



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ

ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΕΠΙΚΟΙΝΩΝΙΩΝ, ΗΛΕΚΤΡΟΝΙΚΗΣ & ΣΥΣΤΗΜΑΤΩΝ

ΠΛΗΡΟΦΟΡΙΚΗΣ

**Προηγμένοι αλγόριθμοι ψηφιακής επεξεργασίας εικόνας και
ταυτοποίησης γραφεία. Μελέτη του λάθους πεπερασμένης
ακρίβειας σε σημαντικούς αλγορίθμους και ευσταθειοποίηση
αυτών**

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

Κωνσταντίνος Ι. Χαλάτσης

Αθήνα, Φεβρουάριος 2021



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ

ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΕΠΙΚΟΙΝΩΝΙΩΝ, ΗΛΕΚΤΡΟΝΙΚΗΣ & ΣΥΣΤΗΜΑΤΩΝ

ΠΛΗΡΟΦΟΡΙΚΗΣ

**Προηγμένοι αλγόριθμοι ψηφιακής επεξεργασίας εικόνας και
ταυτοποίησης γραφέα. Μελέτη του λάθους πεπερασμένης
ακρίβειας σε σημαντικούς αλγορίθμους και ευσταθειοποίηση
αυτών**

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

Κωνσταντίνος Ι. Χαλάτσης

Συμβουλευτική Επιτροπή : Κωνσταντίνος Παπαοδυσσεύς

Ηλίας Κουκούτσης

Ανδρέας Σταφυλοπάτης

Εγκρίθηκε από την επταμελή εξεταστική επιτροπή την 10^η Φεβρουαρίου 2021

.....

Κωνσταντίνος Παπαοδυσσεύς,
Καθηγητής ΣΗΜΜΥ ΕΜΠ

.....

Γεώργιος Φικιώρης,
Καθηγητής ΣΗΜΜΥ ΕΜΠ

.....

Ηλίας Κουκούτσης,
Επ. Καθηγητής ΣΗΜΜΥ ΕΜΠ

.....

Βασίλειος Παπανικολάου,
Καθηγητής ΣΕΜΦΕ ΕΜΠ

.....

Μιχάλης Παναγόπουλος,
Επ. Καθηγητής, Τμήμα Τεχνών Ήχου
και Εικόνας, Ιόνιο Πανεπιστήμιο

.....

Ανδρέας Σταφυλοπάτης,
Καθηγητής ΣΗΜΜΥ ΕΜΠ

.....

Παναγιώτης Τσανάκας,
Καθηγητής ΣΗΜΜΥ ΕΜΠ

Αθήνα, Φεβρουάριος 2021

.....

Κωνσταντίνος Ι. Χαλάτσης

Διδάκτωρ Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών Ε.Μ.Π.

Copyright © Κωνσταντίνος Ι. Χαλάτσης, 2021.

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς τον συγγραφέα.

Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν τον συγγραφέα και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις επίσημες θέσεις του Εθνικού Μετσόβιου Πολυτεχνείου.

Περίληψη

Στη παρούσα εργασία αναπτύσσεται μια νέα μεθοδολογία για τον υπολογισμό των συντελεστών Zernike (Zernike moments), καθώς και μια νέα μέθοδος μελέτης, παρακολούθησης και αντιμετώπισης υπολογιστικών προβλημάτων που οφείλονται στο γεγονός ότι οι υπολογιστές εκτελούν πράξεις με πεπερασμένο μήκος λέξης. Τέλος παρουσιάζεται η υπολογιστική βελτίωση μιας υπάρχουσας μεθοδολογίας αναγνώρισης προτύπων. Ειδικότερα:

Αρχικά αναπτύσσεται μια νέα μεθοδολογία που μελετά τη γένεση, συσσώρευση και διάδοση του σφάλματος πεπερασμένης ακρίβειας στις υπολογιστικές μηχανές, όταν επιτελείται η πράξη της αφαίρεσης. Η ανάγκη μελέτης αυτού του λάθους προέκυψε από το γεγονός ότι πολλοί σύγχρονοι αλγόριθμοι, όπως αρκετοί αλγόριθμοι υπολογισμού των συντελεστών Zernike, υποφέρουν από το σφάλμα πεπερασμένης ακρίβειας. Στη παρούσα εργασία αναπτύσσεται μια νέα μέθοδος επακριβούς εντοπισμού, παρακολούθησης και πρόβλεψης αυτού του σφάλματος, το οποίο μπορεί να γεννηθεί σε μία ή και περισσότερες διαδοχικές αφαιρέσεις. Η μελέτη αυτή είναι ουσιωδώς ανεξάρτητη των εκθετών των ποσοτήτων που συμμετέχουν σε μία αφαίρεση, εξαρτάται δε κατά βάση από τις mantissae των ποσοτήτων αυτών. Ακολούθως εδραιώνεται ότι το σφάλμα πεπερασμένης ακρίβειας είναι ανεξάρτητο του μήκους λέξης. Στη συνέχεια παρουσιάζεται μια μεθοδολογία παρακολούθησης και διόρθωσης του λάθους που οφείλεται στο σφάλμα πεπερασμένης ακρίβειας, στις περιπτώσεις των φημισμένων αλγορίθμων Bernoulli και Baker's Chaotic maps.

Στο επόμενο στάδιο της παρούσης έρευνας αναπτύσσεται μια νέα μέθοδος υπολογισμού των συντελεστών Zernike. Αυτοί οι συντελεστές περιέχουν σημαντική πληροφορία για την εικόνα, όπως για το περίγραμμα και τη λεπτομέρεια. Εξαιτίας αυτού, χρησιμοποιούνται κατά κόρον σε πολλούς τομείς, όπως στην εξαγωγή βιομετρικής πληροφορίας από εικόνες, αναγνώρισης οπτικών χαρακτηριστικών, συμπίεση και κατάτμηση μιας εικόνας κτλ. Όμως, ο μετασχηματισμός μιας εικόνας σε συντελεστές Zernike προσθέτει και σφάλματα που εμποδίζουν την σωστή ανακατασκευή αυτής, καθώς και τη σωστή καταγραφή της σχετικής πληροφορίας. Η χρήση πολικών pixels (polar pixels) αίρει αρκετά από αυτά τα προβλήματα. Προς σε αυτή τη κατεύθυνση, στη παρούσα εργασία παρουσιάζεται μια νέα αναδρομική μεθοδολογία υπολογισμού των Zernike moments πάνω στα πολικά pixels. Με τη χρήση αυτής τη μεθόδου, έγινε η ανακατασκευή πολλών εικόνων, χωρίς αυτές να υποφέρουν από αριθμητικά σφάλματα, όπως το γεωμετρικό και το σφάλμα ολοκλήρωσης, τα οποία συνιστούν πολύ σημαντικό πρόβλημα στον ορθό υπολογισμό των συντελεστών Zernike με τις προϋπάρχουσες μεθόδους. Επιπροσθέτως, αυτή η μεθοδολογία είναι ανεξάρτητη της στροφής, ενώ είναι αρκετά αποδοτική σε χρόνο εκτέλεσης, σε σύγκριση με τις υπάρχουσες τεχνικές. Επιπλέον, η νέα αυτή αναδρομική μέθοδος δεν υποφέρει από λάθη που οφείλονται στο σφάλμα πεπερασμένης ακρίβειας.

Στο τελευταίο κομμάτι της παρούσης εργασίας, παρουσιάζεται μια βελτιωμένη υπολογιστική εκδοχή ενός υπάρχοντα αλγορίθμου ταυτοποίησης γραφεία. Η βελτίωση αυτή αφορά τη ταχύτητα εκτέλεσης του αλγορίθμου και την ακρίβεια των παραγόμενων αποτελεσμάτων. Ακολούθως, γίνεται μια πρώτη προσπάθεια ταυτοποίησης κειμένων, για τα οποία εικάζεται ότι ανήκουν στον Ρήγα Φεραίο Βελεστινλή.

Λέξεις κλειδιά: Αλγόριθμοι Υπολογισμού των Συντελεστών Zernike moments, Polar pixels, Γένεση Λάθους Πεπερασμένης Ακρίβειας, Συσσώρευση Λάθους Πεπερασμένης Ακρίβειας, Ακριβής Προσδιορισμός Λάθους Πεπερασμένης Ακρίβειας στη πράξη της Αφαίρεσης, Αυτόματη Αναγνώριση Γραφεία.

Abstract

In the present thesis, a novel approach for the robust and efficient computation of the Zernike moments / coefficients is introduced, together with a new general methodology for the evaluation of the numerical error due to the finite word length. Moreover, a computational improvement of an existing method for pattern recognition is presented. More specifically:

Initially, in the present work, it is established a new approach for the exact evaluation, accumulation and tracking of the finite precision error associated with the operation of subtraction in an arbitrary computing machine. The necessity of this study has been arisen from the fact that quite many important and frequently employed algorithms seriously suffer from this type of numerical error, including various algorithms performing Zernike moments computation. Here, a new methodology is introduced, which allows for the exact tracking and prediction of the finite precision error generated in one or many successive operation of subtractions. This methodology is practically independent of the order of the quantities involved in the operations of subtraction in the considered algorithm. Furthermore, it is established that the numerical error due to the finite word length does not depend on the exact value of the employed word length. Next, a novel method is proposed for correcting the finite precision error appearing in the well - known Bernoulli and Baker's chaotic maps.

As a next step, a set of novel recursive relations for the evaluation of the Zernike moments is established. These moments incorporate important information concerning an image and, in particular, the one associated with the contour and details of the objects existing in the image. For these reasons, the Zernike coefficients are widely used in biometrics, pattern recognition, image segmentation and / or compression, etc. However, the representation of an image via Zernike moments includes various considerable numerical errors, such as the geometric one and the integration error; these numerical errors highly deteriorate the quality of the reproduced image by means of the Zernike coefficients. Employing polar pixels lifts most of the aforementioned numerical problems. Hence, towards this direction, in the present thesis a new ensemble of iterative relations is introduced and proved for the efficient computation of the Zernike coefficients, on the basis of polar pixels. Application of these recursive formulae in the evaluation of the Zernike moments offers excellent quality of the corresponding introduced images, since they lift the geometric and integration errors. In addition, the presented method is independent of the rotation of the image and quite efficient, while it practically does not suffer from finite precision error.

Finally, in the last section of the present work, it is established a new computational improvement of an already existing method for the identification of the writer of important historical documents. The associated, developed software system has been applied to a first study of the important open question, concerning the writer of the documents "Saganaki tis trelas" and "Fysikis apantisma". The open issue is the eventuality that the great Rigas Feraios has written both documents.

Keywords: Algorithms for the computation of Zernike moments, polar pixels, Finite Precision Error Generation, Finite Precision Error Accumulation, Exact Tracking of Finite Precision Error in the operation of subtraction, Automatic Writer Identification.

Ευχαριστίες

Σε αυτό το σημείο, θα ήθελα να ευχαριστήσω όλους όσους με στήριξαν στην προσπάθειά για ολοκλήρωση της διδακτορικής μου διατριβής και συμμετείχαν με οποιοδήποτε τρόπο στην ολοκλήρωσή της.

Αρχικά θα ήθελα να ευχαριστήσω βαθύτατα τον ουσιαστικό επιβλέποντα της διατριβής αυτής, Καθηγητή της σχολής ΣΗΜΜΥ Ε.Μ.Π., κύριο Κωνσταντίνο Παπαοδυσσέα για την αμέριστη καθοδήγηση που μου παρείχε καθ' όλη τη διάρκεια της ερευνητικής μου πορείας, η οποία και οδήγησε στο αποτέλεσμα που περιγράφεται στην παρούσα διατριβή. Οι βαθιές γνώσεις του, η εμπειρία του και η επιστημονική του κατάρτιση πάνω στο αντικείμενο του Ηλεκτρολόγου Μηχανικού και Μηχανικού Υπολογιστών αλλά και σε όλα τα αντικείμενα της επιστημονικής έρευνας στα οποία έχει εμβαθύνει, είχαν ως αποτέλεσμα στην επίλυση κάθε προβλήματος που προέκυπτε. Επίσης, θα ήθελα να τον ευχαριστήσω για την αστείρευτη διάθεσή του για έρευνα, καθώς και για την προθυμία του να ασχοληθεί ατελείωτα βράδια με οποιοδήποτε πρόβλημα, μέχρι την επίλυσή του. Επιπροσθέτως, πέρα της διάθεσης του για έρευνα, ο ζήλος που έδειχνε για μεταλαμπάδευση της γνώσης του στους συνεργάτες του και στους φοιτητές τον καθιστούν υπόδειγμα καθηγητή και ερευνητή. Χωρίς τη δική του καθοδήγηση, θεωρώ πως η ολοκλήρωση της διατριβής μου θα ήταν αδύνατη. Για όλα τα ανωτέρω, για την ευκαιρία που μου δόθηκε να συμμετέχω επικουρικά στην εκπαιδευτική διαδικασία, για τη εκπαίδευση μου από αυτόν πάνω σε πολλούς επιστημονικούς τομείς, αλλά και για την συναναστροφή και τις συζητήσεις που είχαμε σε προσωπικό επίπεδο όλα αυτά τα χρόνια, τον ευχαριστώ βαθύτατα.

Θα ήθελα επίσης να ευχαριστήσω τον Επίκουρο Καθηγητή της ΣΗΜΜΥ Ε.Μ.Π. κύριο Ηλία Κουκούτση, αρχικά για την επιλογή μου ως υποψήφιο διδάκτορα, αλλά και για την καθοδήγησή του και τη στάση του απέναντί μου καθ' όλη τη διάρκεια της παρούσας διατριβής. Επιπροσθέτως, θα ήθελα να ευχαριστώ τον κύριο Κουκούτση για την άριστη συνεργασία που είχαμε στα πλαίσια της διδασκαλίας μαθημάτων της Σχολής Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών του Ε.Μ.Π. Είναι ένας καθηγητής που δίνει πάντα ιδιαίτερη σημασία στην εκπαιδευτική διαδικασία και στις ανάγκες των φοιτητών. Τέλος τον ευχαριστώ για τις εποικοδομητικές συζητήσεις που είχαμε όλα αυτά τα χρόνια. Η εμπειρία που αποκόμισα από τη συνεργασία μας είναι μοναδική.

Ευχαριστώ επίσης θερμά τον Καθηγητή Ε.Μ.Π. και μέλος της τριμελούς συμβουλευτικής επιτροπής μου κύριο Ανδρέα Σταφυλοπάτη, τον Καθηγητή ΣΕΜΦΕ Ε.Μ.Π. κύριο Βασίλη Παπανικολάου, τον Καθηγητή

ΣΗΜΜΥ Ε.Μ.Π. κύριο Γεώργιο Φικιώρη, τον Καθηγητή ΣΗΜΜΥ Ε.Μ.Π. κύριο Παναγιώτη Τσανάκα και τον Επίκουρο Καθηγητή του Ιονίου Πανεπιστημίου κύριο Μιχάλη Παναγόπουλο, διότι με τίμησαν με τη συμμετοχή τους στην Επταμελή Επιτροπή Εξέτασης της διατριβής μου. Τέλος θα ήθελα να τους ευχαριστήσω για τις όποιες υποδείξεις τους πάνω στη διδακτορική διατριβή μου.

Θα ήθελα να ευχαριστήσω επίσης θερμότατα τα μέλη της ομάδας / συνεργάτες και συγκεκριμένα τους, Δρ. Φώτιο Γιαννόπουλο, τον Δρ. Δημήτρη Αραμπατζή, τον Δρ. Σόλωνα Ζάννο, τον Δρ. Μιχάλη Έξαρχο και τον Δρ. Παναγιώτη Ρουσόπουλο, καθώς και τον υποψήφιο διδάκτορα Αθανάσιο Μαματσή για την άριστη συνεργασία μας όλα αυτά τα χρόνια. Η συνεργασία και η αλληλοβοήθεια που υπήρχε εντός της ομάδας ήταν υποδειγματική.

Ξεχωριστά από τους ανωτέρω θα ήθελα να ευχαριστήσω αρχικά τον Δρ. Δημήτρη Αραμπατζή για τη συνολική προσφορά του στη διαμόρφωση της παρούσας διατριβής. Επιπλέον, ευχαριστώ τον Δρ. Φώτιο Γιαννόπουλο για την εξαιρετική συνεργασία που είχαμε όλα τα χρόνια της πορείας μου ως υποψήφιος διδάκτωρ.

Θα ήθελα επίσης να ευχαριστήσω τους φίλους μου που μου συμπαραστάθηκαν σε αυτή μου την προσπάθειά.

Τέλος, θα ήθελα να απευθύνω τις θερμότερες ευχαριστίες μου και να αφιερώσω την παρούσα διατριβή στους γονείς μου Ιωάννη και Μάρω και στην αδελφή μου Μαρίνα για την αμέριστη ηθική και υλική στήριξη που μου παρείχαν καθ' όλη τη διάρκεια της προσπάθειάς μου αυτής και για την πίστη τους σε εμένα και στις επιλογές μου.

Πίνακας Περιεχομένων

Περίληψη	2
Abstract	3
Ευχαριστίες	4
Πίνακας Περιεχομένων	6
A. ΑΚΡΙΒΗΣ ΑΝΑΛΥΣΗ ΚΑΙ ΑΝΙΧΝΕΥΣΗ ΤΟΥ ΛΑΘΟΥΣ ΠΕΠΕΡΑΣΜΕΝΗΣ ΑΚΡΙΒΕΙΑΣ ΠΟΥ ΓΕΝΝΑΤΑΙ ΣΕ ΜΙΑ Ή ΠΕΡΙΣΣΟΤΕΡΕΣ ΑΦΑΙΡΕΣΕΙΣ. ΕΥΣΤΑΘΕΙΟΠΟΙΗΣΗ ΦΗΜΙΣΜΕΝΩΝ ΑΛΓΟΡΙΘΜΩΝ ΠΟΥ ΥΠΟΦΕΡΟΥΝ ΣΟΒΑΡΑ ΑΠΟ ΤΟ ΛΑΘΟΣ ΑΥΤΟΥ ΤΟΥ ΤΥΠΟΥ.....	11
1 Το state of the art στη μελέτη του σφάλματος πεπερασμένης ακρίβειας	13
2 Συμβολισμοί, σημειογραφία, βασικές έννοιες και ορισμοί σε σχέση με το σφάλμα πεπερασμένης ακρίβειας.	14
3 Το σφάλμα πεπερασμένης ακρίβειας στις τέσσερις βασικές αριθμητικές πράξεις.....	20
3.1 Το λάθος πεπερασμένης ακρίβειας στη πράξη του Πολλαπλασιασμού.	21
3.2 Το λάθος πεπερασμένης ακρίβειας στη πράξη της διαίρεσης	24
3.3 Το λάθος πεπερασμένης ακρίβειας στη πράξη της πρόσθεσης	24
3.4 Το λάθος πεπερασμένης ακρίβειας στη πράξη της αφαίρεσης.....	26
4 Το λάθος πεπερασμένης ακρίβειας στη πράξη της αφαίρεσης αποτελείται από το νομοτελειακό και το τυχαίο λάθος	27
4.1 Το νομοτελειακό λάθος κατά την αφαίρεση	27
4.2 Το τυχαίο / στοχαστικό λάθος πεπερασμένης ακρίβειας κατά την αφαίρεση.	31
4.2.1 Τυχαίο λάθος πεπερασμένης: Συμβολισμοί, βασικές έννοιες και υπολογισμός των θεωρητικών πιθανοτήτων γένεσης αυτού του λάθους.	32
4.2.2 Πειραματική εκτίμηση του τυχαίου λάθους πεπερασμένης ακρίβειας κατά την αφαίρεση.	49
4.2.3 Σύγκριση θεωρητικών προβλέψεων και πειραματικών δεδομένων.	50
4.3 Το ολικό σφάλμα πεπερασμένης ακρίβειας στη πράξη της αφαίρεσης	53
5 Η γένεση και διάδοση του τυχαίου / στοχαστικού λάθους πεπερασμένης ακρίβειας σε ένα αριθμό διαδοχικών αφαιρέσεων.....	56
6 Σύγκριση της γένεσης και διάδοσης του σφάλματος πεπερασμένης ακρίβειας κατά την εκτέλεση του ίδιου αλγορίθμου με διαφορετικά μήκη λέξης.....	59
7 Τα εκτελεσθέντα σχετικά πειράματα και τα αποτελέσματά τους.	63

7.1	Πειραματική επαλήθευση της θεωρητικής πρόβλεψης σε σχέση με τη συσσώρευση του λάθους πεπερασμένης ακρίβειας σε διαδοχικές αφαιρέσεις	63
8	Η γένεση και διάδοση του σφάλματος πεπερασμένης ακρίβειας στον αλγόριθμο Bernoulli chaotic map, καθώς και μια μέθοδος σταθεροποίησης αυτού.	68
8.1	Σύντομη περιγραφή και αριθμητική συμπεριφορά του αλγορίθμου Bernoulli map.....	68
8.2	Τα βαθύτερα αίτια της καταστροφής του Bernoulli map λόγω σφάλματος πεπερασμένης ακρίβειας. ..	69
8.3	Διόρθωση του σφάλματος πεπερασμένης ακρίβειας στον αλγόριθμο Bernoulli map.....	70
9	Μελέτη και ανίχνευση του σφάλματος πεπερασμένης ακρίβειας στον αλγόριθμο Baker's map και μια μέθοδος διόρθωσης του σφάλματος αυτού.	75
9.1	Σύντομη περιγραφή και αριθμητική συμπεριφορά του αλγορίθμου Baker's map.	75
9.2	Τα βαθύτερα αίτια της καταστροφής του Baker's map λόγω σφάλματος πεπερασμένης ακρίβειας.	76
9.3	Διόρθωση του σφάλματος πεπερασμένης ακρίβειας στον αλγόριθμο Baker's map.....	77
10	Βιβλιογραφία Α' Μέρους	80
B.	ΝΕΟΣ, ΑΠΟΔΟΤΙΚΟΣ ΑΛΓΟΡΙΘΜΟΣ ΕΠΕΞΕΡΓΑΣΙΑΣ ΕΙΚΟΝΑΣ ΜΕ ΕΜΦΑΣΗ ΣΤΗ ΧΡΗΣΗ ΠΟΛΥΩΝΥΜΩΝ ΚΑΙ ΣΥΝΤΕΛΕΣΤΩΝ ZERNIKE.	84
1	Ορισμός των Zernike moments και εφαρμοσιμότητα αυτών.	86
2	Η μεγάλη χρησιμότητα των Zernike moments.	87
3	Υπάρχουσες τεχνικές υπολογισμού των Zernike moments.	90
3.1	Τεχνικές μετασχηματισμού μιας εικόνας, εγγεγραμμένης ή περιγεγραμμένης, στον μοναδιαίο κύκλο.90	
3.1.1	Ο μοναδιαίος κύκλος εγγράφεται εντός της τετραγωνικής εικόνας.....	91
3.1.2	Όλη η εικόνα εντάσσεται εντός του περιγεγραμμένου κύκλου.....	91
3.2	Τεχνικές υπολογισμού των συντελεστών Zernike Zpq	92
3.2.1	Τεχνικές υπολογισμού των ποσοτήτων Zernike πάνω στα Cartesian pixels	92
4	Σημαντικά αριθμητικά προβλήματα που παρουσιάζουν οι προαναφερθείσες μέθοδοι υπολογισμού των συντελεστών Zernike.	100
4.1	Γεωμετρικό σφάλμα στον υπολογισμό των συντελεστών Zernike.....	101
4.2	Αριθμητικό σφάλμα ολοκλήρωσης στον υπολογισμό των συντελεστών Zernike.....	101
5	Γένεση και διάδοση του σφάλματος πεπερασμένης ακρίβειας στον υπολογισμό των Zernike moments. 102	
5.1	Γένεση και διάδοση του σφάλματος πεπερασμένης ακρίβειας στον ευθύ τύπο υπολογισμού των ακτινικών πολυωνύμων Zernike $Rpq(\rho)$	102
5.2	Γένεση, συσσώρευση και διάδοση του σφάλματος πεπερασμένης ακρίβειας στους επαναληπτικούς αλγορίθμους υπολογισμού των ακτινικών πολυωνύμων Zernike.	103

5.2.1	Ο υπολογισμός του ακτινικού πολυωνύμου $Rpp(\rho)$: Η βασική πηγή λάθους κοινή για όλους τους σχετικούς αναδρομικούς αλγορίθμους.....	103
5.2.2	Βασικές πηγές f.p. λάθους στον υπολογισμό των ακτινικών πολυωνύμων Zernike $Rpq(\rho)$, $q < p$ με τη μέθοδο “Q-recursive”.....	103
5.2.3	Βασικές πηγές λάθους στον υπολογισμό των ακτινικών πολυωνύμων Zernike $Rpq(\rho)$, $q < p$ με τη μέθοδο “Modified Kintner”.....	105
5.2.4	Βασικές πηγές λάθους στον υπολογισμό των ακτινικών πολυωνύμων Zernike $Rpq(\rho)$, $q < p$ με τη μέθοδο “Modified Prata”.....	107
5.3	Το λάθος πεπερασμένης ακρίβειας στους συντελεστές Zernike.....	108
5.4	Παραδείγματα ανακατασκευής εικόνων με χρήση των τετραγωνικών pixels.....	114
6	Ευθεία Μέθοδος υπολογισμού των ποσοτήτων Zernike που χρησιμοποιεί Polar Pixels.....	121
7	Ένας νέος, εύρωστος και αποδοτικός αναδρομικός αλγόριθμος για τον υπολογισμό των συντελεστών Zernike, κάνοντας χρήση των πολικών (polar) pixels.	125
7.1	Ο υπολογισμός των ακτινικών πολυωνύμων Zernike σύμφωνα με τη νέα αναδρομική μέθοδο.....	126
7.2	Εκμετάλλευση των συμμετριών που εμφανίζονται στο γωνιακό σκέλος των $Zpq(\rho k, \theta l)$, με σκοπό τη μείωση της υπολογιστικής πολυπλοκότητας.....	130
8	Ανάλυση των επιδόσεων του νέου αλγορίθμου.	137
8.1	Ο Προτεινόμενος αλγόριθμος είναι γρήγορος.....	137
8.2	Ο αλγόριθμος που προτείνεται στη παρούσα εργασία εμφανίζει αξιοσημείωτα μικρό γεωμετρικό λάθος, μικρό σφάλμα ολοκλήρωσης και μικρό σφάλμα πεπερασμένης ακρίβειας.....	139
8.3	Η συμπεριφορά του προτεινόμενου αλγορίθμου όσον αφορά το σφάλμα πεπερασμένης ακρίβειας.....	146
9	Βιβλιογραφία Β' Μέρους.....	147
C.	ΥΛΟΠΟΙΗΣΗ ΑΛΓΟΡΙΘΜΩΝ ΠΟΥ ΕΠΙΤΕΛΟΥΝ ΑΥΤΟΜΑΤΗ ΑΝΑΓΝΩΡΙΣΗ ΓΡΑΦΕΑ ΜΕ ΕΦΑΡΜΟΓΗ ΣΕ ΣΗΜΑΝΤΙΚΑ ΠΑΛΑΙΑ ΚΕΙΜΕΝΑ.	151
1	Εισαγωγή στην αυτόματη αναγνώριση γραφέα παλαιών (ιστορικών) κειμένων.	153
1.1	Η σημασία της Αυτόματης Αναγνώρισης Γραφέα παλαιών κειμένων.....	153
1.2	Οι εγγενείς δυσκολίες στη προσπάθεια ταυτοποίησης του γραφέα ενός παλαιού κειμένου.....	154
2	Συνοπτική παρουσίαση μίας νέας μεθόδου ταυτοποίησης του γραφέα παλαιών κειμένων που παρουσιάστηκε για πρώτη φορά από την ερευνητική μας ομάδα.....	156
2.1	Εξαγωγή μεμονωμένων γραμμάτων από το κείμενο, ένθεση αυτών σε μεμονωμένες εικόνες και αυτόματη εξαγωγή των περιγραμμάτων τους.....	157
2.2	Μια πρώτη προσαρμογή ζευγών υλοποιήσεων του ίδιου συμβόλου της αλφαβήτου.....	160
2.2.1	Υπολογισμός της βέλτιστης ομοιοθεσίας που προσαρμόζει την υλοποίηση $\Sigma A2$ στην $\Sigma A1$	161

2.2.2	Ενσωμάτωση της στροφής και της παράλληλης μετατόπισης για μια πρώτη καλή προσαρμογή της καμπύλης Γ_2 στη Γ_1	164
2.3	Το χρησιμοποιηθέν βέλτιστο κριτήριο ομοιότητας μεταξύ δυο υλοποιήσεων της αλφαβήτα.	166
2.4	Τελεσίδικη, βέλτιστη προσαρμογή δυο υλοποιήσεων του ιδίου γράμματος.	168
3	Αλγοριθμική υλοποίηση που επιτυγχάνει μεγάλη επιτάχυνση της διαδικασίας ταυτοποίησης γραφέα..	170
3.1	Διαγράμματα ροής της μεθόδου που παρουσιάζεται στη παρούσα εργασία.	170
3.2	Πλεονεκτήματα της επιλεγείσης υλοποίησης του αλγορίθμου ταυτοποίησης γραφέα.....	175
3.3	Πρώτα αποτελέσματα που ελήφθησαν.....	175
4	Βιβλιογραφία C' Μέρους	177
	Γλωσσάριο Τεχνικών όρων	179
	Συνοπτικό Βιογραφικό Σημείωμα.....	182
	Δημοσιεύσεις	184
	Διεθνή επιστημονικά περιοδικά με κριτές	184
	Διεθνή επιστημονικά συνέδρια με κριτές.....	184
	Υπό κρίση σε διεθνή επιστημονικά περιοδικά με κριτές.....	184

Α. ΑΚΡΙΒΗΣ ΑΝΑΛΥΣΗ ΚΑΙ
ΑΝΙΧΝΕΥΣΗ ΤΟΥ ΛΑΘΟΥΣ
ΠΕΠΕΡΑΣΜΕΝΗΣ ΑΚΡΙΒΕΙΑΣ ΠΟΥ
ΓΕΝΝΑΤΑΙ ΣΕ ΜΙΑ Ή ΠΕΡΙΣΣΟΤΕΡΕΣ
ΑΦΑΙΡΕΣΕΙΣ. ΕΥΣΤΑΘΕΙΟΠΟΙΗΣΗ
ΦΗΜΙΣΜΕΝΩΝ ΑΛΓΟΡΙΘΜΩΝ ΠΟΥ
ΥΠΟΦΕΡΟΥΝ ΣΟΒΑΡΑ ΑΠΟ ΤΟ ΛΑΘΟΣ
ΑΥΤΟΥ ΤΟΥ ΤΥΠΟΥ.

1 To state of the art στη μελέτη του σφάλματος πεπερασμένης ακρίβειας

Στη διεθνή βιβλιογραφία υπάρχουν αρκετές αναφορές σε σχέση με το λάθος πεπερασμένης ακρίβειας, το οποίο μπορεί να είναι είτε λάθος λόγω στρογγυλοποίησης (round – off error) είτε λάθος αποκοπής. Αυτό οφείλεται στο γεγονός ότι κατά την εκτέλεση ενός αλγορίθμου μπορεί να γεννηθεί και να διαδοθεί ένα τέτοιο αριθμητικό λάθος, ώστε να οδηγηθεί ο αλγόριθμος στο να παράγει λανθασμένα αποτελέσματα. Από πολύ νωρίς παρατηρήθηκε η ύπαρξη αυτού του αριθμητικού λάθους σε διάφορους αλγορίθμους [1], [2], [3], [18].

Επίσης στο [4] παρατηρήθηκε η διάδοση του λάθους πεπερασμένης ακρίβειας σε περιοδικές τροχιές. Στο [5] το σφάλμα στρογγυλοποίησης που εμφανίζεται στους chaotic maps αντιμετωπίζεται ως θόρυβος που προστίθεται στα διάφορα αποτελέσματα. Ομοίως, στο [6] μελετάται το σφάλμα στρογγυλοποίησης (round – off error) και η αξιοπιστία των αποτελεσμάτων σε συγκεκριμένα χαοτικά συστήματα. Επιπροσθέτως, στο [7] παρουσιάζεται ένας αλγόριθμος για τη σταθεροποίηση των chaotic maps στους κβαντικούς υπολογιστές. Οι συγγραφείς του [8] προτείνουν μια μεθοδολογία ευσταθειοποίησης του Poincare map. Στο [9] παρουσιάζεται μια προσπάθεια σταθεροποίησης ενός standard map. Επιπλέον στον [10] δίνεται μια σημασιολογική ανάλυση για το round – off error. Στο άρθρο [11] παρουσιάζεται η μελέτη του σφάλματος στρογγυλοποίησης στους Hardy Multiquadric και άλλους παρόμοιους αλγορίθμους που επιτελούν παρεμβολή (interpolators). Ακολούθως, στο [12] τονίζεται ότι η σύγκλιση των Gegenbauer πολυωνύμων επηρεάζεται από το αριθμητικό σφάλμα στρογγυλοποίησης. Επιπλέον στο [13] προτείνεται μια μεθοδολογία για να αντιμετωπιστεί η εμφάνιση του round – off error στον αλγόριθμο Digital Frequency Synthesis (DFS). Τέλος, στα [14], [15] παρουσιάζεται μια πρώιμη μελέτη της προσέγγισης που αναπτύσσεται στη παρούσα εργασία.

