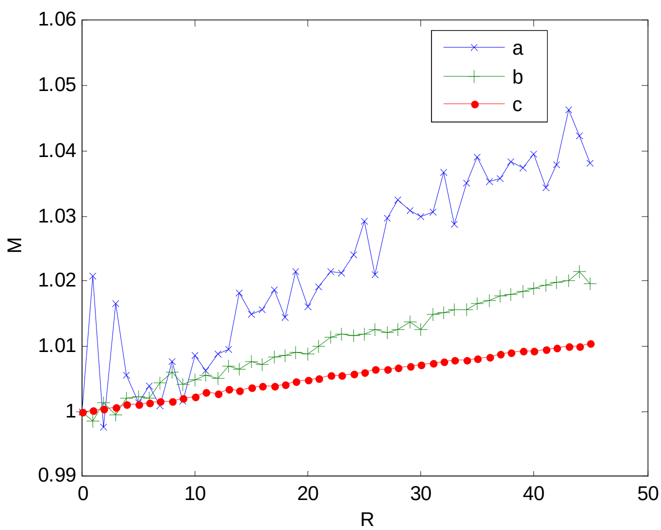


Σχήμα 3.9: Επίδραση περιστροφής (γωνία *r*) στην κανονικοποιημένη τιμή ροπών (*M*) 6ης τάξης συνθετικής εικόνας.(*a*) Διακριτή προσέγγιση, (*b*) γραμμική προσέγγιση, (*c*) κυβική προσέγγιση

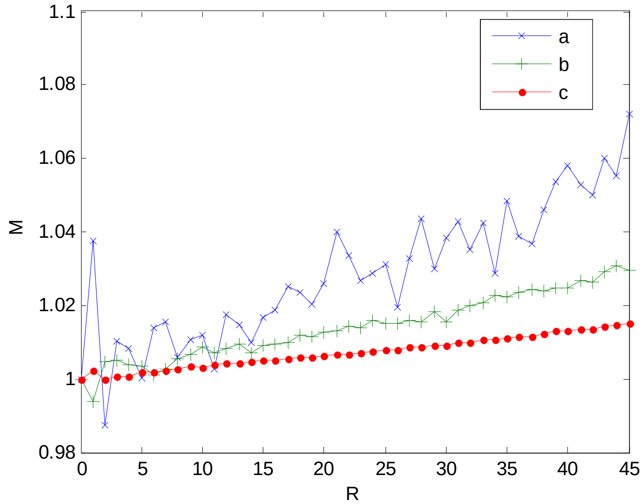


Σχήμα 3.10: Συνθετική εικόνα

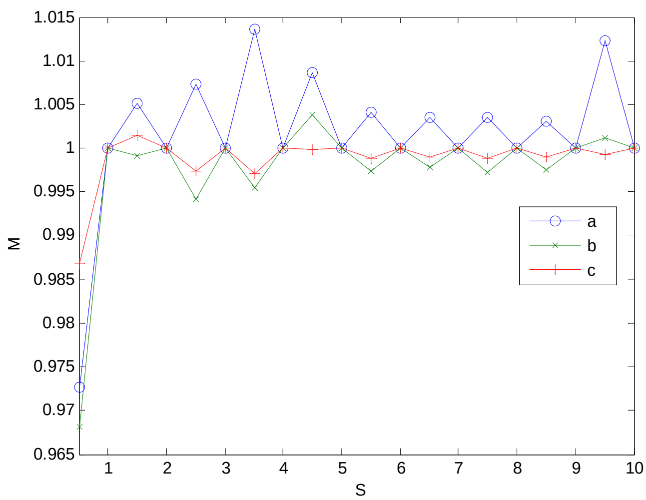
Σχήμα 3.11: Πραγματική εικόνα αποχρώσεων του γκρι



Σχήμα 3.12: Επίδραση περιστροφής (γωνία r) στην κανονικοποιημένη τιμή ροπών (M) 10ης τάξης.(a) Διακριτή προσέγγιση, (b) γραμμική προσέγγιση, (c) κυβική προσέγγιση



Σχήμα 3.13: Επίδραση περιστροφής (γωνία *r*) στην κανονικοποιημένη τιμή ροπών (*M*) 6ης τάξης.(*α*) Διακριτή προσέγγιση, (*b*) γραμμική προσέγγιση, (*c*) κυβική προσέγγιση



Σχήμα 3.14: Επίδραση κλιμάκωσης (λόγος *S*) στην κανονικοποιημένη τιμή ροπών (*M*) 5ης τάξης.(*α*) Διακριτή προσέγγιση, (*b*) γραμμική προσέγγιση, (*c*) κυβική προσέγγιση

Κεφάλαιο 4

Εξελικτική βελτίωση ροπών διακριτών ορθογωνίων πολυωνύμων

Οι ροπές διακριτής ορθογώνιας βάσης παρουσιάζουν τα πλεονεκτήματα της εύκολης υλοποίησης σε υλικό και λογισμικό, των μικρών σφαλμάτων κατά τον υπολογισμό και της μεγάλης ανοχής στο θόρυβο, σε σχέση με τις ροπές συνεχούς βάσης. Στο Κεφάλαιο αυτό προτείνεται μία μεθοδολογία για τη βελτιστοποίηση των χαρακτηριστικών τους, μέσω εξελικτικών τεχνικών. Οι τελικοί περιγραφείς παρουσιάζουν βελτιωμένα χαρακτηριστικά ως προς την ανακατασκευή και την ανάκτηση σημάτων.

Οι ροπές εικόνων χρησιμοποιούνται ευρέως σε διάφορες εφαρμογές υπολογιστικής όρασης και ανάλυσης εικόνων. Οι ροπές με βάση ορθογώνια πολυώνυμα ορισμένα σε διακριτό χώρο προτάθηκαν πρόσφατα, παρουσιάζοντας κάποιες χρήσιμες ιδιότητες. Στο κεφάλαιο αυτό, μελετάται η χρήση εξελικτικά βελτιωμένων ροπών ως περιγραφείς εικόνων. Με τη χρήση μίας εξελικτικής στρατηγικής (ΕΣ), οι παραγόμενες ροπές δοκιμάζονται στην ανακατασκευή εικόνων και στην περιγραφή αντικειμένων, παρουσιάζοντας αυξημένη ακρίβεια σε σχέση με τις αναφερόμενες στη βιβλιογραφία ροπές. Χρησιμοποιώντας την προτεινόμενη μεθοδολογία, μπορούν να εξαχθούν

Όνομα	S	$w(x)$	σχόλια
Chebychev	$N-1$	1	
Krawtchouk	N	$\binom{N}{k} p^k(1-p)^{N-k}$	$0 < p < 1$
Charlier	∞	$e^{-\alpha}\frac{\alpha^k}{k!}$	$\alpha > 0$
Meixner	∞	$c^k \binom{k+\beta}{k}$	$0 < c < 1, \alpha > 0$
Hahn	N	$\binom{\alpha+k}{k}\binom{\beta+N-k}{N-k}$	$\alpha > -1, \beta > -1$

Πίνακας 4.1: Συνήθη διακριτά ορθογώνια πολυώνυμα, με S το διάστημα ορθογωνιότητας και $w(x)$ τη συνάρτηση βάρους

ροπές προσανατολισμένες στην εκάστοτε εφαρμογή. Τα πειραματικά αποτελέσματα έχουν δείξει αυξημένες επιδόσεις στην ανακατασκευή αντικειμένων και τη συσχέτιση εικόνων σε σχέση με άλλες ροπές διακριτής βάσης.

4.1 Εισαγωγή

4.1.1 Ορθοκανονικοποίηση *Gram – Schmidt*

Στη γενική περίπτωση, η διαδικασία Gram-Schmidt δέχεται ένα μη ορθογώνιο σύνολο γραμμικά ανεξαρτήτων συναρτήσεων και κατασκευάζει μία ορθογώνια βάση σε δεδομένο διάστημα και σε σχέση με μία συνάρτηση βάρους $w(x)$. Τα ορθογώνια πολυώνυμα είναι ιδιαίτερα εύκολα στην κατασκευή τους. Χρησιμοποιώντας τις συναρτήσεις βάρους που δίνονται στον πίνακα 4.1 μπορούν να παραχθούν τα περισσότερα κοινά ορθογώνια πολυώνυμα. Οι παρακάτω σχέσεις δίνουν το περίγραμμα της διαδικασίας αυτής. Για $p_0 = 1$:

$$
\begin{aligned}
p_1(x) &= \left(x - \frac{\sum x(p_0(x))^2 w(x)}{\sum (p_0(x))^2 w(x)}\right) p_0(x) \\
p_{i+1}(x) &= \left(x - \frac{\sum x(p_i(x))^2 w(x)}{\sum (p_i(x))^2 w(x)}\right) p_i(x) - \frac{\sum (p_i(x))^2 w(x)}{\sum (p_{i-1}(x))^2 w(x)} p_{i-1}(x)
\end{aligned} \tag{4.1}
$$

Η διαδικασία περιγράφεται αναλυτικά στο [Gautschi, 2004].

4.1.2 Οι εξελικτικές στρατηγικές ως εργαλείο ελαχιστοποίησης συναρτήσεων

Ο τρόπος λειτουργίας των εξελικτικών αλγορίθμων είναι εμπνευσμένος από την βιολογία. Χρησιμοποιεί την ιδέα της εξέλιξης μέσω γενετικής μετάλλαξης, φυσικής επιλογής και διασταύρωσης.

Στην πράξη ο αλγόριθμος ξεκινά μ' ένα σύνολο λύσεων - ονομάζονται γονιδιώματα, δανειζόμενες το όνομά τους από τη βιολογία- οι οποίες συνιστούν τον 'πληθυσμό'. Κατόπιν ζητείται από τον υπολογιστή να δημιουργήσει μια σειρά τυχαίων ανασυνδυασμών και μεταλλάξεων των 'γονιδιωμάτων'.

Οι πιο ικανές λύσεις για ένα συγκεκριμένο πρόβλημα συνεχίζουν να εξελίσσονται και ανασυνδυάζονται τυχαία, μέχρις ότου 'επιβιώσουν' οι καλύτερες. Συνήθως, όσο περισσότερες γενιές περνούν τόσο καλύτερες λύσεις βρίσκονται, μπορεί όμως ο αλγόριθμος να βρεθεί σε σημείο του πεδίου των λύσεων από όπου και δεν μπορεί να προχωρήσει λόγω του ότι βρίσκεται σε τοπικό μέγιστο. Για το λόγο αυτό έχουν υπάρχουν διαφορετικές εκδοχές του αλγόριθμου ανάλογα με τη μορφή του προβλήματος.

Οι εξελικτικές στρατηγικές (ΕΣ) προτάθηκαν στις αρχές της δεκαετίας του 60 από τους Recheberg και Schwefel, και έχουν μελετηθεί εκτενώς τις τελευταίες τρεις δεκαετίες [Schwefel, 1981, Back and Schwefel, 1993]. Ανήκουν στην οικογένεια των εξελικτικών αλγορίθμων, και απευθύνονται στην αντιμετώπιση προβλημάτων βελτιστοποίησης παραμέτρων πραγματικών αριθμών. Εκεί εμφανίζεται και η βασική τους διαφοροποίηση από τους γενετικούς αλγορίθμους, οι οποίοι, γενικά, στοχεύουν στην ελαχιστοποίηση συναρτήσεων με αριθμήσιμο πεδίο ορισμού. Τα βασικά τους χαρακτηριστικά είναι τα παρακάτω [Eiben and Smith, 2003]:

- Οι ΕΣ τυπικά χρησιμοποιούνται για βελτιστοποίηση συνεχών παραμέτρων.
- Δίνεται έμφαση στη μετάλλαξη για την δημιουργία απογόνων.
- Η μετάλλαξη υλοποιείται με την προσθήκη τυχαίου θορύβου κανονικής κατανομής.
- Οι παράμετροι της μετάλλαξης μεταβάλλονται κατά την διάρκεια εκτέλεσης του αλγορίθμου.

Μία ΕΣ είναι ένας στοχαστικός αλγόριθμος αναζήτησης ο οποίος απευθύνεται στο εξής πρόβλημα: Να ελαχιστοποιηθεί μία μη γραμμική συνάρτηση η οποία είναι μία απεικόνιση του χώρου αναζήτησης $S \subseteq R^n$ στον R. Τα βήματα αναζήτησης εισάγονται με στοχαστική μεταβλητότητα, η επονομαζόμενη μετάλλαξη, σημείων ή συνδυασμών σημείων που έχουν παραχθεί κατά την εκτέλεση του αλγορίθμου. Η μετάλλαξη συνήθως πραγματοποιείται με την προσθήκη ενός τυχαίου διανύσματος, προερχόμενου από κανονική κατανομή [Hansen and Ostermeier, 2001] [Hansen et al., 2003] [Hansen and Kern, 2004]. Οι ΕΣ έχουν χρησιμοποιηθεί με επιτυχία σε πλήθος εφαρμογών, όπως ο έλεγχος στροβιλισμού [Koumoutsakos et al., 1998], η λύση αντίστροφων προβλημάτων επεξεργασίας εικόνων και φίλτρα παρακολούθησης radar.

4.2 Ανακατασκευή εικόνων με τη χρήση εξελικτικά βελτιωμένων ροπών

Σε αυτό το κεφάλαιο προτείνεται μία γενική τεχνική για την επιλογή βέλτιστης συνάρτησης βάρους για ανακατασκευή σήματος. Με δεδομένο ένα μονοδιάστατο σήμα μήκους N, η ανακατασκευή του είναι εφικτή με τη χρήση N ορθογώνιων διανυσμάτων μήκους N. Μία προσέγγιση μπορεί, όμως, να εξαχθεί χρησιμοποιώντας μόνο $K, (K < N)$ διανύσματα. Αυτά τα διανύσματα είναι οι τιμές των ορθογωνίων πολυωνύμων με διακριτή βάση. Χρησιμοποιώντας τη διαδικασία ορθοκανονικοποίησης Gram-Schmidt ,αυτά τα διανύσματα μπορούν εύκολα να υπολογιστούν, δεδομένης συνάρτησης βάρους w. Το σφάλμα αυτής της ανακατασκευής είναι:

$$e(x) = f(x) - F^K_{recon}(x) \tag{4.2}$$

Αθροίζοντας τις απόλυτες τιμές από τα παραπάνω, έχουμε ένα απλό μέτρο της ακρίβειας της ανακατασκευής:

$$E = \sum_{x=0}^{N-1} \left| f(x) - F^K_{recon}(x) \right| \tag{4.3}$$

Για εικόνες και, ως εκ τούτου, ορθογώνια πολυώνυμα δύο μεταβλητών, η παραπάνω εξίσωση γίνεται:

$$E_{2D} = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} \left| f(x,y) - F^K_{recon}(x,y) \right| \tag{4.4}$$

υποθέτοντας, χωρίς απώλεια της γενικότητας, τετράγωνες εικόνες. Επιγραμματικά:

- Με δεδομένη μία συνάρτηση βάρους w υπολογίζεται ένας αριθμός πολυωνύμων χρησιμοποιώντας την διαδικασία ορθοκανονικοποίησης Gram-Schmidt.(Εξ. 4.1)
- Με χρήση των παραπάνω πολυωνύμων υπολογίζονται K ροπές του δοσμένου σήματος.
- Πραγματοποιείται η ανακατασκευή του σήματος χρησιμοποιώντας αυτές τις ροπές.
- Η συνάρτηση σφάλματος εξάγεται για το ανακατασκευασμένο σήμα. (Εξ. 4.4)

Η παραπάνω διαδικασία μπορεί να θεωρηθεί μία απεικόνιση της συνάρτησης βάρους $w \in R^N$ στην τιμή σφάλματος $E \in R$. Σε αυτήν την εργασία χρησιμοποιείται μία ΕΣ με προσαρμογή πίνακα συσχέτισης (Covariance Matrix Adaptation, CMA-ES) για τον υπολογισμό της συνάρτησης βάρους w η οποία ελαχιστοποιεί τα σφάλματα ανακατασκευής, όπως αυτά ορίζονται στη σχέση 4.4.

Οι παράμετροι αυτής της εξελικτικής στρατηγικής είναι:

- Τα αρχικά χρωμοσώματα είναι ένας αριθμός τυχαία παραγόμενων συναρτήσεων βάρους, απεικονισμένες ως διανύσματα πραγματικών τιμών μεγέθους N.
- Η συνάρτηση καταλληλότητας είναι το σφάλμα ανακατασκευής.
- 12 χρωμοσώματα βρέθηκαν αρκετά για ταχεία σύγκλιση.

Ο αριθμός των χρωμοσωμάτων επηρεάζει δραστικά τη δυνατότητα και την ταχύτητα σύγκλισης. Λόγω της ενδογενούς τυχαιότητας της παρούσας προσέγγισης, δεν μπορεί να υπάρξει ακριβής αριθμός ο οποίος σε κάθε περίπτωση να δίνει ταχεία και ακριβή λύση. Μικρός αριθμός χρωμοσωμάτων σημαίνει μικρότερο πληθυσμό και πιο αργή σύγκλιση. Επιπλέον, με μικρό πληθυσμό είναι αυξημένη η πιθανότητα (η οποία πάντα υπάρχει) ο αλγόριθμος να τερματίσει σε τοπικό ελάχιστο, με αποτέλεσμα μη επιτυχή ελαχιστοποίηση. Εντούτοις, το υπολογιστικό κόστος του αλγορίθμου είναι ανάλογο του αριθμού των χρωμοσωμάτων. Έτσι, ένα μεγάλο πλήθος χρωμοσωμάτων απαιτεί αντίστοιχο χρόνο εκτέλεσης του αλγορίθμου. Στην περίπτωση αυτή, όμως, η πιθανότητα επιτυχημένης ελαχιστοποίησης αυξάνεται. Αξίζει να σημειωθεί, όμως, ότι μικρές μεταβολές στο μέγεθος του αρχικού πληθυσμού δεν επηρεάζουν το τελικό αποτέλεσμα. Στην παρούσα προσέγγιση, για παράδειγμα, 11 ή 13 χρωμοσώματα θα έδιναν γενικά τα ίδια αποτελέσματα, τόσο σε χρόνο, όσο και σε ακρίβεια.

Τις περισσότερες φορές, ο αλγόριθμος συνέκλινε μετά από 2.000-3.000 υπολογισμούς της συνάρτησης. Τα πειραματικά αποτελέσματα έδειξαν ότι η βέλτιστη συνάρτηση βάρους είναι γενικά συμμετρική ως προς το $N/2$. Έτσι, το μέγεθος των χρωμοσωμάτων μειώθηκε σε $N/2$.

Επιπλέον, δοκιμάστηκαν για αυτόν το σκοπό γενετικοί αλγόριθμοι. Ως συναρτήσεις επιλογής, μετάλλαξης και συνδυασμού επιλέχθηκαν αντίστοιχα οι Stochastic Universal Sampling [Goldberg, 1989], Extended Intermediate Recombination και Breeder Genetic Algorithm [Moehlenbein and Schlierkamp-Voosen, 1993]. Η υλοποίηση βασίστηκε στο Genetic Algorithm Toolbox for MATLAB [Chipperfield et al., 1995]. Ο αλγόριθμος, όμως, συνέκλινε συχνά σε τοπικά ελάχιστα και παρουσίαζε πολύ μεγάλη υπολογιστική πολυπλοκότητα. Έτσι, προτιμήθηκαν οι ΕΣ.

Τα παραγόμενα ορθογώνια πολυώνυμα έδωσαν αυξημένη ακρίβεια ανακατασκευής σε σχέση με τις κοινά χρησιμοποιούμενες ροπές. Από τα πειραματικά αποτελέσματα της παραγράφου 4.6.1 μπορούν να εξαχθούν τα παρακάτω συμπεράσματα:

- Το σύνολο των σημάτων ή των εικόνων είναι αυτό που ορίζει το σύνολο των ορθογωνίων πολυωνύμων που δίνει την μεγαλύτερη ακρίβεια ανακατασκευής. Κανένα συγκεκριμένο σύνολο ροπών δεν παρουσιάστηκε ως βέλτιστο για όλες τις περιπτώσεις.

- Ανάλογα με το σύνολο των εικόνων ή σημάτων, η βελτίωση σε σχέση με τις κοινά χρησιμοποιούμενες ροπές, όπως οι Chebychev, μπορεί να φτάσει το 45%. Έτσι, με δεδομένο ένα αρχικό σύνολο εκπαίδευσης, μπορούν να εξαχθούν ροπές βάσει εφαρμογής.

Προφανώς, περιγραφείς βέλτιστοι για ένα σύνολο εικόνων δεν θα αποδίδουν τόσο σε ένα άλλο. Ανάλογα με το σύνολο εκπαίδευσης, τα εξαγόμενα διανύσματα είναι βελτιστοποιημένα για ένα εύρος εικόνων. Στην περίπτωση που το σύνολο αυτό είναι ανομοιόμορφο, οι περιγραφείς θα είναι βελτιωμένοι για ένα μεγαλύτερο εύρος. Στις περισσότερες εφαρμογές, όμως, η βελτιστοποίηση απαιτείται μόνο για ένα στενό πεδίο. Θα πρέπει να σημειωθεί ότι περιγραφείς βελτιστοποιημένοι για ένα τέτοιο πεδίο παρουσιάζουν πολύ χαμηλή ακρίβεια για διαφορετικού τύπου εικόνες.

4.3 Βελτιστοποιημένοι περιγραφείς σχήματος

Σε αυτή την ενότητα τροποποιείται η παραπάνω τεχνική για την επιλογή μίας βέλτιστης συνάρτησης βάρους για αναγνώριση σχήματος. Με δεδομένη μία συνάρτηση δύο διαστάσεων f, είναι δυνατόν να εξαχθούν N περιγραφείς σχήματος χρησιμοποιώντας N ορθογώνια διανύσματα. Αυτά τα διανύσματα παράγονται από την διαδικασία ορθοκανονικοποίησης Gram-Schmidt, όπως αυτή περιγράφεται στην παράγραφο 4.1.1. Από τι στιγμή που τα διανύσματα βάρους υπολογίζονται με μία προσέγγιση ελαχιστοποίησης συναρτήσεων, είναι εγγυημένη η βελτιωμένη τους απόδοση σε σχέση με άλλους περιγραφείς ροπών με βάση διακριτά πολυώνυμα, όπως οι Chebychev,οι Krawtchouk, και οι Hahn.

4.3.1 Ορισμός προβλήματος

Η εργασία της αναγνώρισης μπορεί να αναχθεί σε ένα πρόβλημα ελαχιστοποίησης. Με δεδομένη συνάρτηση f δύο διαστάσεων, και ένα σύνολο από N ορθογώνια διανύσματα P_i, οι αντίστοιχοι περιγραφείς δίνονται από την παρακάτω σχέση:

$$Z_{m,n} = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} P_m(x) P_n(y) f(x,y) \tag{4.5}$$

Υποθέτοντας ένα σύνολο K εικόνων οι οποίες ανήκουν σε L κλάσεις. Ονομάζουμε F_t την εικόνα t, με $t \in \{1, 2, \ldots, K\}$. Επιπλέον, ορίζουμε συνάρτηση H ως απεικόνιση από τον χώρο $\{1, 2, \ldots, K\}$ στον $\{1, 2, \ldots, L\}$. Η $H(t)$ αναπαριστά την κλάση στην οποία ανήκει η f_t. Σε ένα πρόβλημα αναγνώρισης σχήματος, η επιθυμητή λύση είναι η συνάρτηση H.

4.3.2 Επιλογή ταξινομητή

Παρόλο που στη βιβλιογραφία αναφέρεται πλήθος ταξινομητών, επιλέχθηκε ένας απλός προς απόδειξη των παραπάνω ισχυρισμών. Ένας πιο ισχυρός αλγόριθμος θα ενίσχυε την ακρίβεια των αποτελεσμάτων, αλλά στόχος της εργασίας είναι η επικύρωση των αποτελεσμάτων σχετικά με τους περιγραφείς εικόνων, και όχι με την ταξινόμηση.

Για κάθε εικόνα f_t εξάγεται ένα σύνολο περιγραφέων Z^t_{mn} χρησιμοποιώντας τη σχέση 4.5. Επιπλέον, κάθε κλάση αντιπροσωπεύεται από μία εικόνα, δηλαδή L εικόνες θεωρούνται εκ των προτέρων ταξινομημένες. Η εικόνα t ανήκει στην κλάση q αν οι περιγραφείς της παρουσιάζουν την ελάχιστη Ευκλείδεια απόσταση προς την αντιπροσωπευτική εικόνα της κλάσης q.

4.3.3 Συνάρτηση καταλληλότητας και επιλογή παραμέτρων ΕΣ

Στην προσέγγιση αυτή χρησιμοποιείται μία ΕΣ για τον υπολογισμό των βέλτιστων ορθογώνιων διανυσμάτων για ένα δεδομένο σύνολο εικόνων. Η πιο κρίσιμη διαδικασία αναφορικά με τις εξελικτικές τεχνικές είναι ο καθορισμός της συνάρτησης καταλληλότητας [Eiben and Smith, 2003]. Μία συνάρτηση καταλληλότητας μπορεί να θεωρηθεί ως μία απεικόνιση από το πεδίο των παραμέτρων σε έναν φυσικό αριθμό. Το μέγεθος του τελευταίου αντιστοιχεί στην καταλληλότητα των παραμέτρων. Στην ακραία περίπτωση, όταν η συνάρτηση καταλληλότητας λαμβάνει μηδενική τιμή, οι παράμετροι εισόδου μπορούν να θεωρηθούν ιδανικές. Σε αυτή την Ενότητα παρουσιάζεται μία διερεύνηση σχετικά με διάφορες συναρτήσεις καταλληλότητας και την επίδρασή τους στην ταξινόμηση του συνόλου εκπαίδευσης.

Η πρώτη συνάρτηση που υλοποιήθηκε βασίστηκε στον αριθμό των λανθασμένων αντιστοιχίσεων. Έτσι, μηδενικό αποτέλεσμα σημαίνει 100% ακρίβεια ενώ ένα αποτέλεσμα ίσο με το πλήθος του συνόλου σημαίνει 0% ακρίβεια. Αυτή η προσέγγιση παρόλο που είναι ευθύς και παρουσιάζει μεγάλη απλότητα, αποδείχθηκε μη αποδοτική. Από τη στιγμή που η τιμή της συνάρτησης καταλληλότητας είναι περιορισμένη σε ένα μικρό σύνολο τιμών $(0, 1, \ldots, N)$ η ΕΣ συγκλίνει σε τοπικά, εσφαλμένα ελάχιστα μετά από λίγες επαναλήψεις. Μικρές μεταβολές των παραμέτρων δεν έχουν καμία επίδραση στην έξοδο, αφού η τελευταία λαμβάνει τιμές από ένα πολύ μικρό σύνολο ακεραίων. Η συνάρτηση καταλληλότητας, δηλαδή, παρουσιάζει ένα πολύ ευρύ (μη αριθμήσιμο) πεδίο ορισμό και ένα διακριτό, πεπερασμένο σύνολο τιμών $(f : R^{N/2} \to K \subset Z)$.

Για να αντιμετωπιστεί το πρόβλημα των διακριτών εξόδων, η επόμενη προσέγγιση ήταν η ελαχιστοποίηση της απόστασης μεταξύ εικόνων της ίδιας κλάσης. Οι εικόνες αυτές θα παρουσιάζουν ελάχιστη απόσταση μεταξύ τους και, επομένως, θα αναγνωριστούν με ακρίβεια. Τα προβλήματα της προηγούμενης προσέγγισης αντιμετωπίστηκαν με επιτυχία και η ΕΣ συνέκλινε γρήγορα. Η τελική λύση, όμως, επέφερε μεγάλη απώλεια πληροφορίας και ελαχιστοποίηση όλων των αποστάσεων. Στην ακραία περίπτωση, αν $Z_{mn}^{t} = 0, \forall m, n, t$, όλες οι αποστάσεις μηδενίζονται.

```
Algorithm Fitness Function (Weight Function W)
  Calculate orthogonal vectors using W
  For each image, calculate moment descriptors Z
  Set FV=0
  Repeat for each image
    Find Euclidean distances to all other images
    Rank images by distance
    Calculate the number of incorrect matches T
    Sum the rankings of correct matches S
    FV=FV+T*10.000+S
  End repeat
End algorithm
```

Σχήμα 4.1: Ψευδοκώδικας εξαγωγής βελτιστοποιημένων περιγραφέων

Η προσέγγιση που τελικά επιλέχθηκε ως η πιο αποτελεσματική βασίστηκε στην κατάταξη των αντικειμένων ιδίων κλάσεων μαζί με το πλήθος των λανθασμένων αντιστοιχίσεων. Συγκεκριμένα, το πλήθος των σφαλμάτων πολλαπλασιάστηκε με έναν μεγάλο αριθμό, στην περίπτωση αυτή 10.000, και προστέθηκε με το άθροισμα των κατατάξεων όλων των εικόνων της κλάσης. Ο ψευδοκώδικας για τη συνάρτηση καταλληλότητας δίνεται παρακάτω:

Παρόλο που και αυτή η συνάρτηση καταλληλότητας μπορεί να πάρει μόνο διακριτές τιμές, τα προβλήματα της πρώτης προσέγγισης δεν εμφανίζονται, καθώς το εύρος των τιμών είναι πολύ μεγαλύτερο. Επιπλέον, μικρές μεταβολές των παραμέτρων επηρεάζουν τώρα την έξοδο του συστήματος.

Στόχος της διαδικασίας ταξινόμησης είναι η ελαχιστοποίηση του αριθμού των εσφαλμένων αντιστοιχιών. Η ευθύς αντιμετώπιση, όμως, παρουσίασε το πρόβλημα του περιορισμένου εύρους τιμών εξόδου. Το πρόβλημα αυτό αντιμετωπίστηκε με την εισαγωγή της κατάταξης των αντικειμένων στην αντίστοιχη κλάση ως ενδιάμεσα βήματα προς την βελτιστοποίηση των παραμέτρων.

Οι παράμετροι της ΕΣ, όπως περιγράφηκαν στην παράγραφο 4.2 βρέθηκαν ικανοποιητικές και σε αυτήν την περίπτωση.

4.4 Περιγραφή αλγορίθμου

Παρακάτω παρατίθεται το περίγραμμα του ολοκληρωμένου συστήματος.