Επιπροσθέτως, η προσέγγιση που παρουσιάζεται στο [16], όπως και σε πολλές σχετικές δημοσιεύσεις, βασίζεται στην εξής ιδέα: “Προσπάθησε να εντοπίσεις ένα άνω όριο για την απόλυτη τιμή του λάθους πεπερασμένης ακρίβειας που γεννάται σε έναν αλγόριθμο”. Η προσέγγιση αυτή έχει, μεταξύ άλλων, τα εξής σοβαρά μειονεκτήματα:

1. Με αυτήν δεν καθίσταται δυνατός ο ακριβής υπολογισμός του λάθους πεπερασμένης ακρίβειας που γεννάται και συσσωρεύεται σε κάθε αναδρομή του αλγορίθμου. Αντιθέτως, το υπολογιζόμενο άνω φράγμα είναι, εν γένει, μια πολύ χαλαρή εκτίμηση αυτού του αριθμητικού λάθους .
2. Ο αναλυτικός υπολογισμός αυτού του φράγματος πολλές φορές δεν είναι εφικτός και, κυρίως,

3. Σε πολλούς αλγορίθμους, δεν υπάρχει τέτοιο φράγμα του συσσωρευόμενου αριθμητικού λάθους πρωταρχικά λόγω της ίδιας δομής του αλγορίθμου.

Αντιθέτως, στη παρούσα εργασία, παρουσιάζεται μια πολύ γενική μεθοδολογία ακριβούς εντοπισμού του λάθους πεπερασμένης ακρίβειας που γεννάται τόσο κατά τη πράξη μιας μόνο αφαίρεσης, όσο και κατά την εκτέλεση πολλών διαδοχικών αφαιρέσεων. Αυτή η προτεινόμενη μεθοδολογία δίνει τη δυνατότητα στο χρήστη του αλγορίθμου να γνωρίζει ανά πάσα στιγμή τόσο τα αποτελέσματα του αλγορίθμου, όσο και τον αριθμό των λανθασμένων δεκαδικών ψηφίων με τον οποίο έχουν υπολογιστεί τα αποτελέσματα αυτά. Επιπλέον, όπως θα φανεί και στα αντίστοιχα κεφάλαια της παρούσας διατριβής, η προτεινόμενη μυθολογία μπορεί να οδηγήσει σε άρση του λάθους πεπερασμένης ακρίβειας σε πολλούς σημαντικούς αλγορίθμους.

Επιπλέον, στο [17] παρουσιάζεται μια θεωρητική, μαθηματικοποιημένη προσέγγιση αναφορικά με την ευστάθεια ενός αλγορίθμου. Η προσέγγιση αυτή χρησιμοποιεί την backward error ανάλυση, αλλά και την forward error ανάλυση, οι οποίες εν πολλοίς θεωρούν το λάθος πεπερασμένης ακρίβειας σαν μια διαταραχή (perturbation) της ορθής τιμής κάθε ποσότητας του αλγορίθμου. Ούτε η προσέγγιση αυτή όμως οδηγεί σε ακριβή υπολογισμό του λάθους πεπερασμένης ακρίβειας που γεννάται ή και συσσωρεύεται σε μια οποιαδήποτε ποσότητα που υπολογίζεται σε μια τυχούσα αναδρομή του αλγορίθμου.

Αντιθέτως, στη παρούσα διατριβή, πίσω από την προτεινόμενη μεθοδολογία “υποκρύπτεται σιωπηρά” η εκμετάλλευση των πραγματικών διαδικασιών που λαμβάνουν χώρα εντός του υπολογιστή κατά την εκτέλεση οποιουδήποτε αλγορίθμου. Η συνεπακόλουθη θεωρητική θεμελίωση είναι αποτέλεσμα και συνισταμένη της κατανόησης αυτών των πραγματικών διαδικασιών.

2 Συμβολισμοί, σημειογραφία, βασικές έννοιες και ορισμοί σε σχέση με το σφάλμα πεπερασμένης ακρίβειας.

Στη παρούσα εργασία θα χρησιμοποιηθεί ένα σύνολο από συμβολισμούς και σημειογραφίες, το οποίο παρατίθεται παρακάτω. Συγκεκριμένα:

- Κάθε αριθμητική ποσότητα α εκφράζεται στο δυαδικό format που αναφέρεται ως scientific format ή στα ελληνικά “επιστημονική μορφή”, δηλαδή

$$\alpha = ma \cdot 2^t$$

όπου ο πραγματικός αριθμός ma ανήκει κατά απόλυτο τιμή στο διάστημα $[1,2)$ και ονομάζεται mantissa (συντελεστής), ενώ ο τ είναι ακέραιος και ονομάζεται εκθέτης.

Ένα διεθνώς αποδεκτό πρότυπο αυτής της μορφής είναι το IEEE 754 [19].

- Στο δεκαδικό σύστημα αρίθμησης ο τυχόν πραγματικός αριθμός a θα θεωρείται ότι είναι επίσης εκφρασμένος σε επιστημονική μορφή, δηλαδή θα συμβολίζεται και ως $man(a) \cdot 10^{E(a)}$, όπου βεβαίως ορίζεται ως $man(a)$ η mantissa ενός αριθμού a , με $1 \leq |man(a)| < 10$, και $E(a)$ ως το εκθετικό κομμάτι του αριθμού αυτού.

ΣΥΝΤΟΜΟΓΡΑΦΙΕΣ

- Το ακρωνύμιο f.p.e, το οποίο προέρχεται από τη αγγλική έκφραση “finite precision error”, χρησιμοποιείται για να υποδείξει το σφάλμα πεπερασμένης ακρίβειας. Σημειώνεται ότι πολλές μελέτες αναφέρονται σε αυτό το λάθος ως round - off error, δηλαδή σφάλμα στρογγυλοποίησης ή και ως quantization error, δηλαδή σφάλμα διακριτοποίησης. Παρόλο αυτά, όπως θα αποδειχθεί στη συνέχεια, το σφάλμα που διαδίδεται σε έναν οποιοδήποτε αλγόριθμο και συσσωρεύεται στις ποσότητες αυτού, παύει να είναι σφάλμα στρογγυλοποίησης μετά από έναν αριθμό επαναλήψεων. Για αυτό το λόγο επιλέγεται η χρήση της σημειογραφίας που προτείνεται στη παρούσα εργασία.
- Χρησιμοποιείται το ακρωνύμιο e.d.d. για να εκφράσει το πλήθος των λανθασμένων δεκαδικών ψηφίων μιας ποσότητας. Το ακρωνύμιο αυτό προέρχεται από τη αγγλική έκφραση “erroneous decimal digits”.
- Αναλόγως χρησιμοποιείται το ακρωνύμιο c.d.d. για να συμβολίσει το πλήθος των σωστών δεκαδικών ψηφίων στις MSB θέσεις της εκάστοτε mantissa, εκ του αγγλικού όρου “correct decimal digits”.

□

Έστω ότι όλοι οι αριθμοί ακολουθούν το scientific format, όπως αυτό περιεγράφηκε ανωτέρω, και έχουν n δεκαδικά ψηφία στη mantissa. Έστω επίσης ότι όλες οι πράξεις γίνονται σε ένα σύστημα άπειρης ακρίβειας. Σε αυτή τη περίπτωση η ποσότητα a συμβολίζεται ως a^c . Έστω ακόμα ότι η ίδια ποσότητα παράγεται μετά από υπολογισμούς σε μια μηχανή πεπερασμένης ακρίβειας n ψηφίων στη mantissa. Τότε η σχέση μεταξύ του a και a^c δίνεται από τους παρακάτω ορισμούς.

ΟΡΙΣΜΟΣ 1

Έστω δυο ομόσημες ποσότητες γ_1 και γ_2 , γραμμένες σε επιστημονική μορφή, οι οποίες έχουν το ίδιο πλήθος n ψηφίων στη mantissa. Συγκεκριμένα,

$$\begin{aligned}\gamma_1 &= x_1 \cdot x_2 x_3 \dots x_n \cdot 10^\tau \\ \gamma_2 &= y_1 \cdot y_2 y_3 \dots y_n \cdot 10^\rho\end{aligned}$$

Έστω επίσης ότι ισχύει $\tau \geq \rho$, χωρίς καταστράτηγηση της γενικότητας. Τότε οι δύο αριθμοί διαφέρουν κατά k δεκαδικά ψηφία με $k > 0$ όταν και μόνο όταν ισχύει η παρακάτω σχέση.

$$|\gamma_1 - \gamma_2| = w \cdot 10^{\tau-(n-k)} \quad (2.1)$$

Όπου w συμβολίζεται η mantissa της διαφοράς με $1 \leq w < 10$.

Επισημαίνεται ότι αν ο παραπάνω τύπος δώσει αρνητικό k ή 0 τότε θεωρείται ότι οι ποσότητες γ_1 και γ_2 είναι ίδιες, όσον αφορά τα n δεκαδικά ψηφία της mantissa. Αντιστοίχως, αν ισχύει $k \geq n$, τότε οι δύο αριθμοί γ_1 και γ_2 διαφέρουν ολοκληρωτικά. Επιπλέον τονίζεται ότι, το πλήθος των κοινών δεκαδικών ψηφίων δεν αντιστοιχεί στην ένα προς ένα σύγκριση των ψηφίων των αριθμών που συμμετέχουν.

□

ΠΑΡΑΔΕΙΓΜΑΤΑ ΑΠΟΣΑΦΗΝΙΣΗΣ ΤΟΥ ΑΝΩΤΕΡΩ ΟΡΙΣΜΟΥ

Παράδειγμα 2.1

Σύμφωνα με τον παραπάνω ορισμό οι δυο αριθμοί με 32 δεκαδικά ψηφία στην mantissa,

$$\begin{aligned}\gamma_1 &= 4.5876598765432156763456784678906 \cdot 10^2 \\ \gamma_2 &= 4.5876598765432156763456451234533 \cdot 10^2\end{aligned}$$

Διαφέρουν σε 9 ψηφία, αφού

- Οι εκθέτες των δυο παραπάνω αριθμών αντιστοιχούν σε $\tau = \rho = 2$
- Το πλήθος των δεκαδικών ψηφίων στη mantissa είναι $n = 32$
- $|\gamma_1 - \gamma_2| = 3.334443735413131351234563248954 \cdot 10^{-21}$

Άρα,

$$\tau - (n - k) = -21 \Rightarrow k = 9$$

Αυτό συμπτωματικά συνάγεται και από την απλή παρατήρηση των συγκεκριμένων αριθμών γ_1 και γ_2 , στους οποίους τα διαφορετικά δεκαδικά ψηφία σημειώνονται με bold στη σχέση ορισμού τους. Όπως όμως θα φανεί και στα επόμενα παραδείγματα, αλλά και σε όλη την ανάλυση που θα ακολουθήσει, η οπτική εκτίμηση του πλήθους των δεκαδικών ψηφίων στα οποία δυο ποσότητες διαφέρουν, δεν είναι καθόλου ασφαλής.

Παράδειγμα 2.2

Έστωσαν οι δύο αριθμοί με 12 δεκαδικά ψηφία στη mantissa

$$\gamma_1 = 5.12345678912 \cdot 10^9$$

$$\gamma_2 = 5.12347654321 \cdot 10^9$$

Οι δυο αυτοί αριθμοί διαφέρουν σε 7 ψηφία, αφού σύμφωνα με τον ορισμό 1

- Οι εκθέτες των δυο παραπάνω αριθμών αντιστοιχούν σε $\tau = \rho = 9$
- Το πλήθος των δεκαδικών ψηφίων στη mantissa είναι $n = 12$
- $|\gamma_1 - \gamma_2| = 1.975409 \cdot 10^4$

Άρα,

$$\tau - (n - k) = 4 \Rightarrow k = 7$$

Παράδειγμα 2.3

Έστωσαν οι δύο αριθμοί με 16 δεκαδικά ψηφία στη mantissa

$$\gamma_1 = 7.654789321458968 \cdot 10^3$$

$$\gamma_2 = 7.145289635854741 \cdot 10^3$$

Διαφέρουν σε 15 ψηφία, όπως φαίνεται παραπάνω, αφού

- Οι εκθέτες των δυο παραπάνω αριθμών αντιστοιχούν σε $\tau = \rho = 3$
- Το πλήθος των δεκαδικών ψηφίων στη mantissa είναι $n = 16$
- $|\gamma_1 - \gamma_2| = 5.09499685604226 \cdot 10^2$

Άρα,

$$\tau - (n - k) = 2 \Rightarrow k = 15$$

Παράδειγμα 2.4

Έστωσαν οι δύο αριθμοί με 16 δεκαδικά ψηφία,

$$\gamma_1 = 7.654789321458968 \cdot 10^{12}$$

$$\gamma_2 = 7.654789321458968 \cdot 10^{12}$$

Όπως συνάγεται από τον ορισμό 1 οι γ_1 και γ_2 δε διαφέρουν, αφού

- Οι εκθέτες των δυο παραπάνω αριθμών αντιστοιχούν σε $\tau \geq \rho \Rightarrow \tau = 12$
- Το πλήθος των δεκαδικών ψηφίων στη mantissa είναι $n = 16$
- $|\gamma_1 - \gamma_2| = 0$

Άρα,

$$\tau - (n - k) = 0 \Rightarrow k = -4 < 0 \Rightarrow k = 0$$

Παράδειγμα 2.5

Έστωσαν οι δύο αριθμοί με 8 δεκαδικά ψηφία

$$\gamma_1 = 8.8475963 \cdot 10^{21}$$

$$\gamma_2 = 5.7414789 \cdot 10^{21}$$

Όπως συνάγεται από τον ορισμό 1 οι γ_1 και γ_2 διαφέρουν σε όλα τα δεκαδικά ψηφία, αφού

- Οι εκθέτες των δυο παραπάνω αριθμών αντιστοιχούν σε $\tau = \rho = 21$
- Το πλήθος των δεκαδικών ψηφίων στη mantissa είναι $n = 8$
- $|\gamma_1 - \gamma_2| = 3.1061174 \cdot 10^{21}$

Άρα,

$$\tau - (n - k) = 21 \Rightarrow k = 8 = n$$

Παράδειγμα 2.6

Έστωσαν οι δύο αριθμοί με 10 δεκαδικά ψηφία

$$\gamma_1 = 1.000000000 \cdot 10^\tau$$

$$\gamma_2 = 9.999999999 \cdot 10^{\tau-1}$$

Δε διαφέρουν σε κανένα δεκαδικό ψηφίο, αφού σύμφωνα με τον ορισμό

- Το πλήθος των δεκαδικών ψηφίων στη mantissa είναι $n = 10$
- $|\gamma_1 - \gamma_2| = 1 \cdot 10^{-10+\tau}$
- $\tau \in \mathbb{Z}$

Άρα

$$\tau - (n - k) = -10 + \tau \Rightarrow k = 0$$

Παράδειγμα 2.7

Έστωσαν οι δύο αριθμοί με 16 δεκαδικά ψηφία στη mantissa

$$\gamma_1 = 5.269587962400597 \cdot 10^4$$

$$\gamma_2 = 5.269587962393951 \cdot 10^4.$$

Η απλή οπτική παρατήρηση των δύο ανωτέρω αριθμών, θα μπορούσε να υποδείξει ότι αυτοί διαφέρουν κατά έξι ψηφία. Εν τούτοις, σύμφωνα με τον ορισμό 1, για αυτούς ισχύουν τα κάτωθι:

- Οι εκθέτες των δυο παραπάνω αριθμών αντιστοιχούν σε $\tau = \rho = 4$
- Το πλήθος των δεκαδικών ψηφίων στη mantissa είναι $n = 16$
- $|\gamma_1 - \gamma_2| = 6.6459 \cdot 10^{-8}$

Άρα οι αριθμοί διαφέρουν κατά τέσσερα δεκαδικά ψηφία. Αυτό είναι απολύτως συνεπές, με το γεγονός ότι

$$\gamma_1 = (\text{man}(\gamma_2) + \text{man}(|\gamma_1 - \gamma_2|) \cdot 10^{-12}) \cdot 10^4$$

ΟΡΙΣΜΟΣ 2

Έστω τώρα ότι όλες οι πράξεις ενός αλγορίθμου εκτελούνται με άπειρη ακρίβεια και ότι η ποσότητα γ έχει την ακριβή τιμή γ^c . Έστω επίσης ότι ο ίδιος αλγόριθμος εκτελείται σε μια υπολογιστική μηχανή με

πεπερασμένο πλήθος δεκαδικών ψηφίων, όπου κάθε ποσότητα γ εκφράζεται με n δεκαδικά ψηφία στη mantissa. Τότε ισχύουν τα κάτωθι:

Η ποσότητα γ έχει υπολογιστεί με ακριβώς τα λ τελευταία (LSB) δεκαδικά ψηφία λανθασμένα, αν και μόνο αν ο αριθμός γ διαφέρει από τη προβολή / περιορισμό της ποσότητας γ^c σε ένα μήκος λέξης n δεκαδικών ψηφίων κατά λ , σύμφωνα με το ορισμό 1. Τη προβολή αυτή τη συμβολίζουμε ως γ_n^c . Συνεπώς, για ένα τυχαίο αριθμό γ , με n δεκαδικά ψηφία στη mantissa, ισχύει ότι

$$(\text{πλήθος } c.d.d.) + (\text{πλήθος } e.d.d.) = n \quad (2.2)$$

Τονίζεται ότι στην δυαδική IEEE 754 αναπαράσταση, η mantissa αναπαρίσταται σε μια υπολογιστική μηχανή με ένα συγκεκριμένο αριθμό από bits, έστω m . Τότε ο ίδιος αριθμός αναπαρίσταται στο δεκαδικό σύστημα με n πλήθος δεκαδικών ψηφίων, όπου ο αριθμός n υπολογίζεται με πολύ ικανοποιητική προσέγγιση από τον κάτωθι τύπο

$$n = m \cdot \log_{10} 2 \quad (2.3)$$

Το γεγονός ότι ο αριθμός που υπολογίζεται από τον παραπάνω τύπο δεν είναι ακέραιος, προκαλεί μια αβεβαιότητα στο πλήθος των σωστών δεκαδικών ψηφίων, η ανάλυση της οποίας θα γίνει στα επόμενα κεφάλαια. □

ΟΡΙΣΜΟΣ 3 – ΟΡΟΛΟΓΙΑ

Οι φράσεις “ο αλγόριθμος αποτυγχάνει” ή “ο αλγόριθμος καταστράφηκε λόγω f.p.e” υποδεικνύουν ότι κάποιες κρίσιμες ποσότητες του αλγορίθμου έχουν υπολογιστεί με όλα τα ψηφία στη mantissa λανθασμένα, σύμφωνα με τον ορισμό 2. Συνεπώς, σε αυτή τη περίπτωση ο συγκεκριμένος αλγόριθμος θα δίνει παντελώς αναξιόπιστα αποτελέσματα μετά από ένα σημείο της εκτέλεσής του. □

3 Το σφάλμα πεπερασμένης ακρίβειας στις τέσσερις βασικές αριθμητικές πράξεις.

Η ανάλυση που ακολουθεί αναφέρεται στους αριθμούς οι οποίοι είναι εκφρασμένοι σε επιστημονική μορφή, όπως ισχύει με το πρότυπο IEEE 754. Μπορεί όμως να επεκταθεί ευθύγραμμα και για fixed-point αριθμούς (σταθερής υποδιαστολής), αρκεί να αντικατασταθεί ο όρος e.d.d. (λάθος δεκαδικά ψηφία) με τον όρο lost decimal digits (l.d.d.).

Η φύση του προτύπου IEEE 754 και του αντίστοιχου hardware που το υποστηρίζει υποδεικνύουν ότι το εκθετικό κομμάτι του αριθμού μπορεί να προκαλέσει μόνο υποχειλίση (underflow) ή υπερχειλίση (overflow) και δε συμμετέχει, με άλλο τρόπο, στη γένεση και διάδοση του αριθμητικού σφάλματος. Επομένως, το σφάλμα πεπερασμένης ακρίβειας οφείλεται κατά κύριο λόγο στη mantissa και στις πράξεις που εκτελούνται επ' αυτής από την υπολογιστική μηχανή κάθε φορά. Παρακάτω αναλύεται μια θεωρητική προσέγγιση του σφάλματος πεπερασμένης ακρίβειας που παράγεται στις τέσσερις βασικές αριθμητικές πράξεις, η οποία συνιστά επέκταση και γενίκευση των αποτελεσμάτων που παρατίθενται στα [14], [15] και [20].

3.1 Το Λάθος πεπερασμένης ακρίβειας στη πράξη του Πολλαπλασιασμού.

Έστωσαν οι αριθμοί a_n και b_n εκφρασμένοι σε επιστημονική μορφή, οι οποίοι έχουν n δεκαδικά ψηφία στη mantissa. Επίσης έστω ότι, εξαιτίας υπολογισμών που προηγήθηκαν, ο a_n έχει παραχθεί με τα λ τελευταία (LS) δεκαδικά ψηφία λανθασμένα, ενώ αντιστοίχως ο b_n έχει υπολογιστεί με μ λάθος δεκαδικά ψηφία. Χωρίς καταστρατήγηση της γενικότητας, θεωρούμε ότι ισχύει $\lambda \geq \mu$.

Έστω ότι μια υπολογιστική μηχανή εκτελεί τον πολλαπλασιασμό $p_n = a_n \cdot b_n$. Τότε το αποτέλεσμα p_n υπολογίζεται με $\lambda + 2$ ή $\lambda + 1$ ή λ ή $\lambda - 1$ ή $\lambda - 2$ ή $\lambda - 3$ λάθος δεκαδικά ψηφία, έχοντας τις αντίστοιχες πιθανότητες $P^M(2, \delta)$, $P^M(1, \delta)$, $P^M(0, \delta)$, $P^M(-1, \delta)$, $P^M(-2, \delta)$, $P^M(-3, \delta)$, όπου $\delta = \lambda - \mu$. Οι παραπάνω πιθανότητες εξαρτώνται μόνο από τα κάτωθι:

- Από τη διαφορά των λαθών μ και λ
- Από το εάν επαληθεύεται ή όχι η ανισότητα $|man(a_n) \cdot man(b_n)| \geq 10$
- Από τη στατιστική κατανομή της ποσότητας του σφάλματος. Στις περισσότερες περιπτώσεις το λάθος αυτό ακολουθεί κανονική κατανομή, σύμφωνα με το κεντρικό οριακό θεώρημα.

Σε αυτό το σημείο τονίζεται ότι εάν ισχύει η ανισότητα $|man(a_n) \cdot man(b_n)| \geq 10$, τότε η ποσότητα p_n παράγεται με το πολύ $\lambda + 1$ λανθασμένα δεκαδικά ψηφία, με μεγαλύτερη πιθανότητα να έχει παραχθεί με λ ή $\lambda - 1$ e.d. Αντιστοίχως, αν δεν ισχύει η ανισότητα, τότε η ποσότητα p_n παράγεται με το πολύ $\lambda + 2$ λανθασμένα δεκαδικά ψηφία, με μεγαλύτερη πιθανότητα να έχει παραχθεί με $\lambda + 1$ ή λ e.d.

ΠΑΡΑΔΕΙΓΜΑΤΑ

Παράδειγμα 3.1

Έστωσαν οι δυο αριθμοί z και w εκφρασμένοι τόσο με δεκάξι όσο και με τριάντα δύο δεκαδικά ψηφία στη mantissa. Έστω επίσης ότι οι αναπαραστάσεις αυτών των αριθμών σε μήκος λέξης δεκάξι ψηφίων είναι:

$$z_{16} = 2.543458906364323$$

$$\widetilde{z}_{32} = 2.543458906342428$$

$$w_{16} = 3.808231231387643$$

$$\widetilde{w}_{32} = 3.808231231342319$$

Τονίζεται ότι οι ανωτέρω ποσότητες \widetilde{z}_{32} και \widetilde{w}_{32} προέκυψαν από τη προβολή / περιορισμό σε δεκάξι δεκαδικά ψηφία της αναπαράστασης των αριθμών z και w αντίστοιχα στα τριάντα δύο ψηφία, με στρογγύλευση. Δηλαδή, οι ποσότητες z και w υπολογίστηκαν πρώτα σε τριάντα δύο δεκαδικά ψηφία και εν συνεχεία προεβλήθησαν / περιορίστηκαν σε δεκάξι δεκαδικά ψηφία, προσφέροντας τις αναπαραστάσεις \widetilde{z}_{32} και \widetilde{w}_{32} αντίστοιχα.

Σύμφωνα με τον ορισμό 2, οι ποσότητες z_{16} και w_{16} έχουν τα τελευταία 5 ψηφία λανθασμένα. Εκτελώντας τις πράξεις $p_{16} = z_{16} \cdot w_{16}$ και $p_{32} = z_{32} \cdot w_{32}$ και κρατώντας μόνο τα πρώτα 16 (με στρογγύλευση) ψηφία του p_{32} , έχουμε ότι:

$$p_{16} = 9.686079642967673$$

$$\widetilde{p}_{32} = 9.686079642769013$$

Παρατηρείται ότι οι p_{16} και \widetilde{p}_{32} διαφέρουν κατά 6 δεκαδικά ψηφία, άρα συγκρίνοντας τα λάθη του αποτελέσματος με τον μέγιστο αριθμό λανθασμένων δεκαδικών ψηφίων μεταξύ των όρων του γινομένου, βλέπουμε ότι το λάθος αυξήθηκε κατά ένα. Επισημαίνεται ότι σε αυτή τη περίπτωση ισχύει η ανισότητα $|\text{man}(z_{16}) \cdot \text{man}(w_{16})| < 10$.

Παράδειγμα 3.2

Έστωσαν οι αριθμοί του προηγούμενου παραδείγματος, όπου μετά από μια σειρά πράξεων διαφέρουν κατά 3 δεκαδικά ψηφία. Συγκεκριμένα:

$$z_{16} = 2.543458906342\mathbf{323}$$

$$\widetilde{z}_{32} = 2.543458906342428$$

$$w_{16} = 3.808231231342\mathbf{643}$$

$$\widetilde{w}_{32} = 3.808231231342319$$

Εκτελώντας τη πράξη του πολλαπλασιασμού $z_{16} \cdot w_{16}$ και $z_{32} \cdot w_{32}$ και αφού περιορίσουμε το p_{32} σε 16 δεκαδικά ψηφία λαμβάνουμε:

$$p_{16} = 9.686079642769\mathbf{436}$$

$$\widetilde{p}_{32} = 9.686079642769\mathbf{013}$$

Όπου ως συνήθως, στους αριθμούς με δείκτη 32 και περισπωμένη άνωθεν παρουσιάζονται τα 16 πρώτα δεκαδικά ψηφία του αριθμού. Παρατηρείται ότι τα τρία τελευταία ψηφία είναι λανθασμένα. Στη συγκεκριμένη περίπτωση όμως η πράξη του πολλαπλασιασμού δεν έχει επιφέρει αύξηση του λάθους, μολονότι ισχύει η ανίσωση $|\text{man}(z_{16}) \cdot \text{man}(w_{16})| < 10$, διότι το λάθος πεπερασμένης ακρίβειας που γεννάται κατά τη πράξη του πολλαπλασιασμού εξαρτάται επίσης και από το λάθος με το οποίο εισέρχονται οι πολλαπλασιαστέοι. Με τη γνώση της κατανομής αυτών των λαθών ή με την εκτίμηση αυτής, τότε μπορεί να υπολογιστούν οι αντίστοιχες πιθανότητες για το λάθος του γινομένου [20].

Παράδειγμα 3.3

Έστωσαν οι αριθμοί z_{16}, w_{16}, z_{32} και w_{32} , όπου οι τελευταίο έχουν αναπαράσταση \widetilde{z}_{32} και \widetilde{w}_{32} στα δεκαέξι δεκαδικά ψηφία, όπως φαίνεται κατωτέρω

$$z_{16} = 4.4658941\mathbf{25478965}$$

$$\widetilde{z}_{32} = 4.465894199999584$$

$$w_{16} = 8.123589\mathbf{658748964}$$

$$\widetilde{w}_{32} = 8.123589868231456$$

Εκτελώντας τη πράξη του πολλαπλασιασμού $p_{16} = z_{16} \cdot w_{16}$ και $p_{32} = z_{32} \cdot w_{32}$ και αφού περιορίσουμε το p_{32} σε 16 δεκαδικά ψηφία

$$p_{16} = 3.627909\mathbf{133480866} \cdot 10^1$$

$$\widetilde{p}_{32} = 3.627909287571024 \cdot 10^1$$

Παρατηρείται ότι επαληθεύεται η ανισότητα $|man(z_{16}) \cdot man(w_{16})| \geq 10$. Όπως και στη προηγούμενη περίπτωση η πράξη του πολλαπλασιασμού δεν έχει επιφέρει αύξηση του λάθους. Γνωρίζοντας και την κατανομή των λαθών με το οποίο εισέρχονται οι πολλαπλασιαστέοι, η αντίστοιχη πιθανότητα μπορεί να υπολογιστεί.

3.2 Το Λάθος πεπερασμένης ακρίβειας στη πράξη της διαίρεσης

Έστωσαν δυο αριθμοί a_n και b_n εκφρασμένοι σε επιστημονική μορφή, έχοντας n δεκαδικά ψηφία στη mantissa. Έστω επίσης ότι, εξαιτίας προηγούμενων υπολογισμών η ποσότητα a_n έχει υπολογιστεί με τα λ τελευταία δεκαδικά ψηφία στη mantissa λανθασμένα, ενώ η ποσότητα b_n έχει υπολογιστεί με τα μ LS δεκαδικά ψηφία λανθασμένα, όπου χωρίς βλάβη της γενικότητας, υποθέτουμε ότι $\lambda \geq \mu$.

Ας υποθεθεί ότι μια υπολογιστική μηχανή εκτελεί τη πράξη $d_n = a_n/b_n$. Τότε, η ποσότητα d_n μπορεί να υπολογιστεί είτε με $\lambda + 2$, $\lambda + 1$, λ , $\lambda - 1$, $\lambda - 2$ ή $\lambda - 3$ λάθος δεκαδικά ψηφία, με τις αντίστοιχες πιθανότητες $P^D(2, \delta)$, $P^D(1, \delta)$, $P^D(0, \delta)$, $P^D(-1, \delta)$, $P^D(-2, \delta)$, $P^D(-3, \delta)$, όπου $\delta = \lambda - \mu$. Η γένεση σφάλματος πεπερασμένης ακρίβειας στη πράξη της διαίρεσης οφείλεται εκ νέου στη mantissa και όχι στο εκθετικό κομμάτι των αριθμών που διαιρούνται. Συνεπώς, η γένεση και διάδοση του λάθους εξαρτάται μόνο από τα κάτωθι:

1. Από τη διαφορά μεταξύ των λ και μ
2. Από το εάν επαληθεύεται η ανίσωση $|man(a_n)/man(b_n)| \geq 1$
3. Από τη στατιστική κατανομή του σφάλματος πεπερασμένης ακρίβειας, η οποία συνήθως ακολουθεί κανονική κατανομή.

Τονίζεται ότι όταν επαληθεύεται η ανίσωση $|man(a_n)/man(b_n)| \geq 1$, τότε η ποσότητα d_n μπορεί να υπολογιστεί με το πολύ $\lambda + 1$ λάθος δεκαδικά ψηφία, ενώ με μεγαλύτερη πιθανότητα υπολογίζεται με λ και $\lambda - 1$ e.d.d. Αντιθέτως, εάν η προαναφερθείσα ανίσωση δεν επαληθεύεται τότε η ποσότητα d_n υπολογίζεται με μέγιστο λάθος $\lambda + 2$ e.d.d. Σε αυτή τη περίπτωση, υπάρχει μεγαλύτερη πιθανότητα το αποτέλεσμα d_n να υπολογιστεί με $\lambda + 1$ ή λ λάθος δεκαδικά ψηφία στη mantissa.

3.3 Το Λάθος πεπερασμένης ακρίβειας στη πράξη της πρόσθεσης

Έστωσαν δύο αριθμοί a_n και b_n , εκφρασμένοι σε επιστημονική μορφή, οι οποίοι έχουν n δεκαδικά ψηφία στη mantissa. Εξαιτίας προηγούμενων υπολογισμών η ποσότητα a_n έχει υπολογιστεί με τα λ

τελευταία δεκαδικά ψηφία στη mantissa λανθασμένα, ενώ η ποσότητα b_n έχει υπολογιστεί αντιστοίχως με μ λάθος ψηφία και, χωρίς βλάβη της γενικότητας, έστω $\lambda \geq \mu$.

Έστω ότι μια υπολογιστική μηχανή εκτελεί τη πράξη $t_n = a_n + b_n$, όπου τα a_n, b_n έχουν το ίδιο πρόσημο. Κατά την εκτέλεση της παραπάνω πράξεως, είναι εξαιρετικά απίθανο να παραχθεί ένα επιπλέον λάθος δεκαδικό ψηφίο. Αντιθέτως, μπορεί το λάθος αυτό να ελαττωθεί κατά ένα (1) δεκαδικό ψηφίο όταν ισχύει η παρακάτω ανίσωση

$$E(t_n) > \max\{E(a_n), E(b_n)\} \quad (3.1)$$

Σε αυτή τη περίπτωση η υπολογιστική μηχανή εκτελεί δεξιά ολίσθηση, με σκοπό να φέρει το αποτέλεσμα σε επιστημονική μορφή, αυξάνοντας ταυτόχρονα τον εκθέτη κατά ένα.

Ακόμα και εάν δεν επαληθεύεται η παραπάνω ανίσωση, ο αριθμός λανθασμένων δεκαδικών ψηφίων του αθροίσματος t_n εξαρτάται επίσης από τις τιμές των λαθών πεπερασμένης ακρίβειας των ποσοτήτων a_n και b_n , τα οποία συμβολίζονται με x_n και y_n αντίστοιχα, και από τη διαφορά των εκθετών $E(a_n)$ και $E(b_n)$. Οι προσημασμένες τιμές x_n και y_n , καθώς και το άθροισμά τους έχουν μέση τιμή μηδέν και αρχική τυπική απόκλιση περίπου ίση με $1.14 \cdot 10^{-n}$. Άρα αφού ισχύει $a_n + b_n = a^c + b^c + (x_n + y_n)$, συνεπάγεται ότι στη περίπτωση που ισχύει η ανίσωση $|\text{man}(a_n) + \text{man}(b_n)| < 10$, η πράξη της πρόσθεσης δεν δημιουργεί κατά κανόνα (στατιστικά) παραπάνω λάθος δεκαδικά ψηφία στο αποτέλεσμα.

Αντιστρόφως, εάν ισχύει η ανίσωση $|\text{man}(a_n) + \text{man}(b_n)| \geq 10$, τότε το λάθος πεπερασμένης ακρίβειας του αθροίσματος t_n συνήθως μειώνεται κατά ένα δεκαδικό ψηφίο, λόγω της δεξιάς ολίσθησης που επιτελεί η υπολογιστική μηχανή για να φέρει το t_n σε επιστημονική μορφή.

ΠΑΡΑΔΕΙΓΜΑΤΑ

Παράδειγμα 3.4

Έστωσαν οι αριθμοί $a_n = 6.542942 \cdot 10^1$ και $b_n = 5.217917 \cdot 10^1$, οι οποίοι έχουν κοινό εκθέτη (1), καθώς και οι αντίστοιχες, χωρίς λάθος ποσότητες $a_n^c = 6.542111 \cdot 10^1$ και $b_n^c = 5.219317 \cdot 10^1$. Ο αριθμός a_n έχει τρία λάθος δεκαδικά ψηφία στη mantissa ενώ ο b_n έχει υπολογιστεί με τέσσερα δεκαδικά ψηφία λάθος. Επομένως,

$$t_n^c = a_n^c + b_n^c = 1.176086 \cdot 10^2$$

$$t_n = a_n + b_n = 1.176143 \cdot 10^2$$

Οι αριθμοί t_n^c και t_n διαφέρουν σε δύο δεκαδικά ψηφία αφού

- i. $|t_n^c - t_n| = 5.700000 \cdot 10^{-3}$
- ii. $n = 7$
- iii. $\tau = 2$

Άρα

$$\tau - n + k = -3 \Rightarrow k = 2.$$

Η διόρθωση στον αριθμό των λανθασμένων δεκαδικών ψηφίων του αθροίσματος οφείλεται στο γεγονός ότι α) $|man(a_n) + man(b_n)| \geq 10$ και β) $|x_n + y_n| < 10$.

3.4 Το Λάθος πεπερασμένης ακρίβειας στη πράξη της αφαίρεσης

Έστωσαν δύο αριθμοί a_n και b_n εκφρασμένοι σε επιστημονική μορφή, οι οποίοι έχουν n δεκαδικά ψηφία στη mantissa. Εξαιτίας προηγούμενων υπολογισμών η ποσότητα a_n έχει υπολογιστεί τα λ τελευταία δεκαδικά ψηφία στη mantissa λανθασμένα, ενώ αντιστοίχως η ποσότητα b_n έχει υπολογιστεί με μ λάθος ψηφία όπου, χωρίς βλάβη της γενικότητας, έστω ότι ισχύει $\lambda \geq \mu$.

Ας υποθεθεί ότι μια υπολογιστική μηχανή εκτελεί την πράξη της αφαίρεσης $s_n = a_n - b_n$, όπου τα a_n, b_n έχουν το ίδιο πρόσημο. Τα είδη των λαθών που παράγονται μετά την εκτέλεση της πράξεως είναι τα παρακάτω δύο:

- i. Το νομοτελειακό λάθος (deterministic error).
- ii. Το τυχαίο / στοχαστικό λάθος (erratic error).

Επομένως το συνολικό λάθος που παράγεται είναι το άθροισμα των παραπάνω δύο. Η ανάλυση των δύο ειδών προαναφερθέντων τύπων λαθών ακολουθεί ευθύς αμέσως στο επόμενο κεφάλαιο.

Τέλος, σύμφωνα με τη παραπάνω ανάλυση κάθε βασική αριθμητική πράξη παράγει διαφορετικό λάθος πεπερασμένης ακρίβειας. Συνεπώς, κάθε σχέση ενός αλγορίθμου μπορεί να παράγει ή όχι λάθος δεκαδικά ψηφία ή και να τα ελαττώνει. Εάν μια συγκεκριμένη φόρμουλα παράγει συστηματικά λάθος πεπερασμένης ακρίβειας, τότε αυτή θεωρείται ως μια βασική πηγή λάθους f.p.e. Αντιθέτως, αν μία σχέση δε συμμετέχει στο f.p.e. τότε αυτή θεωρείται αδιάφορη, ενώ εάν το ελαττώνει τότε ο συγκεκριμένος τύπος θεωρείται ως μια βασική αιτία ελάττωσης του σφάλματος. Πρέπει να τονιστεί πως πολλές φορές

η πηγή λάθους μπορεί να εξαρτάται από τη φύση ή και τις στατιστικές ιδιότητες των διάφορων τιμών που λειτουργούν ως είσοδοι στον αλγόριθμο.

4 Το λάθος πεπερασμένης ακρίβειας στη πράξη της αφαίρεσης αποτελείται από το νομοτελειακό και το τυχαίο λάθος

Στο παρόν κεφάλαιο παρουσιάζονται τα είδη των λαθών που γεννώνται κατά τη πράξη της αφαίρεσης. Συγκεκριμένα, τα είδη είναι τα εξής: Το πρώτο είδος f.p.e. εμφανίζεται σε αρκετές περιπτώσεις και μπορεί να υπολογιστεί με ακρίβεια. Αυτό το λάθος το έχουμε ονομάσει νομοτελειακό είτε αιτιοκρατικό λάθος (deterministic / causal error). Το άλλο είδος λάθους που παράγεται κατά την αφαίρεση, είναι άμεσα συνδεδεμένο με τη τυχαιότητα με την οποία μια υπολογιστική μηχανή γεμίζει τις θέσεις των λανθασμένων ψηφίων στη mantissa, όταν ο υπολογιστής εκτελεί αριστερή ολίσθηση, με σκοπό να διατηρήσει την επιστημονική μορφή (π.χ IEEE754) του αποτελέσματος. Αυτό το λάθος το έχουμε ονομάσει τυχαίο / στοχαστικό λάθος (erratic f.p. error).

4.1 Το νομοτελειακό λάθος κατά την αφαίρεση

Στο παρόν κεφάλαιο αποδεικνύεται ότι κατά την αφαίρεση δυο αριθμών ιδίου πρόσημου, είναι πιθανόν να παραχθεί νομοτελειακά σφάλμα πεπερασμένης ακρίβειας. Το f.p.e. που μπορεί να παραχθεί είναι δύναμει αυθαίρετα μεγάλο, με την έννοια ότι κάτω από προϋποθέσεις, το πλήθος των λανθασμένων ψηφίων μπορεί να φτάσει το μήκος λέξης της mantissa. Συγκεκριμένα, το νομοτελειακό λάθος (deterministic error) επισυμβαίνει όταν υπάρχει βύθιση μεταξύ του εκθέτη του αποτελέσματος s_n και του μεγίστου εκθέτη μεταξύ των δύο αφαιρετέων a_n και b_n . Αν με τ συμβολίζεται ο μέγιστος αυτός εκθέτης και με ε συμβολίζεται ο εκθέτης του αποτελέσματος, τότε το νομοτελειακό λάθος αντιστοιχεί σε επιπλέον $d = \tau - \varepsilon$ λανθασμένα δεκαδικά ψηφία στη mantissa. Ο όρος “βύθιση του εκθέτη” χρησιμοποιείται για να υποδηλώσει τη διαφορά αυτή.

$$d = \tau - \varepsilon \quad (4.1)$$

ΠΑΡΑΔΕΙΓΜΑΤΑ

Παράδειγμα 4.1

Έστωσαν δύο θετικοί αριθμοί p_{32}, q_{32} με τριάντα δύο δεκαδικά ψηφία στη mantissa. Έστωσαν επίσης οι προβολές των δύο αυτών αριθμών σε μήκος λέξης δεκαέξι δεκαδικών ψηφίων p_{16}, q_{16} , όπως φαίνεται παρακάτω:

$$p_{16} = 3.145423453312365 \cdot 10^3$$

$$q_{16} = 3.145423453377677 \cdot 10^3$$

$$p_{32} = 3.1454234533123652345678904567894 \cdot 10^3$$

$$q_{32} = 3.1454234533776771231234543212376 \cdot 10^3$$

Τότε οι αντίστοιχες αφαιρέσεις είναι:

$$s_{16} = p_{16} - q_{16} = 6.5312 \cdot 10^{-8}$$

$$s_{32} = p_{32} - q_{32} = 6.53118885555638644482 \cdot 10^{-8}$$

Ο μέγιστος εκθέτης μεταξύ των p_{16} και q_{16} είναι τρία, ενώ ο εκθέτης του αποτελέσματος s_{16} είναι -8. Επομένως η βύθιση του εκθέτη αντιστοιχεί σε $u = \tau - \varepsilon = 11$. Συνεπώς, κατά την αφαίρεση, έντεκα επιπλέον λανθασμένα δεκαδικά ψηφία παρήχθησαν για τον εξής λόγο:

Για να αποκαταστήσει τη διαφορά s_{16} σε επιστημονική μορφή, η υπολογιστική μηχανή πραγματοποιεί αριστερή ολίσθηση, η οποία αντιστοιχεί σε 11 δεκαδικές θέσεις στη mantissa. Αυτή η αριστερή ολίσθηση αφήνει επομένως 11 δεκαδικές θέσεις στην mantissa, το περιεχόμενο των οποίων είναι εντελώς άγνωστο / απροσδιόριστο. Ο υπολογιστής συμπληρώνει επομένως αυθαίρετα – τυχαία τα έντεκα δεκαδικά ψηφία αυτών των θέσεων, τα οποία είναι στη βάση του λανθασμένα.

Όλα τα προηγούμενα επαληθεύονται συγκρίνοντας τους s_{16} και \widetilde{s}_{32} σύμφωνα με τον ορισμό 1. Πράγματι,

1. $s_{16} = 6.531200000000000 \cdot 10^{-8}$
2. $\widetilde{s}_{32} = 6.531188855556386 \cdot 10^{-8}$
3. $\tau = \rho = -8$
4. $n = 16$

$$5. |s_{16} - \widetilde{s}_{32}| = 1.114444361416389 \cdot 10^{-13}$$

$$\tau - (n - k) = -13 \Leftrightarrow k = 11$$

ΠΡΟΤΑΣΗ A.4.1

Έστω ότι όλες οι ποσότητες σε ένα υπολογιστικό περιβάλλον έχουν μήκος λέξης n ψηφίων στη mantissa. Έστω ότι σε αυτό το περιβάλλον ορίζεται ως s_n το αποτέλεσμα της παρακάτω αφαίρεσης:

$$s_n = p_n - q_n, \quad p_n \cdot q_n > 0$$

Όπου,

$$s_n = \gamma_1 \cdot \gamma_2 \dots \gamma_n \cdot 10^\varepsilon, \quad p_n = a_1 \cdot a_2 \dots a_n \cdot 10^\zeta, \quad q_n = \beta_1 \cdot \beta_2 \dots \beta_n \cdot 10^\zeta$$

Έστω επίσης ότι, εξαιτίας προηγούμενων υπολογισμών, οι ποσότητες p_n και q_n έχουν υπολογιστεί με τα πρώτα λ_c δεκαδικά ψηφία ορθά, ή ισοδυνάμως με $\lambda = n - \lambda_c$ τελευταία δεκαδικά ψηφία στη mantissa λανθασμένα. Αν p^c είναι η σωστή τιμή της ποσότητας p_n και q^c είναι η σωστή τιμή του αριθμού q_n , τότε οι p_n, q_n μπορούν να γραφούν στη κάτωθι μορφή:

$$p_n = p^c + x10^{-\lambda_c+\zeta}$$

$$q_n = q^c + y10^{-\lambda_c+\zeta}$$

Όπου x, y οι mantissae των f.p. errors των δυο αριθμών. Έστω ότι ισχύει $1 \leq |x - y| < 10$. Τότε, εάν $d = \zeta - \varepsilon$, η ποσότητα s_n έχει νομοτελειακά τα πρώτα $(\lambda_c - d)$ ψηφία σωστά.

□

Με άλλα λόγια, το αρχικό πλήθος σφάλματος πεπερασμένης ακρίβειας έχει αυξηθεί κατά d με τρόπο νομοτελειακό. Όμως, όπως θα αποδειχθεί και στη συνέχεια, αυτό το λάθος μπορεί να τροποποιηθεί τυχαία.

ΑΠΟΔΕΙΞΗ

Αρχικά, αφού το λ_c το πλήθος των ορθών δεκαδικών ψηφίων των p_n και q_n , τότε ισχύει ότι $p_n = p^c + x10^{-\lambda_c+\zeta}$ και $q_n = q^c + y10^{-\lambda_c+\zeta}$.

Συνεπώς

$$s_n = p_n - q_n = p^c + x10^{-\lambda_c+\zeta} - q^c - y10^{-\lambda_c+\zeta} = (p^c - q^c) + (x - y)10^{-\lambda_c+\zeta} \quad (4.2)$$

Από τη παραπάνω σχέση ο παράγοντας $(p^c - q^c)$ είναι τάξης ε . Η τάξη αυτή είναι μικρότερη ή ίση από την ζ . Έστω επίσης ότι $d = \zeta - \varepsilon$. Επιπροσθέτως, έχει υποτεθεί ότι ισχύει η ανισότητα $1 \leq |x - y| < 10$.
Δηλαδή,

$$s_n = p_n - q_n = \alpha_1 \cdot \alpha_2 \alpha_3 \dots \alpha_n 10^\zeta - \beta_1 \cdot \beta_2 \beta_3 \dots \beta_n 10^\zeta = 0.00 \dots 0 \gamma_1 \dots \gamma_{n-d} 10^\zeta \quad (4.3)$$

Σε αυτό το σημείο η υπολογιστική μηχανή εκτελεί αριστερή ολίσθηση πλήθους d , με σκοπό να διατηρήσει τη επιστημονική μορφή του αποτελέσματος s_n . Αυτή η ολίσθηση παράγει νομοτελειακά το πολύ d λανθασμένα δεκαδικά ψηφία στην “ουρά” της mantissa της διαφοράς.

Επομένως, η ποσότητα s_n έχει υπολογιστεί με d επιπλέον λανθασμένα, ή ισοδυνάμως με $(\lambda_c - d)$ σωστά δεκαδικά ψηφία στη mantissa. Αφαιρώντας τις δύο ποσότητες p_n και q_n , το αποτέλεσμα αυτής της πράξης s_n θα έχει μηδενικά τα πρώτα d δεκαδικά ψηφία στη mantissa

Αφού ισχύει η παραδοχή $1 \leq |x - y| < 10$, οι εκθέτες των λαθών x και y δε συνεισφέρουν στον τελικό υπολογισμό του λάθους πεπερασμένης ακρίβειας.

□

Με εντελώς ανάλογο τρόπο μπορεί να αποδειχθεί το κάτωθι συμπέρασμα:

Έστω οι αριθμοί p_n, q_n , οι οποίοι έχουν n δεκαδικά ψηφία στη mantissa,:

$$p_n = a_1 \cdot a_2 \dots a_n \cdot 10^\zeta$$

$$q_n = \beta_1 \cdot \beta_2 \dots \beta_n \cdot 10^\theta$$

$$p_n \cdot q_n > 0$$

Έστω επίσης ότι όλες οι πράξεις εκτελούνται σε μια υπολογιστική μηχανή που χρησιμοποιεί μήκος λέξης n . Ας θεωρήσουμε ότι η μηχανή αυτή εκτελεί την αφαίρεση

$$s_n = p_n - q_n$$

Όπου χωρίς καμία βλάβη της γενικότητας θεωρούμε ότι $\zeta \geq \theta$.

Έστω επίσης ότι εξαιτίας προηγούμενων υπολογισμών οι ποσότητες p_n, q_n εισέρχονται στην αφαίρεση έχοντας λ_c σωστά δεκαδικά ψηφία στη mantissa (ισοδυνάμως $\lambda = n - \lambda_c$ λανθασμένα δεκαδικά ψηφία).

Σύμφωνα με τους ορισμούς 1 και 2 ισχύει ότι:

$$p_n = p^c + x \cdot 10^{-\lambda_c + \zeta}$$

$$q_n = q^c + y \cdot 10^{-\lambda_c + \theta},$$

Όπου p^c είναι η σωστή τιμή της ποσότητας p_n και q^c είναι η σωστή τιμή του αριθμού q_n και x, y οι mantissae των λαθών πεπερασμένης ακρίβειας.

Έστω επίσης ότι ισχύει η ανισότητα. $1 \leq |x - y \cdot 10^{\zeta - \theta}| < 10$. Σε αυτή την περίπτωση, ας υποθεθεί ότι:

$$d = \max\{\zeta, \theta\} - \varepsilon \Rightarrow$$

$$d = \zeta - \varepsilon$$

Έχοντας υιοθετήσει ότι $\zeta = \max\{\zeta, \theta\}$, χωρίς καμία βλάβη της γενικότητας. Τότε το s_n η ποσότητα s_n έχει υπολογιστεί με d επιπλέον λανθασμένα, ή ισοδυνάμως με $(\lambda_c - d)$ σωστά δεκαδικά ψηφία στη mantissa, δηλαδή με d λιγότερα από τους όρους της αφαίρεσης.

4.2 Το τυχαίο / στοχαστικό λάθος πεπερασμένης ακρίβειας κατά την αφαίρεση.

Χωρίς καταστρατήγηση της γενικότητας, ακολουθεί η ανάλυση της αφαίρεσης των ποσοτήτων p_n και q_n , όταν αυτές είναι της ίδιας τάξης, δεδομένου ότι, όπως θα φανεί κατωτέρω, αυτή είναι συντριπτικά η πιθανότερη περίπτωση που μπορεί να γεννήσει τυχαίο λάθος πεπερασμένης ακρίβειας κατά την πράξη της αφαίρεσης. Επιπλέον, θα δείξουμε παρακάτω ότι το ενδεχόμενο γένεσης τυχαίου λάθους πεπερασμένης ακρίβειας στη περίπτωση των άνισων εκθετών, ανάγεται στη μελέτη των ίσων εκθετών .

Στην παρακάτω ανάλυση οι εκθέτες των p_n και q_n είναι μηδέν, αφού με αυτό τον τρόπο αντιπροσωπεύεται η γενικότερη περίπτωση των ίσων εκθετών. Δηλαδή,

$$p_n - q_n = \text{man}(p_n) \cdot 10^{E(p_n)} - \text{man}(q_n) \cdot 10^{E(p_n)} = 10^{E(p_n)} (\text{man}(p_n) - \text{man}(q_n)).$$

Από τη παραπάνω σχέση αποδεικνύεται ότι το τυχαίο λάθος πεπερασμένης ακρίβειας εξαρτάται μόνο από τη διαφορά των mantissae. Ακριβώς επειδή ο κοινός εκθετικός όρος βγαίνει κοινός παράγοντας κατά την αφαίρεση, μελετάται μόνο η περίπτωση όπου οι αριθμοί που αφαιρούνται είναι μηδενικής τάξης.

Όταν οι εκθέτες είναι άνισοι, ακολουθείται η εξής διαδικασία:

Έστωσαν οι αριθμοί $p_n = a_1 \cdot a_2 \dots a_n \cdot 10^\tau$ και $q_n = \beta_1 \cdot \beta_2 \dots \beta_n \cdot 10^\rho$, όπου χωρίς καμία βλάβη της γενικότητας υποθέτουμε ότι $\tau > \rho$. Εάν $\delta = \tau - \rho$, για να πραγματοποιηθεί η αφαίρεση $s_n = p_n - q_n$, η υπολογιστική μηχανή εκτελεί δεξιά ολίσθηση κατά δ θέσεις στη mantissa του q_n και θέτει τα δ πρώτα δεκαδικά ψηφία του q_n υποχρεωτικά ίσα με το μηδέν. Συνεπώς, προκύπτει ο αριθμός $q_n^\delta = 0.000 \dots \beta_1 \cdot \beta_2 \dots \beta_{n-\delta} \cdot 10^\tau$. Προφανώς, ο εκθέτης του q_n^δ είναι $\rho + \delta = \rho + (\tau - \rho) = \tau$.

Παράδειγμα:

Έστω ότι $n = 8$, $p_n = 4.5672304 \cdot 10^1$ και $q_n = 3.0634568 \cdot 10^{-3}$. Τότε:

$$p_n = 4.5672304 \cdot 10^1$$

$$q_n^\delta = 0.0003063 \cdot 10^1$$

Τότε η αφαίρεση $s_n = p_n - q_n$ επιτελείται ως:

$$s_n = p_n - q_n = p_n - q_n^\delta \Leftrightarrow$$

$$s_n = (4.5672304 - 0.0003063) \cdot 10^1$$

Γεγονός που σημαίνει ότι η αφαίρεση αριθμών με άνισο εκθέτη, ανάγεται πάντα σε αφαίρεση αριθμών με τον ίδιο εκθέτη.

4.2.1 Τυχαίο λάθος πεπερασμένης: Συμβολισμοί, βασικές έννοιες και υπολογισμός των θεωρητικών πιθανοτήτων γένεσης αυτού του λάθους.

Περίπτωση 1

Έστω ότι οι ποσότητες που συμμετέχουν στη πράξη της αφαίρεσης έχουν υπολογιστεί με το ίδιο πλήθος λ_c σωστών δεκαδικών ψηφίων στη mantissa, ή ισοδυνάμως με το ίδιο πλήθος e.d.d. $\lambda = n - \lambda_c$. Τότε,

- a) Η πιθανότητα η ποσότητα $s_n = p_n - q_n$ να έχει $\lambda_c - d - 1$ c.d.d., όπου d η βύθιση του εκθέτη όπως έχει οριστεί στο κεφάλαιο 4.1, συμβολίζεται με P_1^{eq} .
- b) Η πιθανότητα η ποσότητα $s_n = p_n - q_n$ να έχει $\lambda_c - d$ c.d.d., δηλαδή τα σωστά δεκαδικά ψηφία να εξαρτώνται μόνο από τη βύθιση του εκθέτη, συμβολίζεται με P_0^{eq} .
- c) Ομοίως, η ποσότητα s_n έχει $\lambda_c - d + k$ c.d.d, $k > 0$, με πιθανότητα P_{-k}^{eq} . Σε αυτή τη περίπτωση τα σωστά δεκαδικά ψηφία έχουν αυξηθεί με τυχαίο τρόπο κατά ένα παράγοντα k .

Περίπτωση 2

Έστω ότι οι ποσότητες που συμμετέχουν στη πράξη της αφαίρεσης έχουν υπολογιστεί με διαφορετικό πλήθος σωστών δεκαδικών ψηφίων στη mantissa, λ_c και μ_c , όπου χωρίς καμία βλάβη της γενικότητας υποθέτουμε ότι $\lambda_c > \mu_c$ και $d_c = \lambda_c - \mu_c$. Σε αυτή τη περίπτωση οι πιθανότητες που παρουσιάστηκαν παραπάνω εξαρτώνται και από τον παράγοντα d_c . Επομένως:

- a) Η πιθανότητα η ποσότητα $s_n = p_n - q_n$ να έχει $\lambda_c - d - 1$ c.d.d. συμβολίζεται με $P^{un}(1, d_c)$.
- b) Η πιθανότητα η ποσότητα $s_n = p_n - q_n$ να έχει $\lambda_c - d$ c.d.d., συμβολίζεται με $P^{un}(0, d_c)$.
- c) Η ποσότητα s_n έχει $\lambda_c - d + k$ c.d.d, $k > 0$, με πιθανότητα $P^{un}(-k, d_c)$.

□

Όπως αποδεικνύεται στη συνέχεια, το τυχαίο / στοχαστικό λάθος στη πράξη της αφαίρεσης δεν αυξάνει παραπάνω από ένα το πλήθος του σφάλματος πεπερασμένης ακρίβειας.

Ανάλυση της Περίπτωσης 1 όπου οι όροι που αφαιρούνται είναι ίδιας τάξης f.p.e. ($\lambda_c = \mu_c$)

Σε αυτή την ενότητα υπολογίζονται οι πιθανότητες $P_1^{eq}, P_0^{eq}, P_{-k}^{eq}$ που ορίστηκαν παραπάνω. Όπως αποδεικνύεται και στην συνέχεια, η ανισότητα $1 \leq |x - y| < 10$ έχει τη περισσότερη πιθανότητα να ισχύει. Παρόλο αυτά όλες οι ανισότητες που αναφέρονται στην ανάλυση που ακολουθεί πρέπει να λαμβάνονται υπόψη, διότι οι mantissae x, y δεν έχουν συγκεκριμένο πρόσημο και οι απόλυτες τιμές τους κυμαίνονται τυχαία στο διάστημα $[1,10)$. Επομένως:

Έστω η αφαίρεση $s_n = p_n - q_n$. Αν p^c και q^c είναι οι σωστές τιμές των p_n και q_n , ε ο εκθέτης του αποτελέσματος ($p^c - q^c$), μικρότερο από το κοινό εκθέτη ζ των p_n και q_n , τότε $d = \zeta - \varepsilon$ είναι η βύθιση του εκθέτη. Από τους τύπους (4.2) και (4.3), όσον αφορά τη διαφορά των mantissae $|x - y|$, διακρίνονται οι εξής περιπτώσεις:

- Έστω ότι ισχύει $10 \leq |x - y|$. Τότε στο αποτέλεσμα της αφαίρεσης εμφανίζεται ένα επιπλέον λανθασμένο δεκαδικό ψηφίο, δηλαδή το λάθος πεπερασμένης ακρίβειας της διαφοράς s_n αυξάνεται. Αυτό συμβαίνει διότι:

$$s_n = p_n - q_n = (p^c - q^c) + (x - y)10^{-\lambda_c} = (p^c - q^c) + \text{man}(x - y) \cdot 10^1 \cdot 10^{-\lambda_c} =$$

$$s_n = (p^c - q^c) + \text{man}(x - y)10^{-(\lambda_c - 1)}$$

Τότε, σύμφωνα με τους ορισμούς 1 και 2 συνάγεται ότι η διαφορά $x - y$ έχει υπολογιστεί με $\lambda_c - 1$ c.d.d. ή ισοδυνάμως με ένα περισσότερο λανθασμένο δεκαδικό ψηφίο από ότι οι αφαιρετέοι p_n και q_n . Τότε η mantissa που παράγεται έχει $d + 1$ ψηφία που συμπληρώθηκαν τυχαία, επομένως με πολύ μεγάλη πιθανότητα έχει $d + 1$ λανθασμένα δεκαδικά ψηφία.

- Έστω ότι ισχύει $1 \leq |x - y| < 10$. Σε αυτή τη περίπτωση η τάξη του λάθους δεν αλλάζει και το τελικό λάθος που παράγεται εξαρτάται μόνο από το νομοτελειακό λάθος, δηλαδή από τη βύθιση του εκθέτη d , όπως υποδεικνύει ο τύπος (4.2).
- Έστω ότι ισχύει $10^{-1} \leq |x - y| < 1$. Σε αυτή τη περίπτωση το λάθος πεπερασμένης ακρίβειας ελαττώνεται. Αυτό ισχύει διότι τα x, y έχουν ένα κοινό ψηφίο και επομένως η τάξη της ποσότητας $(x - y)10^{-\lambda_c}$ μικραίνει. Συνεπώς το αποτέλεσμα της αφαίρεσης s_n υπολογίζεται με $\lambda_c - d + 1$ σωστά δεκαδικά ψηφία.
- Έστω ότι ισχύει $10^{-2} \leq |x - y| < 10^{-1}$. Σε αυτή τη περίπτωση το λάθος πεπερασμένης ακρίβειας ελαττώνεται κατά 2. Αυτό ισχύει διότι τώρα:

$$s_n = p_n - q_n = (p^c - q^c) + (x - y)10^{-\lambda_c} = (p^c - q^c) + \text{man}(x - y) \cdot 10^{-2} \cdot 10^{-\lambda_c} =$$

$$s_n = (p^c - q^c) + \text{man}(x - y)10^{-(\lambda_c + 2)}$$

Συνεπώς, το αποτέλεσμα της αφαίρεσης s_n υπολογίζεται με $\lambda_c - d + 2$ σωστά δεκαδικά ψηφία.

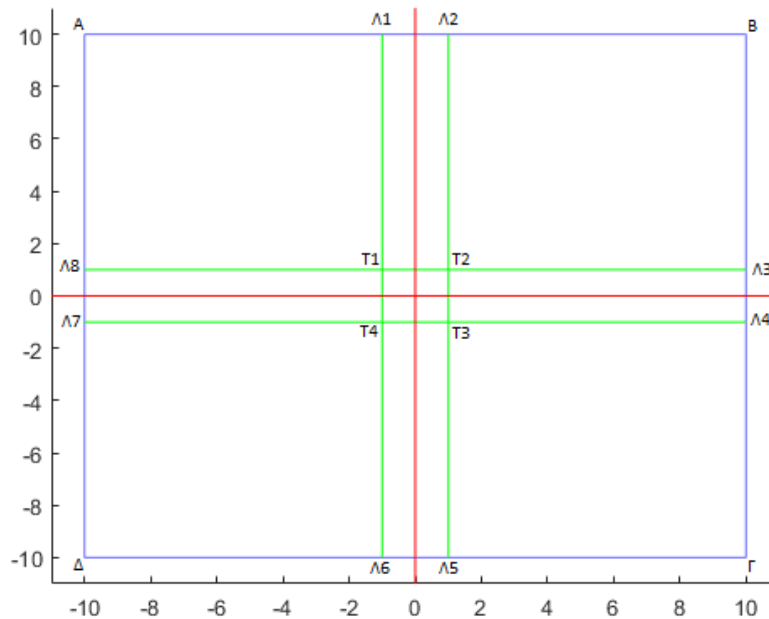
- Στη γενική περίπτωση που ισχύει $10^{-k} \leq |x - y| < 10^{-k+1}$, $k \in \mathbb{N}$, το λάθος πεπερασμένης ακρίβειας της διαφοράς μειώνεται κατά k δεκαδικά ψηφία, διότι όπως προηγουμένως:

$$s_n = p_n - q_n = (p^c - q^c) + \text{man}(x - y)10^{-(\lambda_c + k)}$$

Επομένως, η ποσότητα s_n υπολογίζεται με $\lambda_c + k - d$ σωστά δεκαδικά ψηφία, αρκεί να ισχύει $\lambda_c + k \geq d$. Τελικά, αν ισχύει $\lambda_c - d + k \geq n$, τότε η mantissa της ποσότητας s_n έχει όλα τα δεκαδικά ψηφία σωστά.

Υπολογισμός πιθανοτήτων εμφάνισης συγκεκριμένων ζευγών mantissae (x,y)

Έστω ότι $g_{XY}(x,y)$ είναι η από κοινού συνάρτηση πυκνότητας πιθανότητας γένεσης των σφαλμάτων πεπερασμένης ακρίβειας που εκφράζουν οι mantissae x, y κατά την εκτέλεση μιας αφαίρεσης. Τότε οι πιθανότητες γένεσης συγκεκριμένου αριθμού λανθασμένων δεκαδικών ψηφίων υπολογίζονται γεωμετρικά, μέσω του τετραγώνου Δ , το οποίο φαίνεται στην Εικόνα 4.1 πλευράς μήκους είκοσι (20). Κάθε πλευρά του τετραγώνου αντιστοιχεί στην αλγεβρική τιμή της εκάστοτε mantissa. Το χωρίο $\Delta^R = \{A_2T_2A_3B \cup AA_1T_1A_8 \cup T_3A_4\Gamma A_5 \cup A_7T_4A_6\Delta\}$ αναπαριστά όλους τους πιθανούς συνδυασμούς των mantissae (x,y) των λαθών πεπερασμένης ακρίβειας. Επισημαίνεται ότι το ζεύγος (x,y) δεν μπορεί να βρίσκεται εντός του χωρίου σχήματος σταυρού $A_1A_2T_2A_3A_4T_3A_5A_6T_4A_7A_8T_1A_1$, αφού η κάθε mantissa ανήκει στην ένωση των διαστημάτων $(-10, -1] \cup [1,10)$. Για αυτό το λόγο η από κοινού συνάρτηση πυκνότητας πιθανότητας $g_{XY}(x,y)$ περιορίζεται στο χωρίο Δ^R . Επίσης προφανώς, το περίγραμμα του τετραγώνου δε συμπεριλαμβάνεται στη περιοχή Δ^R που εκφράζει τις πραγματικές τιμές των ζευγών των mantissae (x,y) .