- Στην πρώτη εκτέλεση του αλγορίθμου παράγεται ένας αριθμός τυχαίων διανυσμάτων βάρους.
- Για κάθε ένα από αυτά τα διανύσματα:
  - Τα αντίστοιχα ορθογώνια διανύσματα $P_{m,n}$ υπολογίζονται με τη χρήση της διαδικασίας ορθοκανονικοποίησης Gram-Schmidt.
  - Για κάθε εικόνα υπολογίζονται οι περιγραφείς ροπών χρησιμοποιώντας τα προηγούμενα διανύσματα:

$$Z^t_{m,n} = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} P_m(x) P_n(y) f_t(x,y) \tag{4.6}$$

  - Υπολογίζονται οι ευκλείδειες αποστάσεις μεταξύ των εικόνων:

$$D(t,s) = \sqrt{\sum_m \sum_n (Z^t_{m,n} - Z^s_{m,n})^2} \tag{4.7}$$

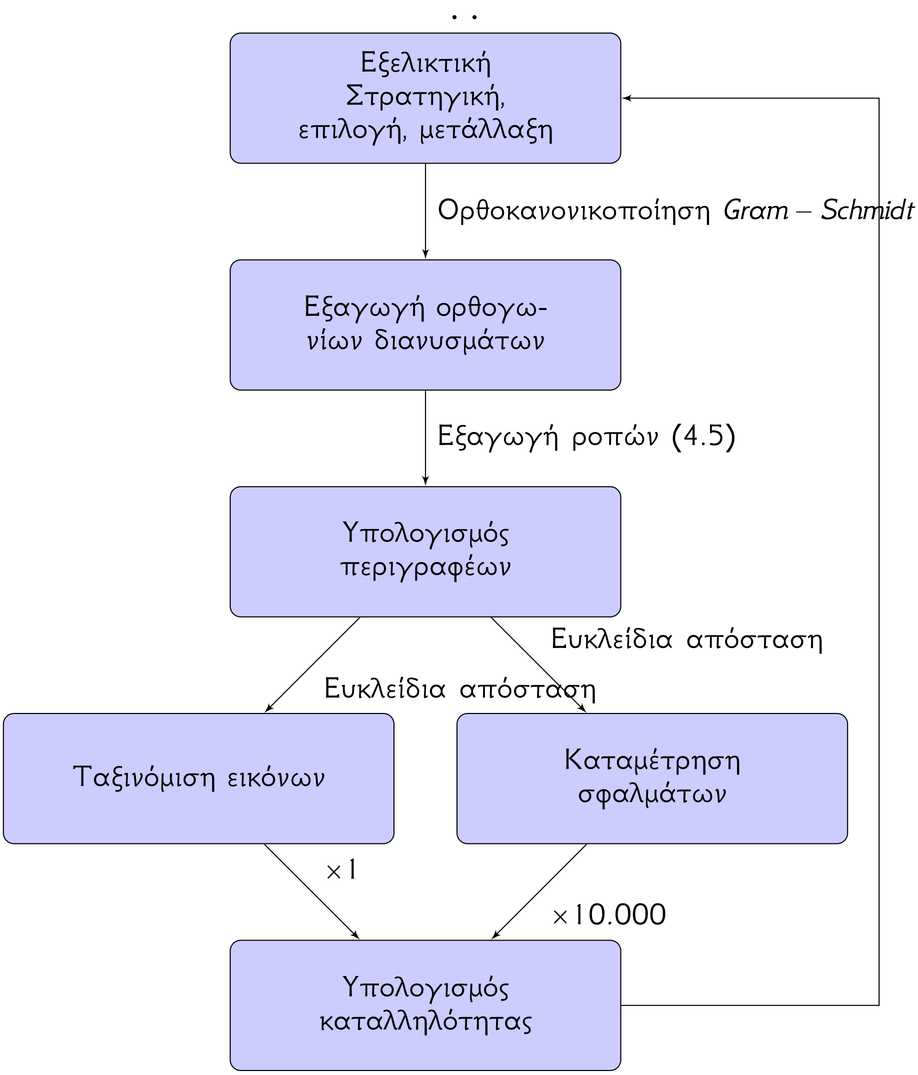
- Ο αριθμός των εσφαλμένων αντιστοιχίσεων T πολλαπλασιάζεται επί 10.000, το άθροισμα των κατατάξεων των αντικειμένων στην ίδια κλάση S υπολογίζεται και αθροίζεται για το σύνολο των εικόνων. Το αποτέλεσμα αντιπροσωπεύει την τιμή καταλληλότητας του συγκεκριμένου διανύσματος βάρους.
- Χρησιμοποιώντας τις υπολογισμένες τιμές καταλληλότητας, τα διανύσματα εισόδου τροποποιούνται και η εκτέλεση επαναλαμβάνεται από το 2ο βήμα, μέχρι να συμπληρωθεί ένας προκαθορισμένος αριθμός επαναλήψεων.

Η διαδικασία παρουσιάζεται σχηματικά στο σχήμα 4.2.

4.5 Ιδιότητες των προτεινόμενων ροπών

4.5.1 Συμμετρία

Από τη στιγμή που η συνάρτηση βάρους είναι συμμετρική ως προς το σημείο $N/2$, η συμμετρία διατηρείται και στα πολυώνυμα. Τα



Σχήμα 4.2: Περίγραμμα διαδικασίας βελτιστοποίησης

πολυώνυμα άρτιου βαθμού είναι συμμετρικά ως προς $N/2$:

$$P_n(x) = P_n(N - x) \tag{4.8}$$

για n άρτιο. Τα πολυώνυμα περιττού βαθμού είναι αντισυμμετρικά:

$$P_n(x) = P_n(N - x) \tag{4.9}$$

Για σήματα δύο διαστάσεων, η συμμετρία επεκτείνεται:

$$\begin{aligned} P_n(x,y) &= (-1)^n P_{n,m}(N - x, y) \\ P_n(x,y) &= (-1)^m P_{n,m}(x, N - y) \\ P_n(x,y) &= (-1)^{m+n} P_{n,m}(N - x, N - y) \end{aligned} \tag{4.10}$$

4.5.2 Αναδρομή

Όπως όλα το ορθογώνια πολυώνυμα, τα προτεινόμενα μπορούν να περιγραφούν με μία αναδρομική σχέση τριών όρων [Gautschi, 2004]:

$$\begin{aligned} P_{k+1}(t) &= (t - \alpha_k)P_k(t) - \beta_k P_{k-1}(t) \\ k &= 0, 1, \ldots, n-1, P_{-1}(t) = 0, P_0(t) = 1 \end{aligned} \tag{4.11}$$

Για τον υπολογισμό των συντελεστών της παραπάνω σχέσης μπορούν να χρησιμοποιηθούν τεχνικές όπως η διακριτή διεργασία Stieltjes [Reichel et al., 1991] ή ο αλγόριθμος του Lanczos [Lanczos, 1988], με δεδομένη τη συνάρτηση βάρους w.

4.5.3 Αναπαράσταση μέσω γεωμετρικών ροπών

Τα παραγόμενα πολυώνυμα είναι δυνατόν να αναπαρασταθούν ως αθροίσματα μονωνύμων. Κατ' επέκταση, η αναπαράσταση των ροπών μέσω γεωμετρικών είναι εφικτή. Με δεδομένη συνάρτηση βάρους w, μπορούν να εξαχθούν οι συντελεστές του πολυωνύμου $\alpha_{i,p}$:

$$P_p(x) = \sum_{i=0}^{p} \alpha_{i,p} x^i \tag{4.12}$$

Τα παραγόμενα πολυώνυμα έχουν την παρακάτω μορφή:

$$Z_{p,q} = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} \sum_{i=0}^{p} \sum_{j=0}^{q} \alpha_{i,p} \alpha_{j,q} x^i y^j f(x,y) \tag{4.13}$$

ή

$$Z_{p,q} = \sum_{i=0}^{p} \sum_{j=0}^{q} \alpha_{i,p}\alpha_{j,q} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} x^i y^j f(x,y) \tag{4.14}$$

Αντικαθιστώντας τον τελευταίο όρο προκύπτει:

$$Z_{p,q} = \sum_{i=0}^{p} \sum_{j=0}^{q} \alpha_{i,p}\alpha_{j,q} m_{i,j} \tag{4.15}$$

με $m_{i,j}$ τη γεωμετρική ροπή τάξης (i,j).

4.5.4 Υπολογιστικό κόστος

Σε αυτή την Ενότητα παρουσιάζονται κάποια χαρακτηριστικά των προτεινόμενων ροπών, σχετικά με την ταχύτητα υπολογισμού τους. Η πιο χρονοβόρα διαδικασία είναι η εξέλιξη των χρωμοσωμάτων (συνάρτηση βάρους). Η διαδικασία αυτή, όμως, απαιτείται να εκτελεστεί μόνο μία φορά, για δεδομένο σύνολο εκπαίδευσης. Χρησιμοποιώντας τις ιδιότητες της συμμετρίας και της αναδρομής, από τη στιγμή που η συνάρτηση βάρους έχει υπολογιστεί, είναι δυνατή η ταχεία εξαγωγή των ροπών. Πιο συγκεκριμένα, τα πολυώνυμα χρειάζεται να υπολογιστούν μόνο σε ένα τεταρτημόριο του χώρου της εικόνας, αφού η εξίσωση 4.10 μπορεί να δώσει της τιμές για τον υπόλοιπο χώρο. Επιπλέον, με τη χρήση της αναδρομικής σχέσης, απαιτούνται μόλις 2 πολλαπλασιασμοί και 2 προσθέσεις ανά στοιχείο και ανά τάξη ροπής. Επομένως, το συνολικό υπολογιστικό κόστος για τον υπολογισμό ενός πολυωνύμου σε μία περιοχή $N \times N$ στοιχείων είναι:

$$\#Mult = \#Add = \frac{N^2}{2} \tag{4.16}$$

Για τον υπολογισμό της αντίστοιχης ροπής:

$$\#Mult = \#Add = N^2 \tag{4.17}$$

Τελικά, για τον υπολογισμό όλων των ροπών τάξης μέχρι M, ο συνολικός αριθμός των απαιτούμενων πράξεων είναι:

$$\#Mult = \#Add = \frac{3N^2}{2} M^2 \tag{4.18}$$

Για τον αποδοτικό υπολογισμό των γεωμετρικών ροπών έχουν προταθεί πλήθος τεχνικών λογισμικού αλλά και υλικού [Kotoulas and Andreadis, 2004] [Li, 1993] [Lin and Wang, 1994] [Chung and Chen, 2005]. Με τη χρήση τέτοιων τεχνικών, ο υπολογισμός των ροπών βάσει των γεωμετρικών μπορεί να αυξήσει σημαντικά την ταχύτητα. Αξίζει να σημειωθεί ότι η σχέση 4.15 είναι γραμμική, οπότε οι ροπές μπορούν να υπολογιστούν αποδοτικά ακόμα και για εικόνες μεγάλου μεγέθους. Το υπολογιστικό κόστος για τον υπολογισμό ροπής τάξης p,q με δεδομένες τις τιμές των γεωμετρικών είναι $2pq$ πολλαπλασιασμοί και pq προσθέσεις.

Μία αρχιτεκτονική υλικού, όπως αυτή που περιγράφεται στην [Kotoulas and Andreadis, 2004], μπορεί να χρησιμοποιηθεί για τον άμεσο υπολογισμό των γεωμετρικών ροπών. Θα πρέπει να αναφερθεί, όμως, ότι για εικόνες μεγάλου μεγέθους ή και για ροπές υψηλής τάξης, οι γεωμετρικές ροπές μπορούν να λάβουν πολύ μεγάλες τιμές. Παρόλο που οι προτεινόμενες ροπές είναι κανονικοποιημένες και δεν παρουσιάζουν αριθμητικές αστάθειες, η εξαγωγή τους από τις γεωμετρικές ροπές μπορεί να οδηγήσει σε σημαντικά σφάλματα.

4.5.5 Εύρος τιμών

Ένα από τα σημαντικότερα προβλήματα των γεωμετρικών ροπών είναι το πολύ μεγάλο εύρος τιμών τους. Για παράδειγμα, για μία εικόνα μεγέθους 256×256 στοιχείων, ένα στοιχείο με συντεταγμένες $(256, 256)$ συμβάλλει σε ροπή τάξης 16 κατά 256^{16} ή 2^{128}. Επομένως, αριθμητικά σφάλματα μπορεί να οδηγήσουν σε ολοκληρωτικά εσφαλμένες τιμές ροπών. Είναι, λοιπόν, προφανές ότι ο υπολογισμός μέσω γεωμετρικών ροπών απαιτεί ιδιαίτερους χειρισμούς. Οι προτεινόμενες ροπές, από την άλλη πλευρά, έχουν ως βάση ορθοκανονικά πολυώνυμα, των οποίων οι τιμές δεν υπόκεινται σε αριθμητικές αστάθειες. Παρακάτω δίνεται μία συνοπτική απόδειξη.

Έστω οικογένεια ορθοκανονικών πολυωνύμων $P_n, n = \{P_0, P_1, \ldots\}$ βαθμού n, με διακριτό πεδίο ορισμού $\{0, 1, \ldots, N-1\}$ Λόγω της κανονικοποίησης θα ισχύει:

$$\sum_{x=0}^{N-1} P_n(x)^2 = 1 \tag{4.19}$$

Ο παραπάνω περιορισμός σημαίνει ότι για κάθε $x \in \{0, 1, \ldots, N-1\}$ θα ισχύει:

$$-1 \leq P_n(x) \leq 1 \tag{4.20}$$

Επομένως, είναι προφανές ότι ακόμα και στην ακραία περίπτωση που για κάποιο x_0, $P_n(x_0) = 1$ και $f(x_0) = 1$ τότε $M_p = 1$.

4.5.6 Αμεταβλητότητα σε γραμμικούς μετασχηματισμούς

Όπως όλες οι αναφερόμενες στη βιβλιογραφία ροπές με βάση διακριτά πολυώνυμα, έτσι και οι προτεινόμενες δεν είναι αμετάβλητες σε γραμμικούς μετασχηματισμούς. Η αμεταβλητότητα, όμως, ίσως είναι και το πιο σημαντικό χαρακτηριστικό των ροπών. Για την εισαγωγή αυτού του χαρακτηριστικού σε τέτοιους περιγραφείς έχουν προταθεί δυο προσεγγίσεις. Η πρώτη βασίζεται στην κανονικοποίηση της εικόνας πριν την ανάλυση. Τεχνικές κανονικοποίησης, όπως αυτές που παρουσιάζονται στις [Gruber and Hsu, 1997] και [Shen and Horace, 1997], μπορούν να χρησιμοποιηθούν με επιτυχία για την κανονικοποίηση της εικόνας ως προς το μέγεθος και την περιστροφή. Μία δεύτερη προσέγγιση είναι αυτή που προτάθηκε από τους Yap et al. [Yap et al., 2003], ως μέσο για την εισαγωγή αμεταβλητότητας στις ροπές Krawtchouk.

Και οι δύο προσεγγίσεις υλοποιήθηκαν και δοκιμάστηκαν, με οριακές διαφορές στην ακρίβεια. Όπως ήταν αναμενόμενο, και οι δύο έδειξαν χαμηλή απόδοση σε μικρές εικόνες και ικανοποιητική σε μεγαλύτερες. Ενώ στη γενική περίπτωση η δεύτερη προσέγγιση είναι πιο γρήγορη από την πρώτη, στη συγκεκριμένη περίπτωση, η ανάγκη για επανειλημμένους υπολογισμούς ροπών στις ίδιες εικόνες συνέβαλε στην επιλογή της πρώτης. Όλες οι εικόνες, λοιπόν, κανονικοποιήθηκαν πριν τη διαδικασία εκπαίδευσης.

4.6 Πειραματικά αποτελέσματα

4.6.1 Ανακατασκευή

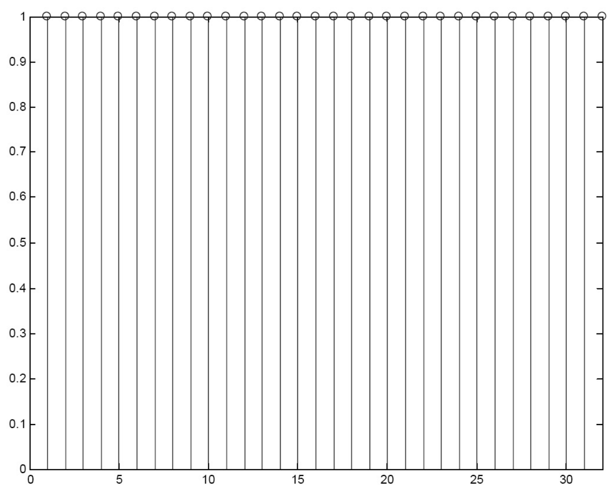
Ανακατασκευή μονοδιάστατων σημάτων

Τα πρώτα αποτελέσματα της προτεινόμενης τεχνικής αναφέρονται στην ανακατασκευή μονοδιάστατων σημάτων.

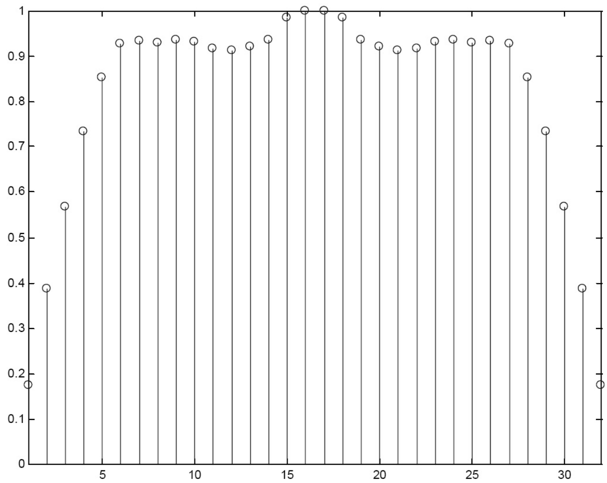
Πειραματικό περιβάλλον Για τη βελτιστοποίηση των συναρτήσεων βάρους διακριτών ορθογωνίων συστημάτων, υλοποιήθηκε το παρακάτω σύστημα:

256 σήματα μήκους 32 χρησιμοποιήθηκαν ως σύνολο εκπαίδευσης σε κάθε δοκιμή. Η ανακατασκευή πραγματοποιήθηκε χρησιμοποιώντας τις 8 πρώτες ροπές. Η ΕΣ συνέκλινε μετά από 1.000 με 10.000 επαναλήψεις. Κάθε δοκιμή πραγματοποιήθηκε δύο φορές, προς επιβεβαίωση του τοπικού ελαχίστου και για κάθε περίπτωση χρησιμοποιήθηκαν 3 διαφορετικά σύνολα εκπαίδευσης. Εκτός αν αναφέρεται διαφορετικά, το τελικό βέλτιστο διάνυσμα παρέμεινε σταθερό και για τα 6 πειράματα κάθε περίπτωσης. Πρέπει να σημειωθεί ότι αυτή η διαδικασία είναι εξαιρετικά χρονοβόρα, απαιτώντας περίπου 24 ώρες ανά πείραμα σε έναν προσωπικό υπολογιστή με κεντρικό επεξεργαστή Pentium 4, 4GHz. Όλα τα αποτελέσματα συγκρίνονται με αυτά που δίνουν οι ροπές Chebychev, αφού τα πολυώνυμα Chebychev μπορούν να εξαχθούν με την μέθοδο ορθοκανονικοποίησης Gram-Schmidt με την απλούστερη συνάρτηση βάρους, συγκεκριμένα την $w(x) = 1$.

Τυχαίες ακολουθίες Δοκιμές σε ομοιόμορφες τυχαίες ακολουθίες από φυσικούς αριθμούς παρουσίασαν ελάχιστο σφάλμα ανακατασκευής για σταθερή συνάρτηση βάρους (Σχήμα 4.3). Η σταθερή συνάρτηση βάρους ($w(x) = c$) αντιστοιχεί σε ροπές Chebychev. Τυχαίες ακολουθίες δυαδικών αριθμών συνέκλιναν στο ίδιο αποτέλεσμα. Οι συναρτήσεις Chebychev, αποτελούν και την πιο φυσική επιλογή, για πλήρως ομοιόμορφο σύνολο εκπαίδευσης, αφού δεν ευνοούν κάποια συγκεκριμένη κατανομή.



Σχήμα 4.3: Σταθερή συνάρτηση βάρους

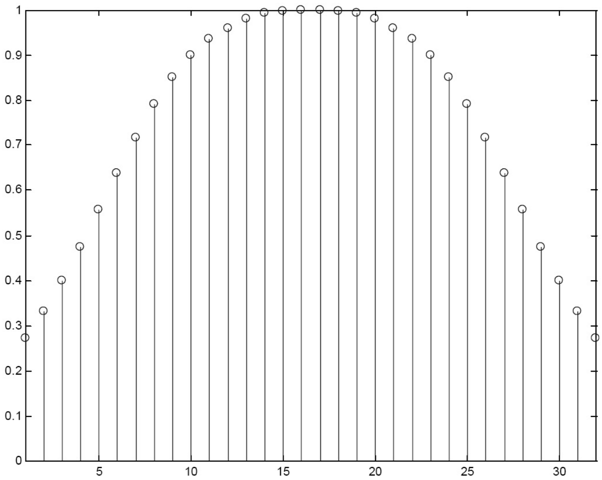


Σχήμα 4.4: Βέλτιστη συνάρτηση βάρους για τυχαία σήματα χαμηλής συχνότητας

Ακολουθίες χαμηλής συχνότητας Η βέλτιστη συνάρτηση βάρους για τυχαία σήματα τα οποία επεξεργάστηκαν με βαθυπερατό φίλτρο παρουσιάζεται στο Σχήμα 4.4. Χρησιμοποιώντας σταθερή συνάρτηση βάρους, η οποία αντιστοιχεί σε ροπές Chebychev, το μέσο σφάλμα ανακατασκευής ανά στοιχείο βρέθηκε 0,0501. Χρησιμοποιώντας την συνάρτηση που προέκυψε από τις ΕΣ το σφάλμα αυτό μειώθηκε κατά 15%, σε 0,0426. Είναι προφανές ότι στην περίπτωση αυτή, η εξελικτική βελτιστοποίηση οδήγησε στην παραγωγή διανυσμάτων βάρους τα οποία δίνουν μεγαλύτερη βαρύτητα στις χαμηλές συχνότητες, μειώνοντας την πληροφορία σχετικά με τις υψηλές.

Ακολουθίες υψηλής συχνότητας Οι υψίσυχνες ακολουθίες δεν μπορούν να περιγραφούν με ακρίβεια από ροπές χαμηλής τάξης. Τα αποτελέσματα της ανακατασκευής τους παρουσίασαν εξαιρετικά υψηλές τιμές σφάλματος. Με τη βελτιστοποίηση της συνάρτησης βάρους από ΕΣ, αυτά τα σφάλματα μειώθηκαν οριακά, αλλά όχι σε βαθμό που θα επέτρεπε σε διακριτά ορθογώνια πολυώνυμα να χρησιμοποιηθούν στην ανακατασκευή τέτοιων σημάτων.

Γραμμές σάρωσης εικόνων Σε αυτό το πείραμα επιλέχθηκε ως σύνολο εκπαίδευσης ένας αριθμός οριζοντίων γραμμών σάρωσης πραγματικών



Σχήμα 4.5: Βέλτιστη συνάρτηση βάρους για γραμμές σάρωσης εικόνων

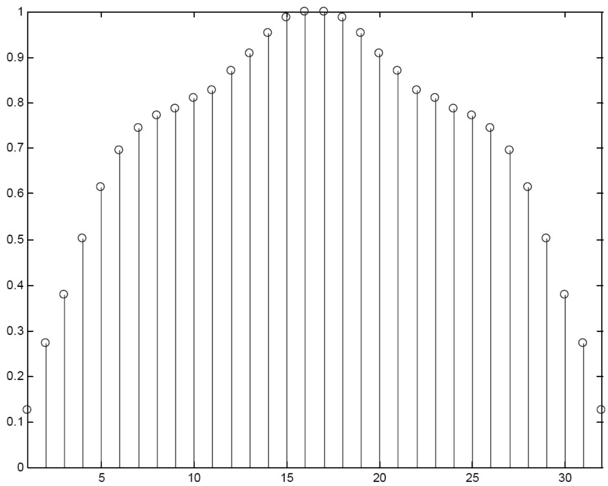
εικόνων. Από τη στιγμή που αυτές είναι συνήθως χαμηλής συχνότητας, σε σχέση με ακολουθίες τυχαίων αριθμών, η βέλτιστη συνάρτηση βάρους βρέθηκε παρόμοια με αυτή της Ενότητας 4.6.1. Οι γραμμές σάρωσης επιλέχθηκαν τυχαία από ένα σύνολο εικόνων. Η βέλτιστη συνάρτηση βάρους παρουσιάζεται στο Σχήμα 4.5.

Ανακατασκευή σημάτων δύο διαστάσεων

Πειραματικό περιβάλλον Το περιβάλλον στο οποίο πραγματοποιήθηκαν τα παρακάτω πειράματα είναι παρόμοιο με το προηγούμενο. Χρησιμοποιήθηκαν ως σύνολα εκπαίδευσης επεξεργασμένες και μη τυχαίες εικόνες, πραγματικές εικόνες και εικόνες χαρακτήρων. Κάθε σύνολο περιείχε 64 εικόνες και κάθε πείραμα εκτελέστηκε 2 φορές. Όλες οι εικόνες είχαν χωρική ανάλυση 4.096 στοιχεία με εξαίρεση τις φωτογραφίες, οι οποίες είχαν 16.384 στοιχεία.

Τυχαίες εικόνες Πάλι, για τυχαίες εικόνες οι ροπές Chebychev έδωσαν τα καλύτερα αποτελέσματα ανακατασκευής. Η συνάρτηση βάρους συνέκλινε στην τιμή 1.

Τυχαίες εικόνες χαμηλών συχνοτήτων Σε αυτό το πείραμα το δοκιμαστικό σύνολο ένας αριθμός από τυχαίες εικόνες αποχρώσεων του γκρι, επεξεργασμένες με ένα φίλτρο μέσης τιμής διαστάσεων 3x3. Η

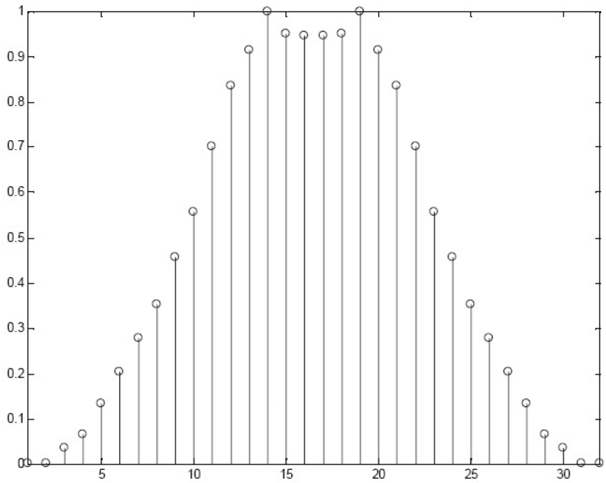


Σχήμα 4.6: Βέλτιστη συνάρτηση βάρους για τυχαίες εικόνες χαμηλής συχνότητας

συνάρτηση βάρους η οποία έδωσε τα πιο ακριβή αποτελέσματα ανακατασκευής παρουσιάζεται στο Σχήμα 4.6. Σε σχέση με τις ροπές Chebychev, η βελτίωση βρέθηκε να είναι $7,2\%$.

Εικόνες Χαρακτήρων Οι δοκιμές σε εικόνες χαρακτήρων έδωσαν πολύ πιο ακριβή αποτελέσματα σε σχέση με τις ροπές Chebychev. Πιο συγκεκριμένα, η βελτίωση της ακρίβειας ανακατασκευής με χρήση ροπών μέχρι 8ης τάξης με βάση συνάρτηση βάρους αυτή που παρουσιάζεται στο σχήμα 4.7 ήταν 35%. Μερικά παραδείγματα φαίνονται στο σχήμα 4.8. Αξίζει να σημειωθεί οτι οι δύο τελευταίες σειρές (Ελληνικοί χαρακτήρες) δεν περιέχονταν σε κανένα σύνολο εκπαίδευσης. Η ανακατασκευή μέσω ροπών Krawtchouk έδειξαν βελτίωση της τάξης του 15% σε σχέση με τις Chebychev.

Πραγματικές Εικόνες Τα πειράματα σε πραγματικές εικόνες παρουσιάζουν αυξημένη ακρίβεια ανακατασκευής για τη χρήση βελτιστοποιημένων ροπών. Για ροπές μέχρι 8ης τάξης η ακρίβεια βελτιώθηκε κατά περίπου 29%. Παραδείγματα τέτοιας ανακατασκευής παρουσιάζονται στο σχήμα 4.9.



Σχήμα 4.7: Βέλτιστη συνάρτηση βάρους για εικόνες χαρακτήρων

4.6.2 Περιγραφή σχήματος

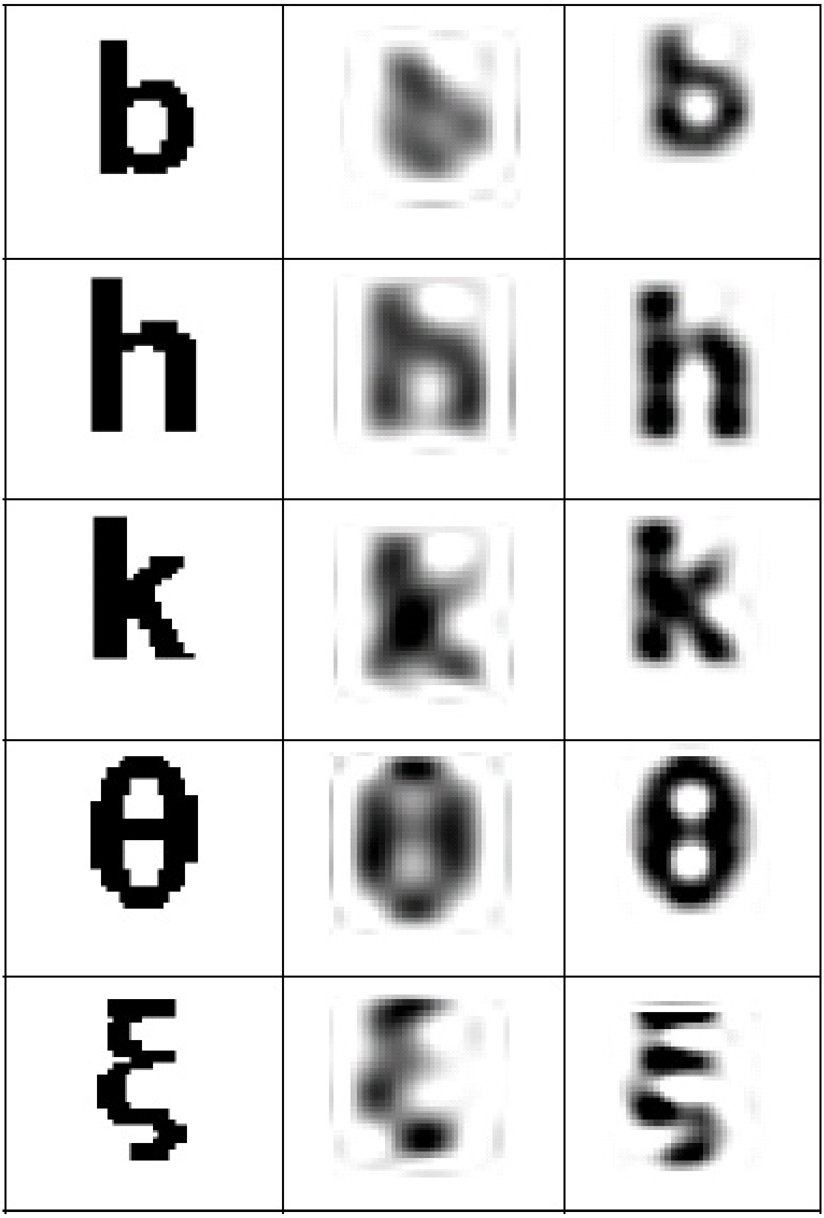
Λατινικοί χαρακτήρες Ως πρώτο βήμα για την δοκιμή της προτεινόμενης προσέγγισης στην περιγραφή του σχήματος εικόνων επιλέχθηκε το σύνολο των Λατινικών μικρών χαρακτήρων. Από το σύνολο των 26 χαρακτήρων κατασκευάστηκαν 130 εικόνες με τυχαία μεγέθυνση και περιστροφή. Επιπλέον, στις μισές από αυτές προστέθηκε Γκαουσσιανός θόρυβος χαμηλού επιπέδου ($\sigma = 0.1$). Για κάθε εικόνα εξάχθηκαν 64 ροπές. Μερικά παραδείγματα φαίνονται στο σχήμα 4.10. Μετά τη σύγκλιση της ΕΣ, δεν παρατηρήθηκαν εσφαλμένες αντιστοιχήσεις και η εργασία βρέθηκε ως η ευκολότερη. Οι ροπές Chebychev, όμως, παρουσίασαν 3 σφάλματα, αντιστοιχώντας σε ακρίβεια 99.6%. Η δοκιμή επαναλήφθηκε για 16 ροπές ανά εικόνα. Η ακρίβεια με τη χρήση ροπών Chebychev βρέθηκε 95%, ενώ της προτεινόμενης τεχνικής 98.7%. Η τελική συνάρτηση βάρους παρουσιάζεται στο σχήμα 4.11

Χειρόγραφοι χαρακτήρες Το δεύτερο πείραμα για την δοκιμή της προτεινόμενης προσέγγισης αποτελούνταν από ένα σύνολο 120 χειρόγραφων ελληνικών χαρακτήρων. Κάθε ένας από τους 24 χαρακτήρες σχεδιάστηκε και σαρώθηκε 5 φορές. Ο θόρυβος που προκλήθηκε από την διαδικασία σάρωσης δεν αφαιρέθηκε. Κάποια παραδείγματα παρουσιάζονται στο σχήμα 4.12 Όπως ήταν αναμενόμενο, τα αποτελέσματα της ταξινόμησης για αυτό το σύνολο ήταν πολύ λιγότερο ακριβή σε σχέση με τα προηγούμενα. Οι ροπές Chebychev παρουσίασαν κατά μέσο όρο 2.5 εσφαλμένες αντιστοιχήσεις, σε ένα σύνολο από 300, δηλαδή μία ακρίβεια 50%. Η προτεινόμενη τεχνική παρήγαγε ένα ακριβές διάνυσμα βάρους μετά από περίπου 3.000 επαναλήψεις. Με τους υπολογισμένους περιγραφείς, το μέσο σφάλμα μειώθηκε σε 2.08, αντιστοιχόντας σε ακρίβεια 58%. Αξίζει να σημειωθεί ότι από τη στιγμή που η αμεταβλητότητα στην περιστροφή αποτελεί βασική ιδιότητα των περιγραφέων ροπών, ένας αριθμός σφαλμάτων είναι αναπόφευκτος. Αυτό συμβαίνει διότι κάποιοι χαρακτήρες προκύπτουν από άλλους μετά από περιστροφή (π.χ. ϵ, ω). Αγνοώντας τέτοια σφάλματα, η ακρίβεια ταξινόμησης αυξάνεται στο 57% για ροπές Chebychev και 67% για τις προτεινόμενες ροπές.