Εικόνα 4.1 Γεωμετρική αναπαράσταση όλων των συνδυασμών των mantissae των σφαλμάτων πεπερασμένης ακρίβειας (x,y) . Δεδομένου ότι αυτοί οι συνδυασμοί δεν ανήκουν στον χωρίο $A_1A_2T_2A_3A_4T_3A_5A_6T_4A_7A_8T_1A_1$, η αντίστοιχη συνάρτηση πυκνότητας πιθανότητας περιορίζεται στο χωρίο $\Delta^R = \{A_2T_2A_3B \cup AA_1T_1A_8 \cup T_3A_4\Gamma A_5 \cup A_7T_4A_6\Delta\}$ της περιμέτρου του τετραγώνου μη συμπεριλαμβανομένης.

Για ευνόητους λόγους θεωρείται ότι η χρησιμοποιούμενη συνάρτηση πυκνότητας πιθανότητας $f_{XY}(x, y)$ είναι ορθώς ανοιγμένη, με την έννοια ότι συνιστά περιορισμό της $g_{XY}(x, y)$ στο χωρίο Δ^R και ότι επί του Δ^R ισχύει η κάτωθι σχέση κανονικοποίησης

$$\iint_{\Delta^R} f_{XY}(x, y) dx dy = 1$$

Σύμφωνα με τα παραπάνω, αρκεί να υπολογιστεί η περιορισμένη συνάρτηση πυκνότητας πιθανότητας $f_{XY}(x, y)$, όπου τα (x, y) ανήκουν στο χωρίο Δ^R του τετραγώνου Δ . Εν συνεχεία σύμφωνα με τη προηγούμενη ανάλυση διακρίνονται τα κάτωθι υποχωρία:

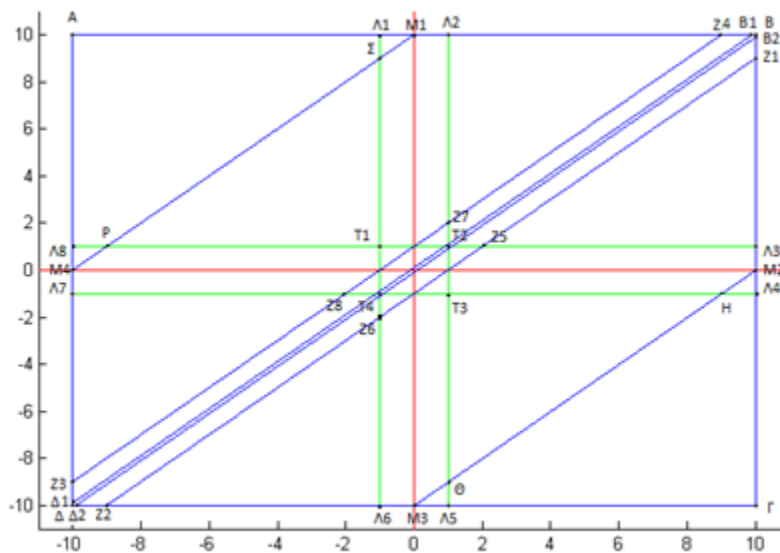
$$\Delta_1 = \{(x, y) \in \Delta : 10 \leq |x - y|\}$$

$$\Delta_0 = \{(x, y) \in \Delta : 1 \leq |x - y| < 10\}$$

$$\Delta_{-1} = \{(x, y) \in \Delta : 10^{-1} \leq |x - y| < 1\}$$

$$\Delta_{-k} = \{(x, y) \in \Delta : 10^{-k} \leq |x - y| < 10^{-k+1}\}$$

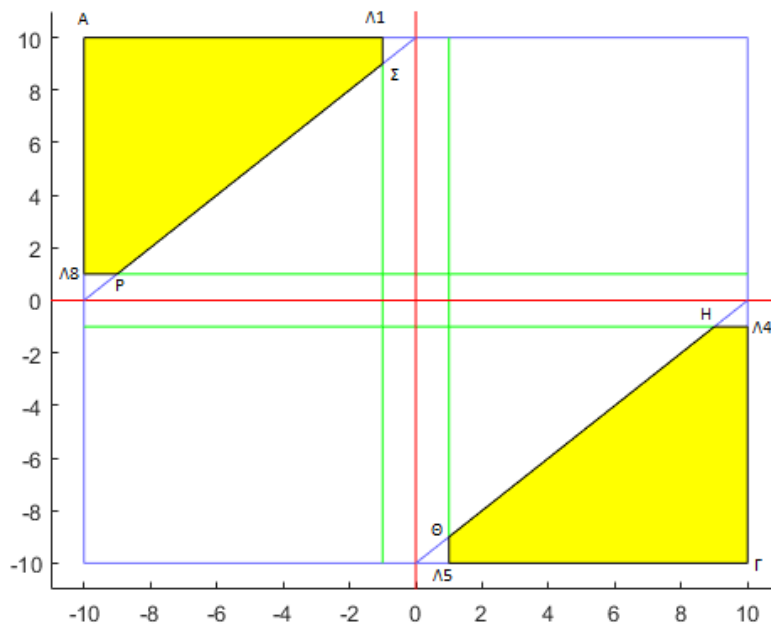
Τα όρια τω ανωτέρω χωρίων φαίνονται στη κάτωθι εικόνα 4.2 .



Εικόνα 4.2 Τα όρια των περιοχών $\Delta_1, \Delta_0, \Delta_{-1}$.

Η πιθανότητα να αυξηθεί κατά ένα το ήδη υπάρχον λάθος πεπερασμένης ακρίβειας κατά την εκτέλεση της αφαίρεσης $s_n = p_n - q_n$, υπολογίζεται για τα ζεύγη που βρίσκονται εντός του χωρίου Δ_1 . Αλλά, τα (x, y) που ανήκουν στο χωρίο Δ_1 κείνται μεταξύ της ευθείας $x - y = 10$ ή της $x - y = -10$ και των πλευρών του τετραγώνου Δ , όπως φαίνεται στην εικόνα 4.3. Δηλαδή, η αφαίρεση των mantissae x, y θα πρέπει να δώσει αποτέλεσμα απολύτως μεγαλύτερου του 10, ώστε ο όρος του σφάλματος στη (4.3) $(x - y)10^{-\lambda_c}$ να αυξηθεί κατά μία τάξη, οπότε και ισχύει η προαναφερθείσα σχέση

$$s_n = p_n - q_n = (p^c - q^c) + \text{man}(x - y)10^{-(\lambda_c - 1)}$$



Εικόνα 4.3 Γεωμετρική αναπαράσταση των συνδυασμών των mantissae των σφαλμάτων πεπερασμένης ακρίβειας (x, y) ανήκουν στο χωρίο Δ_1 , δηλαδή μεταξύ της ευθείας $x - y = 10$ ή της $x - y = -10$ και των πλευρών του τετραγώνου Δ . Στη περίπτωση αυτή το λάθος πεπερασμένης ακρίβειας της διαφοράς αυξάνει κατά ένα σε σχέση με το αιτιοκρατικό λάθος. Δεδομένου ότι αυτοί οι συνδυασμοί δεν ανήκουν στον χωρίο $A_1A_2T_2A_3A_4T_3A_5A_6T_4A_7A_8T_1A_1$, η σχετική πιθανότητα εμφάνισης αυτής της αύξησης f.p.e. υπολογίζεται μέσω του διτλού ολοκληρώματος της συνάρτησης πυκνότητας πιθανότητας στο χωρίο $\Delta_1 = \{H\theta A_5\Gamma A_4H \cup AA_1\Sigma P A_8A\}$

Επομένως η σχετική πιθανότητα υπολογίζεται ως εξής:

$$P_1^{eq} = \iint_{\Delta_1} f_{XY}(x, y) dy dx = \iint_{H\theta A_5\Gamma A_4H} f_{XY}(x, y) dy dx + \iint_{AA_1\Sigma P A_8A} f_{XY}(x, y) dy dx \quad (4.4)$$

Τονίζεται ότι το παραπάνω σφάλμα εν γένει προστίθεται στο σφάλμα που παράγεται εξαιτίας της βύθισης του εκθέτη. Συγκεκριμένα, λόγω της αριστερής ολίσθησης που επιτελεί ο υπολογιστής για να αποκαταστήσει την επιστημονική μορφή, υποχρεωτικά συμπληρώνει **τυχαία** ένα επιπλέον δεκαδικό ψηφίο, το $(d + 1)$ – οστό. Η πιθανότητα το ψηφίο αυτό να συμπληρωθεί τυχαία ορθά θα υπολογιστεί στα επόμενα, αλλά φαίνεται ότι πρακτικά δε παίζει καθόλου σημαντικό ρόλο.

Παράδειγμα 4.2

$$p_n = 5.5312365487\mathbf{96478} \text{ και } p_n^c = 5.5312365487\mathbf{64357}$$

$$q_n = 5.531275555\mathbf{367823} \text{ και } q_n^c = 5.531275555\mathbf{455557}$$

τότε

$$s_n = p_n - q_n = -3.9006\mathbf{57134447982} \cdot 10^{-5}$$

$$s_n^c = p_n^c - q_n^c = -3.9006\mathbf{69119971667} \cdot 10^{-5}$$

Από το παραπάνω παράδειγμα φαίνεται ότι τα p_n και q_n είχαν αρχικά έντεκα σωστά δεκαδικά ψηφία, σύμφωνα με τον ορισμό 1. Το αποτέλεσμα της αφαίρεσης s_n έχει πέντε σωστά δεκαδικά ψηφία. Τα επιπλέον έξι λάθος δεκαδικά ψηφία παρήχθησαν ως εξής. Τα πέντε από αυτά οφείλονται στη βύθιση του εκθέτη αφού $d = \max\{\zeta, \theta\} - \varepsilon = 0 - (-5) = 5$. Το επιπλέον λάθος οφείλεται στο τυχαίο / στοχαστικό λάθος της αφαίρεσης. Πράγματι, το λάθος του τελεστή p_n πριν την αφαίρεση ήταν

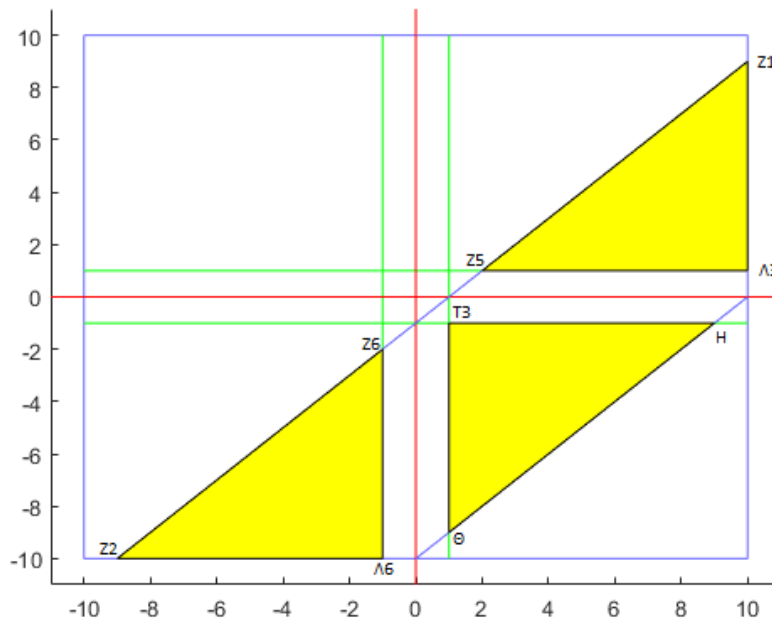
$$x_n = x \cdot 10^{-\lambda_c} = 3.212131231231242 \cdot 10^{-11}$$

ενώ το λάθος του q_n ήταν

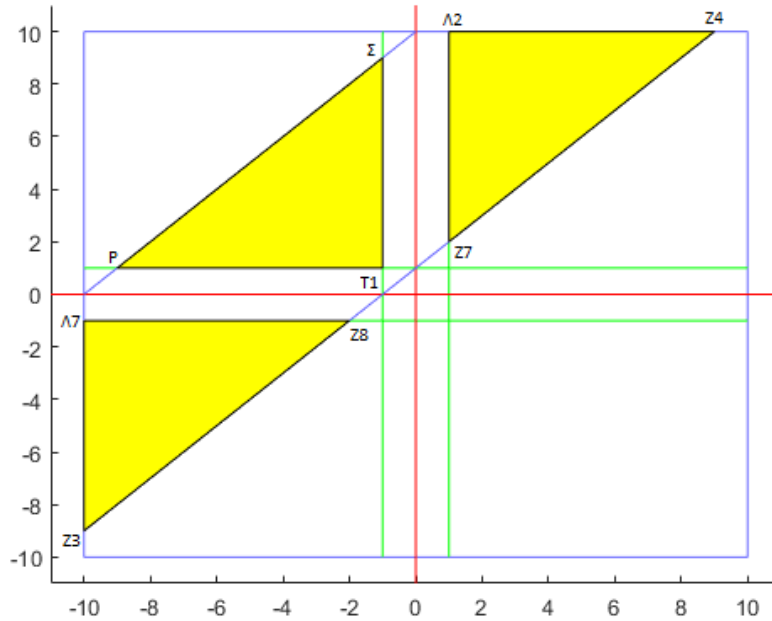
$$y_n = y \cdot 10^{-\mu_c} = -8.773453454353434 \cdot 10^{-11}$$

Συνεπώς το αποτέλεσμα $x_n - y_n = 11.98558468558467 \cdot 10^{-11}$. Η απόλυτη τιμή της mantissa του αποτελέσματος $|x_n - y_n|$ είναι μεγαλύτερη του 10, συνεπώς παράγεται ένα επιπλέον λάθος δεκαδικό ψηφίο όπως επιβεβαιώνει και το παράδειγμα.

Η πιθανότητα να διατηρηθεί μόνο το νομοτελειακό λάθος πεπερασμένης ακρίβειας υπολογίζεται μέσω της απαίτησης τα ζεύγη των mantissae (x, y) να βρίσκονται εντός του Δ_0 . Δηλαδή, τα ζεύγη (x, y) κείνται ή μεταξύ των ευθειών $x - y = 1$ και $x - y = 10$, της δεύτερης ευθείας μη συμπεριλαμβανομένης ή μεταξύ των $x - y = -1$ και $x - y = -10$, της δεύτερης επίσης μη συμπεριλαμβανομένης. Άρα, όπως φαίνεται και στην εικόνα 4.4 και 4.5 η πιθανότητα διατήρησης του λάθους πεπερασμένης ακρίβειας που παράγεται από τη πράξη $s_n = p_n - q_n$ υπολογίζεται με τα (x, y) να βρίσκονται εντός ενός από τα έξι τρίγωνα $Z_1Z_5\Lambda_3$, $T_3H\Theta$, $Z_2Z_6\Lambda_6$, $Z_8Z_3\Lambda_7$, $T_1P\Sigma$, $Z_7Z_3\Lambda_2$. Δηλαδή, η αφαίρεση των mantissae x, y θα πρέπει να δώσει αποτέλεσμα μεταξύ του 1 και μικρότερο του 10, ώστε να διατηρηθεί η τάξη του σφάλματος πεπερασμένης ακρίβειας, μετά τη βύθιση.



Εικόνα 4.4 Γεωμετρική αναπαράσταση των συνδυασμών των mantissae των σφαλμάτων πεπερασμένης ακρίβειας (x, y) που κείνται μεταξύ των ευθειών $x - y = 1$ και $x - y = 10$, της δεύτερης ευθείας μη συμπεριλαμβανομένης. Στη περίπτωση αυτή το λάθος πεπερασμένης ακρίβειας της διαφοράς ισούται με το αιτιοκρατικό λάθος. Δεδομένου ότι αυτοί οι συνδυασμοί δεν ανήκουν στον χωρίο $\Lambda_1\Lambda_2T_2\Lambda_3\Lambda_4T_3\Lambda_5\Lambda_6T_4\Lambda_7\Lambda_8T_1\Lambda_1$, το πρώτο τμήμα της σχετικής πιθανότητα διατήρησης του f.p.e. υπολογίζεται μέσω του διπλού ολοκληρώματος της συνάρτησης πυκνότητας πιθανότητας στο γραμμοσκιασμένο χωρίο.



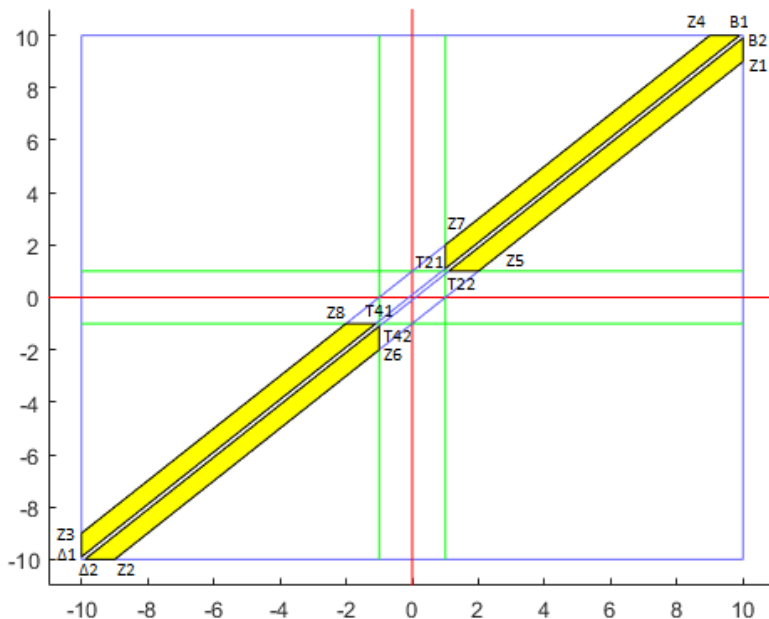
Εικόνα 4.5 Γεωμετρική αναπαράσταση των συνδυασμών των mantissae των σφαλμάτων πεπερασμένης ακρίβειας (x, y) που κείνται μεταξύ των ευθειών $x - y = -1$ και $x - y = -10$, της δεύτερης ευθείας μη συμπεριλαμβανομένης. Στη περίπτωση αυτή το λάθος πεπερασμένης ακρίβειας της διαφοράς ισούται με το αιτιοκρατικό λάθος. Δεδομένου ότι αυτοί οι συνδυασμοί δεν ανήκουν στον χωρίο $A_1A_2T_2A_3A_4T_3A_5A_6T_4A_7A_8T_1A_1$, το δεύτερο τμήμα της σχετικής πιθανότητα διατήρησης του f.p.e. υπολογίζεται μέσω του διπλού ολοκληρώματος της συνάρτησης πυκνότητας πιθανότητας στο γραμμοσκιασμένο χωρίο.

Η πιθανότητα το λάθος πεπερασμένης ακρίβειας της διαφοράς $s_n = p_n - q_n$ να ισούται με το αιτιοκρατικό λάθος είναι:

$$P_0^{eq} = \iint_{A_0} f_{XY}(x, y) dy dx = \iint_{Z_1Z_5A_3} f_{XY}(x, y) dy dx + \iint_{T_3H\theta} f_{XY}(x, y) dy dx + \iint_{Z_2Z_6A_6} f_{XY}(x, y) dy dx + \iint_{Z_8Z_3A_7} f_{XY}(x, y) dy dx + \iint_{T_1P\Sigma} f_{XY}(x, y) dy dx + \iint_{Z_7Z_3A_2} f_{XY}(x, y) dy dx \quad (4.5)$$

Η πιθανότητα το αιτιοκρατικό λάθος της διαφοράς να μειωθεί κατά ένα υπολογίζεται μέσω της απαίτησης τα ζεύγη των mantissae (x, y) να βρίσκονται εντός του Δ_{-1} . Δηλαδή, τα (x, y) κείνται ή μεταξύ των ευθειών $x - y = \frac{1}{10}$ και $x - y = 1$, της δεύτερης ευθείας μη συμπεριλαμβανομένης ή μεταξύ των ευθειών $x - y = -\frac{1}{10}$ και $x - y = -1$, της δεύτερης επίσης μη συμπεριλαμβανομένης. Άρα η

πιθανότητα υπολογίζεται από τα εμβαδά των χωρίων $B_1Z_4Z_7T_{21}$, $B_2Z_1Z_5T_{22}$, $\Delta_1Z_3Z_8T_{41}$, $\Delta_2Z_2Z_6T_{42}$ εξαιρώντας τα χωρία που βρίσκονται μέσα στο “σταυρό” $\Lambda_1\Lambda_2T_2\Lambda_3\Lambda_4T_3\Lambda_5\Lambda_6T_4\Lambda_7\Lambda_8T_1\Lambda_1$, όπως φαίνεται στην εικόνα 4.6



Εικόνα 4.6 Γεωμετρική αναπαράσταση των συνδυασμών των mantissae των σφαλμάτων πεπερασμένης ακρίβειας (x, y) που κείνται ή μεταξύ των ευθειών $x - y = \frac{1}{10}$ και $x - y = 1$, της δεύτερης ευθείας μη συμπεριλαμβανομένης ή μεταξύ των ευθειών $x - y = -\frac{1}{10}$ και $x - y = -1$, της δεύτερης επίσης μη συμπεριλαμβανομένης. Στη περίπτωση αυτή το λάθος πεπερασμένης ακρίβειας της διαφοράς μειώνεται κατά ένα σε σχέση με το αιτιοκρατικό λάθος. Δεδομένου ότι αυτοί οι συνδυασμοί δεν ανήκουν στον χωρίο $\Lambda_1\Lambda_2T_2\Lambda_3\Lambda_4T_3\Lambda_5\Lambda_6T_4\Lambda_7\Lambda_8T_1\Lambda_1$, η σχετική πιθανότητα εμφάνισης αυτής της μείωσης f.p.e. υπολογίζεται μέσω του διπλού ολοκληρώματος της συνάρτησης πυκνότητας πιθανότητας στο γραμμοσκιασμένο χωρίο.

Αναλόγως, και η πιθανότητα να μειωθεί το σφάλμα κατά ένα δίνεται από τον τύπο

$$P_{-1}^{eq} = \iint_{\Delta_{-1}} f_{XY}(x, y) dy dx = \iint_{B_1Z_4Z_7T_{21}} f_{XY}(x, y) dy dx + \iint_{B_2Z_1Z_5T_{22}} f_{XY}(x, y) dy dx + \iint_{\Delta_1Z_3Z_8T_{41}} f_{XY}(x, y) dy dx + \iint_{\Delta_2Z_2Z_6T_{42}} f_{XY}(x, y) dy dx \quad (4.6)$$

Γενικά η πιθανότητα το τελικό σφάλμα να ελαττωθεί κατά $k, k > 0$ υπολογίζεται από το χωρίο Δ_{-k} με εξαίρεση τις περιοχές που βρίσκονται μέσα στο “σταυρό” $\Lambda_1\Lambda_2T_2\Lambda_3\Lambda_4T_3\Lambda_5\Lambda_6T_4\Lambda_7\Lambda_8T_1\Lambda_1$.

Ανάλυση της Περίπτωσης 2 όπου οι όροι που αφαιρούνται έχουν διαφορετικής τάξης f.p.e. ($\lambda_c \neq \mu_c$)

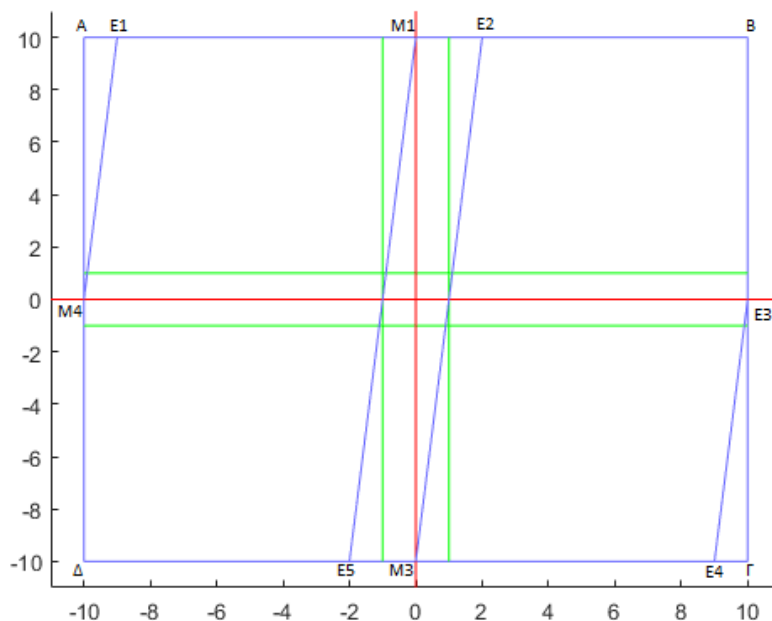
Στη παρούσα έρευνα αναλύεται και η περίπτωση όπου οι δυο όροι που συμμετέχουν στην αφαίρεση έχουν παραχθεί με διαφορετική τάξη f.p.e., δηλαδή έχουν ανόμοιο πλήθος λανθασμένων δεκαδικών ψηφίων στη mantissa. Χωρίς καταστρατήγηση της γενικότητας, μελετάται η χειρότερη περίπτωση όπου τα p_n και q_n είναι ίδιας τάξης, δηλαδή όταν είναι εκφρασμένοι σε επιστημονική μορφή έχουν τον ίδιο εκθέτη. Επειδή δε, σύμφωνα με όλα όσα παρουσιάζονται στη παρούσα εργασία, σε αυτή την περίπτωση, ο εκθέτης δεν παίζει απολύτως κανένα ρόλο στη γένεση του λάθους πεπερασμένης ακρίβειας κατά τη πράξη της αφαίρεσης, για αυτό θα υποθέσουμε ότι ο κοινός εκθέτης των αφαιρετέων είναι μηδέν. Επομένως:

$$p_n = p^c + x10^{-\lambda_c}$$

$$q_n = q^c + y10^{-\mu_c}, \quad \mu_c \neq \lambda_c \tag{4.7}$$

Όπου στη παρούσα ανάλυση υποθέτουμε ότι $\lambda_c < \mu_c$, χωρίς καμία βλάβη της γενικότητας.

Για να γίνει η εκτίμηση των πιθανοτήτων στη προαναφερθείσα περίπτωση, εφαρμόζεται η γεωμετρική προσέγγιση, όπως παρουσιάζεται στις εικόνες 4.7, 4.8 και 4.9.



Εικόνα 4.7 Γεωμετρική αναπαράσταση όλων των συνδυασμών των mantissae και των αντίστοιχων υποχωρίων, των σφαλμάτων πεπερασμένης ακρίβειας (x, y) για τη περίπτωση όπου ο ένας όρος που εισέρχεται στην αφαίρεση έχει ένα λιγότερο σωστό δεκαδικό ψηφίο από τον άλλο όρο, ο οποίος έχει λ_c σωστά δεκαδικά ψηφία.

Συγκεκριμένα, διακρίνονται οι εξής περιπτώσεις

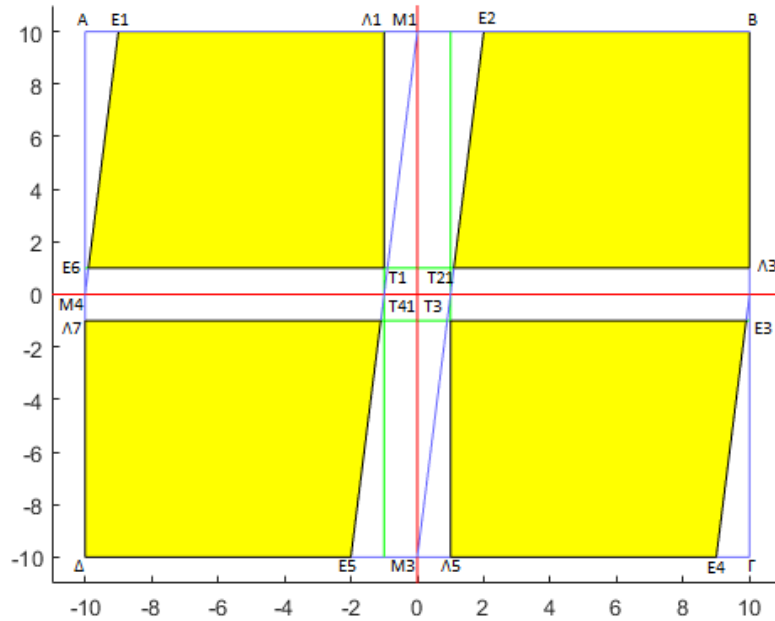
1. Έστω ότι ισχύει $\mu_c = \lambda_c + 1$, δηλαδή ο ένας όρος εισέρχεται στην αφαίρεση με ένα παραπάνω σωστό δεκαδικό ψηφίο από τον άλλο όρο. Τότε ισχύει ότι

$$s_n = p_n - q_n = p^c - q^c + \left(x - \frac{y}{10}\right)10^{-\lambda_c}$$

Η πιθανότητα το s_n να υπολογιστεί με λ_c σωστά δεκαδικά ψηφία προκύπτει από την απαίτηση ότι ισχύει:

$$1 \leq \left|x - \frac{y}{10}\right| < 10 \quad (4.8)$$

Όντως, στην ανωτέρω σχέση (4.8), ο όρος $x - \frac{y}{10}$ υπέχει το ρόλο της mantissa. Άρα, σύμφωνα με τον ορισμό 1 η ποσότητα s_n μετά την αφαίρεση υπολογίζεται με ακριβώς $\lambda_c - d$ (d η βύθιση) σωστά δεκαδικά ψηφία.



Εικόνα 4.8 Το παρόν σχήμα αφορά τη περίπτωση όπου ο ένας όρος που εισέρχεται στην αφαίρεση έχει ένα λιγότερο σωστό δεκαδικό ψηφίο από τον άλλο όρο, ο οποίος έχει λ_c σωστά δεκαδικά ψηφία. Σύμφωνα με όλα τα ανωτέρω, η πιθανότητα το αποτέλεσμα της αφαίρεσης να υπολογιστεί με ακριβώς $\lambda_c - d$ (d η βύθιση) σωστά δεκαδικά ψηφία υπολογίζεται από το διπλό ολοκλήρωμα πάνω στο γραμμοσκιασμένο χωρίο.

Σύμφωνα με τον ορισμό 2 και την εικόνα 4.8 η ανισότητα (4.8) δίνει δύο συμμετρικά χωρία. Το ένα είναι το $D_{0,1}$, το οποίο περιορίζεται από τις ευθείες $x - \frac{y}{10} = 1$ και $x - \frac{y}{10} = 10$ και ορίζεται ως το χωρίο $E_3E_4\Lambda_5T_3$, όπου το Λ_5T_3 ανήκει στο $D_{0,1}$, ενώ το E_3E_4 όχι, σύμφωνα με τη (4.8). Αντιστοίχως, ορίζεται το συμμετρικό χωρίο που οριοθετείται από τις ευθείες $x - \frac{y}{10} = -1$ και $x - \frac{y}{10} = -10$, το οποίο ονομάζεται $\tilde{D}_{0,1}$. Τότε, η πιθανότητα το s_n να παραχθεί και αυτό με $\lambda_c - d$ ορθά δεκαδικά ψηφία ορίζεται ως

$$P^{un}(0,1) = \iint_{D_{0,1}} f_{XY}(x,y) dx dy + \iint_{\tilde{D}_{0,1}} f_{XY}(x,y) dx dy, \quad (4.9)$$

Όπου $f_{XY}(x,y)$ είναι η από κοινού συνάρτηση πιθανότητας της κατανομής του σφάλματος πεπερασμένης ακρίβειας, όπως περιγράφεται στη Περίπτωση 1. Στον τύπο (4.9) το μηδέν που είναι στη πρώτη παράμετρο συμβολίζει τον αριθμό των σωστών δεκαδικών ψηφίων που

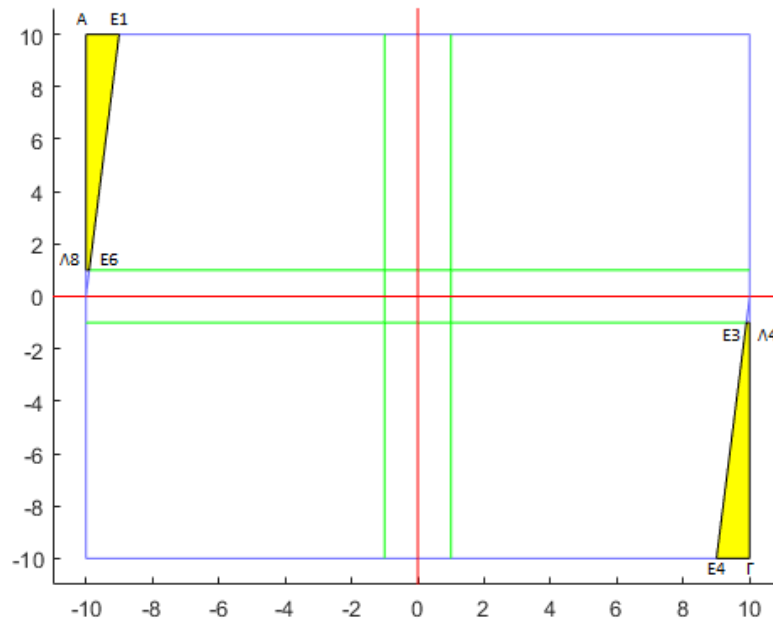
παρέμειναν αναλλοίωτα σε σχέση με το ελάχιστο των $\lambda_c - d, \mu_c - d$. Η δεύτερη παράμετρος συμβολίζει τη διαφορά των $\mu_c - \lambda_c$. Στη προκειμένη περίπτωση ισχύει $\mu_c - \lambda_c = 1$.

2. Έστω εκ νέου ότι ισχύει $\mu_c = \lambda_c + 1$, δηλαδή ο ένας όρος εισέρχεται στην αφαίρεση με ένα παραπάνω σωστό δεκαδικό ψηφίο από τον άλλο όρο. Τότε ισχύει ότι

$$s_n = p_n - q_n = p^c - q^c + \left(x - \frac{y}{10}\right)10^{-\lambda_c}$$

Η πιθανότητα το s_n να υπολογιστεί με $\lambda_c - d - 1$ σωστά δεκαδικά ψηφία ισούται με το να ισχύει ότι

$$10 \leq \left|x - \frac{y}{10}\right| < 11 \tag{4.10}$$



Εικόνα 4.9 Το παρόν σχήμα αφορά τη περίπτωση όπου ο μειωτέος εισέρχεται στην αφαίρεση με ένα λιγότερο σωστό δεκαδικό ψηφίο από τον αφαιρετέος, ο οποίος έχει λ_c σωστά δεκαδικά ψηφία. Σύμφωνα με όλα τα ανωτέρω, η πιθανότητα το αποτέλεσμα της αφαίρεσης να υπολογιστεί με ακριβώς $\lambda_c - d - 1$ σωστά δεκαδικά ψηφία υπολογίζεται από το διπλό ολοκλήρωμα πάνω στο γραμμοσκιασμένο χωρίο.

Πράγματι, σε αυτή τη περίπτωση ισχύει ότι το $x - \frac{y}{10}$ μπορεί να έχει μέγιστη θετική τιμή το 11 ενώ ελάχιστη αρνητική τιμή το -11 , δεδομένου ότι τα x, y είναι και αυτά προσημασμένες mantissae. Στην εικόνα 4.9 η (4.10) αντιστοιχεί σε δύο συμμετρικά χωρία, το $D_{1,1}$, το οποίο περιορίζεται από τις ευθείες $x - \frac{y}{10} = 10$ και $x - \frac{y}{10} = 11$. Όπως και προηγουμένως, ορίζεται το συμμετρικό χωρίο που οριοθετείται από τις ευθείες $x - \frac{y}{10} = -11$ και $x - \frac{y}{10} = -10$, το οποίο συμβολίζεται ως $\tilde{D}_{1,1}$. Επομένως η πιθανότητα να αυξηθεί κατά ένα το τυχαίο λάθος πεπερασμένης ακρίβειας κατά την αφαίρεση δίνεται από το διπλό ολοκλήρωμα

$$P^{un}(1,1) = \iint_{D_{1,1}} f_{XY}(x,y) dx dy + \iint_{\tilde{D}_{1,1}} f_{XY}(x,y) dx dy \quad (4.11)$$

Όπου, σε πλήρη συμφωνία με τη προηγούμενη περίπτωση, η πρώτη παράμετρος εντός της παρενθέσεως συμβολίζει την αύξηση του τυχαίου λάθους πεπερασμένης ακρίβειας κατά 1 και η δεύτερη παράμετρος συμβολίζει τη διαφορά $\mu_c - \lambda_c = 1$.

3. Ομοίως υπολογίζονται και οι πιθανότητες $P^{un}(-k, 1)$, $k = 1,2,3,4$, όπου σε αυτή τη περίπτωση ο s_n υπολογίζεται με $\lambda_c + k - d$ σωστά δεκαδικά ψηφία στη mantissa.
4. Επεκτείνοντας τη παραπάνω ανάλυση, ας υποτεθεί ότι οι όροι p_n, q_n που αφαιρούνται εισέρχονται στην αφαίρεση με ℓ διαφορετικά δεκαδικά ψηφία σωστά, δηλαδή έστω ότι $\mu_c - \lambda_c = \ell$. Τότε, με σχετικά ευθύγραμμη επέκταση της προηγούμενης ανάλυσης, είναι δυνατόν να υπολογιστούν οι παρακάτω πιθανότητες
 - a) $P^{un}(1, \ell)$ ο s_n έχει υπολογιστεί με $\lambda_c + 1 - d$ ορθά δεκαδικά ψηφία.
 - b) $P^{un}(0, \ell)$ ο s_n έχει υπολογιστεί με $\lambda_c - d$ ορθά δεκαδικά ψηφία
 - c) $P^{un}(-k, \ell)$ ο s_n έχει υπολογιστεί με $\lambda_c + k - d$ ορθά δεκαδικά ψηφία, $k > 0$.
5. Ενδεικτικά, θα υπολογίσουμε διεξοδικά μόνο τη πιθανότητα $P^{un}(1, \ell)$. Όντως, έστω ότι ισχύει $\mu_c = \lambda_c + \ell$, δηλαδή ο αφαιρετέος εισέρχεται στην αφαίρεση με ℓ περισσότερα σωστά δεκαδικά ψηφία από τον μειωτέο. Τότε ισχύει:

$$p_n = p^c + x \cdot 10^{-\lambda_c}$$

$$q_n = q^c + y \cdot 10^{-(\lambda_c + \ell)} = q^c + \frac{y}{10^\ell} 10^{-\lambda_c}$$

Συνεπώς,

$$s_n = p_n - q_n = p^c - q^c + \left(x - \frac{y}{10^\ell}\right) 10^{-\lambda_c}$$

Έστω τώρα πως ισχύει ότι

$$10 \leq \left|x - \frac{y}{10^\ell}\right|$$

Ισοδυνάμως

$$\begin{cases} x - \frac{y}{10^\ell} \geq 10 \text{ ή} \\ x - \frac{y}{10^\ell} \leq -10 \end{cases}$$

Σε αυτή τη περίπτωση προκύπτει ότι $s_n = p_n - q_n = p^c - q^c + \left(\text{man}\left(x - \frac{y}{10^\ell}\right) \cdot 10^{-\lambda_c}\right)$.

Επειδή όμως η απόλυτη τιμή της mantissa $\in [1,10)$

$$s_n = p_n - q_n = p^c - q^c + \left(\text{man}\left(x - \frac{y}{10^\ell}\right) \cdot 10^{-\lambda_c - 1}\right),$$

Γεγονός που σημαίνει ότι, σύμφωνα με τον ορισμό 1 το αποτέλεσμα της αφαίρεσης $s_n = p_n - q_n$ έχει υπολογιστεί με ένα λιγότερο ορθό δεκαδικό ψηφίο από το λ_c . Άρα, ορίζοντας τα χωρία

$$D_{1,\ell} = \left\{ \text{Τα σημεία που οριοθετούνται από τις ευθείες } x - \frac{y}{10^\ell} = 10 \text{ και } x - \frac{y}{10^\ell} = 10 + 10^{-\ell} \right\}$$

$$\tilde{D}_{1,\ell} = \left\{ \text{Τα σημεία που οριοθετούνται από τις } x - \frac{y}{10^\ell} = -10 \text{ και } x - \frac{y}{10^\ell} = -(10 + 10^{-\ell}) \right\}$$

Συνάγεται ότι η πιθανότητα το αποτέλεσμα της αφαίρεσης s_n να υπολογιστεί σε αυτή τη περίπτωση με ένα επιπλέον δεκαδικό ψηφίο, λόγω τυχαίου λάθους πεπερασμένης ακρίβειας, ισούται με

$$p^{un}(1, \ell) = \iint_{D_{1,\ell}} f_{XY}(x, y) dx dy + \iint_{\bar{D}_{1,\ell}} f_{XY}(x, y) dx dy \quad (4.12)$$

Εκ νέου αντιπροσωπευτικά παραδείγματα τιμών των προαναφερθέντων πιθανοτήτων παρουσιάζονται στη ενότητα 4.2.3 στην οποία επιτελείται σύγκριση των θεωρητικών τιμών αυτών των πιθανοτήτων με αντίστοιχα πειραματικά αποτελέσματα. Για την από κοινού συνάρτηση πυκνότητας πιθανότητας $g_{XY}(x, y)$ υποθέσαμε ότι ισχύουν τα εξής:

Θεωρούμε ότι οι τυχαίες μεταβλητές X και Y είναι στατιστικά ανεξάρτητες, με από κοινού συνάρτηση πυκνότητας πιθανότητας τη κανονική δυο μεταβλητών (bivariate normal distribution of two independent variables). Δηλαδή θεωρήθηκε ότι η $g_{XY}(x, y)$ είναι ίση με:

$$g_{XY}(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} e^{-\frac{(x-\mu_x)^2}{2\sigma_x^2}} e^{-\frac{(y-\mu_y)^2}{2\sigma_y^2}},$$

Όπου μ_x, μ_y οι μέσες τιμές και σ_x, σ_y οι τυπικές αποκλίσεις των ανεξαρτήτων τυχαίων μεταβλητών X, Y . Ειδικότερα, επειδή τους πληθυσμούς των X, Y γεννά η ίδια υπολογιστική μηχανή, έχει υπετέθει πολύ εύλογα ότι $\mu_x = \mu_y = \mu$ και $\sigma_x = \sigma_y = \sigma$. Λόγω αυτού, η $g_{XY}(x, y)$ γράφεται ως

$$g_{XY}(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{((x-\mu)^2 + (y-\mu)^2)}{2\sigma^2}}$$

Υποθέτοντας, επίσης πολύ εύλογα, ότι η μέση τιμή του λάθους πεπερασμένης ακρίβειας είναι μηδέν, γεγονός που επιβεβαίωσαν όλα τα πειράματα που πραγματοποιήθηκαν, η τελική μορφή της από κοινού συνάρτηση πυκνότητας πιθανότητας των ανεξαρτήτων τυχαίων μεταβλητών X, Y λαμβάνει την τελική μορφή

$$g_{XY}(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{(x^2 + y^2)}{2\sigma^2}} \quad (4.13)$$

Οι υποθέσεις αυτές, όπως θα φανεί στην ενότητα 4.2.3 υποστηρίζεται πλήρως από τα πειραματικά αποτελέσματα.

Στο σημείο αυτό πρέπει να τονίζουμε με έμφαση ότι όλοι η προαναφερθείσα διαδικασία θεωρητικού υπολογισμού των πιθανοτήτων γένεσης τυχαίου λάθους πεπερασμένης ακρίβειας, μπορεί εντελώς ευθύγραμμα να επεκταθεί ώστε να συμπεριλάβει και τη περίπτωση των άνισων εκθετών. Η μόνη (μη ουσιώδης) διαφορά είναι ότι σε όλες τις εξισώσεις των προαναφερθεισών περιπτώσεων, ο όρος y αντικαθίσταται ευθύγραμμα με τον όρο $y \cdot 10^{\theta-\zeta}$.

4.2.2 Πειραματική εκτίμηση του τυχαίου λάθους πεπερασμένης ακρίβειας κατά την αφαίρεση.

Η προσέγγιση που ακολουθήθηκε για τον έλεγχο των αποτελεσμάτων της προηγούμενης ενότητας 4.2.1 είναι η εξής:

1. Επιλέχθηκε ένα σύνολο $C^{16} M = 5 \cdot 10^{11}$ τυχαίων αριθμών p_i^{16} , $i = 1, 2, 3, \dots, M$, εκπεφρασμένων σε μορφή IEEE 754. Αυτοί οι αριθμοί έπαιξαν το ρόλο των σωστών τιμών των προβολών της θεωρητικής άπειρης ακρίβειας p^c . Όλοι οι αριθμοί p_i^{16} είχαν κοινό εκθέτη, έστω τ .
2. Εν συνεχεία, οι αριθμοί αυτού του συνόλου “μολύνθηκαν” με ένα λάθος που ακολουθεί κανονική κατανομή, με μέση τιμή μηδέν και για διάφορες τυπικές αποκλίσεις σ . Με τον όρο “μολύνθηκαν” εννοούμε ότι για κάθε τυπική απόκλιση παράχθηκε ένα λάθος ω_i^N καθώς και ένα νέο σύνολο αριθμών της μορφής $v_i^{16} = p_i^{16} + \omega_i^N$, οι οποίοι αποτελούν τις εκδοχές των “καθαρών” αριθμών p_i^{16} που μολύνθηκαν ο καθένας με το δικό του λάθος πεπερασμένης ακρίβειας ω_i^N .
3. Στη συνέχεια εκτελέστηκαν $1.5 \cdot 10^{13}$ τυχαία επιλεγμένες αφαιρέσεις. Δηλαδή επελέγησαν $1.5 \cdot 10^{13}$ ζεύγη μολυσμένων αριθμών (v_i^{16}, v_j^{16}) και εκτελέστηκαν οι αντίστοιχες αφαιρέσεις $B_{i,j}^{16} = v_i^{16} - v_j^{16}$, $i < j$ διατηρώντας την ακρίβεια στα 16 δεκαδικά ψηφία.
4. Εκτελέστηκαν οι ίδιες ακριβώς αφαιρέσεις μεταξύ των αριθμών p_i^{16}, p_j^{16} , οι οποίοι αντιστοιχούν μονοσήμαντα στους v_i^{16}, v_j^{16} . Δηλαδή, επιτελέστηκαν οι αφαιρέσεις: $A_{i,j}^{16} = p_i^{16} - p_j^{16}$, $i < j$.
5. Πραγματοποιήθηκε η σύγκριση κάθε $A_{i,j}^{16}$ με το αντίστοιχο $B_{i,j}^{16}$ επί τη βάση του ορισμού 1. Επελέγησαν μόνο τα ζεύγη αριθμών $A_{i,j}^{16}, B_{i,j}^{16}$ τα οποία είχαν τα τρία τουλάχιστον περισσότερα σημαντικά δεκαδικά ψηφία κοινά. Εν συνεχεία, υπολογίστηκε το αιτιοκρατικό λάθος πεπερασμένης ακρίβειας, μέσω της βύθισης των εκθετών $d = \max(E(v_i^{16}), E(v_j^{16})) - E(B_{i,j}^{16})$. Εν τέλει, υπολογίστηκε το τυχαίο λάθος.

6. Τέλος πραγματοποιήθηκε στατιστική ανάλυση των λαθών.

Ακολουθώντας την παραπάνω διαδικασία, πραγματοποιήθηκε σύγκριση των θεωρητικών πιθανοτήτων που παρουσιάστηκαν στο κεφάλαιο 4.2.1 με τις αντίστοιχες τις πειραματικές συχνότητες. Τα σχετικά αποτελέσματα φαίνονται στους πίνακες 4.1 ,4.2 , 4.3 , 4.4 και 5.1 του κεφαλαίου.

4.2.3 Σύγκριση θεωρητικών προβλέψεων και πειραματικών δεδομένων.

Περίπτωση 1: Οι όροι που αφαιρούνται εισέρχονται στην αφαίρεση με ίδιας τάξης f.p.e. ($\lambda_c = \mu_c$)

A. Σύγκριση με βάση την υπόθεση ότι οι τυχαίες μεταβλητές X,Y έχουν από κοινού συνάρτηση πυκνότητας πιθανότητας που είναι μια bi variate normal distribution.

Υπενθυμίζεται ότι οι τυχαίες μεταβλητές X,Y εκφράζουν τις mantissae του λάθους πεπερασμένης ακρίβειας με το οποίο οι δυο αφαιρετέοι εισέρχονται στην αφαίρεση. Στον πίνακα 4.1 οι θεωρητικές τιμές των σχετικών πιθανοτήτων υπολογίστηκαν με αριθμητική ολοκλήρωση πάνω στα χωρία που προαναφέρθηκαν στην ενότητα 4.2.1 και για δύο διαφορετικές τιμές του σ . Όπως θα φανεί και στη συνέχεια, το αρχικό σφάλμα πεπερασμένης ακρίβειας ακολουθεί κανονική κατανομή με μηδενική μέση τιμή και τυπική απόκλιση $\sigma 1.45 \pm 0.31$. Ο πίνακας 4.1 θεμελιώνει ότι η θεωρητική προσέγγιση που προηγήθηκε δίνει αποτελέσματα εξαιρετικά κοντά στα πειραματικά, γεγονός που σαφέστατα υποστηρίζει την ορθότητα αυτών.

Πλήθος e.d.d.	Τιμές των θεωρητικών πιθανοτήτων P^{eq} κατά την αφαίρεση σε σύγκριση με τις αντίστοιχες πειραματικές συχνότητες, για δύο τυπικές αποκλίσεις σ του σφάλματος πεπερασμένης ακρίβειας που θεωρήθηκε ότι ακολουθεί bi-variate normal distribution.			
	$\sigma = 1, 2$		$\sigma = 3, 2$	
	Θεωρητικές Πιθαν.	Πειραματικές Συχν.	Θεωρητικές Πιθαν.	Πειραματικές Συχν.
Αύξηση κατά ένα	$2.3430 \cdot 10^{-6}\%$	$7.8 \cdot 10^{-7}\%$	4.34%	5.11%
Διατήρηση	59.71%	59.66%	77.54%	77.28%
Ελάττωση κατά ένα	34.03%	33.82%	15.84%	15.71%
Ελάττωση κατά δυο	5.57%	5.83%	1.72%	1.78%
Ελάττωση κατά τρία	0.18%	0.61%	0.05%	0.002%
Ελάττωση κατά τέσσερα	0.01%	0.06%	0.005%	0.0006%

Πίνακας 4.1 Τιμές των θεωρητικών πιθανοτήτων P^{eq} κατά την αφαίρεση σε σύγκριση με τις αντίστοιχες πειραματικές συχνότητες, για δύο τυπικές αποκλίσεις σ των ανεξαρτήτων τυχαίων μεταβλητών X και Y. Η θεωρητική προσέγγιση που προηγήθηκε δίνει αποτελέσματα εξαιρετικά κοντά στα πειραματικά, γεγονός που σαφέστατα υποστηρίζει την ορθότητα αυτής.

Σημειώνεται ότι όπως θα φανεί και στα επόμενα, τα αρχικά σφάλματα πεπερασμένης ακρίβειας x, y ακολουθούν στατιστικώς ανεξάρτητες κανονικές κατανομές X, Y με μέση τιμή $\mu = 0$ και τυπική απόκλιση $\sigma = 1.45 \cdot \sqrt{2}$. Για την αντίστοιχη από κοινού συνάρτηση πυκνότητας πιθανότητας $f_{XY}(x, y)$ οι αντίστοιχες θεωρητικές πιθανότητες που αφορούν το τυχαίο λάθος πεπερασμένης ακρίβειας είναι οι κάτωθι:

1. $P_1^N = 0.0014$.
2. $P_0^N = 0.7338$.
3. $P_{-1}^N = 0.2361$.
4. $P_{-2}^N = 0.0271$.
5. $P_{-3}^N = 0.0017$.
6. $P_{-4}^N = 1.7303 \cdot 10^{-4}$.

B. Σύγκριση με βάση την υπόθεση ότι οι τυχαίες μεταβλητές X, Y έχουν ομοιόμορφη από κοινού συνάρτηση πυκνότητας πιθανότητας.

Στον πίνακα 4.2 έχει υποτεθεί ότι η $g_{XY}(x, y)$ και συνεπώς ο περιορισμός αυτής $f_{XY}(x, y)$ αντιστοιχούν σε ομοιόμορφη κατανομή. Λόγω της απαίτησης της κανονικοποίησης

$$\iint_{\Delta^R} f_{XY}(x, y) dx dy = 1$$

συνάγεται άμεσα ότι $f_{XY}(x, y) = \frac{1}{(4 \cdot 81)}$.

Επί τη βάση αυτών, κάτωθι πίνακας επιβεβαιώνει για μία φορά ακόμα την εξαιρετική προσέγγιση μεταξύ της θεωρίας που παρουσιάζεται εδώ και της πράξης.

Πλήθος e.d.d.	Τιμές των θεωρητικών πιθανοτήτων P^{eq} κατά την αφαίρεση σε σύγκριση με τις αντίστοιχες πειραματικές συχνότητες. Οι ανεξάρτητες τυχαίες μεταβλητές X, Y ακολουθούν την ίδια ομοιόμορφη κατανομή, κανονικοποιημένη στο χωρίο Δ^R .	
	Θεωρητικές Πιθαν.	Πειραματικές Συχν.
Αύξηση κατά ένα	30.22%	30.66%
Διατήρηση	59.28%	58.77%
Ελάττωση κατά ένα	9.30%	9.52%

Ελάττωση κατά δυο	0.44%	0.2%
Ελάττωση κατά τρία	$5 \cdot 10^{-4}$	$6 \cdot 10^{-4}\%$

Πίνακας 4.2 Τιμές των θεωρητικών πιθανοτήτων P^{eq} κατά την αφαίρεση σε σύγκριση με τις αντίστοιχες πειραματικές συχνότητες. Οι ανεξάρτητες τυχαίες μεταβλητές X, Y ακολουθούν την ίδια ομοιόμορφη κατανομή. Η θεωρητική προσέγγιση που προηγήθηκε δίνει αποτελέσματα εξαιρετικά κοντά στα πειραματικά, γεγονός που σαφέστατα υποστηρίζει την ορθότητα αυτών.

Περίπτωσης 2: Οι όροι που αφαιρούνται εισέρχονται στην αφαίρεση με διαφορετική τάξης f.p.e. ($\mu_c = \lambda_c + 1$)

Πλήθος e.d.d.	Τιμές των θεωρητικών πιθανοτήτων P^{un} κατά την αφαίρεση σε σύγκριση με τις αντίστοιχες πειραματικές συχνότητες, για κανονικές κατανομές και για δύο διαφορετικά σ , όταν ισχύει $\mu_c = \lambda_c + 1$			
	$\sigma = 1.2$		$\sigma = 3.2$	
	Θεωρητικές Πιθαν.	Πειραματικές Συχν.	Θεωρητικές Πιθαν.	Πειραματικές Συχν.
Διατήρηση	91.02%	90.92%	95.01%	95.02%
Ελάττωση κατά ένα	8.97%	9.08%	4.98%	4.92%
Ελάττωση κατά δυο	$1.6557 \cdot 10^{-13}\%$	$8.05 \cdot 10^{-10}\%$	$2.9 \cdot 10^{-7}\%$	$4.09 \cdot 10^{-7}\%$
Ελάττωση κατά τρία	$1.18 \cdot 10^{-14}\%$	$2.4 \cdot 10^{-11}\%$	$4.1 \cdot 10^{-9}\%$	$5.67 \cdot 10^{-9}\%$

Πίνακας 4.3 Τιμές των θεωρητικών πιθανοτήτων P^{un} κατά την αφαίρεση σε σύγκριση με τις αντίστοιχες πειραματικές συχνότητες για δύο διαφορετικά σ , των ανεξαρτήτων τυχαίων μεταβλητών X και Y . Η θεωρητική προσέγγιση που προηγήθηκε δίνει αποτελέσματα εξαιρετικά κοντά στα πειραματικά, γεγονός που σαφέστατα υποστηρίζει την ορθότητα αυτών.

Πλήθος e.d.d.	Τιμές των θεωρητικών πιθανοτήτων P^{un} κατά την αφαίρεση σε σύγκριση με τις αντίστοιχες πειραματικές συχνότητες. Οι τυχαίες μεταβλητές X, Y ακολουθούν ομοιόμορφη κατανομή.	
	Θεωρητικές Πιθαν.	Πειραματικές Συχν.
Διατήρηση	93.81%	94.02%
Ελάττωση κατά ένα	6.07%	5.91%
Ελάττωση κατά δυο	$6 \cdot 10^{-6}\%$	$6.4 \cdot 10^{-6}\%$
Ελάττωση κατά τρία	$6 \cdot 10^{-8}\%$	$6.38 \cdot 10^{-8}\%$

Πίνακας 4.4 Τιμές των θεωρητικών πιθανοτήτων P^{un} κατά την αφαίρεση σε σύγκριση με τις αντίστοιχες πειραματικές συχνότητες, Οι ανεξάρτητες τυχαίες μεταβλητές X, Y ακολουθούν την ίδια ομοιόμορφη κατανομή. Η θεωρητική προσέγγιση που προηγήθηκε δίνει αποτελέσματα εξαιρετικά κοντά στα πειραματικά, γεγονός που σαφέστατα υποστηρίζει την ορθότητα αυτής.

Όπως παρατηρείται από τους παραπάνω πίνακες, η εγγύτητα των θεωρικών προβλέψεων και των πειραματικών δεδομένων είναι εξαιρετική.

4.3 Το ολικό σφάλμα πεπερασμένης ακρίβειας στη πράξη της αφαίρεσης

Η ανάλυση και τα αποτελέσματα που παρουσιάζονται στα κεφάλαια 4.1 και 4.2 συνοψίζονται στη παρακάτω πρόταση

ΠΡΟΤΑΣΗ Α.4.2

Έστω ότι όλες οι ποσότητες υπολογίζονται με ακρίβεια n δεκαδικών ψηφίων στη mantissa και έστω μια οποιαδήποτε ποσότητα s_n που υπολογίζεται ως εξής:

$$s_n = p_n - q_n, \quad p_n \cdot q_n > 0$$

Όπου, όπως πάντα

$$s_n = \gamma_1 \cdot \gamma_2 \dots \gamma_n \cdot 10^\varepsilon, \quad p_n = a_1 \cdot a_2 \dots a_n \cdot 10^\zeta, \quad q_n = \beta_1 \cdot \beta_2 \dots \beta_n \cdot 10^\theta$$

Χωρίς καμία βλάβη της γενικότητας, υποθέτουμε ότι ισχύει η ανίσωση $\zeta > \theta$. Τότε εξετάζουμε τις περιπτώσεις που παρουσιάζονται κατωτέρω, ανάλογα με την τάξη του λάθους πεπερασμένης ακρίβειας με το οποίο οι ποσότητες p_n, q_n έχουν “μολυνθεί” από όλους τους προηγούμενους υπολογισμούς.

Έστω επίσης ότι, εξαιτίας προηγούμενων υπολογισμών, η υψηλότερης τάξης ποσότητα p_n που συμμετέχει στην αφαίρεση έχει υπολογιστεί με λ_c σωστά / $\lambda = n - \lambda_c$ λάθος δεκαδικά ψηφία στη mantissa, ενώ ο άλλος όρος q_n έχει υπολογιστεί με μ_c σωστά / $\mu = n - \mu_c$ λάθος δεκαδικά ψηφία στη mantissa. Σημειώνεται ότι, όταν οι δυο τελεστές q_n, p_n έχουν τον ίδιο εκθέτη, τότε οποιοσδήποτε όρος εκ των δυο μπορεί να αντιπροσωπεύει την ποσότητα με την υψηλότερη τάξη. Συνεπώς, διακρίνονται οι εξής περιπτώσεις:

1. $\lambda_c \leq \mu_c$, δηλαδή η μεγαλύτερης τάξης ποσότητα p_n έχει το ίδιο ή λιγότερο πλήθος σωστών δεκαδικών ψηφίων από τη q_n ποσότητα. Αφού ισχύει ότι $\varepsilon \leq \max\{\zeta, \theta\}$, ορίζεται $d = \max\{\zeta, \theta\} - \varepsilon$ ως η βύθιση του εκθέτη. Τότε, η ποσότητα s_n υπολογίζεται με τα πρώτα $(\lambda_c - d)$ δεκαδικά ψηφία σωστά, λόγω του νομοτελειακού λάθους. Πάνω σε αυτό προστίθεται και το τυχαίο / στοχαστικό λάθος σύμφωνα με τις πιθανότητες που παρουσιάζονται στο κεφάλαιο 4.2 (Είδη σχετικά παραδείγματα 4.3 και 4.4).

2. $\lambda_c > \mu_c$, δηλαδή η υψηλότερης τάξης ποσότητα έχει παραπάνω σωστά δεκαδικά ψηφία. Σε αυτή τη περίπτωση, αφού ισχύει $\varepsilon \leq \max\{\zeta, \theta\}$, διακρίνονται οι παρακάτω δυο υποπεριπτώσεις:
- Αν ισχύει $\lambda_c \leq \mu_c + \max\{\zeta, \theta\} - \min\{\zeta, \theta\}$ τότε η s_n υπολογίζεται με $(\lambda_c - d)$ σωστά δεκαδικά ψηφία εξαιτίας του νομοτελειακού λάθους. Το στοχαστικό λάθος υπολογίζεται σύμφωνα με την ανάλυση στο κεφάλαιο 4.2.
 - Αν ισχύει $\lambda_c > \mu_c + \max\{\zeta, \theta\} - \min\{\zeta, \theta\}$ τότε η ποσότητα s_n εξαιτίας του νομοτελειακού λάθους έχει $e = \mu_c - (\max\{\zeta, \theta\} - \min\{\zeta, \theta\}) - d$ ορθά δεκαδικά ψηφία στη mantissa. Επομένως, η s_n παράγεται με $n - e$ λάθος δεκαδικά ψηφία στη mantissa, λόγω του αιτιοκρατικού λάθους. Πάνω σε αυτό προστίθεται και το τυχαίο / στοχαστικό λάθος σύμφωνα με τις πιθανότητες που παρουσιάζονται στο κεφάλαιο 4.2. Τα παραδείγματα 4.5 και 4.6 επιχειρούν να διασαφηνίσουν και αυτή τη περίπτωση.

□

ΠΑΡΑΔΕΙΓΜΑΤΑ

Παράδειγμα 4.3

Έστω ότι

$$p_n = 1.111112 \cdot 10^{-1} \text{ και } p_n^c = 1.113333 \cdot 10^{-1}$$

$$q_n = 4.222235 \cdot 10^{-2} \text{ και } q_n^c = 4.222345 \cdot 10^{-2}$$

Τότε

$$s_n = p_n - q_n = 6.888885 \cdot 10^{-2} \text{ και } s_n^c = 6.910985 \cdot 10^{-2}$$

Η ποσότητα s_n έχει υπολογιστεί με $\lambda_c - 1 = 2$ c. d. d., αφού $\zeta = -1, \theta = -2, \varepsilon = -2, d = \max\{\zeta, \theta\} - \varepsilon = -1 + 2 = 1$ (βύθιση).

Παράδειγμα 4.4

Έστω ότι

$$p_n = 5.11456112 \cdot 10^{-2} \text{ και } p_n^c = 5.11456001 \cdot 10^{-2}$$

$$q_n = 5.11121555 \cdot 10^{-2} \text{ και } q_n^c = 5.11126666 \cdot 10^{-2}$$

Τότε

$$s_n = p_n - q_n = 3.34557000 \cdot 10^{-5} \text{ και } s_n^c = 3.29334500 \cdot 10^{-5}$$

Η ποσότητα s_n έχει υπολογιστεί με $\lambda_c - 3 = 3$ c.d.d., αφού $\zeta = -2, \theta = -2, \varepsilon = -5, d = \max\{\zeta, \theta\} - \varepsilon = -2 + 5 = 3$ (βύθιση).

Παράδειγμα 4.5

Έστω ότι

$$p_n = 3.133312 \cdot 10^{-1} \text{ και } p_n^c = 3.133301 \cdot 10^{-1}$$

$$q_n = 5.214235 \cdot 10^{-2} \text{ και } q_n^c = 5.211111 \cdot 10^{-2}$$

Τότε

$$s_n = p_n - q_n = 2.611889 \cdot 10^{-1} \text{ και } s_n^c = 2.612190 \cdot 10^{-1}$$

Η ποσότητα s_n έχει υπολογιστεί με $e = 4$ c. d. d., αφού $\zeta = -1, \theta = -2, \varepsilon = -1, d = 0$.

Παράδειγμα 4.6

Έστω ότι

$$p_n = 5.12345678 \cdot 10^5$$

$$q_n = 4.12356567 \cdot 10^3$$

Για την αφαίρεση το p_n παραμένει ως έχει

$$p_n = 5.12345678 \cdot 10^5$$

Το q_n θα υποστεί δεξιά ολίσθηση και θα γίνει

$$q_n^* = 0.04123565 \cdot 10^5$$

Για να γίνει η αφαίρεση πρέπει ο q_n να υποστεί δεξιά ολίσθηση $\alpha = \zeta - \theta = 5 - 3 = 2$

Επομένως η αφαίρεση γίνεται μεταξύ p_n και q_n^* . Η ποσότητα $s_n = p_n - q_n^*$ έχει υπολογιστεί με $e = \mu_c - a - d$ ορθά δεκαδικά ψηφία στη mantissa.