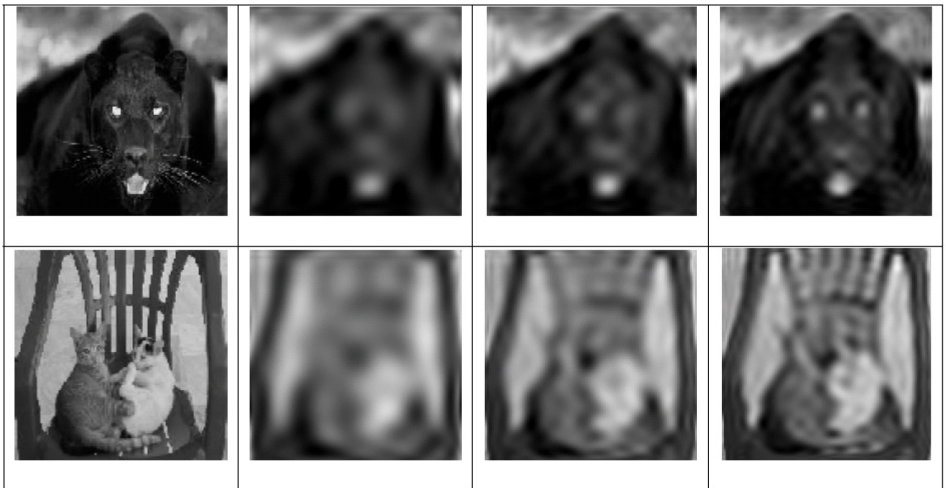
Εικόνες αντικειμένων Το τελευταίο πείραμα δοκιμής της προτεινόμενης μεθοδολογίας ήταν η αναγνώριση αντικειμένων σε μία μεγάλη βάση δεδομένων. Ως σύνολο εκπαίδευσης επιλέχθηκε ένα υποσύνολο από τη συλλογή εικόνων: Amsterdam Library of Object Images [Geusebroek et al., 2005], αποτελούμενου από 120 εικόνες 24 αντικειμένων. Ως πείραμα δοκιμής επιλέχθηκε ένα μεγαλύτερο υποσύνολο, 300 εικόνων. Στόχος του πειράματος ήταν η ταξινόμηση των εικόνων ιδίων αντικειμένων. Οι προτεινόμενες ροπές έδωσαν ακρίβεια 95%, ενώ οι ροπές Chebychev έδωσαν μέση ακρίβεια 85%. Κάποια παραδείγματα παρουσιάζονται στο σχήμα 4.13.

4.7 Συμπεράσματα

Με τη χρήση ΕΣ ως εργαλείο ελαχιστοποίησης συναρτήσεων, κατέστη δυνατή η αντιμετώπιση των περιγραφέων ροπών ως μία εργασία συσχετιζόμενη με την εφαρμογή. Μία τέτοια προσέγγιση εγγυάται βελτιωμένη ακρίβεια σε σχέση με τις προηγούμενα προτεινόμενες ροπές ορθογωνίων πολυωνύμων, όπως οι διακριτές Chebychev, Krawtchouk και Hahn. Παρουσιάστηκαν πειραματικά αποτελέσματα που επιβεβαιώνουν την προσέγγιση αυτή σε εφαρμογές ανακατασκευής και αναγνώρισης σχήματος. Επιπλέον, παρουσιάστηκαν κάποιες ιδιότητες των προτεινόμενων τεχνικών για την αντιμετώπιση του μεγάλου υπολογιστικού κόστους. Οι παρουσιαζόμενοι περιγραφείς μπορούν να χρησιμοποιηθούν σε πλήθος εφαρμογών αναγνώρισης προτύπων, ταξινόμησης αντικειμένων και αντιστοίχισης εικόνων.

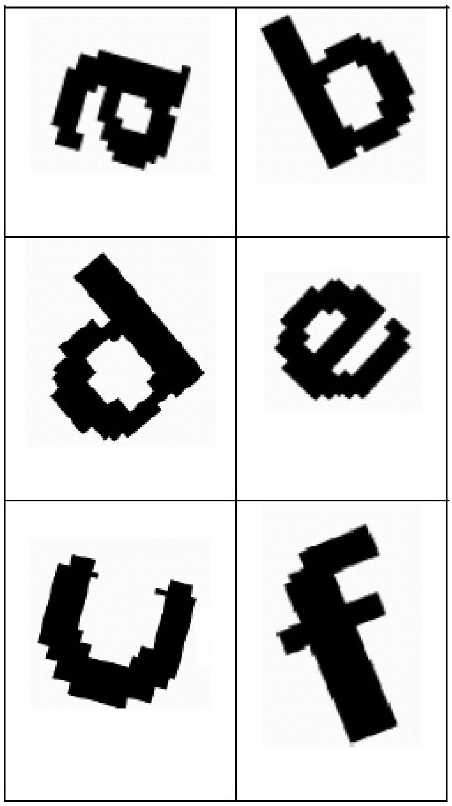


Σχήμα 4.8: Παραδείγματα εικόνων χαρακτήρων

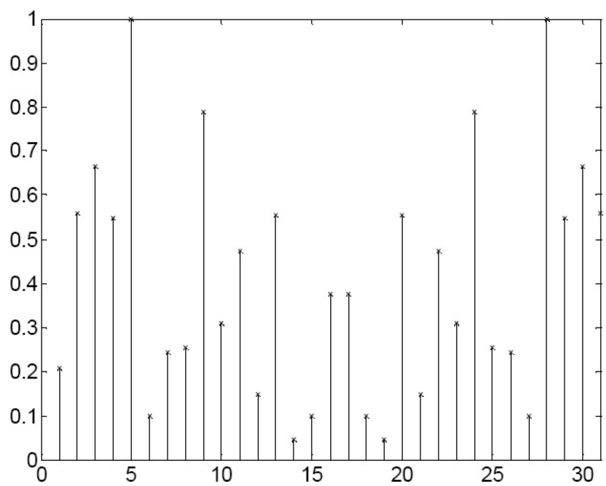


Αρχική	M=16	M=24	M=32

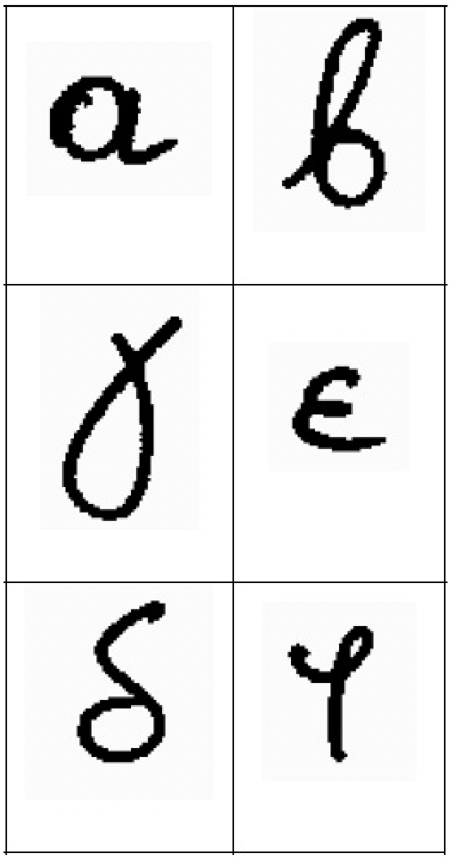
Σχήμα 4.9: Παραδείγματα εικόνων ανακατασκευής



Σχήμα 4.10: Παραδείγματα χαρακτήρων για περιγραφή σχήματος



Σχήμα 4.11: Βέλτιστο διάνυσμα για περιγραφή σχήματος χαρακτήρων



Σχήμα 4.12: Παραδείγματα χειρόγραφων χαρακτήρων για περιγραφή σχήματος

Σχήμα 4.13: Παραδείγματα αντικειμένων για περιγραφή σχήματος

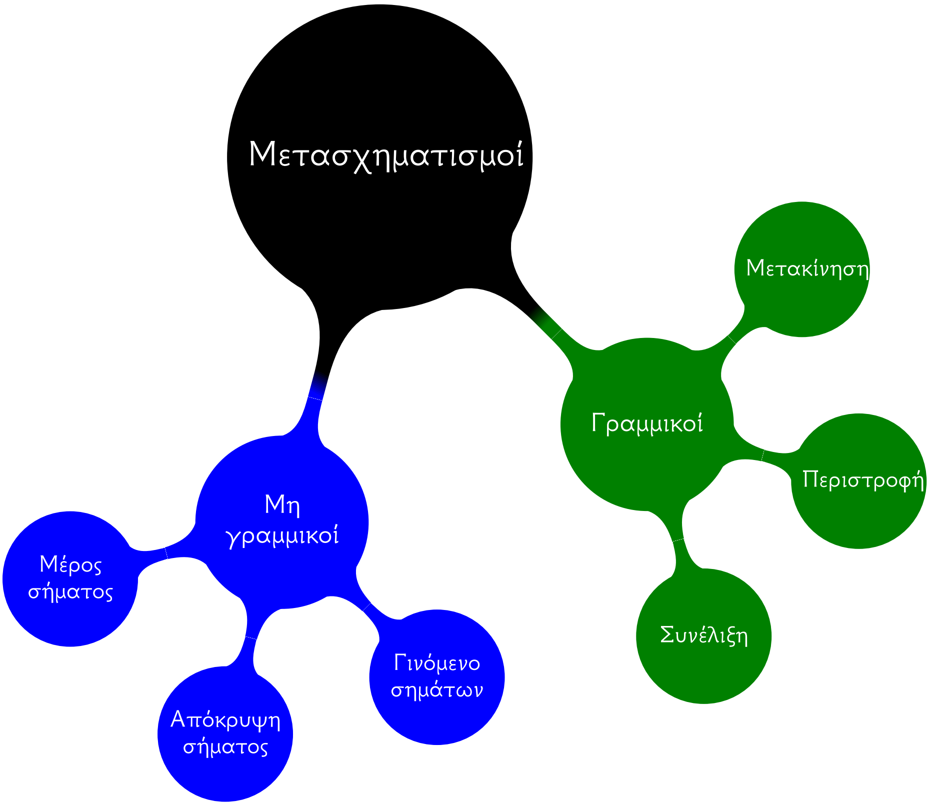
Κεφάλαιο 5

Ιδιότητες των ροπών διακριτής ορθογώνιας βάσης

Οι σχετικά πρόσφατα προταθείσες ροπές διακριτής ορθογώνιας βάσης παρουσιάζουν ιδιαίτερα μικρά σφάλματα ανακατασκευής. Αυτή τους η ιδιότητα δίνει τη δυνατότητα εύκολων μετασχηματισμών μεταξύ του διακριτού χώρου της εικόνας και του χώρου των ροπών. Αυτό επιτρέπει τον υπολογισμό της επίδρασης τυπικών γραμμικών και μη μετασχηματισμών στην εικόνα, με μερικές πολύ χρήσιμες εφαρμογές.

Οι ροπές ορθογώνιας βάσης, όπως οι Chebyshev και οι Krawtchouk χρησιμοποιούνται ολοένα και περισσότερο στο ερευνητικό πεδίο της ανάλυσης εικόνας λόγω των χρήσιμων ιδιοτήτων τους. Πιο συγκεκριμένα, όπως περιγράφηκε στο κεφάλαιο 1, παρουσιάζουν σχετικά μεγάλη ανοχή στο θόρυβο, μεγάλη ακρίβεια στην ανακατασκευή και απλή υλοποίηση. Με βάση τις παραπάνω ιδιότητες, στο κεφάλαιο αυτό παρουσιάζονται εκφράσεις για τον αποδοτικό υπολογισμό ροπών σημάτων που έχουν υποστεί γραμμικούς ή μη μετασχηματισμούς. Στο σχήμα 5.1 παρουσιάζονται οι μετασχηματισμοί στους οποίους αναφέρεται το παρόν κεφάλαιο.

5.1 Γραμμικοί μετασχηματισμοί



Σχήμα 5.1: Επισκόπιση των μετασχηματισμών

5.1.1 Μετακίνηση

Σε αυτή την Ενότητα περιγράφεται η σχέση των ροπών μίας εικόνας που έχει υποστεί μετακίνηση σε σχέση με την αρχική. Πραγματοποιώντας διαδοχικές μετακινήσεις παράλληλα στο σύστημα συντεταγμένων, ο υπολογισμός αυτός γίνεται αποδοτικά, με μικρό υπολογιστικό κόστος.

Για μία εικόνα $f(x,y)$, η οποία μετακινείται κατά τον οριζόντιο άξονα κατά x_0, οι ροπές τις δίνονται από την παρακάτω έκφραση:

$$M_{p,q}^{x_0,0} = \sum_{n=0}^{p} {}_p\alpha_n \sum_{k=0}^{n} \begin{pmatrix} n \\ k \end{pmatrix} \sum_{t=0}^{k} {}_k b_t (-x_0)^{n-k} M_{t,q}^{0,0} \tag{5.1}$$

με $M_{p,q}$ την ροπή τάξης (p,q) της αρχικής. Ομοίως, για κατακόρυφη μετακίνηση κατά y_0 οι νέες ροπές είναι:

$$M^{0,y_0}_{p,q} = \sum_{n=0}^{p} {}_p\alpha_n \sum_{k=0}^{n} \left(\begin{array}{c} n \\ k \end{array} \right) \sum_{t=0}^{k} {}_k b_t (-y_0)^{n-k} M^{0,0}_{p,t} \tag{5.2}$$

Για τον υπολογισμό δεδμοένου μετασχηματισμού x,y, απαιτείται η παρακάτω έκφραση

$$\begin{aligned} M^{x_0,y_0}_{p,q} \quad = \quad & \sum_{n=0}^{p} {}_p\alpha_n \sum_{k=0}^{n} \left(\begin{array}{c} n \\ k \end{array} \right) \sum_{t=0}^{k} {}_k b_t (-x_0)^{n-k} \\ & \sum_{m=0}^{q} {}_q\alpha_m \sum_{l=0}^{m} \left(\begin{array}{c} m \\ l \end{array} \right) \sum_{s=0}^{l} {}_l b_s (-y_0)^{m-l} \\ & M^{0,0}_{t,s} \end{aligned} \tag{5.3}$$

Από την εξίσωση (5.3) φαίνεται ότι πραγματοποιώντας διαδοχικούς απλούς μετασχηματισμούς μειώνεται η υπολογιστική πολυπλοκότητα από $O(n^6)$ σε $O(n^3)$. Τα εξαπλό άθροισμα αντικαθίσταται από δύο τριπλά. Επίσης, ο απ΄ ευθείας υπολογισμός των ροπών για τη νέα εικόνα θα απαιτούσε $O(N \times M)$ πράξεις, με N, M τις διαστάσεις της εικόνας. Επομένως, η χρήση των εκφράσεων (5.1) και(5.2) είναι πλεονεκτική έναντι του άμεσου υπολογισμού για κάθε εικόνα, πέρα από αυτές πολύ μικρών διαστάσεων.

Οι παραπάνω σχέσεις τροποποιούνται εύκολα για την τρίτη διάσταση:

$$M^{x_0,0,0}_{p,q,r} = \sum_{n=0}^{p} {}_p\alpha_n \sum_{k=0}^{n} \left(\begin{array}{c} n \\ k \end{array} \right) \sum_{t=0}^{k} {}_k b_t (-x_0)^{n-k} M^{0,0}_{t,q,r} \tag{5.4}$$

$$M^{0,y_0,0}_{p,q,r} = \sum_{n=0}^{p} {}_p\alpha_n \sum_{k=0}^{n} \left(\begin{array}{c} n \\ k \end{array} \right) \sum_{t=0}^{k} {}_k b_t (-y_0)^{n-k} M^{0,0}_{p,t,r} \tag{5.5}$$

$$M^{0,0,z_0}_{p,q,r} = \sum_{n=0}^{p} {}_p\alpha_n \sum_{k=0}^{n} \left(\begin{array}{c} n \\ k \end{array} \right) \sum_{t=0}^{k} {}_k b_t (-z_0)^{n-k} M^{0,0}_{p,q,t} \tag{5.6}$$

5.1.2 Περιστροφή

Μία περιστροφή κατά θ περιγράφεται από τον πίνακα περιστροφής:

$$\left[\begin{array}{c} u \\ w \end{array} \right] = \left[\begin{array}{cc} cos\theta & sin\theta \\ -sin\theta & cos\theta \end{array} \right] \left[\begin{array}{c} x \\ y \end{array} \right] \tag{5.7}$$

Ακολουθώντας παρόμοια βήματα με αυτά της προηγούμενης παραγράφου, η ροπή τάξης p,q μίας εικόνας μετά από περιστροφή κατά γωνία θ δίνεται από τη σχέση:

$$\begin{aligned} M^{\theta}_{p,q} &= \sum_{n=0}^{p} {}_{p}\alpha_{n} \sum_{t=0}^{n} \binom{n}{t} cos^{n}\theta sin^{n-t}\theta \sum_{m=0}^{q} {}_{q}\alpha_{m} \\ & \sum_{k=0}^{m} \binom{m}{k} cos^{k}\theta sin^{m-k}\theta(-1)^{m-k} \\ & \sum_{\zeta=0}^{n+m-k} {}_{n+m-k}b_{\zeta} \sum_{\xi=0}^{n+k-t} {}_{n+k-t}b_{xi} M^{0}_{\zeta,\xi} \end{aligned} \tag{5.8}$$

Όπως και στην περίπτωση της μετατόπισης, για τη βελτίωση της υπολογιστικής ταχύτητας, η περιστροφή κατά θ μπορεί να περιγραφεί από δύο απλούστερους μετασχηματισμούς:

$$\begin{bmatrix} u \\ w \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -tan\theta & cos^{-1}\theta \end{bmatrix} \begin{bmatrix} cos\theta & sin\theta \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \tag{5.9}$$

Τότε

$$\begin{aligned} \acute{M}_{p,q} &= \sum_{n=0}^{p} {}_{p}\alpha_{n} \sum_{t=0}^{n} \binom{n}{t} cos^{t}\theta sin^{n-t}\theta \sum_{k=0}^{q} {}_{q}\alpha_{k} \\ & \sum_{\zeta=0}^{t} {}_{t}b_{\zeta} \sum_{\xi=0}^{n+k-t} {}_{n+k-t}b_{xi} M^{0}_{\zeta,\xi} \end{aligned} \tag{5.10}$$

και

$$\begin{aligned} M^{\theta}_{p,q} &= \sum_{n=0}^{q} {}_{q}\alpha_{n} \sum_{t=0}^{n} \binom{n}{t} cos^{-n}\theta(-sin\theta)^{t} \sum_{k=0}^{p} {}_{q}\alpha_{k} \\ & \sum_{\zeta=0}^{t+k} {}_{t}b_{\zeta} \sum_{\xi=0}^{n-t} {}_{n+k-t}b_{xi} \acute{M}_{\zeta,\xi} \end{aligned} \tag{5.11}$$

Είναι προφανές ότι για γωνίες κοντά στα $\pi/2$ και $3\pi/2$, η μέθοδος θα παρουσιάζει αριθμητική αστάθεια. Σε αυτές τις περιπτώσεις, όμως, ο υπολογισμός είναι πολύ απλός μέσω της σχέσης 5.8.

5.1.3 Συνέλιξη

Με δεδομένη τη γραμμικότητα των ροπών, σε σχέση με το σήμα εισόδου, η συνέλιξη στο χώρο των ροπών μπορεί να πραγματοποιηθεί άμεσα. Στην περίπτωση σημάτων μίας διάστασης, για σήμα εισόδου $f : Z \to R$ και $h : Z \to R$:

$$(f \otimes g)(x) = \sum_t g(t) f(x-t) \tag{5.12}$$

Συνήθως στην επεξεργασία και ανάλυση εικόνων, το σήμα h είναι ορισμένο μόνο για λίγες τιμές, δηλαδή το μήκος του είναι πολύ μικρότερο αυτού του f. Για παράδειγμα, τα συχνά χρησιμοποιούμενα γραμμικά φίλτρα έχουν διαστάσεις $< 5 \times 5$ [Gonzalez and Woods, 1987], ενώ ο θόρυβος που παρατηρείται λόγω μη σταθερότητας της συσκευής εικοληψίας μοντελοποιείται ως γραμμικό φίλτρο μεγέθους μέχρι 33×33 [Park and Schowengerdt, 1982]. Από την προηγούμενη σχέση, προκύπτει για τις ροπές της συνέλιξης:

$$M_p^{(f \otimes g)} = \sum_{x=0}^{N-1} \sum_t g(t) f(x-t) P_p(x) \tag{5.13}$$

ή

$$M_p^{(f \otimes g)} = \sum_t g(t) \sum_{x=0}^{N-1} f(x-t) P_p(x) \tag{5.14}$$

και τελικά

$$M_p^{(f \otimes g)} = \sum_t g(t) M_p^{f,-t} \tag{5.15}$$

Με $M_p^{f,-t}$ τη ροπή τάξης p του σήματος f, μετά από μετακίνηση κατά t. Στην Ενότητα 5.1.1 περιγράφηκε η επίδραση της μετακίνησης. Παρόλο που η παραπάνω σχέση παρουσιάζει μεγάλη υπολογιστική πολυπλοκότητα για πολύπλοκες μάσκες g, για απλές μάσκες είναι προτιμότερη μία τέτοια υλοποίηση, από πλευράς υπολογιστικού κόστους.

Παράδειγμα

Έστω σήμα διαστάσεων 256×256 στοιχείων και μάσκα μεγέθους 3×3. Οι ροπές μέχρι 4ης τάξης του αποτελέσματος της συνέλιξης των δύο σημάτων, μπορούν να ανακτηθούν με δύο τρόπους:

1. Με τον υπολογισμό της συνέλιξης και την εξαγωγή των ροπών. Ο υπολογισμός της συνέλιξης απαιτεί $256 \times 256 \times 3 \times 3 = 589,824$ πολλαπλασιασμούς και ισάριθμες προσθέσεις. Έστω ότι ο υπολογισμός των ροπών απαιτεί K πολλαπλασιασμούς και L προσθέσεις.

2. Με την εκτέλεση της συνέλιξης στο πεδίο των ροπών, αφού έχουν υπολογισθεί οι ροπές για το αρχικό σήμα. Για τον υπολογισμό όλων των μετατοπισμένων ροπών απαιτούνται 2.700 πολλαπλασιασμοί και 2.700 προσθέσεις, ενώ για τον υπολογισμό της συνέλιξης μέσω αυτών απαιτούνται 300 πολλαπλασιασμοί και 3000 προσθέσεις. Στο παραπάνω υπολογιστικό κόστος θα πρέπει να προστεθούν και οι K πολλαπλασιασμοί και L προσθέσεις.

Ενώ, λοιπόν, στην πρώτη περίπτωση το υπολογιστικό κόστος είναι $K + L + 589,000$ πολλαπλασιασμοί και ισάριθμες προσθέσεις, στη δεύτερη είναι $K + L + 3000$.

5.2 Μη γραμμικοί μετασχηματισμοί

5.2.1 Ροπές γινομένου σημάτων

Με δεδομένα δύο σήματα $f, g : R \to R$, και τις αντίστοιχες ροπές F_p και G_p, είναι δυνατό να υπολογιστούν οι ροπές του σήματος $f \cdot g$ απευθείας από τα F και G. Από τον ορισμό των ροπών:

$$M_{f*g}(q) = \sum_{x=0}^{N} f(x)g(x)P_p(x) \tag{5.16}$$

Αντικαθιστώντας τα $f(x)$ και $g(x)$ μέσω της ιδιότητας ανακατασκευής:

$$M_{f*g}(p) = \sum_{x=0}^{N} \sum_{t=0}^{M} F(t)w(x)P_t(x) \sum_{n=0}^{M} w(x)P_n(x)P_p(x) \tag{5.17}$$

ή

$$M_{f*g}(p) = \sum_{t=0}^{M} F(t) \sum_{n=0}^{M} G(t) \sum_{x=0}^{N} w(x)^2 P_p(x)P_t(x)P_n(x) \tag{5.18}$$

Το τελευταίο άθροισμα της παραπάνω έκφρασης είναι ανεξάρτητο των σημάτων:

$$S(p, t, n) = \sum_{x=0}^{N} w(x)^2 P_p(x)P_t(x)P_n(x) \tag{5.19}$$

και

$$M_{f*g}(p) = \sum_{t=0}^{M} F(t) \sum_{n=0}^{M} G(t)S(p, t, n) \tag{5.20}$$

Σήματα δύο ή τριών διαστάσεων μπορούν να εκφραστούν με παρόμοιο τρόπο:

$$M_{f*g}(p,q) = \sum_{k=0}^{M}\sum_{t=0}^{M} F(t,k) \sum_{m=0}^{M}\sum_{n=0}^{M} G(n,m)S(p,t,n)S(q,k,m) \tag{5.21}$$

$$M_{f*g}(p,q,r) = \sum_{s=0}^{M}\sum_{k=0}^{M}\sum_{t=0}^{M} F(t,k,s) \sum_{u=0}^{M}\sum_{m=0}^{M}\sum_{n=0}^{M} G(n,m,u)S(p,t,n)S(q,k,m)S(r,s,u) \tag{5.22}$$

Παρόλο που οι εκφράσεις 5.20, 5.21 και 5.22 υποδηλώνουν M^2, M^4 και M^6 επαναλήψεις αντίστοιχα, ο όρος S μηδενίζεται σε πολλές περιπτώσεις:

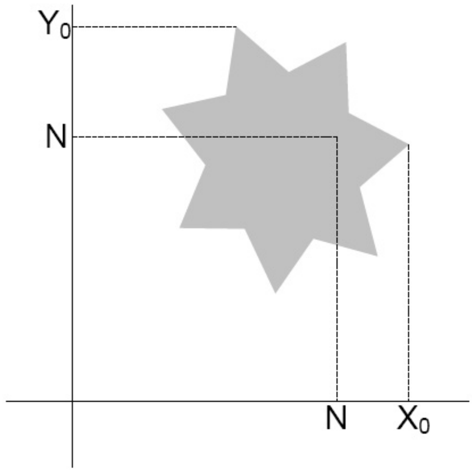
- Για $p > t+n$, το S γίνεται το εσωτερικό γινόμενο ενός ορθογωνίου πολυωνύμου P_p βαθμού p, και ενός πολυωνύμου βαθμού μικρότερου βαθμού. Αυτά τα δύο πολυώνυμα είναι ορθογώνια μεταξύ τους Γαυτσσηι [2004] και επομένως $S = 0$. Το ίδιο ισχύει για $t > p+n$ και $n > p+t$.[1]
- Για διακριτά πολυώνυμα *Chebyshev*, $w(x) = 1$ και μόνο οι μισές τιμές μπορεί να είναι μη μηδενικές.

5.3 Ροπές μέρους σήματος

Σε κάποιες περιπτώσεις είναι χρήσιμη η εξαγωγή των ροπών ενός μέρους σήματος, το οποίο περιορίζεται σε ορθογώνια περιοχή, όπως φαίνεται στο σχήμα 5.2. Οι ροπές του αντικειμένου ως προς την αρχή είναι:

$$M_{p,q} = \sum_{x=0}^{X_0}\sum_{y=0}^{Y_0} f(x,y)P_p(x)P_q(y) \tag{5.23}$$

[1] Η απόδειξη μπορεί να βρεθεί στη βιβλιογραφία, αλλά ακολουθεί μία σύντομη περιγραφή. Μία οικογένεια ορθογωνίων πολυωνύμων αποτελεί ορθοκανονική βάση ως προς κάποια μετρική w. Επομένως κάθε πολυώνυμο μπορεί να εκφραστεί ως γραμμικός συνδυασμός ορθογωνίων πολυωνύμων ίδιου και μικρότερου βαθμού. Έτσι, προκύπτει ότι το πολυώνυμο μικρότερου βαθμού εκφράζεται ως γραμμικός συνδυασμός πολυωνύμων μικρότερου βαθμού από το δεδομένο ορθογώνιο πολυώνυμο. Λόγω της γραμμικότητας του εσωτερικού γινομένου, προκύπτει ότι τα πολυώνυμα είναι ορθογώνια.



Σχήμα 5.2: Ροπές μέρους σήματος

Χωρίζοντας τα αθροίσματα προκύπτει:

$$\begin{aligned} M'_{p,q} &= M_{p,q} - \sum_{x=0}^{N}\sum_{y=N}^{Y_0} f(x,y)P_p(x)P_q(y) \\ &\quad - \sum_{x=N}^{X_0}\sum_{y=0}^{N} f(x,y)P_p(x)P_q(y) \\ &\quad - \sum_{x=N}^{X_0}\sum_{y=N}^{Y_0} f(x,y)P_p(x)P_q(y) \end{aligned} \quad (5.24)$$

5.3.1 Επίδραση απόκρυψης

Με δεδομένη την αρχική εικόνα $f : R^2 \to R$, οι ροπές της δίνονται από τη σχέση:

$$M_{p,q} = \sum_{x=0}^{N-1}\sum_{y=0}^{N-1} P_p(x)P_q(y)f(x,y) \quad (5.25)$$

Έστω $g : R^2 \to R$ η εικόνα με την απόκρυψη, και **C** το αντικείμενο που την προκαλεί. Στη γενική περίπτωση, το αντικείμενο θα είναι της μορφής $c : R^2 :\rightarrow R$. Τότε

$$g(x,y) = \begin{cases} f(x,y) & (x,y) \notin \mathbf{C} \\ c(x,y) & (x,y) \in \mathbf{C} \end{cases} \tag{5.26}$$

Οι αντίστοιχες ροπές είναι:

$$M^{occ}_{p,q} = \sum_{(x,y)\in(R^2 \cap C')} P_p(x)P_q(y)f(x,y) + \sum_{(x,y)\in C)} P_p(x)P_q(y)c(x,y) \tag{5.27}$$

ή

$$M^{occ}_{p,q} = M_{p,q} - \sum_{(x,y)\in C)} P_p(x)P_q(y)f(x,y) + \sum_{(x,y)\in C)} P_p(x)P_q(y)c(x,y) \tag{5.28}$$

5.3.2 Απλή απόκρυψη

Έστω ότι το σήμα f βρίσκεται σε τετράγωνο παράθυρο μεγέθους $N\times N$. Έστω επίσης ότι η απόκρυψη εμφανίζεται για κάθε $x > b$, $b \in C_f$. Τότε

$$M^{occ}_{p,q} = M_{p,q} - \sum_{x=0}^{b}\sum_{y=0}^{N} P_p(x)P_q(y)f(x,y) \tag{5.29}$$

Ανακατασκευάζοντας την εικόνα για την περιοχή απόκρυψης, και υποθέτοντας $w(x,y) = 1$, δηλαδή πολυώνυμα *Chebyshev*:

$$M^{occ}_{p,q} = M_{p,q} - \sum_{x=0}^{b}\sum_{y=0}^{N} P_p(x)P_q(y) \sum_{t=0}^{k}\sum_{s=0}^{k} M_{t,s}P_t(x)P_s(y) \tag{5.30}$$

ή

$$M^{occ}_{p,q} = M_{p,q} - \sum_{t=0}^{k}\sum_{s=0}^{k} M_{t,s} \sum_{x=0}^{b}\sum_{y=0}^{N} P_p(x)P_q(y)P_t(x)P_s(y) \tag{5.31}$$

και

$$M^{occ}_{p,q} = M_{p,q} - \sum_{t=0}^{k}\sum_{s=0}^{k} M_{t,s} \sum_{x=0}^{b} P_p(x)P_t(x) \sum_{y=0}^{N} P_s(y)P_q(y) \tag{5.32}$$

Λόγω της ορθογώνιότητας των πολυωνύμων στο διάστημα $[0, N-1]$, προκύπτει:

$$M^{occ}_{p,q} = M_{p,q} - \sum_{t=0}^{k} M_{t,q} \sum_{x=0}^{b} P_p(x)P_t(x) \tag{5.33}$$

Το τελευταίο άθροισμα είναι πολυώνυμο του παράγοντα b, βαθμού $p+t+1$. Έστω, λοιπόν

$$B_{p,t}(b) = \sum_{x=0}^{b} P_p(x)P_t(x) \tag{5.34}$$

τότε

$$M_{p,q}^{occ} = M_{p,q} - \sum_{t=0}^{k} M_{t,q}B_{p,t}(b) \tag{5.35}$$

Έλεγχος υπόθεσης

Με δεδομένες τις ροπές της αρχική εικόνας f, όπως επίσης και της g, η σχέση (5.35) θα πρέπει να δίνει:

$$M_{p,q}^{g} = M_{p,q}^{f} - \sum_{t=0}^{k} M_{t,q}^{f}B_{p,t}(b) \tag{5.36}$$

για κάποιο b, και για κάθε (p,q). Παρόμοιες εκφράσεις μπορούν να εξαχθούν και για κατακόρυφη απόκρυψη. Η έκφραση 5.36 είναι ένα πολυώνυμο του b. Πιο συγκεκριμένα, ο βαθμός αυτού του πολυωνύμου είναι $p+k+1$. Επομένως, η υλοποίηση της μεθόδου αυτής είναι απλή και υπολογιστικά αποδοτική.

5.3.3 Πειραματικά αποτελέσματα

Στις δοκιμαστικές εικόνες που παρουσιάζονται στο σχήμα 5.3 προστέθηκε απόκρυψη σε διάφορους βαθμούς. Επίσης, η περιοχή της απόκρυψης αντικαταστάθηκε σε κάθε περίπτωση από τυχαίο θόρυβο.

Όπως φαίνεται στο σχήμα 5.4, τα ελάχιστα των συναρτήσεων είναι ευδιάκριτα, και δεν εμφανίζονται τοπικά ελάχιστα.

5.3.4 Γενική περίπτωση απόκρυψης

Η προηγούμενη διερεύνηση ισχύει για απόκρυψη από αντικείμενο με περίγραμμα $x=c$. Στην Ενότητα αυτή εξετάζεται η γενική περίπτωση, στην οποία $x=h(y)$, με $h: R \to R$.

Οι ροπές του μη ορατού μέρους της εικόνας μπορούν να περιγραφούν ως:

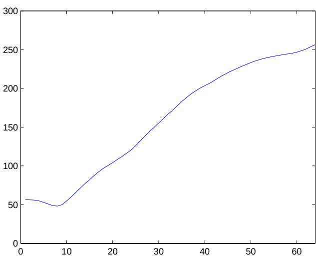
$$M_{p,q}^{o.c.} = \sum_{y=0}^{N} \sum_{x=0}^{h(y)} P_p(x)P_q(y)f(x,y) \tag{5.37}$$

(α′) Απόκρυψη 8-σημείων

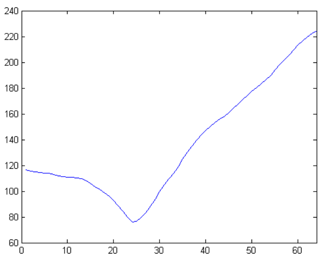
(β′) Απόκρυψη 24-σημείων

(γ′) Απόκρυψη 40-σημείων

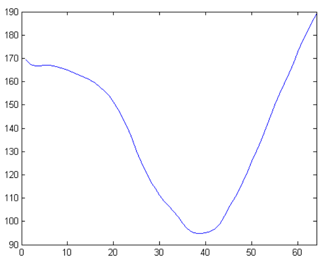
Σχήμα 5.3: Μεταβλητή απόκρυψη δυαδικών εικόνων



(α′) Απόκρυψη 8-σημείων



(β′) Απόκρυψη 24-σημείων



(γ′) Απόκρυψη 40-σημείων

Σχήμα 5.4: Γραφήματα υποθέσεων απόκρυψης και

$$M^{o.c.}_{p,q} = \sum_{y=0}^{N} \sum_{x=0}^{h(y)} P_p(x) P_q(y) \sum_{t=0}^{k} \sum_{s=0}^{k} P_t(x) P_s(y) T_{t,s} \tag{5.38}$$

με παρόμοιο τρόπο

$$M^{o.c.}_{p,q} = \sum_{s=0}^{k} \sum_{t=0}^{k} T_{t,s} \sum_{y=0}^{N} P_q(y) P_s(y) \sum_{x=0}^{h(y)} P_p(x) P_t(x) \tag{5.39}$$

Το τελευταίο άθροισμα αποτελεί και πάλι πολυώνυμο του $h(y)$[2]. Πιο συγκεκριμένα, το πολυώνυμο που προκύπτει είναι πολυώνυμο ως προς το άνω άκρο της άθροισης, βαθμού ίσου με τους αθροισταίους όρους, αυξημένου κατά ένα. Επομένως,

$$M^{o.c.}_{p,q} = \sum_{s=0}^{k} \sum_{t=0}^{k} T_{t,s} \sum_{y=0}^{N} P_q(y) P_s(y) B\left(h(y)\right) \tag{5.40}$$

Για $h(y) = b$, η έκφραση απλουστεύεται στην 5.35.

Πολυώνυμο $h(y)$

Έστω

$$h(y) = \sum_{i=0}^{n} \alpha_i^n y^i \tag{5.41}$$

τότε η σχέση 5.40 γίνεται:

$$M^{o.c.}_{p,q} = \sum_{s=0}^{k} \sum_{t=0}^{k} T_{t,s} \sum_{y=0}^{N} P_q(y) P_s(y) B\left(\sum_{i=0}^{n} \alpha_i^n y^i\right) \tag{5.42}$$

Επειδή ο όρος $B(h(y))$ είναι πολυώνυμο $h(y)$,

$$M^{o.c.}_{p,q} = \sum_{s=0}^{k} \sum_{t=0}^{k} T_{t,s} \sum_{y=0}^{N} P_q(y) P_s(y) \sum_{i=0}^{n} b_i^{p+t+n+1} y^i \tag{5.43}$$

Θέτοντας $w = p + t + n + 1$ προκύπτει:

$$M^{o.c.}_{p,q} = \sum_{s=0}^{k} \sum_{t=0}^{k} T_{t,s} \sum_{i=0}^{n} b_i^{w} \sum_{y=0}^{N} P_q(y) P_s(y) y^i \tag{5.44}$$

[2] Αθροίσματα και γινόμενα πολυωνύμων είναι προφανώς πολυώνυμα.

Θέτοντας

$$S = \sum_{i=0}^{n} b_i^w \sum_{y=0}^{N} P_q(y) P_s(y) y^i \tag{5.45}$$

ισχύει

$$M_{p,q}^{o.c.} = \sum_{s=0}^{k} \sum_{t=0}^{k} T_{t,s} S \tag{5.46}$$

Ο όρος S μηδενίζεται στις περιπτώσεις $q > s + w$ και $s > q + w$, αφού γίνεται το εσωτερικό γινόμενο ορθογωνίου πολυωνύμου με πολυώνυμο μικρότερου βαθμού.

Γραμμική περίπτωση Αν $h(y) = \kappa y + \lambda$, η σχέση 5.44 γίνεται:

$$M_{p,q}^{o.c.} \;=\; \sum_{s=0}^{k} \sum_{t=0}^{k} T_{t,s} \sum_{y=0}^{N} P_q(y) P_s(y) B(\kappa y + \lambda) \tag{5.47}$$

Η παραπάνω σχέση και πάλι θα πρέπει να ισχύει για κάποιο ζεύγος κ, λ, και για κάθε p, q.

Η διερεύνηση σχετικά με την επίδραση της απόκρυψης στις ροπές της εικόνας έδωσε αναλυτικές σχέσεις για τον συσχετισμό τους. Εντούτοις, για την εξέταση του κατά πόσο ένα σήμα μπορεί να έχει προκύψει από απόκρυψη άλλου, οι παραχθείσες σχέσεις παρουσιάζουν μεγάλη υπολογιστική πολυπλοκότητα, γεγονός που μειώνει τη χρησιμότητά τους σε πρακτικές εφαρμογές.

Για τη μείωση της πολυπλοκότητας, μπορεί να εξεταστεί η χρήση τεχνικών ελαχιστοποίησης. Για παράδειγμα, η σχέση 5.47, μπορεί να θεωρηθεί ως πρόβλημα ελαχιστοποίησης 2 διαστάσεων. Η ελαχιστοποίηση θα πρέπει να αντιμετωπισθεί στοχαστικά, με επιπλέον περιορισμό ότι ένα καθολικό ελάχιστο μπορεί και να μην υπάρχει, στην οποία περίπτωση το τελικό σήμα δεν έχει προκύψει από απόκρυψη του αρχικού. Αυτή είναι και η πιο σημαντική δυσκολία στην προσέγγιση αυτή, αφού δεν μπορεί να οριστεί εύκολα ένα καθολικό κατώφλι, το οποίο θα καθορίζει το κατά πόσο η επεξεργαζόμενη εικόνα έχει προκύψει από την αρχική. Επίσης, για να είναι εύρωστη η τεχνική, θα πρέπει να ληφθεί υπόψιν ότι το είδος του περιγράμματος δεν είναι γνωστό εκ των προτέρων. Η *apriori* υπόθεση, όμως, ενός πολυωνυμικού περιγράμματος μεγάλου βαθμού, το οποίο είναι και η πιο γενική περίπτωση, μπορεί να οδηγήσει σε αριθμητικά προβλήματα, ή και προβλήματα σταθερότητας, τα οποία συσχετίζονται με την πολυωνυμική παρεμβολή.

5.4 Εφαρμογή στην ανάκτηση ροπών από τις προβολές σήματος

5.4.1 Σχέση ροπών εικόνας με προβολές της

Παρόλο που η σχέση των ροπών μίας εικόνας με το μετασχηματισμό *Radon* έχει μελετηθεί εκτενώς [Shen et al., 1998] [Rouze et al., 2006][3], δεν έχει πραγματοποιηθεί ως τώρα καμία διερεύνηση σχετικά με σχέση ροπών και προβολών. Με δεδομένη μία μονοδιάστατη προβολή $h(z)$ μίας εικόνας $f(x,y)$, όπως φαίνεται στο σχήμα 5.5 αναπτύσσεται παρακάτω μία τεχνική για τον υπολογισμό των ροπών της εικόνας. Υποθέτοντας το πρότυπο αυτό (πρότυπο *pinhole*), με εστιακό βάθος FL:

$$\frac{y}{x} = \frac{FL}{z} \tag{5.48}$$

$$z = FL \cot(\theta) \tag{5.49}$$

Οι συνεχείς γεωμετρικές ροπές θα δίνονται από τη σχέση:

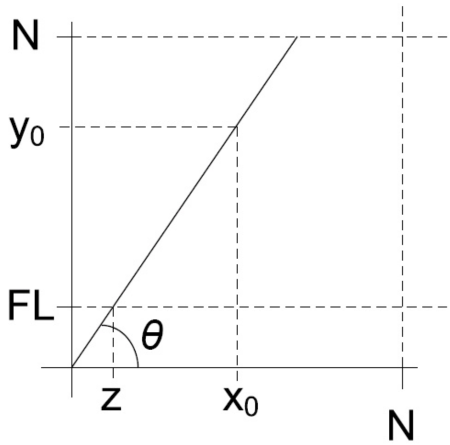
$$M_{p,q} = \int_0^N \int_0^N f(x,y) x^p y^q dxdy \tag{5.50}$$

Μεταφέροντας τα παραπάνω σε πολικό σύστημα συντεταγμένων προκύπτει:

$$m_{p,q} = \int_0^{\pi/4} \int_0^{N/\cos\theta} f(r\cos\theta, r\sin\theta) r^{p+q+1} \cos^p\theta \sin^q\theta drd\theta + \int_{\pi/4}^{\pi/2} \int_0^{N/\sin\theta} f(r\cos\theta, r\sin\theta) r^{p+q+1} \cos^p\theta \sin^q\theta drd\theta \tag{5.51}$$

Η πληροφορία της εικόνας, η οποία εμφανίζεται στην προβολή της

[3] Η ακριβής μαθηματική περιγραφή του μετασχηματισμού Radon δίνεται ως '...ο ολοκληρωτικός μετασχηματισμός του σήματος ο οποίος αποτελείται από το ολοκλήρωμα του σήματος επί όλων των ευθειών'. Πρακτικά, δίνει τις ορθογώνιες προβολές του σήματος σε κάθε ευθεία $y = \alpha x + b$



Σχήμα 5.5: Πρότυπο *Pinhole*

είναι ανεξάρτητη της ακτίνας r (Εξ. 5.49):

$$m_{p,q} = \frac{N^{p+q+2}}{FL(p+q+2)} \left(\int_{0}^{FL} f(z) z^p dz + \int_{FL}^{+\infty} f(z) \frac{1}{z^{q+2}} dz \right) \tag{5.52}$$

Επειδή, πρακτικά, η προβολή είναι πεπερασμένη σε μήκος, ο όρος του απείρου μπορεί να αντικατασταθεί μέσω των παραμέτρων της προβολής. Η παραπάνω έκφραση αντικαθιστά το απαιτούμενο διπλό ολοκλήρωμα υπολογισμού ροπών από δύο απλά. Λόγω της εγγενούς διακριτότητας του υπολογισμού του ολοκληρώματος, η υπολογιστική πολυπλοκότητα μειώνεται σε $O(n)$ από $O(n^2)$.

Κάθε ροπή πολυωνύμου μπορεί να εκφραστεί ως γραμμικός συνδυασμός γεωμετρικών ροπών. Πιο συγκεκριμένα, μία ροπή τάξης (p,q), με πολυώνυμα βάσης:

$$P_p(x) = \sum_{t=0}^{p} {}_p\alpha_t x^t \tag{5.53}$$

δίνεται ως

$$M_{p,q} = \sum_{k=0}^{q} {}_q\alpha_k \sum_{t=0}^{p} {}_p\alpha_t x^t \tag{5.54}$$

Στην περίπτωση υπολογισμού πολυωνυμικών ροπών υψηλής τάξης, ο υπολογισμός των γεωμετρικών δεν μπορεί να πραγματοποιηθεί ξεχωριστά. Κάτι τέτοιο θα οδηγούσε σε αριθμητικές αστάθειες, λόγω του μεγάλου εύρους τιμών των γεωμετρικών ροπών. Επομένως, η άθροιση της σχέσης 5.54 πρέπει να πραγματοποιηθεί μέσα στο ολοκλήρωμα της 5.52. Η σχέση 5.52 αναφέρεται σε προβολή στην αρχή του συστήματος συντεταγμένων. Για τον υπολογισμό των ροπών προβολών σε διαφορετική θέση, μπορούν να χρησιμοποιηθούν οι σχέσεις της παραγράφου 5.1.

5.4.2 Εξαγωγή ροπών εικόνας από πολλαπλές προβολές

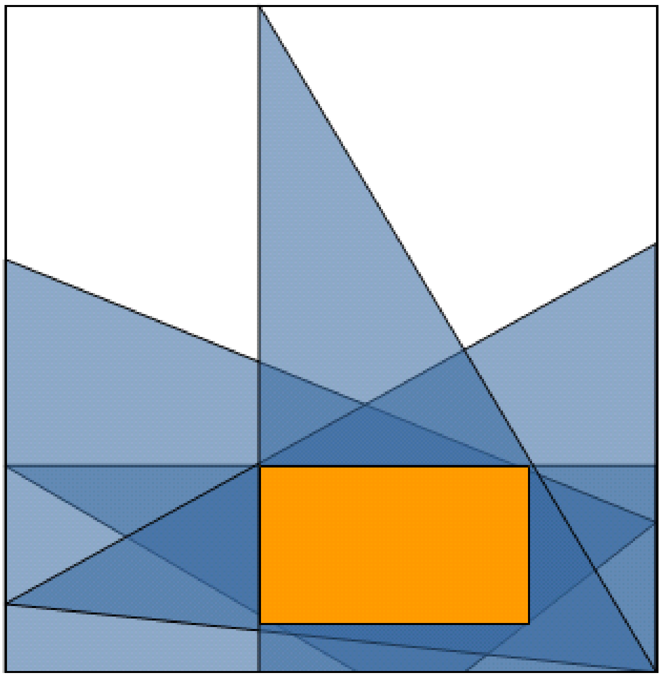
Με την επαναληπτική χρήση της σχέσης 5.52 είναι δυνατός ο υπολογισμός των ροπών της εικόνας από πολλαπλές προβολές:

- Υπολογισμός των ροπών των προβολών μέσω της 5.52.
- Υπολογισμός των ροπών στο μέρος της εικόνας που βρίσκεται εντώς του παραθύρου μέσω της 5.24.
- Μεταφορά των ροπών στην αρχή, μέσω των (5.1-5.11)
- Υπολογισμός των ροπών των γινομένων μέσω της (5.21)

Η αρχή λειτουργίας της τεχνικής αυτής παρουσιάζεται και στο Σχήμα 5.6

5.5 Συμπεράσματα

Στο κεφάλαιο αυτό παρουσιάστηκαν κάποιες χρήσιμες ιδιότητες των ροπών διακριτής πολυωνυμικής ορθογώνιας βάσης. Αναπτύχθηκαν αποδοτικές σχέσεις για τον υπολογισμό ροπών εικόνων που έχουν υποστεί γραμμικούς ή μη μετασχηματισμούς. Πιο συγκεκριμένα, η μελέτη της επίδρασης προβολικού μετασχηματισμού, καθώς και απόκρυψης μέρους της εικόνας αναφέρεται για πρώτη φορά στη βιβλιογραφία. Η χρήση αυτού του συνόλου των εργαλείων μπορεί να βελτιώσει σημαντικά αλγορίθμους ανάλυσης εικόνας και αναγνώρισης προτύπων.



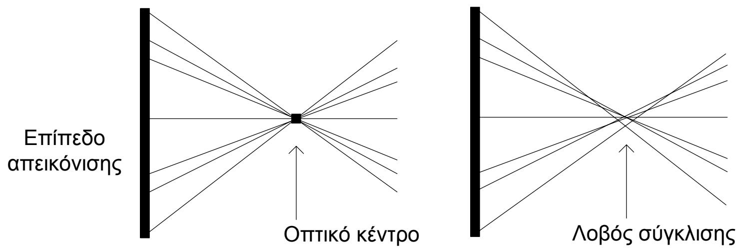
Σχήμα 5.6: Αρχή λειτουργίας εξαγωγής ροπών από προβολές

Κεφάλαιο 6

Ταχύς υπολογισμός εικόνων βάθους

Ένα από τα πιο ενδιαφέροντα προβλήματα στην υπολογιστική όραση είναι η εξαγωγή πληροφορίας βάθους από ένα στερεοσκοπικό ζεύγος εικόνων. Σε αυτό το Κεφάλαιο χρησιμοποιούνται οι ροπές Zernike, όπως περιγράφηκαν στα προηγούμενα Κεφάλαια για την αντιστοίχιση σημείων μεταξύ δύο μη βαθμονομημένων εικόνων. Στη συνέχεια, παρουσιάζεται μία αρχιτεκτονική υλικού για την εξαγωγή πληροφορίας βάθους από βαθμονομημένες εικόνες η οποία παρουσιάζει την μεγαλύτερη ταχύτητα που αναφέρεται στη βιβλιογραφία.

Ένα ενεργό ερευνητικό θέμα στην στερεοσκοπική όραση είναι η εξαγωγή του βάθους σε μια σκηνή, μια διαδικασία που περιλαμβάνει την ακριβή αντιστοίχιση των σημείων δύο εικόνων ενός στερεοσκοπικού ζεύγους. Το βάθος μιας σκηνής επιτρέπει τον υπολογισμό των τρισδιάστατων χαρακτηριστικών ενός αντικειμένου. Εντούτοις, αυτός ο υπολογισμός είναι μια υπολογιστικά απαιτητική διαδικασία η οποία απαιτεί εξειδικευμένο υλικό για την πραγματοποίησή της σε πραγματικό χρόνο.
Η αντιστοίχιση όλων των σημείων των εικόνων εισόδου διευκολύνεται στην περίπτωση που το στερεοσκοπικό σύστημα είναι βαθμονομημένο. Για να πραγματοποιηθεί η βαθμονόμηση, απαιτείται πρώτα σύστημα αραιής αντιστοίχισης σημείων, το οποίο και περιγράφεται παρακά-



Σχήμα 6.1: Το πρότυπο οπής (αριστερά). Στα δεξιά παρουσιάζεται ένα πιο ακριβές πρότυπο για τις συσκευές εικονοληψίας.

τω. Βάσει των σημείων αυτών, εξάγονται οι εσωτερικές και εξωτερικές παράμετροι των συσκευών εικονοληψίας. Τέτοιες παράμετροι είναι η εστιακή απόσταση, το κέντρο εικονοληψίας, το μέγεθος της εικόνας (εσωτερικές), τα σχετικά διανύσματα θέσης των δύο συσκευών καθώς και η σχετική περιστροφή τους (εξωτερικές). Στη συνέχεια, για την ταχεία εξαγωγή πυκνής πληροφορίας βάθους, οι εικόνες υπόκεινται σε παραμόρφωση τέτοια, ώστε οι γραμμές σάρωσης να συμπίπτουν με τις επιπολικές γραμμές.

Το πιο συχνά χρησιμοποιούμενο πρότυπο συσκευής εικονοληψίας είναι το πρότυπο οπής (pinhole model), το οποίο και παρουσιάζεται στο Σχήμα 6.1. Βάσει αυτού του προτύπου ορίζονται οι παράμετροι του συστήματος, με σκοπό την πιο εύκολη επεξεργασία των εικόνων. Για ένα βαθμονομημένο στερεοσκοπικό σύστημα, ισχύουν οι σχέσεις τις Ενότητας 6.2.

6.1 Εφαρμογή ροπών στην αραιή αντιστοίχιση σημείων

Στο πρώτο μέρος του κεφαλαίου αυτού παρουσιάζεται μία νέα αρχιτεκτονική υλικού πραγματικού χρόνου για την αντιστοίχιση χαρακτηριστικών μεταξύ δύο μη βαθμονομημένων εικόνων μίας σκηνής. Η εύρωστη αντιστοίχιση επιτυγχάνεται μέσω της χρήσης γεωμετρικών χαρακτηριστικών και χαρακτηριστικών ροπών. Για την επίτευξη αμεταβλητότητας στην περιστροφή και την μεγαλύτερη ανοχή στο θόρυβο, προτιμήθηκαν οι ροπές *Zernike*. Το προτεινόμενο σύστημα είναι ικανό για την επεξεργασία πάνω από 25 ζευγών εικόνων ανά δευτερόλεπτο, για εικόνες *VGA* [Gonidis, 2007], [Gonidis et al., 2007].

6.1.1 Εισαγωγή

Το πρόβλημα της αντιστοίχισης σημείων μεταξύ δύο εικόνων μίας σκηνής είναι πρωταρχικό στην υπολογιστική όραση [Berthold, 1993]. Στην πρόσφατη βιβλιογραφία προτάθηκαν αρκετές αποδοτικές στρατηγικές για την επίλυση του προβλήματος αυτού [Zhang et al., 1995, Pilu, 1997, Zhao, 2004]. Αυτές βασίζονται σε δύο διαφορετικές προσεγγίσεις. Κατά την πρώτη, ανιχνεύονται χαρακτηριστικά στην εικόνα αναφοράς, και με τη βοήθεια τεχνικών πολλαπλής ανάλυσης (*multiresolution techniques*), αναζητώνται τα αντίστοιχα σημεία στη δεύτερη. Κατά τη δεύτερη, η οποία και υλοποιείται στην παρούσα εργασία, τα χαρακτηριστικά ανιχνεύονται ανεξάρτητα στις δύο εικόνες, και εν συνεχεία αντιστοιχίζονται μέσω κατάλληλου αλγορίθμου. Μία εύρωστη τεχνική για αυτή την αντιστοίχιση προτάθηκε στην [Pilu, 1997], κατά την οποία περιγράφηκε μία άμεση μέθοδος, με τη χρήση τεχνικής *SVD*.

6.1.2 Αλγόριθμος αντιστοίχισης

Ακολουθεί μία σύντομη μαθηματική περιγραφή του αλγορίθμου αντιστοίχισης. Το βασικό του χαρακτηριστικό είναι η εύκολη υλοποίησή του η οποία βασίζεται σε μία καλά ορισμένη λύση ιδιοδιανυσμάτων, η οποία δεν περιλαμβάνει καμία εγγενή επανάληψη.

Ανίχνευση γωνιών

Οι γωνίες χρησιμοποιούνται ευρέως ως σημειακά χαρακτηριστικά σε πολλές εφαρμογές υπολογιστικής όρασης. Μπορούν να ανιχνευθούν αυτόματα χωρίς καμία *apriori* πληροφορία, και είναι γενικά σταθερά χαρακτηριστικά. Ίσως ο πιο δημοφιλής αλγόριθμος είναι αυτός του Harris [Harris and Stephens, 1988], ο οποίος προτιμήθηκε και στην παρούσα υλοποίηση. Βασίζεται στον πίνακα αυτοσυσχέτισης, ο οποίος ορίζεται ως:

$$M_c = e^{\frac{-(x^2+y^2)}{2\sigma}} \otimes \begin{bmatrix} {I_x}^2 & I_x \times I_y \\ I_x \times I_y & {I_y}^2 \end{bmatrix} \tag{6.1}$$

με $\otimes$ τον τελεστή συνέλιξης. Τα I_x και I_y δηλώνουν την οριζόντια και κατακόρυφη παράγωγο, αντίστοιχα. Ο πίνακας αυτοσυσχέτισης εκτελεί μία διεργασία ομαλοποίησης στα γινόμενα των πρώτων παραγώγων, με τη συνέλιξη με ένα Γκαουσσιανό παράθυρο με παράμετρο μεγέθους σ. Δύο μεγάλες ιδιοτιμές του M υποδηλώνουν την παρουσία γωνίας. Προς αποφυγή την αποσύνθεση άμεσων ιδιοτιμών του M, η απόκριση γωνιών δίνεται από την παρακάτω έκφραση:

$$C_H = det(M_c) - \alpha * trace(M_c)^2 \tag{6.2}$$

με α μία μική ποσότητα (στην [Harris and Stephens, 1988] ορίζεται ως 0,04, προς αποφυγή αριθμητικής αστάθειας, λόγω μηδενικής ορίζουσας του M_c. Έτσι, ένα κατώφλι T_h μπορεί να χρησιμοποιηθεί για την επιλογή σημείων γωνιών. Αν $C_H > t_h$, τότε το σημείο επιλέγεται ως γωνία. Στον αρχικό ανιχνευτή Harris , τα I_x και I_y υπολογίζονται με τη συνέλιξη της εικόνας με τις μάσκες $[-1, 0, 1]$ και $[-1, 0, 1]^T$. Στην παρούσα εφαρμογή, η συνέλιξη γίνεται με τις μάσκες Prewitt [Prewitt, 1970], λόγω της αυξημένης σταθερότητάς τους, παρουσία θορύβου:

$$H = \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix} V = \begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix} \tag{6.3}$$

Πίνακας απόστασης

Το δεύτερο βήμα είναι η εξαγωγή ενός πίνακα αποστάσεων ο οποίος περιλαμβάνει τις θέσεις των αντιστοιχιών των σημείων γωνιών. Πιο συγκεκριμένα, υποθέτοντας ότι η πρώτη εικόνα I περιέχει m χαρακτηριστικά $I_i (i = 1...m)$ και η δεύτερη J, n χαρακτηριστικά, $J_j (j = 1...n)$, για την εξαγωγή της ένα προς ένα αντιστοίχισης σημείων μεταξύ των εικόνων, ορίζεται ένας πίνακας αποστάσεων [Zhao, 2004]:

$$G = (C_{ij} + 1)^3 e^{\frac{-r_{ij}}{2\sigma^2}} \quad i = 1...m, j = 1...n \tag{6.4}$$

$r_{ij} = \left\|I_i - J_j\right\|$ είναι η Ευκλείδια απόσταση μεταξύ των προβολών των δύο χαρακτηριστικών, στο επίπεδο της εικόνας. Ο όρος C_{ij} υπολογίζεται ως εξής:

$$C_{ij} = \sum_{u=-w}^{w} \sum_{v=-w}^{w} \left|L(x_i + u, y_i + v) - R(x_j + u, y_j + v)\right| \tag{6.5}$$

με L και R πίνακες οι οποίοι περιέχουν τις ροπές *Zernike* των δύο περιοχών, αντίστοιχα, για παράθυρα μεγέθους $w \times w$, με κέντρο το σημείο ενδιαφέροντος. Ο G είναι ένας θετικά ορισμένος πίνακας. Η παράμετρος σ ελέγχει τον βαθμό της αλληλεπίδρασης μεταξύ των δύο συνόλων χαρακτηριστικών [Pilu, 1997]. Μικρή τιμή του σ επιβάλλει τοπικές αλληλεπιδράσεις, ενώ μεγάλη επιτρέπει και σε πιο καθολικές. Πιο απλά, για μικρές τιμές του σ δεν επιτρέπονται αντιστοιχίες μεταξύ απομακρυσμένων σημείων.