Από όλη τη προηγούμενη ανάλυση, αλλά και όλη αυτή που θα επακολουθήσει, συνάγεται το κάτωθι πολύ σημαντικό συμπέρασμα.

ΠΡΟΤΑΣΗ Α.4.3 ΚΑΙ ΣΧΕΤΙΚΗ ΟΡΟΛΟΓΙΑ

Το πλήθος των σωστών δεκαδικών ψηφίων που παράγεται κατά την αφαίρεση $s_n = p_n - q_n$ είναι πρακτικά ανεξάρτητο από το μήκος λέξης της mantissa, αρκεί τα s_n, p_n, q_n να μην έχουν καταστραφεί. Ισοδυνάμως, έστω ότι p_l και q_l είναι οι προβολές των p^c και q^c σε l δεκαδικά ψηφία και p_m, q_m είναι οι προβολές των p^c και q^c σε m δεκαδικά ψηφία, με $l < m$. Τότε το σφάλμα πεπερασμένης ακρίβειας των ποσοτήτων s_l και s_m έχει τις ίδιες στατιστικές ιδιότητες. \square

5 Η γένεση και διάδοση του τυχαίου / στοχαστικού λάθους πεπερασμένης ακρίβειας σε ένα αριθμό διαδοχικών αφαιρέσεων.

Έστω ότι η διαδικασία των διαδοχικών αφαιρέσεων ξεκινά με ένα αρχικό σύνολο αριθμών S^0 , όπου οι αριθμοί είναι διαφορετικοί μεταξύ τους και όλοι διατυπωμένοι σε επιστημονική μορφή. Έστω επίσης ότι πραγματοποιούνται μόνο οι αφαιρέσεις $p_n - q_n$, θεωρώντας ότι η $q_n - p_n$ έχει ισοδύναμη συμπεριφορά, p_n, q_n ανήκουν στο S^0 και $p_n \neq q_n$. Ως συνήθως, ο κάτω δείκτης n συμβολίζει το χρησιμοποιούμενο μήκος λέξης. Τα αποτελέσματα των αφαιρέσεων συνθέτουν ένα νέο σύνολο αριθμών S^1 . Στη συνέχεια πραγματοποιούνται οι αφαιρέσεις $p_n - q_n$, όπου όμως τώρα τα p_n, q_n λαμβάνονται από το σύνολο S^1 . Με αυτό τον τρόπο δημιουργείται ένα νέο σύνολο S^2 το οποίο περιλαμβάνει όλα τα νέα αποτελέσματα των νέων αφαιρέσεων. Συνεχίζοντας τη παραπάνω διαδικασία των διαδοχικών αφαιρέσεων, έχοντας το τρέχον σύνολο S^γ ως είσοδο παράγεται το σύνολο $S^{\gamma+1}$, όπου γ ο δείκτης της εκάστοτε διαδοχικής επανάληψης.

ΟΡΙΣΜΟΣ 5.1

Έστω ότι έχει πραγματοποιηθεί ένα πλήθος γ διαδοχικών αφαιρέσεων. Τότε, συμβολίζεται με $P^S(\lambda, \gamma)$ η πιθανότητα το αποτέλεσμα μίας τυχούσης αφαίρεσης να παράγεται με τα τελευταία λ δεκαδικά ψηφία στη mantissa λανθασμένα. Αν ο αριθμός λ είναι αρνητικός ή μεγαλύτερος του μήκους λέξης, τότε η προαναφερθείσα πιθανότητα είναι εξ' ορισμού μηδέν. Εάν κατά τη γ -οστή διαδοχική αφαίρεση προκύψει πλήθος λανθασμένων ψηφίων λ αρνητικό, τότε αποδίδεται το ενδεχόμενο αυτό στη

περίπτωση όπου $\lambda = 0$. Δυστυχώς, εάν προκύψει λ μεγαλύτερο του μήκους λέξης, τότε αποδίδεται αυτό το γεγονός στο ενδεχόμενο “η γ -οστή αφαίρεση κατέστησε λανθασμένα όλα τα ψηφία της mantissa / της λέξης”. □

Ας υποθεθεί ότι έχουν ήδη πραγματοποιηθεί γ διαδοχικές αφαιρέσεις και ότι όλες οι πιθανότητες $P^S(\lambda, \gamma)$ είναι γνωστές, για όλες τις αφαιρέσεις που πραγματοποιήθηκαν. Έστω επίσης ότι, χωρίς καταστρατήγηση της γενικότητας, το σύνολο S^γ απαρτίζεται από αριθμούς που έχουν τον ίδιο εκθέτη, δεδομένου ότι αυτή είναι η χειρότερη περίπτωση από πλευράς γένεσης και συσσώρευσης λάθους πεπερασμένης ακρίβειας κατά την αφαίρεση. Τότε, πραγματοποιώντας την $\gamma + 1$ -οστή επανάληψη των αφαιρέσεων, εμφανίζεται η ανάγκη να υπολογιστούν οι αντίστοιχες πιθανότητες $P^S(\lambda, \gamma + 1)$. Επομένως, σύμφωνα με το κεφάλαιο 4.2, η ποσότητα $s_n = p_n - q_n$ υπολογίζεται με λ λανθασμένα δεκαδικά ψηφία στη mantissa εάν

- a. Τα p_n, q_n έχουν λ λανθασμένα δεκαδικά ψηφία και ταυτόχρονα το τυχαίο λάθος πεπερασμένης ακρίβειας δεν αυξάνεται λόγω της αφαίρεσης. Όμως, η πιθανότητα το f.p.e. να μην αυξηθεί κατά την αφαίρεση, έχει ήδη συμβολιστεί ως P_0^{eq} , όπου ο άνω δείκτης eq , κατά τα γνωστά, αναφέρεται στο γεγονός ότι οι αφαιρετέες ποσότητες p_n και q_n υπεισέρχονται στην αφαίρεση με ίσο αριθμό λανθασμένων δεκαδικών ψηφίων, εν προκειμένω λ . Η πιθανότητα να συμβεί το τελευταίο είναι, προφανώς, $P^S(\lambda, \gamma)^2$, δεδομένου ότι οι διαδικασίες παραγωγής των p_n, q_n είναι de-facto στατιστικώς ανεξάρτητες. Άρα η συνολική πιθανότητα διατήρησης των λ ψηφίων σε αυτή τη περίπτωση ισούται με το γινόμενο $P^S(\lambda, \gamma)^2 \cdot P_0^{eq}$, αφού εκ νέου η αφαίρεση $s_n = p_n - q_n$ είναι μια νέα στατιστικώς ανεξάρτητη διαδικασία.
- b. Τα p_n, q_n έχουν $\lambda - 1$ λανθασμένα δεκαδικά ψηφία και ταυτόχρονα το τυχαίο λάθος πεπερασμένης ακρίβειας αυξάνεται κατά ένα λόγω της αφαίρεσης. Η πιθανότητα να συμβεί το πρώτο είναι, προφανώς, $P^S(\lambda - 1, \gamma)^2$. Αλλά, η πιθανότητα το f.p.e. να αυξηθεί κατά ένα κατά την αφαίρεση συμβολίζεται ως P_1^{eq} . Άρα η συνολική πιθανότητα διατήρησης των λ ψηφίων σε αυτή τη περίπτωση ισούται με το γινόμενο $P^S(\lambda - 1, \gamma)^2 \cdot P_1^{eq}$.
- c. Τα p_n, q_n έχουν $\lambda + 1$ λανθασμένα δεκαδικά ψηφία και ταυτόχρονα το τυχαίο λάθος πεπερασμένης ακρίβειας μειώνεται κατά ένα λόγω της αφαίρεσης. Η πιθανότητα να συμβεί το πρώτο είναι, $P^S(\lambda + 1, \gamma)^2$, ενώ η πιθανότητα το f.p.e. να μειωθεί κατά ένα κατά την αφαίρεση συμβολίζεται ως P_{-1}^{eq} . Άρα η συνολική πιθανότητα διατήρησης των λ ψηφίων σε αυτή τη περίπτωση ισούται με το γινόμενο $P^S(\lambda + 1, \gamma)^2 \cdot P_{-1}^{eq}$.

- d. Ομοίως τα p_n, q_n έχουν αμφότερα υπολογιστεί με $\lambda + k$ λανθασμένα δεκαδικά ψηφία στη mantissa, $k > 0$, με πιθανότητα $P^S(\lambda + k, \gamma)^2$ και η αφαίρεση μειώνει το πλήθος των λανθασμένων ψηφίων του αποτελέσματος κατά k . Η σχετική συνολική πιθανότητα ισούται τώρα με $P^S(\lambda + k, \gamma)^2 \cdot P_{-k}^{eq}$, $k > 0$.
- e. Η ποσότητα p_n έχει υπολογιστεί με $\lambda + 1$ e.d.d., η ποσότητα q_n έχει υπολογιστεί με λ e.d.d. και το αποτέλεσμα της αφαίρεσης παράγεται με λ λανθασμένα δεκαδικά ψηφία στη mantissa. Η πιθανότητα να συμβεί το πρώτο είναι $P^S(\lambda + 1, \gamma) \cdot P^S(\lambda, \gamma)$, ενώ η πιθανότητα μείωσης του λάθους κατά ένα κατά την νέα αφαίρεση γράφεται ως $P^{un}(-1, 1)$. Συνεπώς τώρα η συνολική σχετική πιθανότητα ισούται με $P^S(\lambda + 1, \gamma) \cdot P^S(\lambda, \gamma) \cdot P^{un}(-1, 1)$.
- f. Η ποσότητα p_n έχει υπολογιστεί με $\lambda + 1$ e.d.d., η ποσότητα q_n έχει υπολογιστεί με $\lambda - \delta$ e.d.d., $0 \leq \delta < \lambda$, και το αποτέλεσμα της αφαίρεσης παράγεται με λ λανθασμένα δεκαδικά ψηφία στη mantissa. Η πιθανότητα να συμβεί το πρώτο είναι $P^S(\lambda + 1, \gamma) \cdot P^S(\lambda - \delta, \gamma)$, ενώ η πιθανότητα μείωσης του λάθους κατά ένα κατά την νέα αφαίρεση γράφεται ως $P^{un}(-1, \delta + 1)$. Συνεπώς η συνολική σχετική πιθανότητα ισούται με $P^S(\lambda + 1, \gamma) \cdot P^S(\lambda - \delta, \gamma) \cdot P^{un}(-1, \delta + 1)$.
- g. Η ποσότητα p_n έχει υπολογιστεί με $\lambda + k$ e.d.d., $k > 0$, η ποσότητα q_n έχει υπολογιστεί με $\lambda - \delta$ e.d.d., $0 \leq \delta < \lambda$ και το αποτέλεσμα της αφαίρεσης παράγεται με λ λανθασμένα δεκαδικά ψηφία στη mantissa. Η πιθανότητα να συμβεί το πρώτο είναι $P^S(\lambda + k, \gamma) \cdot P^S(\lambda - \delta, \gamma)$, ενώ η πιθανότητα μείωσης του λάθους κατά k κατά την νέα αφαίρεση γράφεται ως $P^{un}(-k, \delta + k)$. Συνεπώς η συνολική σχετική πιθανότητα ισούται με $P^S(\lambda + k, \gamma) \cdot P^S(\lambda - \delta, \gamma) \cdot P^{un}(-k, \delta + k)$.

ΘΕΩΡΗΜΑ 5.1

Επομένως, σύμφωνα με τα παραπάνω, η πιθανότητα $P^S(\lambda, \gamma + 1)$ υπολογίζεται ως εξής

$$P^S(\lambda, \gamma + 1) = \sum_{k=-4}^1 (P^S(\lambda - k, \gamma)^2 \cdot P_k^{eq}) + 2 \sum_{k=1}^4 \sum_{\delta=0}^4 P^S(\lambda + k, \gamma) \cdot P^S(\lambda - \delta, \gamma) \cdot P^{un}(-k, \delta + k) \quad (5.1)$$

Όπου θέσαμε το τέσσερα ως άνω όριο των k και δ , δεδομένου ότι η πιθανότητα διόρθωσης του λάθους πεπερασμένης ακρίβειας κατά περισσότερο από 5 ψηφία είναι πρακτικά αμελητέα. Εν τούτοις,

μπορούμε να αλλάξουμε τα όρια της άθροισης με οποιοδήποτε επιθυμητό αριθμό, οπότε και έχουμε μέγιστη ακρίβεια στην τιμή των πιθανοτήτων, αλλά εις βάρος του χρόνου υπολογισμού. □

Σημειώνεται ότι η διακύμανση του λάθους πεπερασμένης ακρίβειας σε κάθε διαδοχική αφαίρεση αυξάνεται. Αυτό οφείλεται στο ότι εάν οι τυχαίες μεταβλητές X, Y έχουν διακυμάνσεις σ_X^2 και σ_Y^2 , τότε ισχύει $Var(X - Y) = \sigma_X^2 + \sigma_Y^2$.

Ειδικότερα, αν $\sigma_X = \sigma_Y = \sigma$, τότε η τυπική απόκλιση της τυχαίας μεταβλητής X, Y είναι $\sigma\sqrt{2}$. Αυτό σημαίνει ότι σε κάθε διαδοχική αφαίρεση μεγαλώνει η πιθανότητα να αυξηθεί το λάθος πεπερασμένης ακρίβειας σε μια αφαίρεση κατά ένα δεκαδικό ψηφίο.

Αξιοποιώντας τον παραπάνω τύπο υπολογίζονται διαδοχικά όλες οι πιθανότητες $P^S(\lambda, \gamma)$ με $\lambda = 0, 1, \dots, n$. Ένα παράδειγμα των ανωτέρω παρουσιάζεται στο κάτωθι πίνακα 5.1.

Πλήθος e.d.d.	Τιμές των τελικών θεωρητικών πιθανοτήτων P κατά την αφαίρεση, για δύο τυπικές αποκλίσεις σ .	
	$\sigma = 1.2$ στην 8 ^η επανάληψη	$\sigma = 3.2$ στη 2 ^η επανάληψη
	Θεωρητικές Πιθανότητες	Θεωρητικές Πιθανότητες
Αύξηση κατά ένα	1.46%	51.76%
Διατήρηση	7.38%	34.57%
Ελάττωση κατά ένα	22.84%	2.53%
Ελάττωση κατά δυο	33.77%	0.25%
Ελάττωση κατά τρία	23.68%	0.02%
Ελάττωση κατά τέσσερα	8.88%	0.0008%

Πίνακας 5.1 Τιμές των τελικών θεωρητικών πιθανοτήτων P κατά τη μετάβαση από τη γ -οστή διαδοχική αφαίρεση στην $(\gamma + 1)$ -οστή, για δύο διαφορετικές τυπικές αποκλίσεις σ και δυο διαφορετικές αναδρομές των ανεξαρτήτων τυχαίων μεταβλητών X και Y .

6 Σύγκριση της γένεσης και διάδοσης του σφάλματος πεπερασμένης ακρίβειας κατά την εκτέλεση του ίδιου αλγορίθμου με διαφορετικά μήκη λέξης.

Έστω ένας τυχαίος επαναληπτικός αλγόριθμος \mathcal{A} , ο οποίος αποτελείται από προσθέσεις και αφαιρέσεις. Αυτός ο αλγόριθμος εκτελείται τόσο με n (έστω $n \geq 7$) όσο και με $m \geq 2n$ δεκαδικά ψηφία ακρίβεια, χρησιμοποιώντας το ίδιο σύνολο αριθμών ως είσοδο και στις δυο περιπτώσεις. Έστω επίσης x μια

οποιαδήποτε ποσότητα του αλγορίθμου και x_n^Y η τιμή αυτής της ποσότητας στη $\gamma - \text{οστή}$ επανάληψη, όταν οι πράξεις πραγματοποιούνται έχοντας n δεκαδικά ψηφία ακρίβεια. Ομοίως, x_m^Y ορίζεται η τιμή της ποσότητας x στη $\gamma - \text{οστή}$ επανάληψη του αλγορίθμου, όταν όλες οι πράξεις γίνονται με m ψηφία στη mantissa.

Σύμφωνα με το κεφάλαιο 4, το λάθος πεπερασμένης ακρίβειας που παράγεται κατά την αφαίρεση είναι το αλγεβρικό άθροισμα ενός νομοτελειακού και ενός τυχαίου λάθους. Όπως προαναφέρθηκε, το αιτιοκρατικό λάθος (deterministic error) υπολογίζεται από τη διαφορά του εκθέτη του αποτελέσματος και του μεγίστου εκθέτη μεταξύ των δυο όρων που αφαιρούνται. Επομένως, αν ο αλγόριθμος δεν έχει καταστραφεί, το νομοτελειακό λάθος για τα x_m^Y και x_n^Y είναι το ίδιο, αφού οι εκθέτες είναι ανεξάρτητοι της ακρίβειας που εφαρμόζεται κάθε φορά. Συνεπώς μόνο το τυχαίο / στοχαστικό λάθος μπορεί να οδηγήσει τα x_m^Y και x_n^Y να υπολογιστούν με διαφορετικό πλήθος ορθών δεκαδικών ψηφίων (c.d.d.). Σύμφωνα και με το κεφάλαιο 4, στους πίνακες 6.1 και 6.2 παρουσιάζονται οι πιθανότητες το πλήθος των e.d.d. των όρων x_m^Y και x_n^Y να διαφέρουν κατά diff, εκ του difference, σε μια επανάληψη. Στη προκειμένη περίπτωση ο πίνακας 6.1 παρουσιάζει τη μέγιστη τιμή αυτής της διαφοράς και ο 6.2 τη μέση τιμή αυτής της διαφοράς, μετά από 23 επαναλήψεις.

Με $P^{max}(\text{diff} \geq k)$ συμβολίζεται η μέγιστη πιθανότητα ή συχνότητα οι ποσότητες x_{16}^Y και x_{32}^Y να διαφέρουν τουλάχιστον κατά k δεκαδικά ψηφία.	Μέγιστη θεωρητική πιθανότητα (Άνω φράγμα)	Μέγιστη πειραματική συχνότητα
$P^{max}(\text{diff} \geq 0)$	1.00	1.00
$P^{max}(\text{diff} \geq 1)$	0.7518	0.7849
$P^{max}(\text{diff} \geq 2)$	0.3641	0.0623
$P^{max}(\text{diff} \geq 3)$	0.1555	0.0055
$P^{max}(\text{diff} \geq 4)$	0.0645	$5.3079 \cdot 10^{-4}$
$P^{max}(\text{diff} \geq 5)$	0.0220	$7.1357 \cdot 10^{-5}$
$P^{max}(\text{diff} \geq 6)$	0.0056	$6.5868 \cdot 10^{-6}$
$P^{max}(\text{diff} \geq 7)$	0.0011	$8.2335 \cdot 10^{-7}$
$P^{max}(\text{diff} \geq 8)$	0.0001	–
$P^{max}(\text{diff} \geq 9)$	$3.3486 \cdot 10^{-5}$	–
$P^{max}(\text{diff} \geq 10)$	$4.8216 \cdot 10^{-6}$	–
$P^{max}(\text{diff} \geq 11)$	$5.8147 \cdot 10^{-7}$	–
$P^{max}(\text{diff} \geq 12)$	$6.6462 \cdot 10^{-8}$	–

$P^{max}(\text{diff} \geq 13)$	$7.6124 \cdot 10^{-9}$	—
$P^{max}(\text{diff} \geq 14)$	$5.5846 \cdot 10^{-10}$	—

Πίνακας 6.1 Οι θεωρητικές πιθανότητες και πειραματικές συχνότητες εμφάνισης της μέγιστης διαφοράς του πλήθους των e.d.d. μεταξύ δυο αριθμών σε μια επανάληψη, όταν ο αλγόριθμος εκτελείται παράλληλα με 16 και 32 δεκαδικά ψηφία ακρίβεια. Η δεύτερη στήλη παρουσιάζει τη θεωρητική πρόβλεψη, ενώ η τρίτη τα σχετικά πειραματικά αποτελέσματα. Η εγγύτητα των θεωρητικών προβλέψεων και των πειραματικών αποτελεσμάτων είναι εξαιρετική.

Με $P(\text{diff} \geq k)$ συμβολίζεται η μέση πιθανότητα ή συχνότητα οι ποσότητες x_{16}^Y και x_{32}^Y να διαφέρουν τουλάχιστον κατά k δεκαδικά ψηφία.	Μέση θεωρητική πιθανότητα (Άνω φράγμα)	Μέση πειραματική συχνότητα
$P(\text{diff} \geq 0)$	0.9977	1.0000
$P(\text{diff} \geq 1)$	0.6921	0.4713
$P(\text{diff} \geq 2)$	0.2817	0.0316
$P(\text{diff} \geq 3)$	0.0968	0.0025
$P(\text{diff} \geq 4)$	0.0281	$2.2693 \cdot 10^{-4}$
$P(\text{diff} \geq 5)$	0.0070	$1.9164 \cdot 10^{-5}$
$P(\text{diff} \geq 6)$	0.0015	$1.4438 \cdot 10^{-6}$
$P(\text{diff} \geq 7)$	0.0002	$1.5512 \cdot 10^{-7}$
$P(\text{diff} \geq 8)$	$5.3414 \cdot 10^{-5}$	—
$P(\text{diff} \geq 9)$	$8.5184 \cdot 10^{-6}$	—
$P(\text{diff} \geq 10)$	$5.1422 \cdot 10^{-7}$	—
$P(\text{diff} \geq 11)$	$1.1140 \cdot 10^{-7}$	—
$P(\text{diff} \geq 12)$	$1.1066 \cdot 10^{-8}$	—
$P(\text{diff} \geq 13)$	$8.1498 \cdot 10^{-10}$	—
$P(\text{diff} \geq 14)$	$5.8454 \cdot 10^{-11}$	—

Πίνακας 6.2 Οι θεωρητικές πιθανότητες και πειραματικές συχνότητες εμφάνισης της μέσης διαφοράς του πλήθους των e.d.d. μεταξύ δυο αριθμών σε μια επανάληψη, όταν ο αλγόριθμος εκτελείται παράλληλα με 16 και 32 δεκαδικά ψηφία ακρίβεια. Η δεύτερη στήλη παρουσιάζει τη θεωρητική πρόβλεψη, ενώ η τρίτη τα σχετικά πειραματικά αποτελέσματα. Εκ νέου, οι θεωρητικές προβλέψεις και τα πειραματικά αποτελέσματα είναι εξαιρετικά κοντά.

Από αυτούς του πίνακες παρατηρείται ότι οι θεωρητικές πιθανότητες να εμφανίζεται διαφορά $\text{diff} \geq k$ δεκαδικών ψηφίων μεταξύ x_{16}^Y και x_{32}^Y είναι σαφώς μεγαλύτερη της πειραματικής συχνότητας. Αυτό συμβαίνει για τους εξής δύο λόγους:

- Τα τυχαία λάθη λ_{16} και λ_{32} που συσσωρεύονται στις ποσότητες x_{16}^Y και x_{32}^Y δεν είναι στατιστικώς ανεξάρτητα. Όντως, όπως προκύπτει από τον τύπο (4.2), αλλά και από το γεγονός

ότι το χρησιμοποιούμενο υπολογιστικό σύστημα παράγει τα λανθασμένα ψηφία με παρόμοιο τρόπο, προκύπτει άμεσα ότι τα λάθη πεπερασμένης ακρίβειας των x_{16}^{γ} και x_{32}^{γ} είναι συσχετισμένα.

- b. Ο υπολογισμός των μέγιστων και των μέσων θεωρητικών πιθανοτήτων έγινε επί τη βάση της ομοιόμορφης κατανομής που δίνει υποχρεωτικά τα χειρότερα αποτελέσματα όσον αφορά τη λάθη πεπερασμένης ακρίβειας (Είδη τους υπολογισμούς των διπλών ολοκληρωμάτων στο 4.2.3 κεφάλαιο). Για μικρότερες τυπικές αποκλίσεις οι τιμές των θεωρητικών πιθανοτήτων και των πρακτικών συχνοτήτων πλησιάζουν πολύ.

Βέβαια για τους ανωτέρω λόγους, οι θεωρητικές πιθανότητες είναι ένα πολύ ασφαλές άνω φράγμα διαφοράς του πλήθους των λανθασμένων δεκαδικών ψηφίων που συσσωρεύονται στη mantissa των τυχαίων ποσοτήτων x_{16}^{γ} και x_{32}^{γ} του τυχόντος αλγορίθμου \mathcal{A} . Αυτό το γεγονός μας επιτρέπει να διατυπωθεί το κάτωθι συμπέρασμα.

ΘΕΩΡΗΜΑ 6.1

Έστω ένα πλήθος από γ διαδοχικές αφαιρέσεις, οι οποίες εκτελούνται έχοντας αρχικά ως είσοδο ένα σύνολο αριθμών που είναι εκφρασμένοι σε επιστημονική μορφή. Έστω ότι αυτές οι αφαιρέσεις εκτελούνται τη μία φορά με n ($n \geq 7$) δεκαδικά ψηφία ακρίβεια και με m δεκαδικά ψηφία την άλλη, όπου $m > 2n + 7$. Αν s_n^{γ} και s_m^{γ} είναι τα αποτελέσματα της γ -οστής αφαίρεσης, όταν αυτή υλοποιείται με n και m ακρίβεια αντιστοίχως, τότε:

Η πιθανότητα $P_g^D = \{|\lambda_n - \lambda_m| = g\}$ δίνεται από το τύπο

$$P_g^D = 2 \sum_{i=0}^n P^S(i, \gamma) \cdot P^S(i + g, \gamma) \quad (6.1)$$

Η πιθανότητα P_g^{TD} να ισχύει $|\lambda_n - \lambda_m| \geq g$ δίνεται από την παρακάτω εξίσωση:

$$P_g^{TD} = \sum_{j=0}^n \left(P^S(j, \gamma) \sum_{i=j+g}^m P^S(i, \gamma) \right) + \sum_{j=0}^{n-g} \left(P^S(j, \gamma) \sum_{i=j+g}^n P^S(i, \gamma) \right) \quad (6.2)$$

Σύμφωνα με τα κεφάλαια 4 και 5 η πιθανότητα $|(\text{Πλήθος e. d. d. του } s_n) - (\text{Πλήθος e. d. d. του } s_m)| \leq 10$ έχει θεωρητική τιμή ίση ή μεγαλύτερη της $1 - O(10^{-7})$. Πρακτικά όμως, σύμφωνα με τα πειράματα που εκτελέστηκαν, ισχύει ότι

$$|(\text{Πλήθος e. d. d. του } s_n) - (\text{Πλήθος e. d. d. του } s_m)| \leq 8 \quad (6.3)$$

Με πιθανότητα πάρα πολύ κοντά στο 1 (πρακτικά ίση με 1). □

Υπενθυμίζεται ότι, οι πιθανότητες που παρουσιάζονται στο κεφάλαιο 4 είναι ανεξάρτητες από το μήκος λέξης με το οποίο εκτελούνται οι αφαιρέσεις. Τέλος, όσον αφορά το πρώτο διπλό άθροισμα στο τύπο (6.2) το j αναφέρεται στην n ακρίβεια και το i στην m , ενώ το αντίστροφο ισχύει για το δεύτερο διπλό άθροισμα.

7 Τα εκτελεσθέντα σχετικά πειράματα και τα αποτελέσματά τους.

Στο παρόν κεφάλαιο παρουσιάζονται πειράματα που σχεδιάστηκαν και πραγματοποιήθηκαν ειδικά για τον έλεγχο των προαναφερθεισών θεωρητικών προβλέψεων, καθώς και τα αποτελέσματα των πειραμάτων αυτών.

7.1 Πειραματική επαλήθευση της θεωρητικής πρόβλεψης σε σχέση με τη συσσώρευση του λάθους πεπερασμένης ακρίβειας σε διαδοχικές αφαιρέσεις.

Η επαλήθευση των θεωρητικών αναλύσεων και προβλέψεων των κεφαλαίων 5 και 6 έγινε ως εξής:

Αρχικά επελέγει ένα σύνολο S_{n_2} , στο οποίο κάθε αριθμός ήταν εκφρασμένος σε επιστημονική μορφή, με 88 δεκαδικά ψηφία στη mantissa. Οι αριθμοί αυτοί προήλθαν από μια γεννήτρια τυχαίων αριθμών που ακολουθούν ομοιόμορφη κατανομή.

Στη συνέχεια έγινε προβολή των αριθμών του παραπάνω συνόλου σε 40 και 16 δεκαδικά ψηφία ακρίβεια, οπότε με αυτό τον τρόπο δημιουργήθηκαν τα αντίστοιχα σύνολα S_{n_1} και S_n . Συμβολίσαμε με $\alpha_{n_2}^i$ τα στοιχεία του S_{n_2} , με $\alpha_{n_1}^i$ τα στοιχεία του S_{n_1} και α_n^i τα στοιχεία που αποτελούν το S_n .

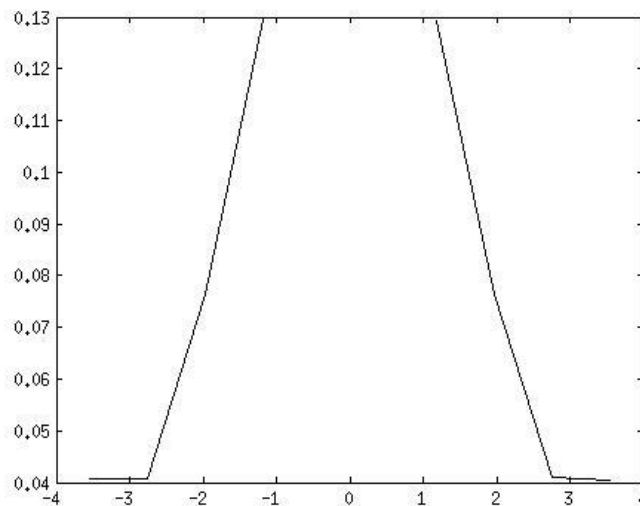
Εν συνεχεία, υπολογίστηκε το λάθος πεπερασμένης ακρίβειας που παράχθηκε από τη παραπάνω διαδικασία προβολής του $\alpha_{n_2}^i$ σε $\alpha_{n_1}^i$ και α_n^i . Αυτό πραγματοποιήθηκε με σύγκριση των n πρώτων δεκαδικών ψηφίων των $\alpha_{n_2}^i$ και των α_n^i , σύμφωνα με τους ορισμούς του κεφαλαίου Α. Το λάθος της α_n^i

αναπαράστασης συμβολίζεται ως x_n^i . Ομοίως, συγκρίνοντας τα n_1 πρώτα δεκαδικά ψηφία των ποσοτήτων $\alpha_{n_2}^i$ και $\alpha_{n_1}^i$, υπολογίστηκε το λάθος πεπερασμένης ακρίβειας $x_{n_1}^i$ της αναπαράστασης $\alpha_{n_1}^i$.

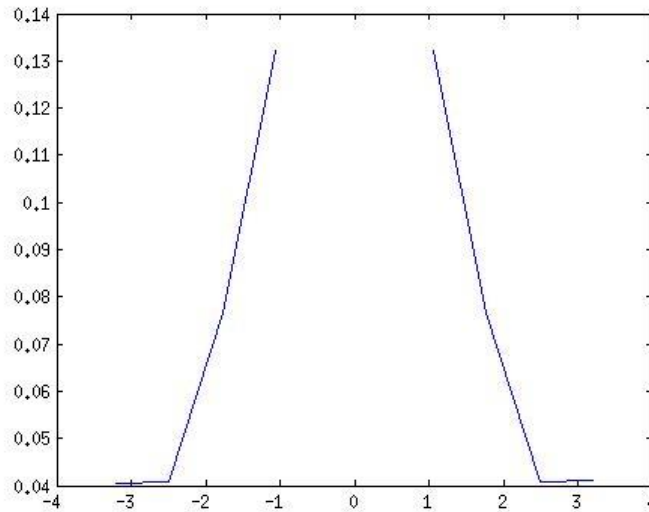
Στη συνέχεια ελέγχθηκε η υπόθεση ότι τα x_n^i και $x_{n_1}^i$ ακολουθούν κανονική κατανομή με μέση τιμή μ_n και μ_{n_1} και τυπική απόκλιση σ_n και σ_{n_1} , σύμφωνα με τα τεστ των Sapiro – Wilk και Kolmogorov [21], [22]. Η υπόθεση αυτή δεν απορρίφθηκε σε επίπεδο σημαντικότητας $\alpha = 0.001$. Πρακτικά, η μέση τιμή των δυο κατανομών ήταν μηδέν, ενώ η τυπική απόκλιση εξαρτιόταν ελάχιστα από τους επιλεγμένους αριθμούς. Επαναλαμβάνοντας τη παραπάνω διαδικασία για διαφορετικά σύνολα S_{n_2} κάθε φορά, παρατηρήθηκε ότι οι μέσες τιμές και τυπικές αποκλίσεις, με βαθμό εμπιστοσύνης 99%, ήταν:

- a) Για το x_n^i : $\mu_n = 0$ και $\sigma_n = (1.49 \pm 5 \cdot 10^{-3}) \cdot 10^{-15}$
- b) Για το $x_{n_1}^i$: $\mu_{n_1} = 0$ και $\sigma_{n_1} = (1.64 \pm 5 \cdot 10^{-3}) \cdot 10^{-39}$

Όπως φαίνεται και από τις εικόνες 7.1 και 7.2 η mantissa του λάθους πεπερασμένης ακρίβειας των $\alpha_{n_1}^i$ και α_n^i ακολουθεί τη κανονική κατανομή με εξαιρετική προσέγγιση.



Εικόνα 7.1 Η mantissa του σφάλματος πεπερασμένης ακρίβειας των α_n^i ακολουθεί κανονική κατανομή με μέση τιμή $\mu_n = 0$ και τυπική απόκλιση σ_n περίπου ίση με $(1.49 \pm 5 \cdot 10^{-3}) \cdot 10^{-15}$. Το σφάλμα αυτό αφορά τη περίπτωση που χρησιμοποιούνται $n = 16$ δεκαδικά ψηφία στη mantissa για την αναπαράσταση ενός τυχαίου αριθμού.



Εικόνα 7.2 Η mantissa του σφάλματος πεπερασμένης ακρίβειας των $\alpha_{n_1}^i$ ακολουθεί κανονική κατανομή με μέση τιμή $\mu_{n_1} = 0$ και τυπική απόκλιση σ_{n_1} περίπου ίση με $(1.64 \pm 5 \cdot 10^{-3}) \cdot 10^{-39}$. Το σφάλμα αυτό αφορά τη περίπτωση που χρησιμοποιούνται $n_1 = 40$ δεκαδικά ψηφία στη mantissa για την αναπαράσταση ενός τυχαίου αριθμού εξ' αυτών που αρχικά σχημάτισαν το σύνολο S_{n_2} .

Στο επόμενο στάδιο του πειράματος εκτελέστηκαν $1.5 \cdot 10^{13}$ τυχαίες αφαιρέσεις ανάμεσα στα στοιχεία του S_n , με τον περιορισμό ότι τα εκάστοτε στοιχεία α_n^i δεν αφαιρούνται με τον εαυτό τους. Ταυτόχρονα εκτελέστηκαν οι ακριβώς αντίστοιχες αφαιρέσεις μεταξύ των στοιχείων $\alpha_{n_2}^i$ αφενός και $\alpha_{n_1}^i$ αφετέρου.

Ακολούθως, από τα $1.5 \cdot 10^{13}$ αποτελέσματα των τυχαίων αφαιρέσεων έγινε επιλογή $5 \cdot 10^{11}$ αριθμών $d_n^{1,i}$ εκφρασμένων σε επιστημονική μορφή, οι οποίοι σχημάτισαν το σύνολο S_n^1 . Έπειτα επιλέχθηκαν τα αντίστοιχα αποτελέσματα από τις αφαιρέσεις των n_1 και n_2 αναπαραστάσεων τα οποία δημιούργησαν τα σύνολα $S_{n_1}^1$ και $S_{n_2}^1$ αντιστοίχως. Στη συνέχεια εκτελέστηκαν $1.5 \cdot 10^{13}$ τυχαίες αφαιρέσεις μεταξύ των στοιχείων του συνόλου S_n^1 . Κρατώντας $5 \cdot 10^{11}$ αποτελέσματα σχηματίστηκε, όπως και προηγουμένως, το S_n^2 . Επιλέγοντας πάλι τα αντίστοιχα αποτελέσματα από τις ίδιες αφαιρέσεις των n_1 και n_2 αναπαραστάσεων των αυτών αφαιρετέων, δημιουργήθηκαν τα σύνολα $S_{n_1}^2$ και $S_{n_2}^2$. Η διαδικασία αυτή επαναλήφθηκε συνολικά 23 φορές.

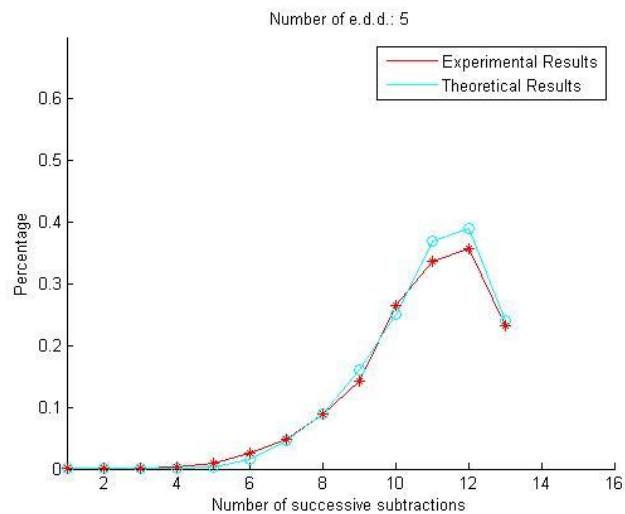
Πλήθος d.d.	Τελικές τιμές των θεωρητικών πιθανοτήτων και πειραματικών συχνοτήτων τροποποίησης του λάθους πεπερασμένης ακρίβειας κατά την $(\gamma + 1) - \text{οστή}$ αφαίρεση, σε σχέση με το λάθος των αφαιρετέων, για δύο τυπικές αποκλίσεις σ .			
	$\sigma = 1.2$ στην 8 ^η επανάληψη		$\sigma = 3.2$ στη 2 ^η επανάληψη	
	Θεωρητικές Πιθ.	Πειραματικές Συχν.	Θεωρητικές Πιθ.	Πειραματικές Συχν.
Αύξηση κατά ένα	1.46%	0.31%	51.76%	56.18%

Διατήρηση	7.38%	5.21%	34.57%	38.99%
Ελάττωση κατά ένα	22.84%	24.84%	2.53%	4.11%
Ελάττωση κατά δυο	33.77%	36.06%	0.25%	0.52%
Ελάττωση κατά τρία	23.68%	21.39%	0.02%	0.13%
Ελάττωση κατά τέσσερα	8.88%	8.84%	0.0008%	0.03%

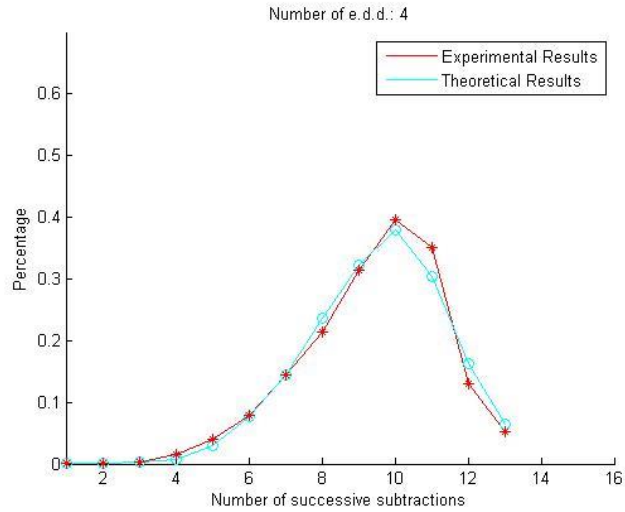
Πίνακας 7.1 Τελικές τιμές των θεωρητικών πιθανοτήτων και πειραματικών συχνοτήτων τροποποίησης του λάθους πεπερασμένης ακρίβειας κατά την $(\gamma + 1)$ -οστή αφαίρεση, σε σχέση με το λάθος των αφαιρετέων, για δύο τυπικές αποκλίσεις σ . Τα λάθη πεπερασμένης ακρίβειας των αφαιρετέων γεννήθηκαν από τις προηγούμενες γ διαδοχικές αφαιρέσεις. Οι θεωρητικές πιθανότητες και οι πειραματικές συχνότητες αφορούν τη 2^η αναδρομή αφενός και την 8^η αφετέρου. Η θεωρητική προσέγγιση που προηγήθηκε δίνει αποτελέσματα εξαιρετικά κοντά στα πειραματικά, γεγονός που σαφέστατα υποστηρίζει την ορθότητα αυτής.

Η σύγκλιση των πειραματικών και θεωρητικών αποτελεσμάτων είναι εξαιρετική, γεγονός που υποστηρίζει σαφώς την ορθότητα της θεωρητικής προσέγγισης.

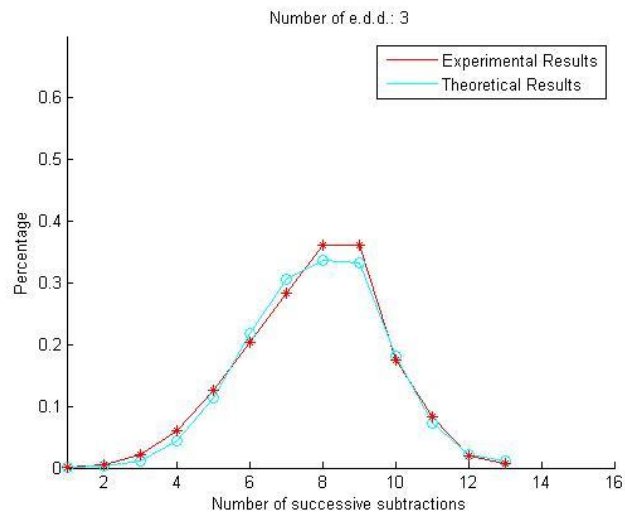
Σε κάθε επανάληψη της προαναφερθείσας διαδικασίας γινόταν καταγραφή των λαθών πεπερασμένης ακρίβειας, κρατώντας τόσο το ολικό σφάλμα, όσο και το νομοτελειικό και το στοχαστικό. Τα θεωρητικά και πειραματικά αποτελέσματα που σχετίζονται με το ολικό σφάλμα φαίνονται στις εικόνες 7.3, 7.4, 7.5 και 7.6.



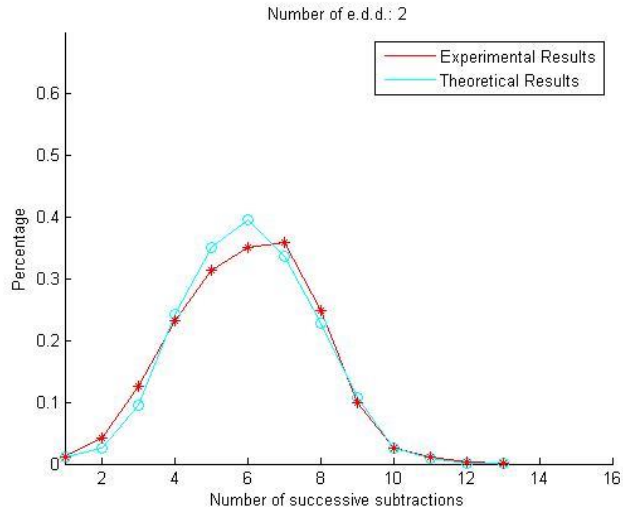
Εικόνα 7.3 Γράφημα του ποσοστού $P(5, \gamma)$ κατά τη γ -οστή διαδοχική αφαίρεση, το αποτέλεσμα των οποίων παρήχθη με ακριβώς πέντε (5) λανθασμένα δεκαδικά ψηφία στη mantissa του. Ο άξονας x αντιστοιχεί στον αύξοντα αριθμό γ της επαναλαμβανόμενης διαδικασίας, ενώ ο άξονας y το προαναφερθέν ποσοστό $P(5, \gamma)$.



Εικόνα 7.4 Γράφημα του ποσοστού $P(4, \gamma)$ κατά τη γ – οστική διαδοχική αφαίρεση, το αποτέλεσμα των οποίων παρήχθη με ακριβώς τέσσερα (4) λανθασμένα δεκαδικά ψηφία στη mantissa του. Ο άξονας x αντιστοιχεί στον αύξοντα αριθμό γ της επαναλαμβανόμενης διαδικασίας, ενώ ο άξονας y το προαναφερθέν ποσοστό $P(4, \gamma)$.



Εικόνα 7.5 Γράφημα του ποσοστού $P(3, \gamma)$ των αφαιρέσεων το αποτέλεσμα των οποίων παρήχθη με ακριβώς τρία (3) λανθασμένα δεκαδικά ψηφία στη mantissa του. Ο άξονας x αντιστοιχεί στον αύξοντα αριθμό γ της επαναλαμβανόμενης διαδικασίας, ενώ ο άξονας y το προαναφερθέν ποσοστό $P(3, \gamma)$.



Εικόνα 7.6 Γράφημα του ποσοστού $P(2, \gamma)$ των αφαιρέσεων το αποτέλεσμα των οποίων παρήχθη με ακριβώς δύο (2) λανθασμένα δεκαδικά ψηφία στη mantissa του. Ο άξονας x αντιστοιχεί στον αύξοντα αριθμό γ της επαναλαμβανόμενης διαδικασίας, ενώ ο άξονας y το προαναφερθέν ποσοστό $P(2, \gamma)$.

8 Η γένεση και διάδοση του σφάλματος πεπερασμένης ακρίβειας στον αλγόριθμο Bernoulli chaotic map, καθώς και μια μέθοδος σταθεροποίησης αυτού.

8.1 Σύντομη περιγραφή και αριθμητική συμπεριφορά του αλγορίθμου Bernoulli map.

Έστω μια ποσότητα x_0 , εκφρασμένη στο δεκαδικό σύστημα με κινητή υποδιαστολή, για την οποία ισχύει $0 \leq x_0 < 1$. Εκκινώντας από τη τιμή x_0 , ο αλγόριθμος Bernoulli map ορίζεται ως εξής:

$$x_{k+1} = \begin{cases} 2x_k & \text{όταν } 0 \leq x_k < \frac{1}{2} \\ 2x_k - 1 & \text{όταν } \frac{1}{2} \leq x_k < 1 \end{cases} \quad (8.1)$$

Η γένεση, διάδοση και συσσώρευση του σφάλματος πεπερασμένης ακρίβειας κατά την εκτέλεση της (8.1) είναι άρρηκτα συνδεδεμένη με το πλήθος των επαναλήψεων του αλγορίθμου ($2^N 10^{-Dec.Prec} = 1$) [23]. Η συνεχής αυτή συσσώρευση και διάδοση του σφάλματος πεπερασμένης ακρίβειας έχει ως

αποτέλεσμα τη καταστροφή του Bernoulli map. Όπως έχει πειραματικά αποδειχθεί [23], το πλήθος των λανθασμένων δεκαδικών ψηφίων πρακτικά διπλασιάζεται σε κάθε αναδρομή.

8.2 Τα βαθύτερα αίτια της καταστροφής του Bernoulli map λόγω σφάλματος πεπερασμένης ακρίβειας.

Η προσέγγιση που εφαρμόζεται στη παρούσα εργασία, επιτρέπει τον εντοπισμό των βαθύτερων και πραγματικών αιτίων για τα οποία κάθε αναδρομικός αλγόριθμος καταστρέφεται λόγω λάθους πεπερασμένης ακρίβειας και εν προκειμένω ο Bernoulli map. Όντως, τα βαθύτερα αυτά αίτια θα αναδειχθούν μέσω του κάτωθι πολύ αντιπροσωπευτικού παραδείγματος.

Καταρχάς, έστω ότι όλοι οι υπολογισμοί εκτελούνται με δεκαξέρι δεκαδικά ψηφία ακρίβεια στη mantissa. Έστω η αρχική ποσότητα $x_0 = 5.055022710000000 \cdot 10^{-1}$. Στην επόμενη αναδρομή η τιμή που υπολογίζεται είναι η $x_1 = 1.100454200000001 \cdot 10^{-2}$. Όπως παρατηρείται, το τελευταίο ψηφίο (ο τελευταίος άσσος) είναι λάθος, σύμφωνα με την ανάλυση των κεφαλαίων 2 και 4 του παρόντος πονήματος. Συγκεκριμένα, αφού το $x_0 > \frac{1}{2}$, εκτελείται ο πολλαπλασιασμός $2x_0$, παράγοντας έναν αριθμό της τάξης 10^0 και έπειτα από αυτόν τον αριθμό αφαιρείται μία μονάδα. Το αποτέλεσμα είναι τάξης 10^{-2} . Συνεπώς, ο εκθέτης του αποτελέσματος βυθίζεται κατά 2, γεννώντας νομοτελειακά 2 λανθασμένα δεκαδικά ψηφία, σύμφωνα με την ανάλυση του κεφαλαίου 4.1. Το λάθος αυτό μπορεί να τροποποιηθεί λόγω της εμφάνισης του στοχαστικού λάθους πεπερασμένης ακρίβειας, σύμφωνα με την ανάλυση της παραγράφου 4.2. Τα σχετικά αποτελέσματα υποδεικνύουν ότι η μεγιστοπιθήνη περίπτωση είναι να παραχθεί ένα επιπλέον λανθασμένο δεκαδικό ψηφίο, το δεύτερο πιθανότερο ενδεχόμενο είναι το νομοτελειακό λάθος πεπερασμένης ακρίβειας να παραμείνει αμετάβλητο, ενώ υπάρχει μικρή πιθανότητα, αλλά όχι αμελητέα, το νομοτελειακό λάθος να μειωθεί κατά ένα ή δύο δεκαδικά ψηφία κτλ. (κεφάλαιο 4.2).

Στις επόμενες αναδρομές παράγονται οι εξής τιμές της ποσότητας x_k :

$$\begin{aligned}x_2 &= 2.200908400000001 \cdot 10^{-2} \\x_3 &= 4.401816800000002 \cdot 10^{-2} \\x_4 &= 8.803633600000005 \cdot 10^{-2}\end{aligned}\tag{8.2}$$

Οι τιμές x_2 , x_3 και x_4 υπολογίζονται διπλασιάζοντας τη τιμή της προηγούμενης αναδρομής. Με αυτό το τρόπο διπλασιάζεται και η τιμή του λάθους. Στην επόμενη αναδρομή υπολογίζεται η τιμή $x_5 = 1.760726720000001 \cdot 10^{-1}$. Σε αυτή τη περίπτωση η x_4 ποσότητα πολλαπλασιάζεται με τον

αριθμό 2. Η τιμή του λάθους διπλασιάζεται, παράγοντας ένα επιπλέον λανθασμένο δεκαδικό ψηφίο. Όμως, η υπολογιστική μηχανή εκτελεί δεξιά ολίσθηση για να διατηρήσει την επιστημονική μορφή του αποτελέσματος, οδηγώντας στην ελάττωση του σφάλματος κατά ένα. Επομένως, το πλήθος των e.d.d. διατηρείται. Οι τιμές x_6, x_7 και x_8 παράγονται με το διπλασιασμό της τιμής της προηγούμενης αναδρομής.

Στην ένατη αναδρομή παράγεται ένα επιπλέον λανθασμένο ψηφίο, αφού το λάθος διπλασιάζεται, ο εκθέτης του αποτελέσματος δεν αυξάνεται και δεν υπάρχει δεξιά ολίσθηση. Επομένως, υπολογίζεται η τιμή $x_9 = 8.171627520000015 \cdot 10^{-1}$. Στις επόμενες αναδρομές το πλήθος των λανθασμένων δεκαδικών ψηφίων παραμένει σταθερό και ίσο με δύο. Δηλαδή

$$\begin{aligned} x_{10} &= 6.343255040000031 \cdot 10^{-1} \\ x_{11} &= 2.686510080000062 \cdot 10^{-1} \end{aligned} \quad (8.3)$$

Συνεχίζοντας, η x_{11} διπλασιάζεται με σκοπό τον υπολογισμό της τιμής x_{12} , δεδομένου ότι $x_{11} < \frac{1}{2}$. Η x_{12} ισούται με:

$$x_{12} = 5.373020160000124 \cdot 10^{-1} \quad (8.4)$$

Όπως παρατηρείται, το πλήθος των λανθασμένων δεκαδικών ψηφίων και ο εκθέτης του λάθους πεπερασμένης ακρίβειας αυξάνονται συστηματικά κατά ένα, είτε λόγω του διπλασιασμού του λάθους αυτού, είτε λόγω του νομοτελειακού λάθους που παράγεται κατά την αφαίρεση $2x_k - 1$. Προχωρώντας στην εκτέλεση των επόμενων αναδρομών, ο αλγόριθμος Bernoulli map καταστρέφεται εν τέλει λόγω f.p.e. σε 53 επαναλήψεις. Το γεγονός αυτό, καθώς και ο πολύ μικρός αριθμός αναδρομών όπου το f.p.e. γίνεται καθοριστικό, καθιστά απαγορευτική τη χρήση του φημισμένου αυτού αλγορίθμου.

8.3 Διόρθωση του σφάλματος πεπερασμένης ακρίβειας στον αλγόριθμο Bernoulli map

Έστω ότι η αρχική ποσότητα x_0 είναι εκφρασμένη σε επιστημονική μορφή με n δεκαδικά ψηφία στη mantissa και έχει στον εκθέτη τη τιμή -1 . Τότε όλες οι ποσότητες x_k του Bernoulli map υπολογίζονται έχοντας n δεκαδικά ψηφία στη mantissa. Επιπροσθέτως, όλοι οι ενδιάμεσοι υπολογισμοί γίνονται με το πολύ $(n + 1)$ δεκαδικά ψηφία ακρίβεια.

Έστω ότι η x_k έχει υπολογιστεί με n δεκαδικά ψηφία στη mantissa και $E(x_k) = -1$. Τότε,

- i. Αν $x_k < \frac{1}{2}$ τότε διπλασιάζοντας τη ποσότητα x_k δεν αλλάζει ο εκθέτης του αποτελέσματος.

- ii. Αν $\frac{1}{2} \leq x_k < 1$ τότε ο εκθέτης του αποτελέσματος θα έχει τη τιμή μηδέν. Δηλαδή $E(2x_k) = 0$ με τη ποσότητα $2x_k$ να έχει $n + 1$ δεκαδικά ψηφία στη mantissa. Σε αυτό το σημείο όταν αφαιρείται η μονάδα από το $2x_k$ για να παραχθεί η ποσότητα x_{k+1} , τότε η τάξη του αποτελέσματος γίνεται μικρότερη από ή ίση με 10^{-1} . Συνεπώς, η x_{k+1} έχει το πολύ n δεκαδικά ψηφία στην mantissa.

Γενικά, αν x_0 είναι τάξης 10^{-p} , με $p \geq 1$ και n δεκαδικά ψηφία στη mantissa, τότε όλες οι ποσότητες x_k του Bernoulli map υπολογίζονται με το πολύ $(n + p - 1)$ δεκαδικά ψηφία στη mantissa, ενώ όλοι οι ενδιάμεσοι υπολογισμοί γίνονται με το πολύ $(n + p)$ ακρίβεια.

Σε αυτό το σημείο, επιλέγεται να γίνονται όλοι οι υπολογισμοί των ποσοτήτων έχοντας μήκος λέξης $\tau = n + p + 2$. Σύμφωνα με τα αποτελέσματα που παρουσιάζονται στο κεφάλαιο 4, αν p είναι ο αριθμός των λανθασμένων ψηφίων στη mantissa του x_k , τότε το x_{k+1} θα έχει το πολύ $p + 2 + E(x_k) - E(x_{k+1})$ e.d.d. Ο προσθετός +2 στον αριθμό των e.d.d οφείλεται σε δυο παράγοντες:

1. Στην ενδεχόμενη αύξηση του λάθους πεπερασμένης ακρίβειας λόγω διπλασιασμού του είτε
2. Στο μέγιστο στοχαστικό λάθος που προστίθεται στο νομοτελικό κατά την εκτέλεση της αφαίρεσης $2x_k - 1$.

Συνεπώς, αν $\psi = p + 2 + E(x_k) - E(x_{k+1})$, τότε όλα τα λανθασμένα δεκαδικά ψηφία στη mantissa της x_{k+1} βρίσκονται μεταξύ της $\psi - \text{οστής}$ θέσης και της $\tau - \text{οστής}$.

Στη συνέχεια, εξαναγκάζοντας τα ψηφία μετά τη $\psi - \text{οστή}$ θέση της mantissa της ποσότητας x_{k+1} να είναι μηδέν σε κάθε αναδρομή, οι τιμές που παράγονται από τον αλγόριθμο του Bernoulli map είναι πάντα σωστές, χωρίς να γεννάται λάθος πεπερασμένης ακρίβειας. Αυτό επιτυγχάνεται μετατρέποντας προγραμματιστικά τον αριθμό σε κείμενο, εξασφαλίζοντας με αυτό τον τρόπο τη mantissa και τον εκθέτη του αριθμού x_{k+1} . Έπειτα μηδενίζονται τα τελευταία $\tau - \psi + 1$ ψηφία της mantissa, όντας σε μορφή κειμένου (string) και στη συνέχεια το κείμενο αυτό μαζί με τον εκθέτη, προβάλλονται ξανά σε αριθμό κινητής υποδιαστολής. Σε αυτό το σημείο πρέπει να τονιστεί ότι λαμβάνεται υπόψη και η ισότητα ενός δεκαδικού αριθμού με την αντίστοιχη ποσότητα της οποίας τα τελευταία ψηφία στη mantissa έχουν τον αριθμό 9. Για παράδειγμα, η ποσότητα $y = 1.2399999$ και $z = 1.24$ είναι ίσες. Σε αυτή τη περίπτωση στρογγυλοποιείται ο αριθμός z μετά την όγδοη θέση. Η θέση στρογγυλοποίησης γίνεται με βάση λογικών κανόνων, μεθόδων στρογγυλοποίησης και κανόνων που παρουσιάζονται στο κεφάλαιο A (Ορισμός 1).

Η υλοποίηση της παραπάνω μεθοδολογίας προσθέτει ελάχιστα στο χρόνο εκτέλεσης του αλγορίθμου. Στη περίπτωση όπου η ακρίβεια είναι $\tau \leq 18 \Leftrightarrow n + \rho \leq 16$, τότε η προτεινόμενη διόρθωση του αλγορίθμου μπορεί να υλοποιηθεί χρησιμοποιώντας τις δυνατότητες της υπολογιστικής μηχανής. Αντιθέτως, αν $\tau > 18$, τότε θα πρέπει να γίνει χρήση ενός λογισμικού που προσφέρει την εκάστοτε επιθυμητή ακρίβεια, όπως είναι η βιβλιοθήκη Multiple-Precision Floating point computations with correct Rounding (MPFR).

Με τη χρήση της παραπάνω βιβλιοθήκης επιτυγχάνεται αύξηση της ακρίβειας με την οποία υπολογίζονται οι τιμές του Bernoulli map. Παρόλο που για τους υπολογισμούς της (8.1) χρησιμοποιούνται διάφορα μήκη λέξης, έχοντας το ίδιο x_0 ως είσοδο, ο αλγόριθμος του Bernoulli map καταστρέφεται λόγω f.p.e. Αυξάνοντας την ακρίβεια, αυξάνεται και ο αριθμός των αναδρομών εκτέλεσης της (8.1) χωρίς αυτή να καταστρέφεται. Αντίθετα, με τη προτεινόμενη μεθοδολογία ο αλγόριθμος του Bernoulli map εκτελείται χωρίς περιορισμούς στο πλήθος των αναδρομών, δίνοντας πάντα αξιόπιστα αποτελέσματα. Αυτό επιβεβαιώνεται ακολουθώντας την εξής μεθοδολογία.

Ο αλγόριθμος εκτελείται παράλληλα δυο φορές, τη μία φορά με τη προτεινόμενη μεθοδολογία έχοντας n δεκαδικά ψηφία στη mantissa ($n > 7$) και παράλληλα με τη κλασσική προσέγγιση έχοντας $m = 2n + 8$ δεκαδικά ψηφία στη mantissa. Μόλις διαπιστώνεται απόκλιση στα πρώτα n ψηφία μεταξύ των δύο εκτελέσεων, σύμφωνα με όσα αναπτύχθηκαν στα προηγούμενα κεφάλαια, τότε διπλασιάζεται η ακρίβεια της κλασσικής προσέγγισης και ξανά εκτελούνται οι αλγόριθμοι έχοντας το x_{k-1} ως είσοδο. Η σύγκριση επιτυγχάνεται προβάλλοντας τις τιμές της κλασσικής μεθόδου σε n ψηφία. Όλα τα πειράματα κατέδειξαν ότι, η προτεινόμενη μεθοδολογία δίνει πάντα σωστά αποτελέσματα, ενώ η κλασσική μέθοδος καταστρέφεται πάντα και γρήγορα.

ΠΕΙΡΑΜΑ 8.1

Πράγματι, για τη θεμελίωση της δύναμης και ευστάθειας της μεθόδου που προτείνεται στα πλαίσια της παρούσης εργασίας, υλοποιήθηκε και πραγματοποιήθηκε το εξής πείραμα:

Ο αλγόριθμος Bernoulli Chaotic Map εκτελέστηκε παράλληλα με τέσσερα διαφορετικά μήκη λέξης και πάντα την ίδια είσοδο x_0 . Τα μήκη λέξης που επελέγησαν ήταν δεκάξι, χίλια, δυο χιλιάδες και τέσσερις χιλιάδες δεκαδικά ψηφία. Ο λόγος που επελέγησαν τόσο μεγάλα μήκη λέξεως οφείλεται στο γεγονός ότι ο συγκεκριμένος αλγόριθμος Bernoulli “μολύνεται” ταχύτατα και ραγδαία από το λάθος πεπερασμένης ακρίβειας. Επομένως, οι λόγοι για τους οποίους επελέγησαν τα τρία τελευταία πολύ μεγάλα μήκη λέξεως είναι οι εξής:

1. Όταν τόσο μεγάλα μήκη λέξεως μολύνονται σε σχετικά μικρό αριθμό αναδρομών, αναδεικνύεται άμεσα το πόσο ευάλωτος είναι ο αλγόριθμος Bernoulli map στο συγκεκριμένο αριθμητικό λάθος.
2. Παρ' όλα αυτά, ένα μήκος λέξης που περιλαμβάνει χίλια δεκαδικά ψηφία εγγυάται ότι τα δεκάξι πρώτα δεκαδικά ψηφία των αποτελεσμάτων θα είναι ορθά για ικανό πλήθος αναδρομών.
3. Προβάλλοντας έναν τυχόντα αριθμό x με μήκος λέξης χίλια σε δεκάξι δεκαδικά ψηφία λαμβάνεται η αναπαράσταση x_{16}^{1000} αυτού. Επαναλαμβάνοντας τη διαδικασία αυτή της προβολής του x όταν αυτός αναπαρίσταται με 2000 και 4000 δεκαδικά ψηφία, λαμβάνονται οι αναπαραστάσεις x_{16}^{2000} και x_{16}^{4000} του αριθμού x . Συγκρίνοντας τα x_{16}^{2000} και x_{16}^{4000} με χρήση του ορισμού 1, είναι δυνατόν να υπολογιστεί ο ακριβής αριθμός των ορθών ψηφίων του x στα πρώτα 16 ψηφία αυτού. Εάν τα x_{16}^{2000} και x_{16}^{4000} είναι ίσα, τότε υπάρχει βεβαιότητα ότι τα 16 πρώτα ψηφία του x είναι απολύτως ορθά. Συγκρίνοντας εκ νέου σε αυτή τη περίπτωση το x_{16}^{1000} με αυτά τα 16 ορθά ψηφία, συνάγεται άμεσα ο αριθμός των λανθασμένων δεκαδικών ψηφίων του x_{16}^{1000} .

Με εφαρμογή της προαναφερθείσας λογικής, κατέστησαν σαφή τα κάτωθι:

- a. Ο κλασικός αλγόριθμος Bernoulli καταστρεφόταν εντυπωσιακά γρήγορα, σε μερικές εκατοντάδες αναδρομές.
- b. Ο προτεινόμενος αλγόριθμος έδινε πάντα απολύτως ορθά τα δεκάξι πρώτα δεκαδικά ψηφία, όσο τα x_{16}^{2000} και x_{16}^{4000} ήταν ταυτόσημα.
- c. Μόλις τα x_{16}^{2000} και x_{16}^{4000} διαφοροποιούνται στα δυο τουλάχιστον τελευταία δεκαδικά ψηφία τους, τότε το μήκος λέξης διπλασιαζόταν και οι υπολογισμοί επαναλαμβάνονταν από το σημείο που τα x_{16}^{2000} και x_{16}^{4000} δεν ήταν μολυσμένα. Με αυτό τον τρόπο, εξασφαλιζόταν ότι τα δεκάξι πρώτα δεκαδικά ψηφία των δύο μεγάλων αυτών μηκών λέξεως ήταν πάντα ορθά.