Μέθοδος Ανάλυσης Μοναδιαίας Τιμής

Το τελευταίο βήμα της αντιστοίχισης είναι η εφαρμογή *SVD* στον *G*:

$$G = UDV^T \tag{6.6}$$

με $U \in M_{m,m}$ και $V \in M_{n,n}$ ορθογώνιους πίνακες και $D \in M_{m,n}$ διαγώνιος πίνακας ο οποίος περιέχει τις μοναδιαίες τιμες στα στοιχεία του D_{ii}, σε φθίνουσα σειρά ($\sigma_1 > \sigma_2 > \ldots > \sigma_m$). Για την ενίσχυση των ισχυρών αντιστοιχίσεων, ο D μετατρέπεται σε ένα νέο πίνακα F, αντικαθιστώντας κάθε στοιχείο τις διαγωνίου του με 1. Ο νέος πίνακας P προκύπτει ως εξής:

$$P = UEV^T \tag{6.7}$$

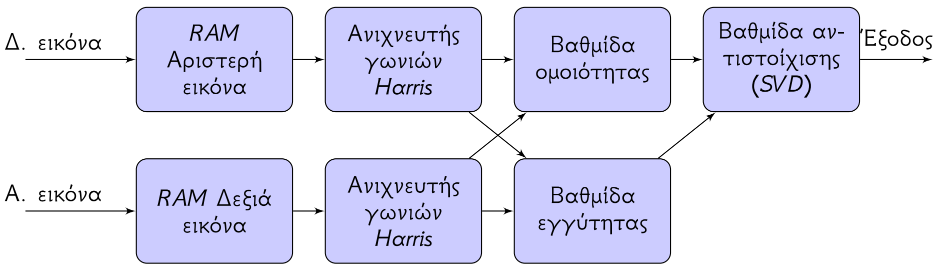
Ο P έχει την ενδιαφέρουσα ιδιότητα ότι αν το στοιχείο του P_{ij} είναι το μεγαλύτερο στοιχείο στην γραμμή i και στη στήλη j, $((max_i(P_{ij}, i = 1 \ldots n) = max_j(P_{ij}, j = 1 \ldots m)))$, τότε τα δύο χαρακτηριστικά I_i και J_j μπορούν να θεωρηθούν ότι αντιστοιχούν. Σε διαφορετική περίπτωση, αυτό σημαίνει ότι το I_i ανταγωνίσθηκε ανεπιτυχώς με άλλα χαρακτηριστικά, για την αντιστοίχιση με το J_j. Παρατηρείται ότι αυτή η απλή μέθοδος συμπεριλαμβάνει αμφότερα τα κριτήρια απόστασης και αποκλεισμού. Το πρώτο υλοποιείται λόγω της φύσης του πίνακα αποστάσεων και το δεύτερο λόγω της ορθογωνιότητας του P [Scott and Longuet-Higgins, 1991] Πέρα από τη χωρική απόσταση η οποία υπολογίζεται μέσω των Ευκλείδειων αποστάσεων, ο αλγόριθμος συμπεριλαμβάνει πληροφορία ομοιότητας των στοιχείων. Αυτή περιγράφεται μέσω των ροπών *Zernike*, ως τοπικών περιγραφέων των σημείων ενδιαφέροντος.

6.1.3 Υλοποίηση σε υλικό

Η προτεινόμενη αρχιτεκτονική εξάγει σημειακές αντιστοιχίες μεταξύ δύο εικόνων στερεοσκοπικού ζεύγους, μεγέθους 640×480 στοιχείων. Η είσοδος τροφοδοτείται σειριακά με δεδομένα φωτεινοτήτων 8 ψηφίων, ενώ η έξοδος δηλώνει τις θέσεις των σημείων αντιστοίχισης. Το γενικό διάγραμμα της αρχιτεκτονικής παρουσιάζεται στο σχήμα 6.2.

Ανίχνευση γωνιών

Το πρώτο μέρος του κυκλώματος είναι το υποσύστημα ανίχνευσης γωνιών, το οποίο βασίζεται στη μέθοδο *Harris* [Harris and Stephens, 1988]. Οι εικόνες τροφοδοτούνται σειριακά, ενώ μία δυαδική έξοδος δηλώνει την ύπαρξη ή μη γωνίας. Ο αλγόριθμος αποτελείται από τον



Σχήμα 6.2: Γενικό διάγραμμα αρχιτεκτονικής αντιστοίχισης σημείων

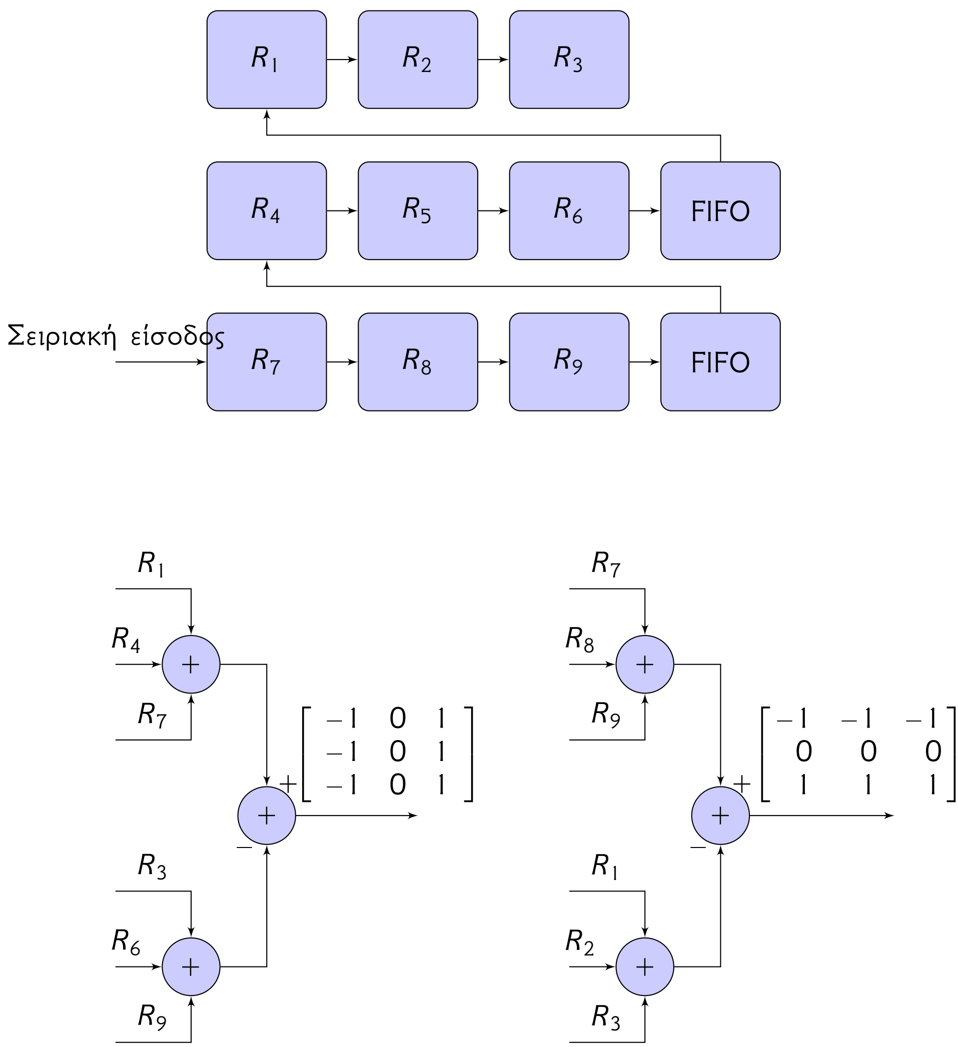
υπολογισμό των προσεγγιστικών κλίσεων της εικόνας, I_x και I_y, οι οποίες προκύπτουν από τη συνέλιξή της με τις μάσκες *Prewitt* [Zhao, 2004](Σχέση 6.3). Για την εκτέλεση της λειτουργίας αυτής, τα στοιχεία της γειτονιάς-8 διατηρούνται σε καταχωρητές, εκμεταλλευόμενοι μία αρχιτεκτονική τύπου *σερπαντίνας*. Τα παραγόμενα αποτελέσματα συνδυάζονται για την εξαγωγή των: $I_x^2 = I_x I_x$, $I_y^2 = I_y I_y$ και $I_{xy} = I_x I_y$, και εν συνεχεία τροφοδοτούνται σε ένα βαθυπερατό, Γκαουσσιανό φίλτρο. Στην παρούσα υλοποίηση, η τιμή της 'γωνιότητας ' c, υπολογίζεται βάσει της σχέσης (6.8):

$$c = \frac{I_{x2} I_{y2} - I_{xy2}}{I_{x2} + I_{y2}} \tag{6.8}$$

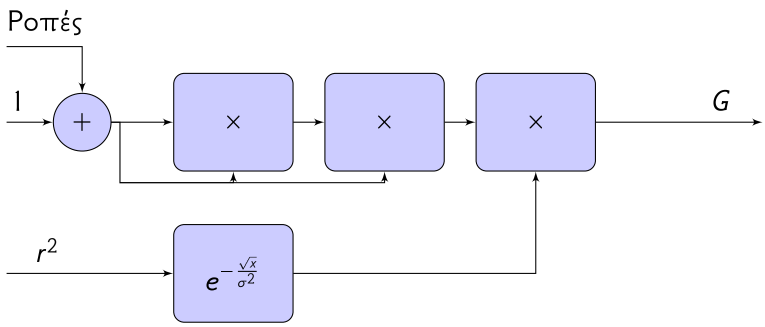
Για την υπό εξέταση γειτονιά, κάθε τιμή η οποία δεν αποτελεί τοπικό μέγιστο υποβαθμίζεται. Αυτό οδηγεί σε έναν αραιό χάρτη σημείων ενδιαφέροντος για κάθε εικόνα. Τέλος, από τη στιγμή που η τιμή c παρέχει ένα μέτρο της ποιότητας κάθε γωνίας, ορίζεται και ένα κατώφλι πάνω από το οποίο το σημείο θεωρείται γωνία. Τα παραγόμενα σημεία τροφοδοτούνται ως είσοδος στα υποσυστήματα Εγγύτητας και Ομοιότητας, τα οποία λειτουργούν παράλληλα.

Κριτήριο εγγύτητας

Το κριτήριο εγγύτητας το οποίο χρησιμοποιείται αποτελεί την Ευκλείδεια απόσταση μεταξύ των υπό εξέταση σημείων, τα οποία ανιχνεύθηκαν κατά το προηγούμενο βήμα. Αρχικά, δύο μετρητές υλοποιούνται προς μετατροπή της θέσης του κάθε στοιχείου, σε καρτεσιανές συντεταγμένες. Μέσω καταχωρητών οι οποίοι ενεργοποιούνται από το σήμα ύπαρξης γωνίας, η τρέχουσα θέση *x,y* μπορεί να αποθηκευτεί.



Σχήμα 6.3: Υλοποίηση φίλτρου *Prewitt*



Σχήμα 6.4: Υλοποίηση πίνακα εγγύτητας

Τα τετράγωνα των Ευκλειδείων αποστάσεων υπολογίζονται βάσει της σχέσης (6.9), η οποία υλοποιείται κατά άμεσο τρόπο μέσω δύο πολλαπλασιαστών, ενός αθροιστή και δύο αθροιστών.

$$r^2 = (x_1 - x_2)^2 + (y_1 - y_2)^2 \tag{6.9}$$

Κριτήριο ομοιότητας

Η εισαγωγή κριτηρίου ομοιότητας είναι απαραίτητη για την απόρριψη εσφαλμένων αντιστοιχίσεων. Ένα τέτοιο κριτήριο, παρέχει ένα μέτρο τοπικής ομοιότητας των περιοχών ενδιαφέροντος. Ως τοπικός περιγραφέας, επιλέχθηκαν οι ροπές *Zernike* [Kan and Srinath, 2002], για τους παρακάτω λόγους:

1. Παρέχουν αμεταβλητότητα στην περιστροφή.
2. Έχουν ορθογώνια βάση, και επομένως παρουσιάζουν ελάχιστο πλεονασμό πληροφορίας.
3. Είναι αποδοτικές στην υλοποίηση, με τη μέθοδο που περιγράφηκε στο κεφάλαιο 2.

6.1.4 Κύκλωμα αντιστοίχισης σημείων

Η τελική βαθμίδα της αρχιτεκτονικής εκτελεί την αντιστοίχιση σημείων ενδιαφέροντος, τα οποία έχουν ανιχνευθεί σε προηγούμενη βαθμίδα, βάσει των κριτηρίων εγγύτητας και ομοιότητας.

Αρχικά, ο πίνακας εγγύτητας G υπολογίζεται βάσει της (6.4). Για τον σκοπό αυτό, χρησιμοποιείται ένας πίνακας αναζήτησης μεγέθους

$m \times n \times (x-bits)$ για τον υπολογισμό του όρου $e^{-r_{ij}^2/2\sigma^2}$. Για $n \times m = 10000$, απαιτούνται $12kBytes$ μνήμης.

Το επόμενο βήμα του αλγορίθμου είναι η εκτέλεση της *SVD* του *G*. Αυτός ο υπολογισμός είναι χρονοβόρος για πίνακες μεγάλου μεγέθους. Η παρούσα υλοποίηση βασίζεται στην παραλλαγή των Brent και Luk [Bobda and Steenbock, 2001, Brent and Luk, 1985] της μεθόδου *Hestenes* [Hestenes, 1958], η οποία βασίζεται στη μονόπλευρη επανάληψη *Jacobi* για τη διαγωνιοποίηση πραγματικών, συμμετρικών πινάκων. Η τεχνική αυτή βασίζεται στον υπολογισμό διαγωνίου πίνακα *V* τέτοιου ώστε ο πίνακας $GV = W$ να έχει ορθογώνιες στήλες. Έχοντας τον πίνακα *W*, το ευκλείδειο μέτρο κάθε μη μηδενικής στήλης $W_{(:,i)}$ θα είναι μοναδιαία κανονικοποιημένο. Οι μοναδιαίες τιμές και διανύσματα υπολογίζονται ακολούθως:

$$\sigma_i = \left\|W_{(:,i)}\right\|, U_{(:,i)} = \frac{W_{(:,i)}}{\sigma_i} \; and \, W = UD \tag{6.10}$$

Η *SVD* του πίνακα *G* δίνεται:

$$GV = W \rightarrow GV = UD \rightarrow G = UDV^T \tag{6.11}$$

Οι περιστροφές που περιγράφονται από τον πίνακα *Q* εφαρμόζονται διαδοχικά στον πίνακα *A* για την εξαγωγή του *W*. Στο βήμα *k*, με πίνακα περιστροφής $Q^{(k)}$, ισχύει:

$$G^{(k+1)} = Q^{(k)} G^{(k)}, \; 0 \leq k \leq k_r \tag{6.12}$$

Αν θεωρήσουμε ένα βήμα ως μία σειρά $\frac{n(n-1)}{2}$ ορθογωνιοποιήσεων του πίνακα $G^{(k)}$, η σύγκλιση του πίνακα $G^{(k)}$ στον *W* επιτυγχάνεται σε ένα βήμα, με $G^{(0)} = G$ και $W = G^{(k_r)}$

Έστω $G^{(k)} = (\alpha_1^{(k)}, \ldots, \alpha_n^{(k)})$ και $Q^{(k)}$ η περιστροφή στο επίπεδο (i,j):

$$\begin{array}{ll} q_{ii}{}^{(k)} = cos\theta & q_{ij}{}^{(k)} = sin\theta \\ q_{ji}{}^{(k)} = -sin\theta & q_{jj}{}^{(k)} = cos\theta \end{array} \tag{6.13}$$

Αξίζει να σημειωθεί οτι ο πολλαπλασιασμός με τον $Q^{(k)}$ επηρεάζει μόνο δύο στήλες:

$$\left(\alpha_i^{(k+1)}, \alpha_j^{(k+1)}\right) = \left(\alpha_i^{(k)}, \alpha_j^{(k)}\right) \begin{pmatrix} cos\theta & sin\theta \\ -sin\theta & cos\theta \end{pmatrix} \tag{6.14}$$

Η γωνία περιστροφής θ επιλέγεται έτσι ώστε τα νέα ζεύγη στηλών να είναι ορθογώνια. Χρησιμοποιώντας τις σχέσεις του Rustishauser [Rutishauser, 1971] , θέτοντας:

$$\alpha = \alpha_i^{(k)T} \alpha_i^{(k)} \; , \; \beta = \alpha_j^{(k)T} \alpha_j^{(k)} \; , \; \gamma = \alpha_i^{(k)T} \alpha_j^{(k)} \tag{6.15}$$

Προκύπτει

$$\xi = \frac{\beta - \alpha}{\gamma} \; , \; t = \frac{sign(\xi)}{|\xi| + \sqrt{1 + \xi^2}} \; ,$$
$$cos\theta = \frac{1}{\sqrt{1 + t^2}} \; , \; sin\theta = tcos\theta \tag{6.16}$$

Εντούτοις, παραμένει το πρόβλημα της επιλογής των (i,j) , το οποίο συνήθως αντιμετωπίζεται μέσω κάποιας δεδομένης επανάληψης. Ένας στόχος είναι η διέλευση από όλα τα πιθανά ζεύγη στηλών ακριβώς μία φορά σε κάθε επανάληψη. Μία τέτοια επανάληψη αποτελείται από κυκλικές ανά γραμμή ανακατατάξεις:

$$(1,2),(1,3),...,(1,n),(2,3),...,(2,n),...,(n-1,n) \tag{6.17}$$

Απαιτούνται, λοιπόν δύο τοπικές μνήμες (L,R), αρκετά μεγάλες για την αποθήκευση μίας στήλης του G, δηλαδή μεγέθους ίσο με τον αριθμό των γωνιών που ανιχνεύονται στην εικόνα. Είναι προφανές ότι το στοιχείο 1 παραμένει στον L, ενώ τα $2,...,n$ ανακατατάσσονται κυκλικά, με μήκος $n-1$. Κατά δύο οποιαδήποτε διαδοχικά βήματα, ένα ζεύγος (i,j) μπορεί να εμφανιστεί στο (L,R) το πολύ μία φορά. Επίσης, τα ζεύγη (i,j) δεν (j,i) μπορούν να εμφανιστούν ταυτόχρονα.

Ο πίνακας V ενημερώνεται με τον ίδιο τρόπο με τον πίνακα G:

$$V^{(k+1)} = Q^{(k)}V^{(k)}, \; 0 \leq k \leq k_r \tag{6.18}$$

με $V^{(0)} = I_n$ και $V = V^{(k_r)}$.

Μετά τη μετατροπή του D σε έναν νέο πίνακα E, αντικαθιστώντας κάθε διαγώνιο στοιχείο με 1, υπολογίζεται ο νέος πίνακας P (6.7) μέσω πολλαπλασιαστών. Η εύρεση του μεγίστου κατά γραμμή και στήλη υλοποιείται μέσω συγκριτών, μετά από κατάλληλη σάρωση του πίνακα P. Η δυαδική έξοδος αναφέρεται στη διεύθυνση των γωνιακών σημείων στα οποία βρέθηκε αντιστοιχία.

6.1.5 Συμπεράσματα

Στο πρώτο μέρος από το παρόν κεφάλαιο παρουσιάστηκε μία αρχιτεκτονική υλικού για την εξαγωγή σημειακών αντιστοιχιών μεταξύ δύο μη βαθμονομημένων εικόνων. Με τη χρήση της αρχιτεκτονικής που περιγράφηκε στο προηγούμενο κεφάλαιο, έγινε δυνατή η ταχεία εξαγωγή τον ροπών *Zernike* γύρω από γωνιακά σημεία. Οι ροπές *Zernike* παρουσιάζουν αμεταβλητότητα στην περιστροφή, καθώς και σχετικά μεγάλη ανοχή στο θόρυβο. Με βάση αυτή την περιγραφή, καθώς και ενός κριτηρίου εγγύτητας, κατέστη δυνατή η εύρωστη αντιστοίχιση σημείων. Η αρχιτεκτονική υλικού που παρουσιάστηκε, αποτελεί την πιο γρήγορη στη βιβλιογραφία, με δυνατότητα επεξεργασίας ζευγών εικόνων σε πραγματικό χρόνο, ακόμα και για μεγάλα μεγέθη. Τυπική εφαρμογή της τεχνικής αυτής αποτελούν συστήματα αυτόνομης πλοήγησης, στα οποία γενικά απαιτείται απόκριση πραγματικού χρόνου.

6.2 Εισαγωγή

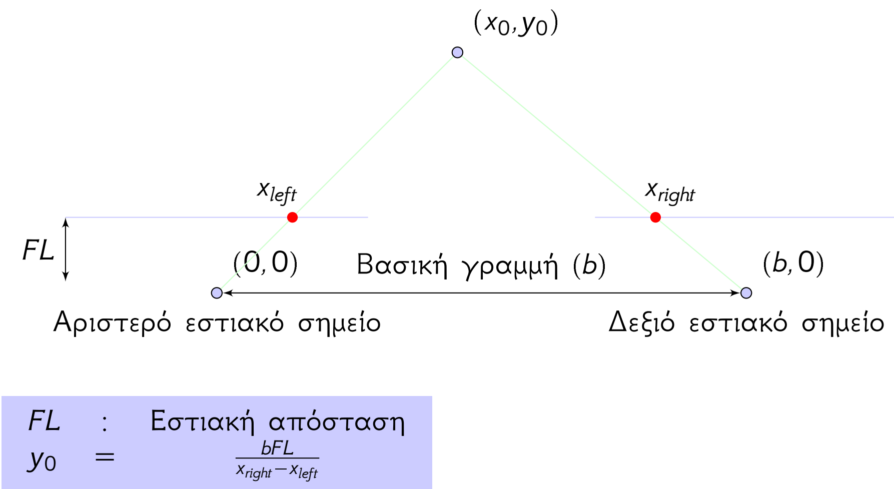
Στις περισσότερες εφαρμογές ρομποτικής όρασης, ή υπολογιστικής όρασης γενικότερα είναι αναγκαία η ανάκτηση 3Δ πληροφορίας σχετικά με τις επεξεργαζόμενες σκηνές. Αυτή η ανάκτηση μπορεί να επιτευχθεί με διάφορους τρόπους. Εφ'οσον χρησιμοποιείται μόνο οπτική πληροφορία, οι πιο συχνά χρησιμοποιούμενοι είναι οι παρακάτω:

- Χρήση πολλαπλών συσκευών απεικόνισης
- Χρήση κινούμενης συσκευής απεικόνισης
- Επεξεργασία της διεύθυνσης σκίασης των αντικειμένων
- Χρήση δομημένου φωτισμού

Ο πρώτος από αυτούς, και πιο συγκεκριμένα η χρήση ενός στερεοσκοπικού συστήματος όρασης, ίσως να είναι και ο πιο συνήθης. Η αρχή λειτουργίας ενός τέτοιου συστήματος παρουσιάζεται στο σχήμα 6.5. Η πιο κρίσιμη διαδικασία στην στερεοσκοπική όραση είναι η αντιστοίχιση σημείων των δύο εικόνων. Με δεδομένη την παραλλαγή και την παρουσία θορύβου, η αντιστοίχηση αποτελεί το πιο δύσκολο και το πιο χρονοβόρο στάδιο στην εξαγωγή 3Δ πληροφορίας από ένα στερεοσκοπικό σύστημα.

Η διαδικασία αυτή είναι ισοδύναμη με την αντιστοίχιση κάθε σημείου της εικόνας αναφοράς σε ένα σημείο της δεύτερης εικόνας. Παρόλο που το πρόβλημα αυτό παρουσιάζεται ως δισδιάστατη αναζήτηση, στην πράξη εκφυλίζεται σε μονοδιάστατη, υποθέτοντας ένα βαθμονομημένο στερεοσκοπικό σύστημα, στο οποίο οι επιπολικές ευθείες βρίσκονται στις οριζόντιες γραμμές σάρωσης. Στην περίπτωση αυτή, ο χάρτης ανομοιομορφίας είναι μία διακριτή συνάρτηση $d(x,y)$ οι οποία αντιστοιχεί σημεία μεταξύ των δύο εικόνων:

$$I_{left}(x,y) = I_{right}(x - d(x,y), y) \tag{6.19}$$



Σχήμα 6.5: Αρχή λειτουργίας στερεοσκοπικής όρασης

με I_{left} και I_{right} τις δύο εικόνες. Η παραπάνω σχέση ποσοτικοποιεί την απόσταση κάθε ορατού από το στερεοσκοπικό σύστημα σημείου, σύμφωνα με την παρακάτω εξίσωση:

$$z = \frac{b\alpha}{d(d,y)} \tag{6.20}$$

με b τη βασική γραμμή, α το εστιακό βάθος και z την ζητούμενη απόσταση.

Η εξαγωγή, λοιπόν, του χάρτη βάθους, με δεδομένες τις παραμέτρους του συστήματος και την ανομοιομορφία, είναι μία πολύ απλή διαδικασία. Ως εκ τούτου, τέτοιες τεχνικές έχουν χρησιμοποιηθεί ευρέως σε εφαρμογές σχετιζόμενες με την υπολογιστική όραση, όπως η 3Δ ανακατασκευή, η εύρεση θέσης, η πλοήγηση κινουμένων ρομπότ,η αποφυγή εμποδίων, η παρακολούθηση κ.α. [Corso et al., 2003, Burschka and Hager, 2002, Sun et al., 2004, Atzpadin et al., 2004] Εντούτοις, η εξαγωγή της ανομοιομορφίας είναι μία απαιτητική εργασία, ιδιαίτερα όταν οι υπό επεξεργασία εικόνες δεν είναι ιδιαίτερα μικρές.

Έχουν προταθεί πολλοί αλγόριθμοι για την αντιμετώπιση αυτού του προβλήματος. Οι πιο πολλοί εμπίπτουν σε δύο βασικές κατηγορίες. Οι αλγόριθμοι περιοχής (Area based algorithms) παράγουν πυκνούς χάρτες ανομοιομορφίας χρησιμοποιώντας τις φωτεινότητες των εικόνων ως μετρικές απόστασης. Απόλυτες αποστάσεις, μέτρα συσχέτισης και άλλες μετρικές έχουν χρησιμοποιηθεί εκτενώς σε τεχνικές βασισμένες σε παράθυρα. Οι τεχνικές αυτές μειονεκτούν σε σκηνές με περιοχές χωρίς υφή και σε συστήματα με μεγάλη βασική απόσταση.

Από την άλλη μεριά, οι αλγόριθμοι που βασίζονται σε χαρακτηριστικά, βασίζονται σε σημεία τα οποία έχουν εξαχθεί από ανιχνευτές χαρακτηριστικών και έχουν περιγραφεί από περιγραφείς χαρακτηριστικών. Παρόλο που γενικά παράγουν πιο ακριβή αποτελέσματα, είναι πιο αργοί και αρχικά δίνουν αραιούς χάρτες οι οποίοι απαιτούν τεχνικές παρεμβολής με επιπλέον υπολογιστικό κόστος. Έτσι, οι τεχνικές βασισμένες σε περιοχές προτιμώνται όταν απαιτείται ταχεία επεξεργασία. Εντούτοις, μόνο τα τελευταία χρόνια έχουν προταθεί τεχνικές πραγματικού χρόνου, λόγω της αύξησης των επιδόσεων των επεξεργαστών.

Τεχνικές λογισμικού έχουν προταθεί με ταχύτητες μέχρι 17 πλαισίων ανα δευτερόλεπτο για εικόνες *VGA*. Με χρήση λογισμικού, οι Woodfill και Von Herzen [Woodfill and Von Herzen, 1997] πρότειναν ένα σύστημα ικανό για επεξεργασία εικόνων μεγέθους 320×240 στοιχείων με ρυθμό 42 πλαισίων ανά δευτερόλεπτο. Εντούτοις, το σύστημά τους αποτελείται από 16 συσκευές FPGA καθώς και από 16 συσκευές μνήμης*SRAM* χωρητικότητας από 1*MByte*, καταλήγοντας σε ένα εξαιρετικά ακριβό, από άποψη κόστους και επιφάνειας πυριτίου, σύστημα.

Σε αυτό το κεφάλαιο παρουσιάζεται μία αρχιτεκτονική κατάλληλη για τη γρήγορη διεκπεραίωση αυτού του σταδίου επεξεργασίας. Αυτή τη στιγμή είναι η ταχύτερη στην βιβλιογραφία, κατάλληλη για επεξεργασία 30 σκηνών 3MPixel ανά δευτερόλεπτο. Η αρχιτεκτονική αυτή, μπορεί να χωρέσει σε μία συσκευή, και λειτουργεί με παράλληλο, πλήρως διοχετευμένο τρόπο. Οι εικόνες αρχικά φιλτράρονται από ένα απλό γραμμικό φίλτρο προς μείωση του θορύβου. Ο αρχικός χάρτης ανομοιομορφίας υπολογίζεται από μία παράλληλη μονάδα, χρησιμοποιώντας μία απλή μέθοδο βασισμένη σε περιοχή. Για τη μείωση τον σφαλμάτων, υλοποιήθηκε ένα Κυψελιδωτό Αυτόματο ως φίλτρο, το οποίο εφαρμόζεται στην τελική έξοδο του συστήματος.

6.3 Προτεινόμενος αλγόριθμος

6.3.1 Γραμμικό φίλτρο

Για την ταχεία προ-επεξεργασία των εικόνων εισόδου επιλέχθηκε ένα μονοδιάστατο γραμμικό φίλτρο, προς μείωση του επιπέδου θορύβου. Επιπλέον, με τη χρήση του φίλτρου, αυξάνεται η συσχέτιση μεταξύ των γειτονικών στοιχείων των εικόνων, δίνοντας, έτσι, καλύτερης ποιότητας χάρτες ανομοιομορφίας. Το φίλτρο μπορεί να περιγραφεί με την παρακάτω σχέση:

$$F(x,y) = \frac{1}{4}\left(f(x-1,y) + f(x+1,y)\right) + \frac{1}{2}f(x,y) \tag{6.21}$$

Υλοποιώντας ένα φίλτρο δύο διαστάσεων, τα αποτελέσματα βελτιώνονται. Εντούτοις, η αυξημένη απαιτούμενη επιφάνεια πυριτίου μπορεί να μην δικαιολογεί την βελτίωση αυτή.

Σε πολλές πρακτικές εφαρμογές, οι απόλυτες τιμές των φωτεινοτήτων είναι αναξιόπιστες, και κάποιο είδος κανονικοποίησης πρέπει να προηγηθεί. Από τη στιγμή που το παρόν σύστημα στηρίζεται σε μονοδιάστατη επεξεργασία, είναι δυνατόν να υλοποιηθεί ένα Λαπλασιανό φίλτρο με μεγάλη οριζόντια μάσκα. Για την υλοποίηση αυτή, αρκεί ένα φίλτρο μέσης τιμής, η έξοδος από το οποίο θα αφαιρείται από την αρχική τιμή του στοιχείου.

6.3.2 Εκτίμηση ανομοιομορφίας

Ο αλγόριθμος εκτίμησης ανομοιομορφίας αποτελεί ακραία περίπτωση ενός απλού αλγορίθμου περιοχής, στον οποίο τα παράθυρα αναζήτησης αποτελούνται από ένα στοιχείο. Παρόλο που ένας τέτοιος αλγόριθμος δεν παρουσιάζει μεγάλη ακρίβεια, είναι εξαιρετικά γρήγορος. Για κάθε σημείο στη δεξιά εικόνα, πραγματοποιείται έρευνα σε ένα παράθυρο στην αριστερή. Το παράθυρο έχει μέγεθος $1 \times D$ στοιχεία, με D τον αριθμό των επιπέδων ανομοιομορφίας. Ο αριθμός αυτός εξαρτάται από τη γεωμετρία του συστήματος, όπως επίσης και από τις παραμέτρους των καμερών. Πρέπει, λοιπόν, να είναι γνωστός εκ των προτέρων. Η σχετική θέση του στοιχείου που παρουσιάζει τη μικρότερη απόσταση αποθηκεύεται ως η τιμή ανομοιομορφίας. Το μέτρο που επιλέχθηκε για αυτή την απόσταση είναι η απόλυτη διαφορά των φωτεινοτήτων. Ο αλγόριθμος περιγράφεται από την παρακάτω σχέση:

$$d(x,y) = \min_{d}\left(\left|I^{filt}_{Left}(x,y) - I^{filt}_{Right}(x-d,y)\right|\right) \tag{6.22}$$

Χρησιμοποιόντας τη σχέση 6.21, προκύπτει:

$$\begin{aligned} d(x,y) \;&=\; \min_{d}\Bigg(\Bigg|\frac{1}{4}\left(I_{Left}(x-1,y) + I_{Left}(x+1,y)\right) \\ &\quad + \frac{1}{2}I_{Left}(x,y) - \frac{1}{4}I_{Right}(x-1-d,y) \end{aligned}$$

$$+ I_{Right}(x-d+1,y)) - \frac{1}{2}I_{Right}(x-d,y)\Big|\Big) \tag{6.23}$$

Είναι προφανές, ότι με τη χρήση αυτού του φίλτρου, ο αλγόριθμος συμπεριφέρεται ως αλγόριθμος περιοχής, αντί ως σημείου.

6.3.3 Οριζόντιο φίλτρο ΚΑ

Λόγω του μεγάλου αριθμού των εσφαλμένων αντιστοιχίσεων, είναι αναγκαία μία μέθοδος αποτελεσματικής μετέπειτα επεξεργασίας. Τυπικά γραμμικά φίλτρα, ή ταξινομητές δεν έδωσαν ικανοποιητικά αποτελέσματα, παραμορφώνοντας ακμές ή εξαλείφοντας λεπτομέρειες. Μία προσέγγιση βασισμένη σε ΚΑ προτιμήθηκε.

Τα ΚΑ είναι μοντέλα φυσικών συστημάτων, στα οποία ο χώρος και ο χρόνος είναι διακριτά και οι αλληλεπιδράσεις είναι τοπικές [Sirakoulis, 2004, Sirakoulis et al., 2003]. Λόγω των τοπικών αλληλεπιδράσεων, η υλοποίηση των ΚΑ σε υλικό είναι αποδοτική, τόσο σε ταχύτητα, όσο και σε απαιτήσεις. Επίσης, οι πράξεις οι οποίες εκτελούνται από ένα ΚΑ είναι γενικά (όπως και την περίπτωσή αυτή) πολύ απλές, υλοποιήσιμες με απλούς συγκριτές, αθροιστές ή μετρητές. Ένα ΚΑ χαρακτηρίζεται από 5 ιδιότητες: τον αριθμό των χωρικών διαστάσεων (στην περίπτωση αυτή 1Δ), το πλάτος κάθε πλευράς του πίνακα (w), με w_j το πλάτος της jστής πλευράς του ($j = 1, 2, \dots, n$), το πλάτος της γειτονιάς της κυψελίδας (d), όπου D_j είναι το πλάτος της γειτονιάς κατά την jστή πλευρά του πίνακα, τις καταστάσεις των κυψελίδων (συνήθως δυαδικές) και τον κανόνα του ΚΑ, ο οποίος είναι μία συνάρτηση F (στην περίπτωση αυτή δυαδική) κατά την οποία η κατάσταση μίας κυψελίδας, στο χρόνο $t+1$ υπολογίζεται βάση της F. Η F είναι συνάρτηση της κατάστασης της κυψελίδας στο χρόνο (t) και των καταστάσεων των γειτονικών κυψελίδων κατά τον ίδιο χρόνο. Οι τέσσερις κανόνες που υλοποιήθηκαν για αυτή την εφαρμογή είναι:

- Για κάθε σημείο της εικόνας ανομοιομορφίας, έλεγξε αν κανένα από τα στοιχεία στη γειτονιά 3 δεν παρουσιάζουν την ίδια τιμή με το τρέχον. Σε αυτή την περίπτωση, θέσε την τιμή του τρέχοντος σημείου σε -1(αδιευκρίνιστο). Πρέπει να σημειωθεί ότι η γειτονιά 3 αποτελείται από το τρέχον σημείο καθώς και από τα δύο άμεσα γειτονικά του ($(i-1), i, (i+1)$).

- Για κάθε σημείο της εικόνας ανομοιομορφίας, έλεγξε αν και τα δύο άλλα σημεία της γειτονιάς 3 του τρέχοντος παρουσιάζουν την ίδια τιμή. Σε αυτή την περίπτωση, θέσε το τρέχον σημείο στην επικρατούσα τιμή για την γειτονιά 3 του.

- Για κάθε σημείο της εικόνας ανομοιομορφίας, έλεγξε αν λιγότερα από 2 σημεία της γειτονιάς 3 του τρέχοντος σημείου παρουσιάζουν την ίδια τιμή με το τρέχον. Επιπλέον, εξέτασε αν λιγότερα από 2 στοιχεία της γειτονιάς 5 παρουσιάζουν την ίδια τιμή με το τρέχον. Σε αυτή την περίπτωση, θέσε την τιμή ανομοιομορφίας σε -1. Η γειτονιά 5 αποτελείται από τις κυψελίδες $(i-2),(i-1),i,(i+1),(i+2)$.

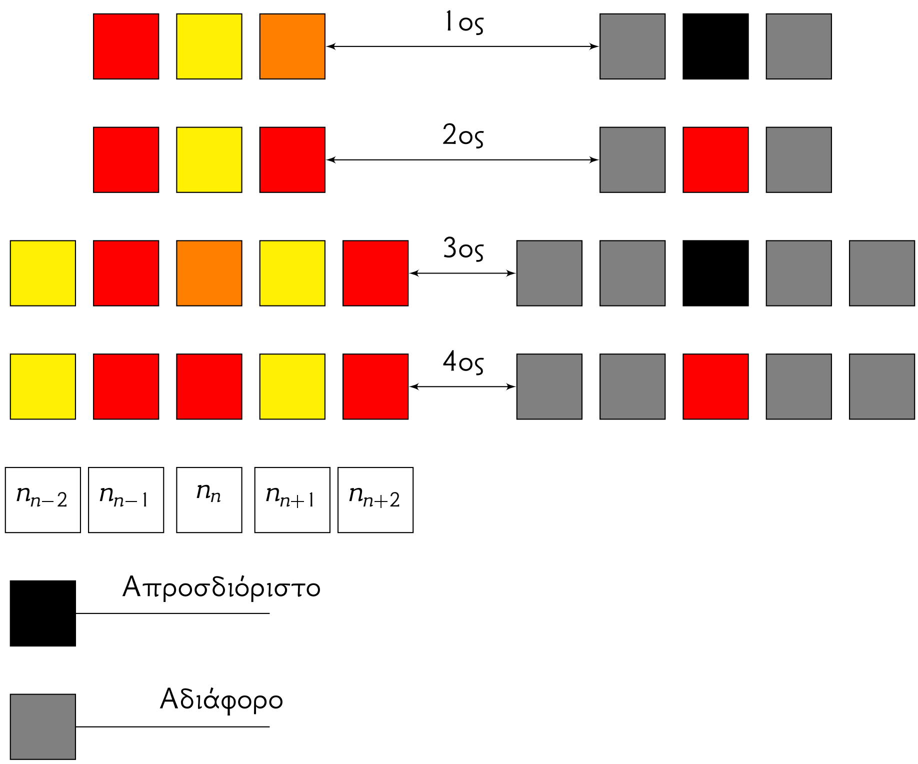
- Για κάθε σημείο της εικόνας ανομοιομορφίας, έλεγξε αν περισσότερα από 2 στοιχεία της γειτονιάς 5 παρουσιάζουν την ίδια τιμή. Σε αυτή την περίπτωση, θέσε την τιμή του τρέχοντος σημείου στην επικρατούσα για την γειτονιά τιμή.

Η διαδικασία αυτή παρουσιάζεται και στο σχήμα 6.5.

Οι κανόνες αυτοί βασίστηκαν στην υπόθεση ότι οι χάρτες ανομοιομορφίας είναι σημειακά συνεχείς. Στη συνεχή περίπτωση αυτό σημαίνει ότι δηλαδή παρουσιάζουν αριθμήσιμο πλήθος σημείων ασυνέχειας. Στη διακριτή περίπτωση, θεωρήθηκε ότι σε μία μικρή περιοχή, οι τιμές του βάθους μπορούν να μεταβάλλονται μόνο μία φορά. Το βασικό αποτέλεσμά τους είναι η απόρριψη ενός μεγάλου αριθμού εσφαλμένων ανακατασκευών. Αυξάνοντας την μέγιστη τιμή ανομοιομορφίας, ο αριθμός των σφαλμάτων αυξάνεται. Εντούτοις, από τη στιγμή που τα σημεία αυτά δεν έχουν συνοχή, απορρίπτονται από τα φίλτρα. Το τελικό αποτέλεσμα είναι η μείωση της πυκνότητας του χάρτη, αλλά όχι και της ακρίβειας.

6.3.4 Κατακόρυφα και 2Δ ΚΑ

Ένα κατακόρυφο ή και 2Δ φίλτρο ΚΑ μπορεί επίσης να υλοποιηθεί για την αύξηση της ακρίβειας του τελικού χάρτη βάθους. Η τελική βελτίωση γενικά δικαιολογεί την αύξηση της πολυπλοκότητας του συστήματος. Ένα κατακόρυφο φίλτρο χρησιμοποιεί ακριβώς τους ίδιους κανόνες με το οριζόντιο, με τη διαφορά ότι η γειτονιά γίνεται τώρα $(x,y-2,y-1,\ldots,y+2)$ αντί της $(x-2,x-1,\ldots,x+2,y)$. Ένα φίλτρο δύο διαστάσεων απαιτεί ελαφρώς τροποποιημένους κανόνες στη γειτονιά $(x-2,\ldots,x+2,y-2,\ldots,y+2)$. Και στις δύο περιπτώσεις, η απαιτούμενη επιφάνεια αυξάνεται δραματικά, προκειμένου να υπάρχει πρόσβαση σε περισσότερες από μία γραμμές σάρωσης.Εντούτοις, τα φίλτρα αυτά δεν μειώνουν την ταχύτητα του συστήματος, με εξαίρεση μία μικρή αρχική καθυστέρηση (latency). Η βέλτιστη υλοποίηση, από άποψη απλότητας



Σχήμα 6.6: Κανόνες ΚΑ

αλλά και ακρίβειας, βρέθηκε να είναι μία που υλοποιεί το πλήρες οριζόντιο ΚΑ, όπως περιγράφηκε στην προηγούμενη παράγραφο, μαζί με ένα απλουστευμένο κατακόρυφο, το οποίο υλοποιεί μόνο τους δύο πρώτους κανόνες.

6.3.5 Πειραματικά αποτελέσματα

Σε αυτή την παράγραφο παρουσιάζονται μερικά πειραματικά αποτελέσματα της προτεινόμενης μεθόδου[1]. Το πρώτο ζεύγος εικόνων προέρχεται από το πανεπιστήμιο Tsukuba, όπως φαίνεται στο Σχήμα 6.7. Τα αποτελέσματα τις μεθόδου για αυτό το ζεύγος φαίνονται στο Σχήμα 6.8.

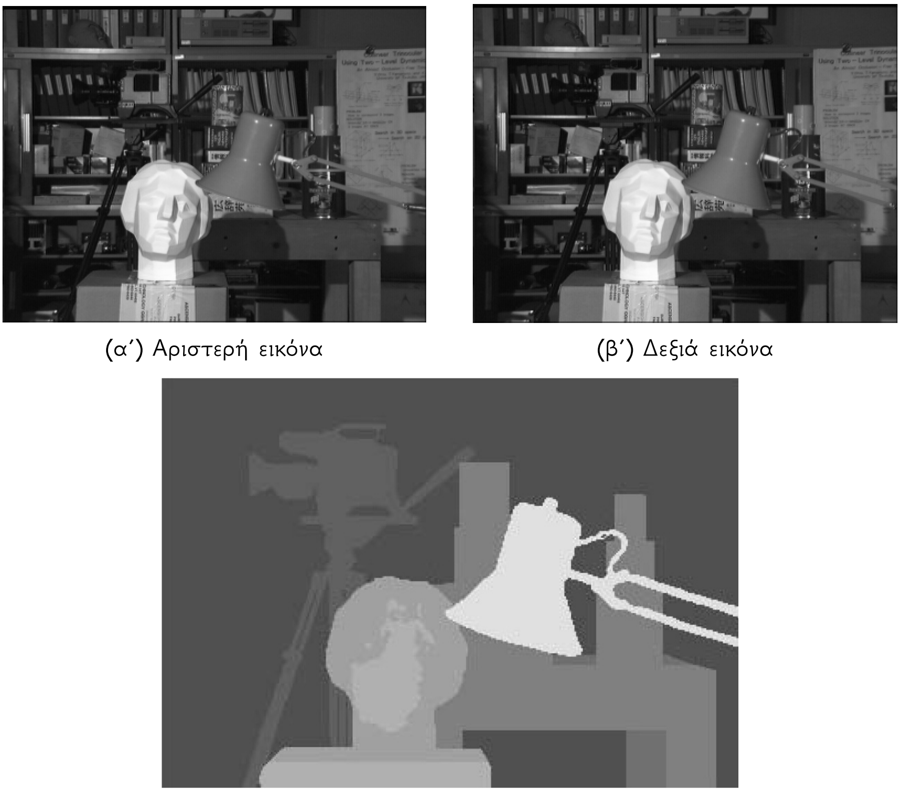
Το δεύτερο ζεύγος εικόνων είναι το ζεύγος Venus, όπως φαίνεται στο Σχήμα 6.9. Τα αποτελέσματα τις μεθόδου για αυτό το ζεύγος φαίνονται στο Σχήμα 6.10. Τέλος, παρατίθονται τα αποτελέσματα για το ζεύγος Cones (Σχήματα 6.11 και 6.12).

Είναι προφανές ότι πριν την επεξεργασία με τα ΚΑ, οι χάρτες βάθους παρουσιάζουν πολύ υψηλό επίπεδο θορύβου. Εντούτοις, μετά την επεξεργασία, οι περιοχές των σκηνών γίνονται αντιληπτές.

Όπως με όλες τις τεχνικές περιοχής, η προτεινόμενη παρουσιάζει βέλτιστα αποτελέσματα σε περιοχές με παρουσία υφής. Επιπλέον, από τη στιγμή που ο αλγόριθμος χρησιμοποιεί αντιστοιχήσεις σημείων αντί παραθύρων, οι ανομοιομορφίες γύρω από ασυνέχειες υπολογίζονται με ακρίβεια. Γενικά, οι τιμές σε περιοχές ορατές μόνο από τη μία κάμερα δεν έχουν συνοχή και απορρίπτονται από το φίλτρο ΚΑ. Στον πίνακα 6.1 παρουσιάζονται ποσοτικά αποτελέσματα σχετικά με την κάλυψη και την ακρίβεια του προτεινόμενου συστήματος, για τα ζεύγη εικόνων του Middlebury. Για την ανάκτηση πυκνών χαρτών ανομοιομορφίας απαιτείται κάποιας μορφής παρεμβολή. Οι περισσότερες τεχνικές που έχουν προταθεί πρόσφατα, παρουσιάζουν μεγαλύτερη ακρίβεια και ποσοστό κάλυψης από την προτεινόμενη. Η τελευταία, όμως, όταν υλοποιείται σε υλικό, παρουσιάζει πολύ μεγαλύτερη ταχύτητα επεξεργασίας.

Συγκριτικά αποτελέσματα παρουσιάζονται στον Πίνακα 6.3. Αξίζει να σημειωθεί, ότι τα αποτελέσματα των Gong καιYang, όπως και του Veksler βασίστηκαν στις αντίστοιχες εργασίες. Οι αναγραφόμενες ταχύτητες επεξεργασίας αναφέρονται σε υπολογιστές 2*GHzP*4 και 600*MHzPIII*, αντίστοιχα.

[1] Οι δοκιμαστικές εικόνες εμφανίζονται στη ιστοσελίδα του πανεπιστημίου *Middleburry*, http://stereo.middlebury.edu



(α′) Αριστερή εικόνα

(β′) Δεξιά εικόνα

(γ′) Χάρτης βάθους

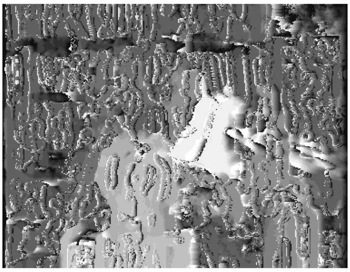
Σχήμα 6.7: Το ζεύγος εικόνων *Tsukuba*

6.4 Υλοποίηση σε υλικό

Το βασικό πλεονέκτημα του αλγορίθμου που περιγράφηκε στις προηγούμενες ενότητες είναι η δυνατότητα για απλή και αποδοτική υλοποίηση σε υλικό. Η αρχιτεκτονική μπορεί να χωριστεί σε τρία μέρη:

- Φίλτρα προεπεξεργασίας
- Υπολογισμός ανομοιομορφίας
- Τελική επεξεργασία με Κυψελιδωτά Αυτόματα

Όλα τα παραπάνω ικανοποιούν τις συνθήκες διοχέτευσης, δηλαδή παρουσιάζεται έγκυρη έξοδος σε κάθε κύκλο ρολογιού. Οι καθυστερήσεις



(α΄) Χάρτης πριν την επεξεργασία με ΚΑ

(β΄) Τελικός χάρτης

Σχήμα 6.8: Αποτελέσματα για το ζεύγος εικόνων *Tsukuba*

	Tsukuba		Venus		Cones		Teddy	
	Acc	Cov	Acc	Cov	Acc	Cov	Acc	Cov
Χωρίς ΚΑ ή προεπεξεργασία	44%	100%	24%	100%	13%	100%	16%	100%
Χωρίς ΚΑ, με 1Δ γραμ. φίλτρο	47%	100%	30%	100%	16%	100%	19%	100%
Χωρίς ΚΑ, με 2Δ γραμ. φίλτρο	52%	100%	33%	100%	18%	100%	21%	100%
Με ΚΑ	95%	53%	92%	33%	90%	20%	92%	22%

Πίνακας 6.1: Ποσοτικά αποτελέσματα

στην έξοδο δεν παρουσιάζουν προβλήματα, αφού είναι της τάξης των μικροδευτερολέπτων. Ένας επιπλέον περιορισμός που πρέπει να τεθεί για την αυτονομία του συστήματος είναι ότι δεν πρέπει να στηρίζεται σε υπολογισμούς λογισμικού από εξωτερικό υπολογιστή. Το μπλοκ διάγραμμα του συστήματος παρουσιάζεται στο Σχήμα 6.13.

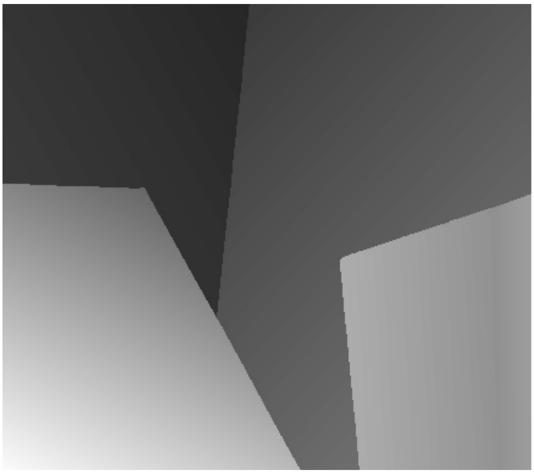
Η αρχιτεκτονική υλοποιήθηκε σε μία συσκευή FPGA της οικογένειας Cyclone II της *Altera* χρησιμοποιώντας VHDL. Η μέγιστη συχνότητα λειτουργίας της συσκευής βρέθηκε να είναι 115*MHz*.

6.4.1 Προεπεξεργασία

Η αρχιτεκτονική που υλοποιεί το γραμμικό φίλτρο είναι πολύ απλή. Απαιτεί μόνο 3 καταχωρητές 8 ψηφίων, 5 ολισθητές και 2 αθροιστές.

(α′) Αριστερή εικόνα

(β′) Δεξιά εικόνα



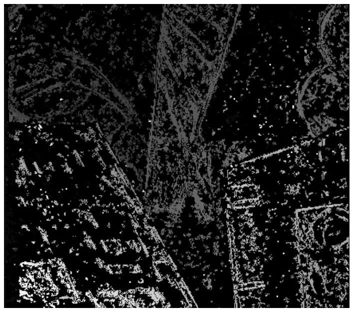
(γ′) Χάρτης βάθους

Σχήμα 6.9: Το ζεύγος εικόνων *Venus*

Το κύκλωμα παρουσιάζεται στο σχήμα 6.14. Οι δύο εικόνες επεξεργάζονται ταυτόχρονα, οπότε απαιτούνται δύο τέτοια κυκλώματα. Τα βασικά του χαρακτηριστικά είναι το μικρό μέγεθος και η αρχιτεκτονική διαδοχικής διοχέτευσης.

Για την υλοποίηση ενός 2Δ φίλτρου, πρέπει να χρησιμοποιηθούν επιπλέον στοιχεία. Μία τέτοια συσκευή παρουσιάζεται στο σχήμα 7. Αποτελείται από $3N$ ψηφία σειριακής μνήμης *FIFO*, με N το πλάτος των εικόνων, καθώς και κάποια λογική ελέγχου. Το τελικό φίλτρο είναι όμοιο με αυτό του σχήματος 6.15. Δύο τέτοια φίλτρα πρέπει να υλοποιηθούν προκειμένου να επιτευχθεί βέλτιστη απόδοση. Ένα φίλτρο κανονικοποίησης μέσης τιμής μπορεί να υλοποιηθεί με μικρό επιπλέον κόστος, συγκεκριμένα 16 καταχωρητές, 15 αθροιστές, έναν αφαιρέτη

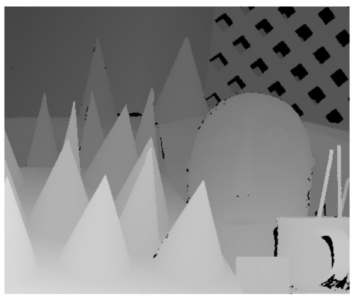
(α΄) Χάρτης πριν την επεξεργασία με ΚΑ



(β΄) Τελικός χάρτης

Σχήμα 6.10: Αποτελέσματα για το ζεύγος εικόνων *Venus*

(α΄) Αριστερή εικόνα



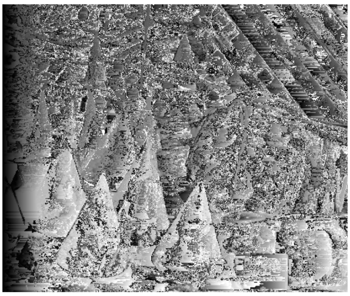
(β΄) Χάρτης βάθους

Σχήμα 6.11: Το ζεύγος εικόνων *Cones*

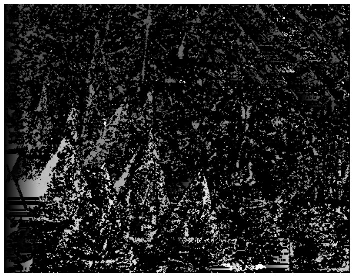
και ένα στοιχείο ολίσθησης. Από τη στιγμή που και αυτή η μονάδα επεξεργάζεται ένα εικονοστοιχείο ανά κύκλο, δεν επιφέρει καμία επιβράδυνση στο σύστημα.

6.4.2 Εκτίμηση ανομοιομορφίας

Η μονάδα αυτή παρουσιάζει την μεγαλύτερη υπολογιστική πολυπλοκότητα σε κάθε σύστημα στερεοσκοπικής όρασης. Προκειμένου να καλύπτει τις επιβαλλόμενες ανάγκες για υψηλή ταχύτητα, υλοποιήθηκε ένα παράλληλο σύστημα. Οι απαιτήσεις της μονάδας αυτής σε υλικό παρουσιάζονται στον πίνακα 6.2. Το λογικό διάγραμμα παρουσιάζεται



(α′) Χάρτης πριν την επεξεργασία με ΚΑ



(β′) Τελικός χάρτης

Σχήμα 6.12: Αποτελέσματα για το ζεύγος εικόνων *Cones*

Καταχωρητές 8 ψηφίων	Καταχωρητές $\log_2 D$ ψηφίων	Συγκριτές	Αφαιρέτες	Μετρητές
$4n$	$2N$	N	N	1

Πίνακας 6.2: Κόστος μονάδας εκτίμησης ανομοιομορφίας (N πλάτος εικόνας και D τα επίπεδα ανομοιομορφίας)

στο σχήμα 6.16

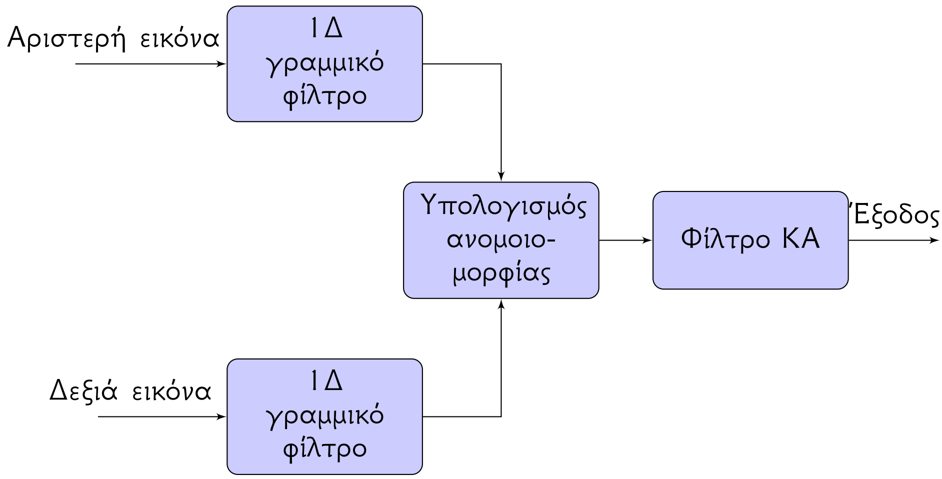
Οι καταχωρητές είναι οργανωμένοι ως εξής:

1. Μία συστοιχία από N καταχωρητές 8 ψηφίων χρησιμοποιείται ως μνήμη εισόδου. Οι φωτεινότητες των στοιχείων της πρώτης εικόνας αποθηκεύονται σειριακά. Μόλις διεκπαιρεωθεί αυτή η εργασία, οι έξοδοι δίνονται παράλληλα.

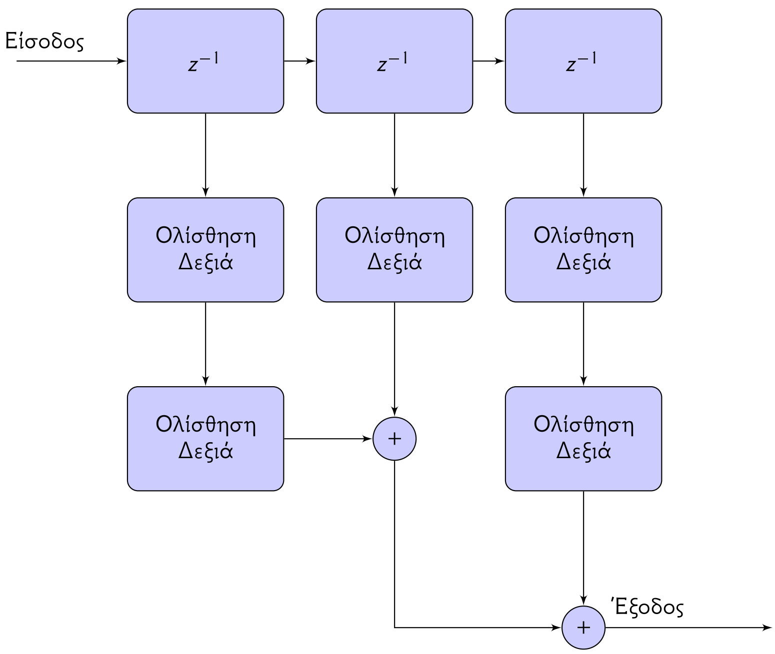
2. Μία δεύτερη συστοιχία N καταχωρητών 8 ψηφίων διατηρεί τα παραπάνω δεδομένα όσο υπολογίζεται η ανομοιομορφία.

3. Μία τρίτη, όμοια, συστοιχία χρησιμοποιείται για κάθε γραμμή σάρωσης της δεύτερης εικόνας. Όσο η εικόνα φορτώνεται, υπολογίζονται και οι διαφορές φωτεινότητας από N στοιχεία μη προσημασμένης αφαίρεσης.

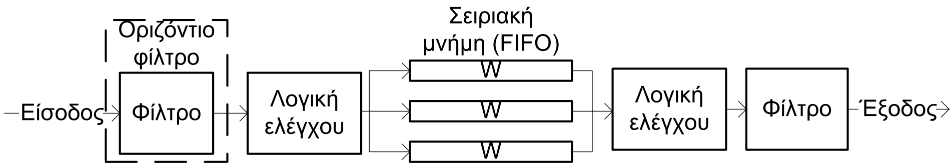
4. Μία τέταρτη συστοιχία N καταχωρητών 8 ψηφίων χρησιμοποιείται για να αποθηκεύει τα ελάχιστα των διαφορών φωτεινοτήτων μεταξύ των δύο εικόνων, με τη βοήθεια N συγκριτών.



Σχήμα 6.13: Μπλοκ διάγραμμα συστήματος



Σχήμα 6.14: Μπλοκ διάγραμμα γραμμικού φίλτρου



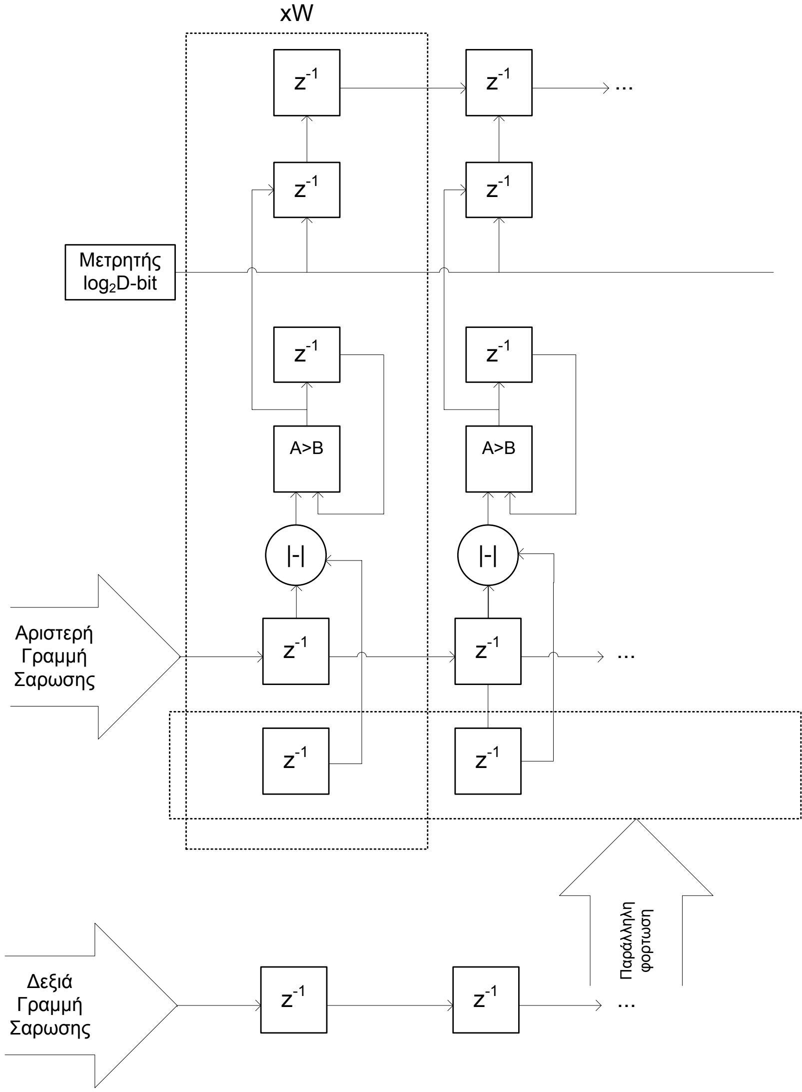
Σχήμα 6.15: Μπλοκ διάγραμμα γραμμικού φίλτρου 2Δ, για εικόνα πλάτους W στοιχείων. Κάθε στοιχείο μνήμης έχει μέγεθος λέξης 1 *Byte*

5. Μία συστοιχία N καταχωρητών, πλάτους $\log_2 D$ ψηφίων υλοποιείται για την αποθήκευση των θέσεων των ελαχίστων. Αυτές αποτελούν και τις τιμές της ανομοιομορφίας.
6. Η τελευταία συστοιχία δρα ως μνήμη εξόδου, προκειμένου να επιτευχθεί σειριακή έξοδος των εξόδων προς την επόμενη μονάδα.

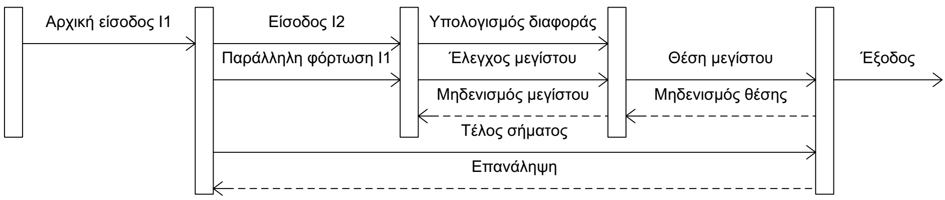
Όλα τα παραπάνω χρησιμοποιούν N καταχωρητές, με N το πλάτος των εικόνων. Παρακάτω δίνεται μία περιγραφή της λειτουργίας της συσκευής:

1. Κατά τη διάρκεια των πρώτων N κύκλων η πρώτη γραμμή σάρωσης της δεξιάς εικόνας φορτώνεται στην πρώτη συστοιχία καταχωρητών.
2. Η δεύτερη συστοιχία καταχωρητών φορτώνεται παράλληλα από την πρώτη. Ταυτόχρονα, η δεύτερη γραμμή σάρωσης της δεξιάς εικόνας φορτώνεται στην πρώτη συστοιχία, ενώ η πρώτη γραμμή της αριστερής εικόνας φορτώνεται σειριακά στην τρίτη συστοιχία καταχωρητών. Κατά τη διάρκεια αυτής της φάσης, υπολογίζονται οι απόλυτες διαφορές των φωτεινοτήτων από τα στοιχεία απόλυτης αφαίρεσης και οι τιμές τους αποθηκεύονται στην τέταρτη συστοιχία καταχωρητών. Οι αντίστοιχες θέσεις τους δίνονται από το μετρητή και αποθηκεύονται στην πέμπτη συστοιχία καταχωρητών.
3. Όταν η δεύτερη γραμμή σάρωσης της δεξιάς εικόνας καταλαμβάνει την πρώτη συστοιχία καταχωρητών, η δεύτερη γραμμή της αριστερής εικόνας αρχίζει να παρουσιάζεται στην τρίτη συστοιχία, η τέταρτη τίθεται σε μέγιστες τιμές, εξέρχονται οι τιμές εξόδου από την τελική συστοιχία και επαναλαμβάνεται η διαδικασία.

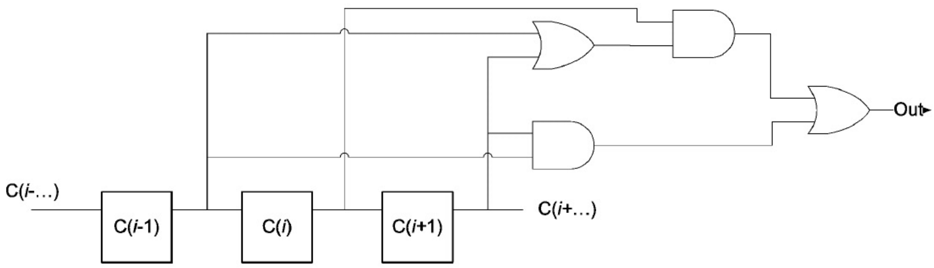
Ολόκληρη η διαδικασία αντιστοίχισης λειτουργεί σε πλήρη διαδοχική διοχέτευση. Μετά από μία αρχική καθυστέρηση $2N$ κύκλων, παρουσιάζεται έγκυρη έξοδος σε κάθε κύκλο.



Σχήμα 6.16: Λογικό διάγραμμα μονάδας εκτίμησης ανομοιομορφίας



Σχήμα 6.17: Ακολουθιακό διάγραμμα κυκλώματος



Σχήμα 6.18: Λογικό διάγραμμα 1Δ φίλτρου ΚΑ γειτονιάς 3

Στο σχήμα 6.17 παρουσιάζεται το ακολουθιακό διάγραμμα του κυκλώματος.

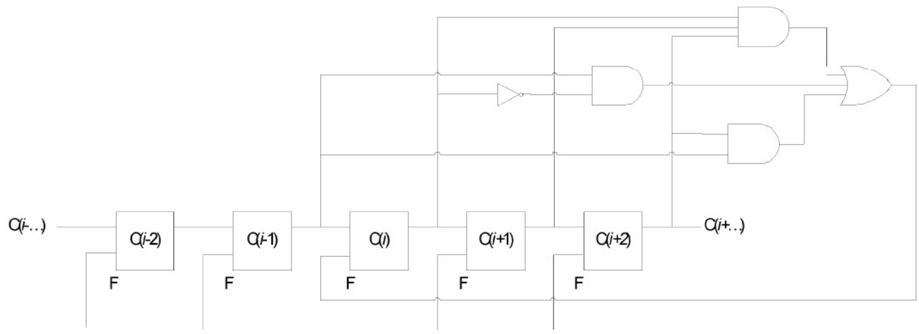
6.4.3 Φίλτρα ΚΑ

Τα φίλτρα ΚΑ συνήθως παρουσιάζουν μεγάλη υπολογιστική πολυπλοκότητα όταν υλοποιούνται σε λογισμικό, γεγονός που καθιστά τη χρήση τους μη πρακτική. Εντούτοις, λόγω της παράλληλης αρχιτεκτονικής τους, είναι δυνατή η υλοποίησή τους σε υλικό χωρίς επιβάρυνση στην ταχύτητα. Μία τέτοια αρχιτεκτονική παρουσιάζει το πλεονέκτημα της απλής υλοποίησης, όπως φαίνεται και στα σχήματα 6.18 και 6.19.

Είναι δυνατόν να υλοποιηθεί και ένα κατακόρυφο ή και δισδιάστατο τέτοιο φίλτρο, με επιπλέον κόστος $3N$ ή $5N$ καταχωρητών.

6.5 Ανάλυση απόδοσης

Όλες οι μονάδες που περιγράφηκαν λειτουργούν σε διαδοχική διοχέτευση. Μετά από μία αρχική περίοδο καθυστέρησης, δίνεται έξοδος σε κάθε κύκλο ρολογιού. Χρησιμοποιώντας μία συσκευή *FPGA* με τυπική συχνότητα λειτουργίας 100*MHz*, αυτό σημαίνει ότι ένας χάρτης βάθους,

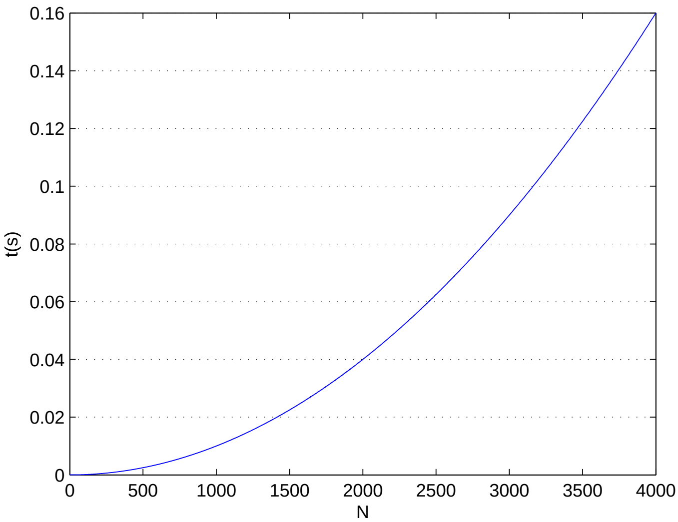


Σχήμα 6.19: Λογικό διάγραμμα 1Δ φίλτρου ΚΑ γειτονιάς 5

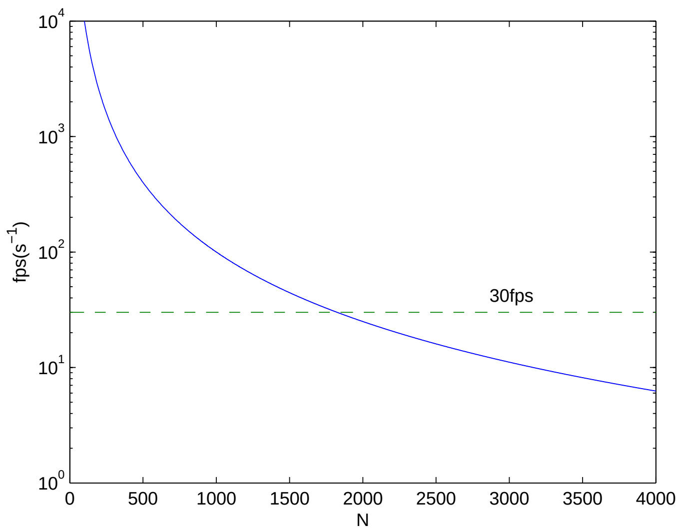
μεγέθους 1*MPixel* μπορεί να ανακτηθεί σε 10*ms*. Η σχέση μεταξύ του πλάτους των εικόνων (θεωρώντας τετράγωνες εικόνες) και του απαιτούμενου για την επεξεργασία χρόνου φαίνεται στο Σχήμα 6.20. Στο σχήμα 6.21 παρουσιάζονται ο αριθμός των επεξεργαζόμενων ζευγών ανά δευτερόλεπτο, σε σχέση με το μέγεθος της εικόνας.

Εικόνες *VGA* επεξεργάζονται με ρυθμό 325*fps*, εικόνες 1*Mpixel* με 100*fps* και 4*Mpixel* με 25*fps*. Η καθυστέρηση εξόδου του συστήματος καθορίζεται από το πλάτος των εικόνων και τις αρχιτεκτονικές που επιλέχτηκαν για τα γραμμικά φίλτρα και τα φίλτρα ΚΑ. Στην πιο απλή περίπτωση μονοδιάστατων φίλτρων, αυτή είναι $2N + 15$ κύκλοι ρολογιού, ενώ στην περίπτωση των φίλτρων δύο διαστάσεων γίνεται $12N + 15$ κύκλοι ρολογιού. Σε κάθε περίπτωση, ο χρόνος αυτός είναι μεταξύ 20 και 120*μsec*, ασήμαντος για σχεδόν κάθε εφαρμογή.

Μετρικές όπως η *PDS*(Pixels × Disparity Range/Seconds), χρησιμοποιούνται συχνά για την εκτίμηση της ταχύτητας τεχνικών στερεοσκοπικής ανάκτησης βάθους. Εντούτοις, στη συγκεκριμένη περίπτωση, η ταχύτητα επεξεργασίας δεν εξαρτάται από τα επίπεδα ανομοιομορφίας και επομένως δεν δίνει αντιπροσωπευτικά αποτελέσματα. Για παράδειγμα, για τον υπολογισμό 64 επιπέδων η παρούσα αρχιτεκτονική παρουσιάζει $6,400,000,000PDS$ ενώ για 32 επίπεδα $6,400,000,000PDS$. Γενικά $Speed_{PDS} \approx 10^8 \times D$. Στον πίνακα 6.3 παρουσιάζονται συγκριτικά αποτελέσματα για τους χρόνους επεξεργασίας. Αξίζει να σημειωθεί ότι λόγω της προαναφερθείσας ιδιότητας της υλοποίησης, η βελτίωση της ταχύτητας αυξάνει με τα επίπεδα ανομοιομορφίας, αφού οι μέθοδοι των Gong& Yang και του Veksler παρουσιάζουν υπολογιστική πολυπλοκότητα $O(D)$ σε σχέση με το εύρος αναζήτησης.



Σχήμα 6.20: Απαιτούμενος χρόνος επεξεργασίας για ζεύγος εικόνων πλάτους N



Σχήμα 6.21: Επεξεργαζόμενα ζεύγη ανά δευτερόλεπτο (*fps*), για πλάτος N

	Tsukuba			Venus		
	Ακρ.	Πυκν.	Χρόνος	Ακρ.	Πυκν.	Χρόνος
Προτεινόμενη	95%	53%	1,1	92%	33%	1,7
Gong&Wang(03)	99%	71%	4,7	99,08%	62%	109
Gong&Wang(05)	99,3%	77%	6,2	99,7%	73%	156
Veksler(02)	99,64%	75%	6.000	99,84%	73%	13.000
Veksler(03)	99,62%	66%	1.000	98,17%	68%	5.000

Πίνακας 6.3: Σύγκριση επιδόσεων τεχνικών

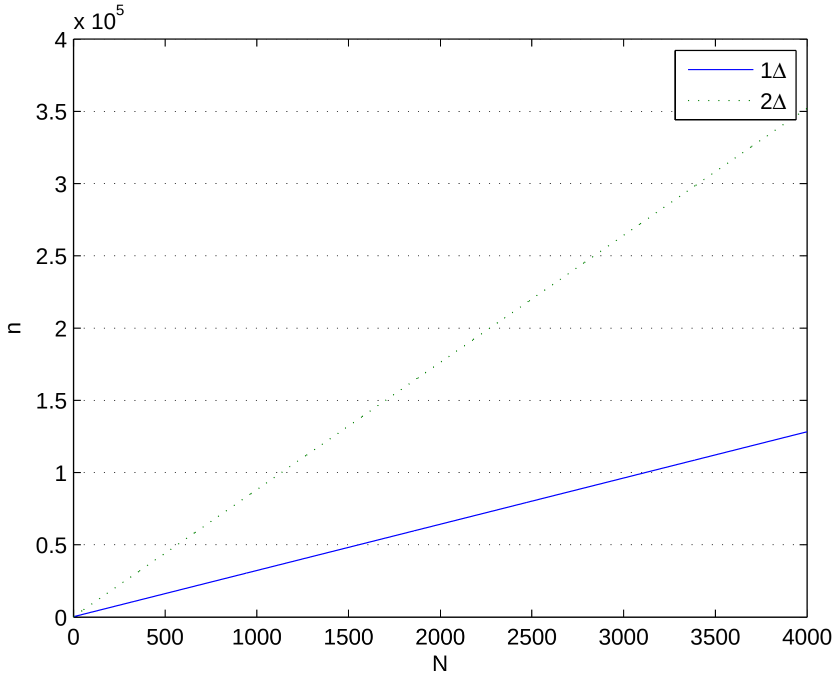
Η ταχύτητα της προτεινόμενης τεχνικής είναι η μεγαλύτερη που αναφέρεται στη βιβλιογραφία.

6.6 Αποτίμηση κόστους

Οι τελικές απαιτήσεις υλικού του συστήματος εξαρτώνται από τις επιλεχθείσες αρχιτεκτονικές για τα γραμμικά φίλτρα και τα φίλτρα ΚΑ. Με τη χρήση μονοδιάστατων φίλτρων, απαιτούνται $3N+19$ καταχωρητές 8 ψηφίων, $2N+14$ καταχωρητές $\log_2 D$ ψηφίων, N συγκριτές, N αφαιρέτες, 20 ολισθητές 8 ψηφίων, 3 αθροιστές 8 ψηφίων, καθώς και ένας μικρός αριθμός λογικών πυλών δύο εισόδων. Η σχέση μεταξύ του μεγέθους των εικόνων και του αριθμού των απαιτούμενων ψηφίων καταχώρησης παρουσιάζεται στο σχήμα 6.22 Στον πίνακα 6.4 παρουσιάζεται το κόστος για διάφορες υλοποιήσεις. Παρόλο που η παρούσα τεχνολογία *FPGA* μπορεί να καλύψει τις ανάγκες αυτές, το σύστημα μπορεί να επεκταθεί και σε ένα σύνολο από συσκευές, προκειμένου να μειωθεί το κόστος επεξεργασίας πολύ μεγάλων εικόνων. Κάτι τέτοιο είναι εφικτό λόγω της μικρής ποσότητας πληροφορίας που χρειάζεται να μεταβιβασθεί. Τέτοιες υλοποιήσεις παρουσιάζονται και στο σχήμα 6.23. Τα στάδια προ-επεξεργασίας και μετά-επεξεργασίας συμπεριλαμβάνουν τα γραμμικά φίλτρα και τα φίλτρα ΚΑ, όπως αυτά περιγράφηκαν σε προηγούμενες ενότητες.

6.7 Περαιτέρω διερεύνηση

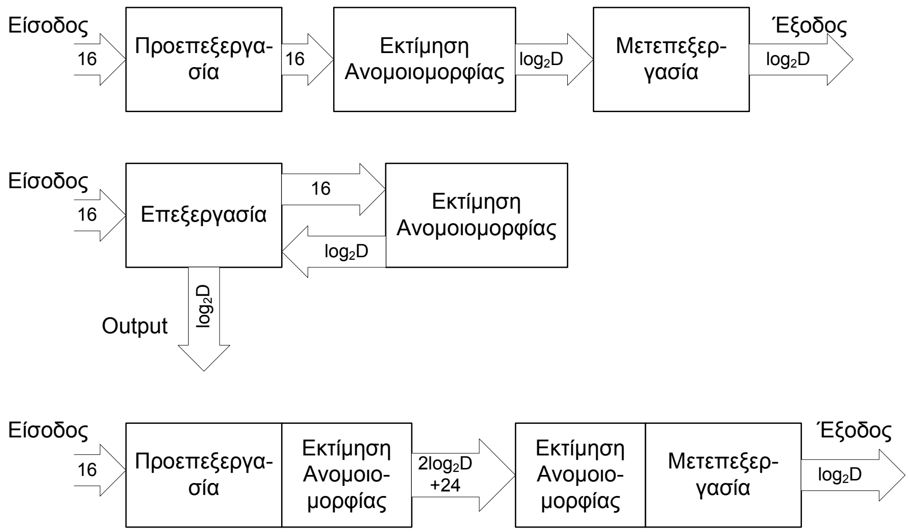
Οι παραγόμενοι χάρτες ανομοιομορφίας, παρόλο που εξάγονται πολύ γρήγορα, σε πολλές εφαρμογές δεν ενδείκνυνται λόγω της σχετικά χαμηλής τους ακρίβειας. Εντούτοις, η παραπάνω αρχιτεκτονική είναι χρήσιμη και για τέτοιες απαιτήσεις, ως στάδιο προεπεξεργασίας για άλλες μεθόδους, ή με κάποιες παραλλαγές.



Σχήμα 6.22: Απαιτούμενα ψηφία καταχωρητών n σε σχέση με το πλάτος N της επεξεργαζόμενης εικόνας

	Καταχωρητές 8 ψηφίων	Καταχωρητές $\log_2 D$ ψηφίων
Φίλτρα 1Δ	$3N+19$	$8N+14D$
1Δ γραμμικό, 2Δ ΚΑ	$3N+19$	$8N+52D$
1Δ γραμμικό, Οριζ.& Κατ. ΚΑ (Προτεινόμενο)	$3N+19$	$6N+23D$
Φίλτρα 2Δ	$7N+19$	$8N+52D$

Πίνακας 6.4: Κόστος επεξεργασίας



Σχήμα 6.23: Παραδείγματα αρχιτεκτονικών πολλαπλών συσκευών

6.7.1 Αύξηση ακρίβειας μεθόδου

Στις περισσότερες τεχνικές εξαγωγής χαρτών ανομοιομορφίας θεωρείται η παρακάτω παραδοχή:

'Κάθε σημείο μίας εικόνας, αντιστοιχίζεται το πολύ σε ένα σημείο της άλλης '(Κριτήριο μοναδικότητας, *Uniqueness constraint*)

Κάτι τέτοιο μπορεί να υλοποιηθεί με την επανάληψη της αναζήτησης που περιγράφηκε στο υπόλοιπο κεφάλαιο με αντίστροφη φορά. Δηλαδή, την εύρεση των αντιστοιχιών της δεξιάς εικόνας στην αριστερή. Κάθε σημείο για το οποίο δεν υπάρχει και αντίστροφη αντιστοιχία, απορρίπτεται ως μη αξιόπιστο. Το αποτέλεσμα της παραλλαγής αυτής είναι ένας πολύ πιο ακριβής χάρτης, με πολύ μικρότερη κάλυψη. Είναι προφανές ότι μία τέτοια υλοποίηση απαιτεί διπλάσιο χρόνο από ότι αυτή της μονής κατεύθυνσης.

6.7.2 Η μέθοδος ως στάδιο άλλων τεχνικών

Οι πιο ακριβείς τεχνικές που μπορούν να βρεθούν στη βιβλιογραφία είναι αυτές του Yoon [Yoon and Kweon, 2006] , της Veksler [Veksler, 2003] κ.α. Όλες οι παραπάνω τεχνικές χρησιμοποιούν την απόσταση μεταξύ των στοιχείων των εικόνων του στερεοσκοπικού ζεύγους. Αυτό αποτελεί και ένα ιδιαίτερα χρονοβόρο στάδιο για τη συνολική επεξεργασία. Παρόλο που συνήθως δεν παρουσιάζει την μεγαλύτερη υπολογιστική πολυπλοκότητα στο σύστημα, η χρήση της προτεινόμενης τεχνικής μπορεί να οδηγήσει σε σημαντική βελτίωση της απόδοσης.

6.8 Συμπεράσματα

Σε αυτό το κεφάλαιο παρουσιάστηκε μία νέα τεχνική ανάκτησης πληροφορίας βάθους, καθώς και η υλοποίησή της σε υλικό. Αυτή η τεχνική μπορεί να επεξεργασθεί εικόνες μεγάλου μεγέθους σε πραγματικό χρόνο, εκμεταλλευόμενη μία παράλληλη δομή, υλοποιημένη σε μία συσκευή *FPGA*. Με τη χρήση φίλτρων ΚΑ, οι πυκνοί χάρτες βάθους παρουσιάζουν επαρκή ακρίβεια για πολλές εφαρμογές όρασης μηχανών. Η ταχύτητα επεξεργασίας είναι η μεγαλύτερη που αναφέρεται στη βιβλιογραφία, ικανή να επεξεργασθεί ζεύγη εικόνων μεγέθους μέχρι και 3*MPixel* σε πραγματικό χρόνο.

Κεφάλαιο 7

Συμπεράσματα-Προτάσεις για μελλοντική έρευνα

Οι ροπές εικόνων αποτελούν συχνά χρησιμοποιούμενους περιγραφείς σχήματος σε πληθώρα εφαρμογών, σε όλο το εύρος του ερευνητικού πεδίου της επεξεργασίας και ανάλυσης εικόνας. Σε αυτή την διατριβή, παρουσιάστηκαν οι βασικοί τύποι τους, και αναλύθηκαν τα χαρακτηριστικά κάθε ενός.

Βασικός στόχος της διερεύνησης που πραγματοποιήθηκε στο μεγαλύτερο μέρος της διατριβής ήταν η βελτίωση των χαρακτηριστικών των ροπών για την πιο αποδοτική χρήση τους. Πιο συγκεκριμένα, αρχικά παρουσιάστηκαν τεχνικές για την ταχεία υλοποίησή τους τόσο σε υλικό όσο και σε λογισμικό. Αναπτύχθηκαν υλοποιήσεις που στοχεύουν στις ροπές *Zernike*, *Legendre*, *Chebyshev* καθώς και στις γεωμετρικές ροπές. Η ταχύτητα των προτεινόμενων τεχνικών είναι η μεγαλύτερη της βιβλιογραφίας, και επιτρέπει επεξεργασία πραγματικού χρόνου για εικόνες μεγάλου μεγέθους. Η βελτίωση ταχύτητας των προτεινόμενων αρχιτεκτονικών σε σχέση με άλλες τεχνικές είναι πάνω από μία τάξη μεγέθους.

Επιπλέον, περιγράφηκε μία νέα τεχνική για τον υπολογισμό των ροπών αυτών σε δυαδικές εικόνες, η οποία στηρίζεται στις ιδιότητες των παραγοντικών πολυωνύμων στο διακριτό χώρο, και στην περιγραφή των δυαδικών εικόνων μέσω του περιγράμματός τους. Εκτός των συνήθων τύπων μελετήθηκε και ο πρόσφατα προταθείς Γωνιακός Ακτινικός Μετασχηματισμός, ο οποίος και υλοποιήθηκε σε μία αποδοτική αρχιτεκτονική. Και σε αυτές τις περιπτώσεις, η βελτίωση στην ταχύτητα είναι μεγαλύτερη από μία τάξη μεγέθους. Επιπλέον, οι προτεινόμενες τεχνικές έχουν ουσιαστικά μικρότερες απαιτήσεις σε μνήμη, γεγονός που τις καθιστά κατάλληλες για χρήση σε ενσωματωμένα συστήματα.

Πέρα από τον ταχύ υπολογισμό των ροπών, παρουσιάζεται και μεθοδολογία για την μείωση των σφαλμάτων που εισάγονται κατά τον υπολογισμό τους. Λόγω της διακριτής φύσης των ψηφιακών εικόνων, ο υπολογισμός των ροπών επηρεάζεται από σφάλματα διακριτοποίησης. Πιο συγκεκριμένα, εμφανίζονται σφάλματα προσέγγισης τόσο στον υπολογισμό των ολοκληρωμάτων, όσο και στην προσέγγιση του μοναδιαίου δίσκου από καρτεσιανό πλέγμα. Η τεχνική που αναπτύχθηκε αντιμετωπίζει την πρώτη πηγή σφαλμάτων, με τη χρήση πολυωνυμικής παρεμβολής στο χώρο των ροπών. Τα πειραματικά αποτελέσματα έδειξαν ότι με τη χρήση της προτεινόμενης μεθοδολογίας, όχι μόνο είναι πιο ακριβής ο υπολογισμός των ροπών, αλλά βελτιώνονται και τα χαρακτηριστικά αμεταβλητότητας τους, τα οποία είναι ίσως το πιο σημαντικό πλεονέκτημα στη χρήση των ροπών. Η μείωση του σφάλματος σε σχέση με άλλες τεχνικές εξαρτάται από τον τύπο και την τάξη της υπολογιζόμενης ροπής, αλλά γενικά, για τις πιο συχνά χρησιμοποιούμενες, το σφάλμα μειώνεται κατα 50%.

Οι πρόσφατα προταθείσες ροπές διακριτής ορθογώνιας βάσης, παρόλο που δεν παρουσιάζουν εγγενή αμεταβλητότητα, δεν εμφανίζουν σφάλματα κατά τον υπολογισμό τους, αφού είναι ορισμένες στο διακριτό πεδίο της εικόνας. Οι συναρτήσεις βάσης για αυτές τις ροπές είναι πολυώνυμα ορθογώνια στο διακριτό χώρο της εικόνας. Τέτοια πολυώνυμα μπορούν να προκύψουν εύκολα μέσω της τεχνικής $Gram-Schmidt$, με δεδομένο ένα διάνυσμα βάρους w. Στην αναζήτηση για ένα τέτοιο διάνυσμα, το οποίο θα βελτιστοποιεί τα χαρακτηριστικά των παραγόμενων ροπών χρησιμοποιήθηκαν εξελικτικές τεχνικές. Πιο συγκεκριμένα, αντιμετωπίσθηκε η αναζήτηση αυτή ως πρόβλημα ελαχιστοποίησης, για το οποίο χρησιμοποιήθηκαν εξελικτικές στρατηγικές. Οι παραγόμενες ροπές έδωσαν σημαντικά βελτιωμένα χαρακτηριστικά ως προς την ανακατασκευή και την ανάκτηση σχημάτων.

Παρόλα τα μειονεκτήματα των ροπών διακριτής βάσης, αυτές παρουσιάζουν και ένα σύνολο πολύ χρήσιμων ιδιοτήτων, λόγω της ευκολίας στην ανακατασκευή του σήματος. Προς εκμετάλλευση αυτών των ιδιοτήτων, παρουσιάζονται τεχνικές για τον υπολογισμό της επίδρασης γραμμικών και μη μετασχηματισμών στο χώρο των ροπών. Με βάση τα αποτελέσματα αυτής της διερεύνησης, αναπτύσσονται δύο τεχνικές για τον υπολογισμό των ροπών αντικειμένου το οποίο αποκρύπτεται μερικώς, καθώς και για τον υπολογισμό ροπών εικόνας από μονοδιάστατες προβολές της. Η πρώτη τεχνική δίνει τη δυνατότητα μερικής αμεταβλητότητας στην απόκρυψη, η οποία επηρεάζει τις τιμές των ροπών σε πολύ μεγάλο βαθμό. Η δεύτερη αποτελεί μία πολύ αποδοτική μέθοδο για το χαρακτηρισμό μίας εικόνας, όταν είναι γνωστές μόνο οι προβολές της.

Ένα βασικό πρόβλημα που απαντάται στις περισσότερες εφαρμογές υπολογιστικής όρασης, και ιδιαίτερα σε αυτές της ρομποτικής όρασης, είναι η ανάκτηση πληροφορίας βάθους. Ίσως ο πιο συχνά χρησιμοποιούμενος τρόπος για την ανάκτηση της πληροφορίας αυτής είναι η χρήση ζεύγους συσκευών απεικόνισης. Η εύρεση των αντιστοιχιών σημείων μεταξύ των δύο εικόνων αποτελεί το βασικό πρόβλημα σε τέτοιου είδους τεχνικές. Εντούτοις, το πρόβλημα συχνά διαχωρίζεται σε δύο μέρη: Στην εύρεση ενός σχετικά μικρού αριθμού σημείων αντιστοιχιών στις αρχικές, μη βαθμονομημένες εικόνες, προκειμένου να πραγματοποιηθεί η βαθμονόμηση, και στην αντιστοίχιση της πλειοψηφίας των σημείων των βαθμονομημένων εικόνων. Το πρώτο μέρος αντιμετωπίζεται με μία αρχιτεκτονική υλικού, η οποία εκμεταλλεύεται τις ροπές *Zernike* ως τοπικούς περιγραφείς. Η τεχνική που προτάθηκε για τον ταχύ υπολογισμό τους, υλοποιείται για την ανάπτυξη συστήματος πραγματικού χρόνου. Για την αντιμετώπιση του δεύτερου μέρους του προβλήματος περιγράφεται μία αρχιτεκτονική, υλοποιημένη σε υλικό, η οποία παρουσιάζει τη μεγαλύτερη ταχύτητα στη βιβλιογραφία, αφού μπορεί να επεξεργαστεί σε πραγματικό χρόνο ζεύγη εικόνων μεγέθους μέχρι και 4 εκατομμυρίων στοιχείων. Παρόλο που οι παραγόμενοι χάρτες βάθους δεν παρουσιάζουν μεγάλη ακρίβεια, η εξαιρετικά υψηλή ταχύτητα επεξεργασίας καθιστά τη μέθοδο χρήσιμη σε πολλές εφαρμογές.

Επιγραμματικά, στην παρούσα διατριβή παρουσιάζονται τεχνικές για τον ταχύ και ακριβή υπολογισμό ροπών εικόνων, για την βελτιστοποίηση των χαρακτηριστικών τους, και για την εφαρμογή γραμμικών και μη μετασχηματισμών. Με βάση αυτές τις τεχνικές, υλοποιούνται μέθοδοι για την αντιμετώπιση προβλημάτων που σχετίζονται με την υπολογιστική όραση, όπως η ανάκτηση πληροφορίας βάθους, και η ανακατασκευή από προβολές.

Προτάσεις για μελλοντική έρευνα

Η μελέτη που πραγματοποιήθηκε στα πλαίσια της παρούσας διατριβής θέτει τις βάσεις για περαιτέρω έρευνα στην οπτική ανάκτηση πληροφορίας με τη χρήση των ροπών. Μερικές τέτοιες προτάσεις είναι οι παρακάτω:

- Επέκταση της τεχνικής ανάκτησης ροπών από προβολές σε τρεις διαστάσεις.
- Περαιτέρω μελέτη της επίδρασης της απόκρυψης στις ροπές, και πιθανή εξαγωγή ροπών με ανοχή σε αυτή.
- Εφαρμογή των Εξελικτικών Στρατηγικών για την εξαγωγή ροπών αμετάβλητων σε μετασχηματισμούς.
- Μελέτη τεχνικών για την βελτίωση της ακρίβειας της τεχνικής ανάκτησης χαρτών βάθους.

Παραρτήματα

Παράρτημα Α΄

Αριθμητικά στοιχεία και παραδείγματα ροπών

Τυπικές τιμές ροπών

Λόγω των συχνών προβλημάτων που εμφανίζονται κατά τον υπολογισμό και τη χρήση ροπών, εξαιτίας του δυναμικού εύρους τιμών τους, παρατίθενται κάποια αριθμητικά στοιχεία για διαφόρους τύπους ροπών που μελετήθηκαν στη Διδακτορική Διατριβή. Από την παρακάτω διερεύνηση, γίνεται σαφές ότι οι γεωμετρικές ροπές επηρεάζονται περισσότερο, ενώ για άλλους τύπους το πρόβλημα εμφανίζεται μόνο κατά τον υπολογισμό τους, αν δεν ληφθούν κατάλληλα μέτρα. Συγκεκριμένα, στην περίπτωση που χρησιμοποιηθούν οι γεωμετρικές ροπές ως ενδιάμεσο βήμα, ή αν χρησιμοποιηθούν αναδρομικές μέθοδοι με βοηθητικές μεταβλητές μεγάλου εύρους τιμών [Papakostas et al., 2006].

Όπως περιγράφηκε στην Ενότητα 1.3 η γεωμετρική ροπή τάξης p,q δίνεται από τη σχέση:

$$\mu_{p,q} = \sum_{x}^{N} \sum_{y}^{N} x^p y^q f(x,y) \qquad (\text{Α}'.1)$$

Ενώ οι κεντρικές ροπές:

$$\mu_{p,q}^{\bar{x},\bar{y}} = \sum_{x}^{N} \sum_{y}^{N} (x-\bar{x})^p (y-\bar{y})^q f(x,y) \qquad (\text{Α}'.2)$$

N	p	q	$M_{p,q}$	N	p	q	$M_{p,q}$
128	0	0	16.384	128	1	1	66.064.384
128	2	1	$5,615\,10^{9}$	128	1	3	$536,971\,10^{9}$
128	3	3	$4,364\,10^{15}$	128	10	3	$8,6909\,10^{29}$
128	10	5	$9,4183\,10^{33}$	128	10	30	$7,898\,10^{85}$
128	30	30	$3,6044\,10^{127}$	128	40	30	$3,0884\,10^{148}$
256	0	0	65,536	256	1	1	$1,0654\,10^{9}$
256	2	1	$1,8147\,10^{11}$	256	1	3	$3,4774\,10^{13}$
256	3	3	$1,135\,10^{18}$	256	10	3	$2,9334\,10^{34}$
256	10	5	$1,2766\,10^{39}$	256	10	30	$3,7790\,10^{98}$
256	30	30	$1,8836\,10^{146}$	256	40	30	$1,6877\,10^{170}$
1.024	0	0	1,048,576	1.024	1	1	$2,7434\,10^{11}$
1.024	2	1	$1,8719\,10^{14}$	1.024	1	3	$1,4369\,10^{17}$
1.024	3	3	$7,5263\,10^{22}$	1.024	10	3	$3,2201\,10^{43}$
1.024	10	5	$2,2488\,10^{49}$	1.024	10	30	$7,7789\,10^{123}$
1.024	30	30	$4,3923\,10^{183}$	1.024	40	30	$4,1892\,10^{213}$

Πίνακας Α′.1: Τιμές γεωμετρικών ροπών για λευκές εικόνες ($f(x,y) = 1$) μεγέθους $N \times N$

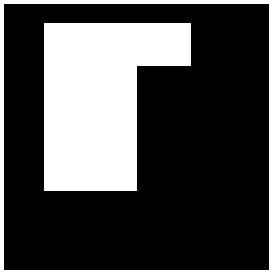
Αρχικά, ας υποθέσουμε μία λευκή εικόνα μεγέθους 128×128 στοιχείων. Επίσης, ας υποθέσουμε ότι $p = 5$ και $q = 0$. Παρόλο που ούτε το μέγεθος της εικόνας μπορεί να θεωρηθεί μεγάλο, ούτε η τάξη της ροπής, η τιμή που προκύπτει για την ροπή (5,0) είναι:

$$\mu_{5,0} = \sum_{x}^{N} \sum_{y}^{N} x^5 y^0 1 = 9.164 \cdot 10^{13} \tag{Α′.3}$$

Για ροπή 10ης τάξης, το αποτέλεσμα γίνεται $1.6839 \cdot 10^{24}$. Στον Πίνακα Α′.1 δίνονται τιμές για λευκές εικόνες διαφόρων μεγεθών.

Παρόλο που οι τιμές αυτές μπορούν να αναπαρασταθούν μέσω τυπικής αναπαράστασης πλωτού σημείου, συνήθως χρησιμοποιούνται για την εξαγωγή άλλων αμετάβλητων χαρακτηριστικών, όπως είναι οι σταθερές του *Hu*, οι ροπές *Zernike* ή κάποιοι τύποι σταθερών αμετάβλητων σε άλλους μετασχηματισμούς [Rahtu et al., 2006]. Η τελική τιμή των παραπάνω περιγραφέων παρουσιάζει μικρό δυναμικό εύρος, αφού προκύπτει από αφαιρέσεις γεωμετρικών ροπών παραπλήσιου βαθμού. Αυτές οι αφαιρέσεις εισάγουν σημαντικά σφάλματα, τα οποία μπορεί να καταστήσουν τα τελικά αποτελέσματα μη πρακτικά.

Η ίδια διερεύνηση δε μπορεί να πραγματοποιηθεί για κεντρικές ροπές περιττών τάξεων, αφού για $f(x,y) = c$ δίνουν μηδενικό αποτέλεσμα.



Εικόνα παραδείγματος

Τάξη		Γεωμετρικές ροπές	Κεντρικές-γεωμετρικές ροπές	Ροπές *Zernike*	Ροπές *Chebyshev*
0	0	4.191	4.191	4.191	32,742
1	0	191.214	$-3,3083\,10^{-11}$	-	-16,312
0	1	188.979	$-4,5475\,10^{-11}$	-	-15,8398
1	1	8.116.416	$-5,0573\,10^{5}$	$-725,66+741,35i$	4,998
2	0	10.354.695	$1,2440\,10^{6}$	-1.018,9	-20,083
2	1	11.343.520	$2,44\,10^{6}$	-	11,115
2	2	$2,2908\,10^{10}$	$7,8474\,10^{8}$	$-287,82-351,77i$	5,896
3	0	$6,088\,10^{8}$	$1,0024\,10^{7}$	-	20,404
3	1	$2,260\,10^{10}$	$-5,2827\,10^{8}$	$1.152,3-343,39i$	-5,599
3	2	$1,173\,10^{12}$	$8,5742\,10^{9}$	-	-13,23
3	3	$7,460\,10^{13}$	$-4,4778\,10^{11}$	$232,24-162,16i$	3,969

Πίνακας Α′.2: Παραδείγματα τιμών ροπών

Παράρτημα Β′

Φίλτρα πόλων

Παραγοντικά πολυώνυμα

Ένα αύξων παραγοντικό πολυώνυμο βαθμού m ορίζεται ως:

$$
\begin{aligned}
P_m(x) &= \prod_{k=0}^{m-1}(x+k) \\
&= x(x+1)\dots(x+m-1) \qquad \text{(Β′.1)}
\end{aligned}
$$

και συμβολίζεται $x^{\overline{m}}$. Αντίστοιχα, ένα φθίνον παραγοντικό πολυώνυμο, ορίζεται ως:

$$
\begin{aligned}
P_m(x) &= \prod_{k=0}^{m-1}(x-k) \\
&= x(x+1)\dots(x-m+1) \qquad \text{(Β′.2)}
\end{aligned}
$$

και συμβολίζεται $x^{\underline{m}}$. Τα πολυώνυμα αυτά παρουσιάζουν κάποιες πολύ χρήσιμες ιδιότητες στον διακριτό λογισμό. Είναι προφανές ότι κάθε παραγοντικό πολυώνυμο μπορεί να εκφραστεί σε κανονική μορφή πολυωνύμου. Ισχύει, όμως, και το αντίστροφο. Δηλαδή, κάθε δύναμη μπορεί να εκφραστεί μέσω παραγοντικών πολυωνύμων. π.χ.:

$$
\begin{aligned}
x^0 &= x^{\underline{0}} \\
x^1 &= x^{\underline{1}}
\end{aligned}
$$

$$
\begin{aligned}
x^2 &= x^{\underline{2}} - x^{\underline{1}} \\
x^3 &= x^{\underline{3}} - 3x^{\underline{2}} + x^{\underline{1}}
\end{aligned}
\tag{B′.3}
$$

Οι σχέσεις που εκφράζει τις ακέραιες δυνάμεις με τα παραγοντικά πολυώνυμα και αντίστροφα είναι οι:

$$
x^{\underline{m}} = \sum_{n=0}^{m} s_m^n x^n \tag{B′.4}
$$

$$
x^m = \sum_{n=0}^{m} \dot{s}_m^n x^{\underline{n}} \tag{B′.5}
$$

Οι αριθμοί s_m^n και $\dot{s}_m^n$ ονομάζονται αριθμοί του *Stirling* πρώτου και δεύτερου τύπου αντίστοιχα. Όμοιες εκφράσεις ισχύουν και για αύξοντα παραγοντικά πολυώνυμα.

Συσσωρευτικές ροπές

Ως συσσωρευτική ροπή τάξης m ενός σήματος $f(x)$, μπορεί να ορισθεί η εξής παράσταση:

$$
M_m = \sum_x f(x) x^{(\overline{m})} \tag{B′.6}
$$

Ομοίως, για σήματα δύο διαστάσεων παίρνει την παρακάτω μορφή:

$$
M_{m,n} = \sum_x \sum_y f(x,y) x^{(\overline{m})} y^{(\overline{n})} \tag{B′.7}
$$

Στην προηγούμενη Ενότητα σημειώθηκε ότι οι δυνάμεις ακεραίων μπορούν να εκφραστούν μέσω παραγοντικών πολυωνύμων. Επομένως, λόγω της γραμμικότητας των ροπών, και οι γεωμετρικές ροπές μπορούν να εκφραστούν μέσω των ροπών συσσώρευσης.

Θεώρημα 1 *Για κάθε σήμα $f(x)$, η γεωμετρική ροπή P οποιασδήποτε τάξης m μπορεί να εκφραστεί ως γραμμικός συνδυασμός ροπών συσσώρευσης τάξης μέχρι και m.*

Απόδειξη Η γεωμετρική ροπή τάξης m ενός σήματος $f : \{0, \ldots, N\} \to R$ δίνεται από τη σχέση:

$$
P_m = \sum_{x=0}^{N} f(x) x^m \tag{B′.8}
$$

Από την Β΄.4 ισχύει:

$$P_m = \sum_{x=0}^{N} f(x) \sum_{n=0}^{m} \dot{s}_m^n x^n \tag{Β΄.9}$$

Άρα

$$P_m = \sum_{n=0}^{m} \dot{s}_m^n \sum_{x=0}^{N} f(x) x^n \tag{Β΄.10}$$

Και τελικά

$$P_m = \sum_{n=0}^{m} \dot{s}_m^n M_n \tag{Β΄.11}$$

Με M_n τη συσσωρευτική ροπή τάξης n.

Φίλτρα πόλων και παραγοντικά πολυώνυμα

Ένα φίλτρο ενός πόλου, $(z-1)^{-1}$ αποτελεί ουσιαστικά ένα συσσωρευτή. Η απόκρισή του στο πεδίο του χρόνου δίνεται, δηλαδή, ως το άθροισμα των προηγουμένων εισόδων. Ένα φίλτρο n πόλων, αντίστοιχα, περιγράφεται στο πεδίο του μετασχηματισμού z ως $(z-1)^{-n}$.

Στο πεδίο του χρόνου, ένα φίλτρο n πόλων παρουσιάζει στην έξοδό του τη συσσωρευτική ροπή τάξης n. Αυτή η ιδιότητα τους καθιστά τα φίλτρα αυτά πολύ χρήσιμα στην ανάλυση εικόνων βάσει των χαρακτηριστικών ροπών.

Βιβλιογραφία

N. Atzpadin, P. Kauff, and O. Schreer. Stereo analysis by hybrid recursive matching for real-time immersive video conferencing. *IEEE Transactions on Circuits and Systems for Video Technology*, 14(3):321–334, March 2004.

S. B. Belkasim, M. Ahmadi, and M. Shridar. Efficient algorithm for fast computation of Zernike moments. *Journal of the Franklin Institute*, 333: 577–581, 1996.

T. Back and H. Schwefel. An overview of evolutionary algorithms for parameter optimization. *Evolutionary Computation*, 1(1):1–23, January 1993.

K. P. H. Berthold. *Robot Vision*. The MIT Press, McGraw Book Company, Ninth printing, 1993.

S. Bezdidko. Optimization of optical systems using orthogonal polynomials. *Optics and Spectroscopy*, 48:670, October 1980.

C. Bobda and N. Steenbock. Singular value decomposition on distributed reconfigurable systems. In *IEEE International Workshop on Rapid System Prototyping*, pages 38–43, 2001.

M. Bober. MPEG-7 Visual Shape Descriptors. *IEEE Transactions on Circuits and Systems for Video Technology*, 11(6):717, 2001.

R. P. Brent and F. T. Luk. The solution of singular-value and eigen-value problems on multiprocessor arrays. *SIAM J. Sci. Star. Comput.*, 6:69–84, 1985.

D. Burschka and G. Hager. Scene classification from dense disparity maps in indoor environments. In *Proc. of the 17th Intl. Conf. Pattern Recognition*, volume 3, pages 708–712, 2002.

A. Chipperfield, P. Fleming, H. Pohlheim, and C. Fonseca. The MATLAB genetic algorithm toolbox. *IEE Colloqium on Applied Control Techniques Using MATLAB*, Digest 1995(014), 1995.

K. Chung and P. Chen. An efficient algorithm for computing moments on a block representation of a grey-scale image. *Pattern Recognition*, 38(12): 2578–2586, December 2005.

J. Corso, D. Burschka, and G. Hager. Direct plane tracking in stereo images for mobile navigation. In *IEEE International Conference on Robotics and Automation, 2003 ICRA'03.*, volume 1, 2003.

S. Dudani, K. Breeding, and R. McGhee. Aircraft identification by moment invariants. *IEEE Transactions on Computers*, 26:39–46, January 1977.

A. Eiben and J. Smith. *Introduction to Evolutionary Computing*. Springer-Verlag, Berlin Heidelberg, 2003.

J. Flusser. Fast calculation of geometric moments of binary images. In *22nd OAGM'98 Workshop Pattern Recognition Medical Computer Vision*, pages 265–274, Illmitz, Austria, 1998.

W. Gautschi. *Orthogonal Polynomials: Computation and Approximation*. Calendon Press, Oxford, 2004.

J. M. Geusebroek, G. J. Burghouts, and A. W. M. Smeulders. The Amsterdam library of object images. *Int. J. Comput. Vision*, 61(1):103–112, Jan 2005.

D. Goldberg. *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison Wesley Publishing Company, 1989.

P. Gonidis. *Αρχιτεκτονικές υλικού πραγματικού χρόνου για την εξαγωγή αντιστοιχιών σημείων σε στερεοσκοπικά ζεύγη πληροφοριών*. Δημοκρίτειο Πανεπιστήμιο Θράκης, Ξάνθη, 2007.

P. Gonidis, L. Kotoulas, and I. Andreadis. A new hardware module for stereo matching using Zernike moments. *The Third International Conference on Autonomic and Autonomous Systems, ICAS 2007, accepted for presentation*, 2007.

R. Gonzalez and R. Woods. *Digital image processing*. Addison-Wesley Reading, Mass, 1987.

M. Gruber and K. Hsu. Moment-based image normalization with high noise tolerance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(2):136–139, February 1997.

J. Gu, H. Shu, C. Toumoulin, and L. Luo. A novel algorithm for fast computation of Zernike moments. *Pattern Recognition*, 35(12):2905–2911, December 2002.

N. Hansen and S. Kern. Evaluating the CMA evolution strategy on multimodal test functions. In *Eighth International Conference on Parallel Problem Solving from Nature PPSN VIII*, pages 282–291, 2004.

N. Hansen and A. Ostermeier. Completely derandomized self-adaptation in Evolution Strategies. *Evolutionary Computation*, 9(2):159–195, February 2001.

N. Hansen, S. Mueller, and P. Koumoutsakos. Reducing the time complexity of the derandomized evolution strategy with covariance matrix adaptation (CMA-ES). *Evolutionary Computation*, 11(1):1–18, January 2003.

C. Harris and M. Stephens. A combined corner and edge detector. In *Fourth Alvey Vision Conference*, pages 148–151, 1988.

M. R. Hestenes. Inversion of matrices by biorthogonalization and related results. *J. Soc. Indust. Appl. Math.*, 20(1):51–90, Nov. 1958.

K. Hoffman and R. Kunze. *Linear Algebra, 2nd ed.* Prentice Hall, Englewood Cliffs, NJ, 1971.

M. Hu. Visual pattern recognition by moment invariants. *IRE Transactions on Information Theory*, 8(2):179–187, February 1962.

S. Hwang and W. Kim. Fast and efficient method for computing ART. *IEEE Transactions on Image Processing*, 15(1):112–117, 2006.

C. Kan and M. Srinath. Invariant character recognition with Zernike and orthogonal fourier-mellin moments. *Pattern Recognition*, 35:143–154, November 2002.

W. Kim and Y. Kim. A new region-based shape descriptor: The ART (Angular Radial Transform) Descriptor. *ISO/IEC JTC1/SC29/WG11/ M*, 2001.

L. Kotoulas. *Αρχιτεκτονικές υλικού πραγματικού χρόνου για την εξαγωγή χαρακτηριστικών σχήματος και εφαρμογές στην οπτική ανάκτηση πληροφοριών*. Δημοκρίτειο Πανεπιστήμιο Θράκης, Ξάνθη, 2004.

L. Kotoulas and I. Andreadis. Fast computation of Chebyshev moments. *IEEE Transactions on Circuits and Systems for Video Technology*, 16(7): 884–888, July 2006.

L. Kotoulas and I. Andreadis. Efficient hardware architectures for computation of image moments. *Real Time Imaging*, 10(6):371–378, December 2004.

L. Kotoulas and I. Andreadis. Real-time computation of Zernike moments. *IEEE Transactions on Circuits and Systems for Video Technology*, 15(6): 801–809, June 2005.

P. Koumoutsakos, J. Freund, and D. Parekh. Evolution strategies for parameter optimization in jet flow control. In *Proceedings of the Summer Program, Center for Turbulence Research*, pages 121–132, 1998.

C. Lanczos. *Applied Analysis*. Dover Publications, 1988.

A. M. Legendre. Sur l'attraction des spheroides. *Mem. Math. et Phys. presentes α l'Ac. r. des. sc. par divers savants*, 10, 1785.

B. Li. A new computation of geometric moments. *Pattern Recognition*, 26 (1):109–113, January 1993.

S. Liao and M. Pawlak. On image analysis by moments. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(3):254–266, March 1996.

S. Liao and M. Pawlak. On the accuracy of Zernike moments for image analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(12):1358–1364, December 1998.

W. Lin and S. Wang. A note on the calculation of moments. *Pattern Recognition Letters*, 15(11):1065–1070, November 1994.

J. Mason. *Chebyshev Polynomials*. Chapman & Hall/CRC, 2003.

H. Moehlenbein and D. Schlierkamp-Voosen. Predictive models for the breeder genetic algorithm. *Evolutionary Computation*, 1(1):25–49, 1993.

R. Mukundan and K. Ramakrishnan. Computation of Legendre and Zernike moments. *Pattern Recognition*, 28(9):1433–1442, September 1995.

R. Mukundan, S. Ong, and P. Lee. Image analysis by Tchebichef moments. *IEEE Transactions on Image Processing*, 10(9):1357–1364, September 2001.

M. V. Online. www.machinevisiononline.org. web resource, 2007.

G. Papakostas, Y. Boutalis, C. Papaodysseus, and D. Fragoulis. Numerical error analysis in Zernike moments computation. *Image and Vision Computing*, 24(9):960–969, September 2006.

S. Park and R. Schowengerdt. Image sampling, reconstruction, and the effect of sample-scene phasing. *Applied Optics*, 21(17):3142–3151, April 1982.

W. Phillips. A new fast algorithm for moment computation. *Pattern recognition*, 26(11):1619–1621, 1993.

M. Pilu. A direct method for stereo correspondence based on singular value decomposition. In *IEEE CVPR*, pages 261–266, 1997.

J. Prewitt. *Object enhancement and extraction*. B. S. Lipkin and A. Rosenfeld, editors, Picture Processing and Psychopictories, Academic Press, 1970.

R. Prokop and A. Reeves. A survey of moment-based techniques for unoccluded object representation and recognition. *CVGIP: Graphical Models and Image Processing*, 54(5):438–460, 1992.

E. Rahtu, M. Salo, J. Heikkil, and J. Flusser. Generalized affine moment invariants for object recogn. In *Pattern Recognition, 2006. ICPR 2006. 18th International Conference on*, volume 2, pages 634–637, 20-24 Aug. 2006. doi: 10.1109/ICPR.2006.599.

L. Reichel, G. Ammar, and W. Gragg. Discrete Least Squares Approximation by Trigonometric Polynomials. *Mathematics of Computation*, 57(195): 273–289, July 1991.

N. Rouze, V. Soon, and G. Hutchins. On the connection between the Zernike moments and Radon transform of an image. *Pattern Recognition Letters*, 27(6):636–642, April 2006.

H. Rutishauser. *The Jacobi method for real symmetric matrices*. J. H. Wilkinsonand C.Reinsch, eds., Handbook for Automatic Computation, 1971.

F. Scheid. *Schaum's Outline of Theory and Problems of Numerical Analysis*. McGraw Hill, New York, 1968.

H. Schwefel. *Numerical optimization of computer models*. John Wiley & Sons, Inc. New York, NY, USA, 1981.

G. Scott and H. Longuet-Higgins. An algorithm for associating the features of two patterns. In *Proceedings of the Royal Statistical Society of London*, volume B244, pages 21–26, April 1991.

D. Shen and H. Horace. Generalized affine invariant image normalization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(5): 431–440, May 1997.

T. Shen, D. Lun, and W. Siu. On the efficient computation of 2-d image moments using the discrete Radon-transform. *Pattern Recognition*, 31(2): 115–120, February 1998.

G. Sirakoulis. A TCAD system for VLSI implementation of the CVD process using VHDL. *Integration, the VLSI Journal*, 37(1):63–81, January 2004.

G. Sirakoulis, I. Karafyllidis, and A. Thanailakis. A CAD system for the construction and VLSI implementation of Cellular Automata algorithms using VHDL. *Microprocessors and Microsystems*, 27(8):381–396, August 2003.

Z. Sun, G. Bebis, and R. Miller. On-road vehicle detection using optical sensors: a review. In *Proc 7th Intl. IEEE Conf. on Intelligent Transportation Systems*, pages 585–590, 2004.

M. Teague. Image analysis via the general theory of moments. *JOSA*, 70 (8):920–930, August 1980.

O. Veksler. Fast variable window for stereo correspondence using integral images. *Computer Vision and Pattern Recognition, 2003. Proceedings. IEEE Computer Society Conference on*, 1, January 2003.

G. Wang and S. Wang. Recursive computation of Tchebichef moment and its inverse transform. *Pattern Recognition*, 39(1):47–56, January 2006.

J. Woodfill and B. Von Herzen. Real-Time Stereo Vision on the PARTS Reconfigurable Computer. In *IEEE Symposium on FPGAs for Custom Computing Machines*, volume 4. IEEE Computer Society Press, 1997.

P. Yap, R. Paramesran, and S. Ong. Image analysis by Krawtchouk moments. *IEEE Transactions on Image Processing*, 12(11):1367–1377, November 2003.

K. Yoon and I. Kweon. Adaptive Support-Weight Approach for Correspondence Search. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 4(28):650–656, April 2006.

F. Zernike. Beugungstheorie des Schneidenverfahrens und seiner verbesserten Form, der Phasenkontrastmethode. *Physica*, 1:689–704, 1934.

Z. Zhang, R. Deriche, O. Faugeras, and Q. Luong. A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry. *Artificial Intelligence Journal*, 78:87–119, October 1995.

F. Zhao. Image matching based on singular value decomposition. In *PCM 2004*, pages 119–126, 2004.

Δημοσιεύσεις βασιζόμενες στην παρούσα έρευνα

Δημοσιεύσεις σε διεθνή επιστημονικά περιοδικά με κριτές

1. L. Kotoulas and I. Andreadis. Accurate calculation of image moments. *IEEE Transactions on Image Processing*, accepted for publication.

2. L. Kotoulas and I. Andreadis. Fast computation of Chebyshev moments. *IEEE Transactions on Circuits and Systems for Video Technology*, 16(7):884–888, July 2006.

3. L. Kotoulas and I. Andreadis. Evolutionary enhanced image moment descriptors. *Pattern Recognition*, submitted for publication.

4. L. Kotoulas and I. Andreadis. Fast moment generating architectures. *IEEE Transactions on Circuits and Systems for Video Technology*, submitted for publication.

5. L. Kotoulas and I. Andreadis. Fast computation of ART. *IEEE Transactions on Circuits and Systems for Video Technology*, submitted for publication.

6. L. Kotoulas, G. Sirakoulis, I. Andreadis and A. Gasteratos. A hardware architecture for real time extraction of disparity maps from large images. *IEEE Transactions on Instrumentation and Measurements* , submitted for publication.

7. C. Georgoulas, L. Kotoulas, G. Sirakoulis, I. Andreadis and A. Gasteratos. Real time disparity map computation module. *Microprocessors and Microsystems*, submitted for publication.

Δημοσιεύσεις σε διεθνή επιστημονικά συνέδρια με κριτές

1. L. Kotoulas and I. Andreadis. Discrete orthogonal moments in image analysis *The Fourth IASTED International Conference on Signal Processing, Pattern Recognition, and Applications*, Innsbruck, Austria, February 2007, pp. 18–21.

2. L. Kotoulas and I. Andreadis. Image analysis using moments. *5th conference of Technology and Automation*, Thessalonica, Greece, October 2005, pp. 360–364.

3. L. Kotoulas, C. Georgoulas, A. Gasteratos, G. Ch. Sirakoulis and I. Andreadis. A novel three-stage algorithm for accurate disparity maps. *5th IASTED Conf., VIIP*, Benidorm, Spain, September 2005, pp. 355–359.

4. L. Kotoulas, A. Gasteratos, G. Ch. Sirakoulis, C. Georgoulas and I. Andreadis. A novel three-stage algorithm for accurate disparity maps. *EOS Conference of Machine Vision and Industrial Imaging*, Munich, Germany, June 2005, pp. 13–14.

Ευρετήριο