ΠΕΙΡΑΜΑ 8.2

Είσοδος στον αλγόριθμο:

$$x_0 = 1.375675568 \cdot 10^{-1}$$

Παρατηρείται ότι η ακρίβεια της mantissa του x_0 είναι $n = 10$, καθώς και ο εκθέτης αυτού είναι $p = -1$. Τότε:

1. Χρησιμοποιώντας παράλληλα με το προαναφερθέν μήκος λέξης τη διπλή ακρίβεια της υπολογιστικής μηχανής (IEEE 754 double-precision format), ο κλασικός αλγόριθμος Bernoulli map δίνει αναξιόπιστα αποτελέσματα (για $n = 10$) μετά από 53 επαναλήψεις. Συγκεκριμένα, η τιμή που υπολογίζεται στην πεντηκοστή τρίτη επανάληψη είναι $x_{53} = 1.250 \cdot 10^{-1}$, με τη σωστή τιμή να είναι $x_{53correct} = 8.62483456 \cdot 10^{-2}$
2. Χρησιμοποιώντας την MPFR βιβλιοθήκη για την εξασφάλιση ακρίβειας 20000 δεκαδικών ψηφίων, ο κλασικός υπολογισμός του Bernoulli map δίνει αναξιόπιστα αποτελέσματα μετά από 66387 επαναλήψεις. Πράγματι, αναφέροντας τα 16 πρώτα δεκαδικά ψηφία:

$$x_{66386} = 7.809960427520000 \cdot 10^{-1} \text{ και } x_{66387} = 5.619920855039999 \cdot 10^{-1}$$

Αντιθέτως, η μεθοδολογία διόρθωσης που προτείνεται εδώ δίνει σωστά αποτελέσματα και μετά από την επανάληψη που αναφέρεται παραπάνω, δηλαδή την 66387, η οποία επιβεβαιώνεται και όταν όλοι οι υπολογισμοί γίνονται με μήκος λέξης 40000 δεκαδικών ψηφίων, σύμφωνα με το θεώρημα 6.1. Πράγματι:

$$x_{66386} = 7.809960427520000 \cdot 10^{-1} \text{ και } x_{66387} = 5.619920855040000 \cdot 10^{-1}.$$

ΠΕΙΡΑΜΑ 8.3

Είσοδος στον αλγόριθμο:

$$x_0 = 2.9845177519611193657895434567893 \cdot 10^{-1}$$

Παρατηρείται ότι η ακρίβεια της mantissa του x_0 είναι $n = 32$, καθώς και ο εκθέτης αυτού είναι $p = -1$. Τότε,

1. Χρησιμοποιώντας 32 δεκαδικά ψηφία ακρίβεια που προσφέρεται από την MPFR βιβλιοθήκη ο αλγόριθμος Bernoulli map έχει καταστραφεί μετά από 108 επαναλήψεις. Μετά από αυτήν την επανάληψη ο μη διορθωμένος Bernoulli δίνει ολικά αναξιόπιστα αποτελέσματα και δη μηδενικές τιμές.
2. Χρησιμοποιώντας τη MPFR βιβλιοθήκη για την εξασφάλιση ακρίβειας 40000 δεκαδικών ψηφίων, έγινε σύγκριση των τιμών που δίνονται από τον διορθωμένο αλγόριθμο και την προβολή των αριθμών με 40000 ακρίβεια σε 32 δεκαδικά ψηφία. Σε αυτή τη περίπτωση και οι δυο εκτελέσεις έδιναν τον ίδιο αποτέλεσμα $5.4320510855416601846604107874304 \cdot 10^{-1}$ στην 108 – οστή επανάληψη.

Σε αυτό το σημείο τονίζεται ότι είναι αδύνατη η σταθεροποίηση ενός αλγορίθμου, στον οποίο γεννάται ένα τέτοιο είδος σφάλματος πεπερασμένης ακρίβειας όταν η αρχική τιμή εισόδου x_0 είναι μια πεπερασμένη αναπαράσταση ενός άρρητου αριθμού. Για παράδειγμα, εάν κρατηθούν μόνο τα n πρώτα ψηφία του άρρητου αριθμού $\sqrt{2}$, μια παρόμοια ανάλυση όπως τη παραπάνω δείχνει ότι το σφάλμα πεπερασμένης ακρίβειας θα διαδοθεί τελικά γρήγορα και θα καταστρέψει τον αλγόριθμο. Αυτό οφείλεται και στο γεγονός ότι εκτός από το λάθος πεπερασμένης ακρίβειας προστίθεται και το αριθμητικό λάθος που οφείλεται στη πεπερασμένη / ρητή αναπαράσταση του $\sqrt{2}$. Δηλαδή, έχοντας την αναλυτική λύση, ο αλγόριθμος Bernoulli map και οποιοσδήποτε άλλος παρόμοιος, όταν εκτελείται στην υπολογιστική μηχανή, θα δίνει αποτελέσματα εντελώς διαφορετικά της αναλυτικής λύσης. Γενικά τα παραπάνω ισχύουν για οποιεσδήποτε τιμές εισόδου οι οποίες είναι άρρητες και πόσο μάλλον τις υπερβατικές.

9 Μελέτη και ανίχνευση του σφάλματος πεπερασμένης ακρίβειας στον αλγόριθμο Baker's map και μια μέθοδος διόρθωσης του σφάλματος αυτού.

9.1 Σύντομη περιγραφή και αριθμητική συμπεριφορά του αλγορίθμου Baker's map.

Έστωσαν x_0, y_0 δυο αριθμοί εκφρασμένοι στο δεκαδικό σύστημα με κινητή υποδιαστολή, για τις οποίες ισχύει $0 \leq x_0 < 1$ και $0 \leq y_0 < 1$. Εκκινώντας από τιμές x_0, y_0 ο αλγόριθμος Baker's map ορίζεται ως εξής:

$$(x_{k+1}, y_{k+1}) = \begin{cases} \left(2x_k, \frac{y_k}{2}\right) & \text{όταν } 0 \leq x_k < \frac{1}{2} \\ \left(2(1-x_k), 1 - \frac{y_k}{2}\right) & \text{όταν } \frac{1}{2} \leq x_k < 1 \end{cases} \quad (9.1)$$

Όπως συμβαίνει και με το Bernoulli map, ο ανωτέρω αλγόριθμος του Baker, εμφανίζει επίσης σοβαρότατο λάθος πεπερασμένης ακρίβειας, όταν εκτελείται στον υπολογιστή. Η συσσώρευση αυτού

του λάθους καθιστά των Baker's map μη εφαρμόσιμο σε πολλά πρακτικά προβλήματα, στα οποία θα μπορούσε αυτός να παίξει καταλυτικό ρόλο.

9.2 Τα βαθύτερα αίτια της καταστροφής του Baker's map λόγω σφάλματος πεπερασμένης ακρίβειας.

Ο λόγος γένεσης και συσώρευσης του σφάλματος πεπερασμένης ακρίβειας, όσον αφορά τη ποσότητα x_{k+1} , είναι παρόμοιος με αυτόν του Bernoulli map, όπως παρουσιάστηκε στο προηγούμενο κεφάλαιο 8. Συγκεκριμένα:

1. Η εκτέλεση της πράξης $(1 - x_k)$ προκαλεί βύθιση του εκθέτη του αποτελέσματος κατά μία τάξη έστω m . Σύμφωνα με την ενότητα 4.1 αυτή η βύθιση παράγει m e.d.d. νομοτελειακά.
2. Ο πολλαπλασιασμός με το δύο μπορεί να προκαλέσει δεξιά ολίσθηση, όταν ο αριθμός επαναδιατυπώνεται σε επιστημονική μορφή, η οποία όταν επισυμβαίνει, ελαττώνει το πλήθος του σφάλματος πεπερασμένης ακρίβειας κατά ένα.
3. Σε αντίθετη περίπτωση με το (2), ο πολλαπλασιασμός της ποσότητας x_k ή της $1 - x_k$ με τον αριθμό δυο διπλασιάζει και το οποιοδήποτε λάθος το οποίο έχει παραχθεί και διαδοθεί λόγω προηγούμενων πράξεων. Όταν μόνο η τάξη του f.p.e. μεγαλώνει, λόγω του πολλαπλασιασμού με το 2, τότε συνήθως η τάξη του λάθους να μεγαλώνει συνεχώς. Συνεπώς, το λάθος γίνεται κυρίαρχο και ο αλγόριθμος Baker's map καταστρέφεται μετά από έναν αριθμό επαναλήψεων.
4. Αντιθέτως, ο υπολογισμός του y_k πρακτικά δε παράγει λάθος πεπερασμένης ακρίβειας για του κάτωθι λόγους:
 - i. Αρχικά τονίζεται ότι η πράξη $(1 - \frac{y_k}{2})$ μπορεί να οδηγήσει στη γένεση σφάλματος πεπερασμένης ακρίβειας. Όμως, εν αντιθέσει με τον υπολογισμό του x_k , ο εκθέτης του αποτελέσματος y_{k+1} βυθίζεται κατά ένα (1), διότι $0 \leq y < 1$ και επομένως ισχύει η συνθήκη $0 < \frac{y_k}{2} < \frac{1}{2}$. Επομένως, η ποσότητα y_{k+1} υπολογίζεται με το πολύ ένα λανθασμένο ψηφίο λόγω νομοτελειακού λάθους πεπερασμένης ακρίβειας, όπως περιγράφεται στην ενότητα 4.1. Επιπλέον, σύμφωνα με την ενότητα 4.2, η πράξη $(1 - \frac{y_k}{2})$ δε παράγει κάποιο στοχαστικό / τυχαίο λάθος, αφού ο όρος 1 στη πράξη αυτή είναι καθαρός αριθμός.
 - ii. Επιπροσθέτως, διαιρώντας τον αριθμό y_k με το 2, υποδιπλασιάζεται και το οποιοδήποτε λάθος με το οποίο εισέρχεται η ποσότητα y_k στη πράξη της διαίρεσης. Εκτιμάται ότι δύο ή τρεις συνεχόμενες διαιρέσεις με τον αριθμό δύο μπορούν να

ελαττώσουν κατά ένα το πλήθος των λανθασμένων ψηφίων λόγω πεπερασμένης ακρίβειας.

Επομένως, με μεγάλη πιθανότητα το y_k παράγεται συστηματικά με ένα ή δύο ή κανένα λάθος δεκαδικά ψηφία στη mantissa.

9.3 Διόρθωση του σφάλματος πεπερασμένης ακρίβειας στον αλγόριθμο Baker's map.

Η μεθοδολογία που ακολουθείται για τη διόρθωση του σφάλματος πεπερασμένης ακρίβειας στον αλγόριθμο Baker's map είναι ανάλογη αυτής που παρουσιάζεται στην ενότητα 8.3. Όταν η αρχική τιμή x_0 έχει n δεκαδικά ψηφία ακρίβεια στη mantissa και ρ είναι η τάξη του εκθέτη, τότε, κατά την k -οστή αναδρομή του αλγορίθμου η ποσότητα x_k υπολογίζεται πάντα με $n + \rho \leq 2n$ δεκαδικά ψηφία στη mantissa, για τους ίδιους λόγους που παρουσιάζονται στο κεφάλαιο 8 και συγκεκριμένα λόγω του νομοτελειακού λάθους. Επομένως:

Έστω ότι η ποσότητα x_k υπολογίζεται με n δεκαδικά ψηφία στη mantissa, ενώ ο εκθέτης αυτής ισούται με $E(x_k) = -1$. Τότε:

- i. Αν $x_k < \frac{1}{2}$, τότε ο διπλασιασμός του x_k για τον υπολογισμό του x_{k+1} , διπλασιάζει το λάθος πεπερασμένης ακρίβειας, χωρίς όμως να μεγαλώνει την τάξη του αποτελέσματος. Προφανώς, εάν η ανίσωση $0 \leq x_k < \frac{1}{2}$ εξακολουθεί να ισχύει για μια συνεχή ακολουθία υπολογισμών, τότε το λάθος πεπερασμένης ακρίβειας μεγαλώνει συχνά ανεξέλεγκτα.
- ii. Σε αυτό το σημείο τονίζεται ότι, η ποσότητα $2(1 - x_k)$ γράφεται και ως $-(2x_k - 2)$. Στη περίπτωση όπου το $\frac{1}{2} \leq x_k < 1$, ο εκθέτης του αποτελέσματος θα έχει τη τιμή μηδέν, δηλαδή $E(2x_k) = 0$, ενώ η πλήρης mantissa θα καταλαμβάνει $n + 1$ δεκαδικά ψηφία. Συνεπώς, όταν αφαιρείται ο αριθμός δύο από το $2x_k$ για να παραχθεί η ποσότητα x_{k+1} , τότε η τάξη του αποτελέσματος γίνεται μικρότερη από ή ίση με 10^{-1} . Με αυτό τον τρόπο η x_{k+1} έχει το πολύ n δεκαδικά ψηφία στην mantissa. Το αρνητικό πρόσημο δεν επηρεάζει το μήκος λέξης.

Συνοπτικά, αν ο εκθέτης της ποσότητας x_0 είναι ρ , $\rho > 1$ και η ακρίβεια της mantissa είναι n δεκαδικά ψηφία, τότε όλες οι ποσότητες x_k του Baker's map υπολογίζονται με το πολύ $(n + \rho - 1)$ δεκαδικά ψηφία στη mantissa.

Όλοι οι υπολογισμοί των ποσοτήτων που παράγονται από τον αλγόριθμο του Baker's map έχουν στη mantissa το πολύ $\tau = n + \rho + 2$ δεκαδικά ψηφία.

Επιπλέον, σύμφωνα με την ανάλυση του κεφαλαίου 4 και της ενότητας 8.3, αν p είναι το πλήθος των e.d.d. του x_k , τότε το x_{k+1} θα έχει το πολύ $\psi = p + 2 + E(x_k) - E(x_{k+1})$ λανθασμένα δεκαδικά ψηφία. Άρα, όλα τα λάθος ψηφία της ποσότητας x_{k+1} βρίσκονται μεταξύ της ψ -οστής θέσης και της τ -οστής.

Η διόρθωση μιας ποσότητας x_k γίνεται μηδενίζοντας τα ψηφία της mantissa αυτής μετά τη ψ -οστή θέση, ακολουθώντας παρόμοια λογική με αυτή της ενότητας 8.3. Με αυτόν τον τρόπο υπολογίζονται όλα τα x_{k+1} έχοντας μηδενικό πλήθος λανθασμένων δεκαδικών ψηφίων λόγω πεπερασμένης ακρίβειας. Όπως έχει ήδη αναφερθεί παραπάνω, η διόρθωση επιτυγχάνεται αν ο κώδικας μετατρέπει τον εκάστοτε αριθμό σε κείμενο για να μηδενίσει να επιθυμητά ψηφία, και έπειτα τον επαναφέρει τον αριθμό αυτό στον αρχικό του τύπο δεδομένων. Τονίζεται ότι και σε αυτή τη περίπτωση πρέπει να λαμβάνεται υπόψη η υπολογιστική ισότητα ενός αριθμού με τη ποσότητα της οποίας τα τελευταία ψηφία στη mantissa, αντί να είναι μηδέν, έχουν τον αριθμό 9.

Στη περίπτωση όπου η ακρίβεια που απαιτείται είναι $\tau \leq 18 \Leftrightarrow n + \rho \leq 16$, τότε η διόρθωση του αλγορίθμου επιτυγχάνεται χωρίς κάποιο επιπλέον λογισμικό, διότι, σύμφωνα με τα προαναφερθέντα, επαρκεί η standard long double precision των υπολογισμών. Αντιθέτως, αν $\tau > 18$, τότε για τους υπολογισμούς χρειάζεται μια βιβλιοθήκη που επεκτείνει την ακρίβεια των πράξεων, όπως είναι η Multiple-Precision Floating point computations with correct Rounding (MPFR).

ΠΕΙΡΑΜΑ 9.1

Επί τη βάση του κλασσικού αλγορίθμου έγινε ο υπολογισμός των ποσοτήτων x_k και y_k για έναν αριθμό επαναλήψεων. Τα σχετικά αποτελέσματα παρουσιάζονται στη πρώτη στήλη του πίνακα 9.1. Για το συγκεκριμένο πείραμα, εκτελέστηκε η κλασσική προσέγγιση του αλγορίθμου Baker's map με δεκάξι ψηφία ακρίβεια. Επίσης εκτελέστηκε η κλασσική προσέγγιση του αλγορίθμου Baker's map και με ακρίβεια 4000 δεκαδικών ψηφίων, εξασφαλίζοντας με αυτό τον τρόπο τις σωστές τιμές των δεκαέξι πρώτων ψηφίων, μέχρι τουλάχιστον τη 54 επανάληψη, ούτως ώστε να εφαρμοστεί ο ορισμός 1 για τον υπολογισμό των λανθασμένων λ.ψ της ποσότητας x_k . Παράλληλα, εκτελέστηκε και ο προτεινόμενος αλγόριθμος διόρθωσης του f.p.e με δεκάξι δεκαδικά ψηφία ακρίβεια. Όπως μπορεί να παρατηρηθεί στον πίνακα 9.1, όταν οι υπολογισμοί γίνονται με δεκαέξι δεκαδικά ψηφία ακρίβεια ο κλασσικός, χωρίς

διόρθωση, αλγόριθμος του Baker's map καταστρέφεται μετά από 54 αναδρομές. Δηλαδή, τα αποτελέσματα που παράγονται είναι αναξιόπιστα.

Αντιθέτως, η προτεινόμενη μέθοδος συνέχισε να δίνει ορθά αποτελέσματα για εκατοντάδες αναδρομές, όπως φαίνεται στον πίνακα 9.1, μέχρι $k = 54$. Με αυτόν τον τρόπο επιβεβαιώνεται ότι η προσέγγιση της παρούσης έρευνας δίνει πρακτικά σωστά αποτελέσματα για τεράστιο πλήθος αναδρομών, ενώ η κλασική μέθοδος αποτυγχάνει τάχιστα.

A.A. αναδρομής k	x_k μέσω της Κλασικής προσέγγισης	x_k μέσω της Προτεινόμενης προσέγγισης	Σωστή τιμή της ποσότητας x_k
1	$6.666713337760000 \cdot 10^{-1}$	$6.666713337760000 \cdot 10^{-1}$	$6.666713337760000 \cdot 10^{-1}$
2	$6.666573324480000 \cdot 10^{-1}$	$6.666573324480000 \cdot 10^{-1}$	$6.666573324480000 \cdot 10^{-1}$
3	$6.666853351040001 \cdot 10^{-1}$	$6.666853351040000 \cdot 10^{-1}$	$6.666853351040000 \cdot 10^{-1}$
53	$7.500000000000000 \cdot 10^{-1}$	$8.359693271040000 \cdot 10^{-1}$	$8.359693271040000 \cdot 10^{-1}$
54	$5.000000000000000 \cdot 10^{-1}$	$3.280613457920000 \cdot 10^{-1}$	$3.280613457920000 \cdot 10^{-1}$

Πίνακας 9.1 Πίνακας που θεμελιώνει την απόλυτη υπεροχή της προτεινόμενης μεθόδου για την εκτέλεση του αλγορίθμου Baker's map, σε σχέση με την κλασική μέθοδο εκτέλεσης.

□

ΠΕΙΡΑΜΑ 9.2

Είσοδος στον αλγόριθμο:

$$x_0 = 4.7563847567324536215463456738971 \cdot 10^{-1}$$

$$y_0 = 2.9987654678598764532546756889092 \cdot 10^{-1}$$

Παρατηρείται ότι η ακρίβεια της mantissa του x_0, y_0 είναι $n = 32$, καθώς και ο εκθέτης είναι $\rho = -1$.

Τότε:

1. Χρησιμοποιώντας 32 δεκαδικά ψηφία ακρίβεια που προσφέρεται από την mpfr βιβλιοθήκη και ταυτόχρονα εκτελώντας με την ίδια είσοδο x_0, y_0 τον ίδιο αλγόριθμο με 30000 δεκαδικά ψηφία στη mantissa, αποδεικνύεται ότι ο κλασικός τρόπος εκτέλεσης του αλγορίθμου Baker's map οδηγεί σε καταστροφή μετά από 102 επαναλήψεις. Μετά από αυτήν την επανάληψη ο μη διορθωμένος αλγόριθμος του Baker's map αδυνατεί παντελώς να δώσει σωστές τιμές.

2. Χρησιμοποιώντας εκ νέου τη MPFR βιβλιοθήκη για την εξασφάλιση ακρίβειας 30000 δεκαδικών ψηφίων, έγινε σύγκριση των τιμών που δίνονται από τον διορθωμένο αλγόριθμο και της προβολής σε ακρίβεια 32 δεκαδικών ψηφίων των τιμών που προέρχονται από του υπολογισμούς με 30000 ακρίβεια. Σε αυτή τη περίπτωση και οι δυο εκτελέσεις έδιναν τον ίδιο αποτέλεσμα $4.6237598415521391911969312735232 \cdot 10^{-1}$ στην 102 – οστή επανάληψη. Σύγκριση των αποτελεσμάτων με χρήση μήκους λέξης πολύ μεγαλύτερου του 30000 δεκαδικών ψηφίων, απέδειξε ότι η εκτέλεση του αλγορίθμου Baker’s map συνέχισε να δίνει απολύτως ορθά αποτελέσματα για χιλιάδες αναδρομές. Τονίζεται εκ νέου ότι οι πράξεις στον διορθωμένο αλγόριθμο γινότουσαν πάντα και μόνο με πολύ μικρό μήκος λέξης σύμφωνα με την ανάλυση που προηγήθηκε.
3. Σε όλες τις περιπτώσεις, το πλήθος των λανθασμένων δεκαδικών ψηφίων στον υπολογισμό του y_k δε ξεπερνά ποτέ το ένα. Συνεπώς δεν απαιτείται διόρθωση η μέθοδος υπολογισμού του y_k .

□

Όπως ισχύει και στην περίπτωση του Bernoulli map, εάν οι ποσότητες x_k και y_k είναι άρρητοι αριθμοί, τότε η σταθεροποίηση του αλγορίθμου Baker’s map είναι αδύνατη, αφού οι υπολογιστικές μηχανές εκτελούν τους υπολογισμούς με πεπερασμένο πλήθος δεκαδικών ψηφίων και η αρχική διαφοροποίηση των πεπερασμένων αναπαραστάσεων των x_0, y_0 από τις πραγματικές άρρητες τιμές τους συνεχώς μεγαθύνεται.

10 Βιβλιογραφία Α' Μέρους

- [1] A. C. De Oliveira Miranda, M. A. Meggiolaro, L. F. Martha, and J. T. P. De Castro, “Stress intensity factor predictions: Comparison and round-off error,” *Comput. Mater. Sci.*, vol. 53, no. 1, pp. 354–358, 2012.
- [2] F. Jézéquel, J. M. Chesneaux, and J. L. Lamotte, “A new version of the CADNA library for estimating round-off error propagation in Fortran programs,” *Comput. Phys. Commun.*, vol. 181, no. 11, pp. 1927–1928, 2010.
- [3] Z. Bai, “Error Analysis of the Lanczos Algorithm for the Nonsymmetric Eigenvalue Problem,” *Math. Comput.*, vol. 62, no. 205, p. 209, 1994.
- [4] J. H. Lowenstein and F. Vivaldi, “Anomalous transport in a model of Hamiltonian round-off,”

- Nonlinearity*, vol. 11, no. 5, pp. 1321–1350, 1998.
- [5] G. Turchetti, S. Vaienti, and F. Zanlungo, “Relaxation to the asymptotic distribution of global errors due to round off,” *Epl*, vol. 89, no. 4, 2010.
- [6] R. Schmitz, “Use of chaotic dynamical systems in cryptography,” *J. Franklin Inst.*, vol. 338, no. 4, pp. 429–441, 2001.
- [7] B. Georgeot and D. L. Shepelyansky, “Stable Quantum Computation of Unstable Classical Chaos,” *Phys. Rev. Lett.*, vol. 86, no. 23, 2001.
- [8] M. Bonakdar, M. Samadi, H. Salarieh, and A. Alasty, “Stabilizing periodic orbits of chaotic systems using fuzzy control of Poincaré map,” *Chaos, Solitons and Fractals*, vol. 36, no. 3, pp. 682–693, 2008.
- [9] O. J. Kwon, “Targeting and stabilizing chaotic trajectories in the standard map,” *Phys. Lett. Sect. A Gen. At. Solid State Phys.*, vol. 258, no. 4–6, pp. 229–236, 1999.
- [10] M. Martel, “Semantics of roundoff error propagation in finite precision calculations,” *Higher-Order Symb. Comput.*, vol. 19, no. 1, pp. 7–30, 2006.
- [11] A. H. D. Cheng, “Multiquadric and its shape parameter - A numerical investigation of error estimate, condition number, and round-off error by arbitrary precision computation,” *Eng. Anal. Bound. Elem.*, vol. 36, no. 2, pp. 220–239, 2012.
- [12] A. Gelb, “Parameter optimization and reduction of round off error for the Gegenbauer reconstruction method,” *J. Sci. Comput.*, vol. 20, no. 3, pp. 433–459, 2004.
- [13] A. Kountouris, “A randomized algorithm for controlling the round-off error accumulation in recursive digital frequency synthesis (DFS),” *Digit. Signal Process. A Rev. J.*, vol. 19, no. 4, pp. 534–544, 2009.
- [14] C. N. Papaodysseus, E. B. Koukoutsis, and C. N. Triantafyllou, “Error Sources and Error Propagation in the Levinson-Durbin Algorithm,” *IEEE Trans. Signal Process.*, vol. 41, no. 4, pp. 1635–1651, 1993.
- [15] C. Papaodysseus, E. Koukoutsis, and C. Vassilatos, “Error Propagation and Methods of Error Correction in LS FIR Filtering and I-Step Ahead Linear Prediction,” *IEEE Trans. Signal Process.*, vol. 42, no. 5, pp. 1097–1108, 1994.

- [16] N. J. Higham, *Accuracy and stability of numerical algorithms*. 2nd Ed., SIAM, 2002.
- [17] L. N. Trefethen and D. Bau, *Numerical linear algebra*, SIAM, 1997, Chapter III.
- [18] M. Buhmann, J. Stoer, and R. Bulirsch, *Introduction to Numerical Analysis*, Springer 1991, Chapter 1.
- [19] Κ. Παπαοδυσσεύς, Μ. Έξαρχος, Δ. Αραμπατζής, Φ. Γιαννόπουλος, “Λογική Σχεδίαση Ψηφιακών Συστημάτων”, 2018.
- [20] Φ. Γιαννόπουλος, “Πρωτότυπες Μέθοδοι Αυτόματης Αναγνώρισης Γραφέα και Αντιμετώπισης Υπολογιστικών Προβλημάτων Υλοποίησης” Διδακτορική Διατριβή, Εθνικό Μετσόβιο Πολυτεχνείο, Αθήνα, 2018
- [21] N. Mohd Razali and Y. Bee Wah, “Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests,” *J. Stat. Model. Anal.*, vol. 2, no. 1, pp. 21–33, 2011.
- [22] M. F.J.Jr., “The Kolmogorov-Smirnov test for goodness of fit,” *J. Am. Stat. Assoc.*, vol. 56, no. 1951, pp. 68–78, 1951.
- [23] A.M Selvam, *Chaotic Climate Dynamics*. 2007.

**B. ΝΕΟΣ, ΑΠΟΔΟΤΙΚΟΣ ΑΛΓΟΡΙΘΜΟΣ
ΕΠΕΞΕΡΓΑΣΙΑΣ ΕΙΚΟΝΑΣ ΜΕ
ΕΜΦΑΣΗ ΣΤΗ ΧΡΗΣΗ ΠΟΛΥΩΝΥΜΩΝ
ΚΑΙ ΣΥΝΤΕΛΕΣΤΩΝ ΖΕΡΝΙΚΕ.**

1 Ορισμός των Zernike moments και εφαρμοσιμότητα αυτών.

Μια οποιαδήποτε συνάρτηση δυο μεταβλητών $f(x, y)$ γράφεται σαν άθροισμα πλήρων ορθογωνίων συναρτήσεων (complete orthogonal functions). Συχνά, από αρκετούς συγγραφείς, ένα τέτοιο σύνολο πλήρων ορθογωνίων συναρτήσεων ονομάζεται και “ορθογώνια συναρτησιακή βάση” (orthogonal base functions). Μια τέτοια κλάση είναι και τα Zernike moments ή Zernike coefficients, τα οποία αποτελούν ένα σετ πλήρων ορθογωνίων πολυωνύμων μέσα στο μοναδιαίο κύκλο. Τα Zernike moments παρουσιάστηκαν για πρώτη φορά από τον Teague το 1980 και είναι της μορφής

$$Z_p^q = \frac{p+1}{\pi} \iint_{x^2+y^2 \leq 1} V_p^{q*}(x, y) f(x, y) dx dy \quad (1.1)$$

Όπου p η τάξη και q η επαναληψιμότητα του εκάστοτε συντελεστή Zernike.

Η συνάρτηση $V_p^q(x, y)$ των πολυωνύμων Zernike που εμφανίζεται στο διπλό ολοκλήρωμα στη σχέση (1.1) ορίζεται ως

$$V_p^q(x, y) = R_p^q(\rho) e^{jq\theta}, \rho \in [-1, 1] \quad (1.2)$$

Όπου στον ανωτέρω τύπο δίνεται ότι:

- ρ το μέτρο της επιβατικής ακτίνα του τυχόντος σημείου $M(x, y)$ μέσα στον μοναδιαίο κύκλο, το οποία ισούται με

$$\rho = \sqrt{x^2 + y^2} \quad (1.3)$$

- $\theta = \tan^{-1}\left(\frac{y}{x}\right)$ είναι η γωνία που σχηματίζεται μεταξύ του άξονα του x και της επιβατικής ακτίνας $\vec{\rho} = xi + yj$.
- $R_p^q(\rho)$ ορίζεται ως το ακτινικό πολυώνυμο που υπολογίζεται από τον κάτωθι τύπο.

$$R_p^q(\rho) = \sum_{k=0}^{\frac{p-q}{2}} \frac{(-1)^k (p-k)!}{k! \left(\frac{p+q}{2} - k\right)! \left(\frac{p-q}{2} - k\right)!} \rho^{p-2k} \quad (1.4)$$

Όπου το $p - |q|$ πρέπει να είναι πάντα άρτιο, το $p \geq 0$ και $0 \leq q \leq p$.

Τα Zernike moments έχουν σημαντικά πλεονεκτήματα σε σχέση με τα άλλα σύνολα ορθογωνίων συναρτήσεων. Πράγματι:

- Τα Zernike moments είναι ανεξάρτητα από τη στροφή της εικόνας.
- Ενώ έχουν παρόμοια υπολογιστική πολυπλοκότητα με άλλα είδη moments, όπως τα geometric και Legendre moments, υπερτερούν στην αναπαράσταση της πληροφορίας και στην αποτελεσματική ανακατασκευή της εικόνας, έχοντας αντοχή στον θόρυβο, όπως θα αναπτυχθεί εκτενέστερα στα αμέσως επόμενα κεφάλαια.
- Επιπλέον, κάθε συντελεστής Zernike Z_p^q έχει ανεξάρτητο ρόλο στη διαδικασία της ανακατασκευής της εικόνας. Εξαιτίας αυτού, η ανακατασκευή επιτυγχάνεται πιο εύκολα, συγκριτικά με τα άλλα είδη πλήρων ορθογωνίων συναρτήσεων.
- Η ανεξαρτησία των Zernike moments στη μετατόπιση και το scaling επιτυγχάνεται γραμμικά, με την κανονικοποίηση των ορθογώνιων ακτινικών πολυωνύμων. Σημειώνεται ότι η ανεξαρτησία από τη παράλληλη μετατόπιση εξασφαλίζεται θέτοντας το κέντρο του μοναδιαίου κύκλου στο κέντρο βάρους της εικόνας.

2 Η μεγάλη χρησιμότητα των Zernike moments.

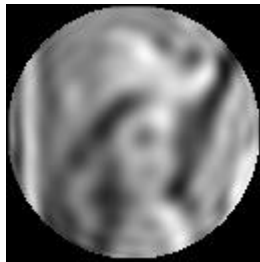
Κάθε εικόνα μπορεί να αναπαρασταθεί από ένα πλήθος συντελεστών Zernike. Οι χαμηλής τάξης συντελεστές Zernike περιέχουν τη πληροφορία για τα περιγράμματα της εικόνας που αντιπροσωπεύουν, ενώ οι υψηλότερης τάξης περιγράφουν τη λεπτομέρεια. Η επιλογή της μέγιστης τάξης Zernike p_{max} για την ικανοποιητική ανακατασκευή της εικόνας εξαρτάται από τα κάτωθι:

- Από το μέγεθος της εικόνας που ανακατασκευάζεται.
- Από τη λεπτομέρεια του περιεχομένου της εικόνας.
- Από το επιθυμητό λάθος ανακατασκευής και την ποιότητα της ανακατασκευασμένης εικόνας.

Μία βασική χρήση των συντελεστών Zernike είναι η συμπίεση της εικόνας. Όντως, επιλέγοντας το κατάλληλο p_{max} μπορεί να ανακατασκευαστεί ικανοποιητικά η αρχική εικόνα, όπως φαίνεται και από στις παρακάτω εικόνες 2.1 και 2.2 .



Αρχική εικόνα



$p_{max} = 25$



$p_{max} = 50$



$p_{max} = 75$



Εικόνα 2.1 Ανακατασκευή της εικόνας Lena, για διάφορα p_{max} , με χρήση της μεθόδου που παρουσιάζεται στην παρούσα εργασία



Αρχική εικόνα



$p_{max} = 25$



$p_{max} = 50$



$p_{max} = 75$



Εικόνα 2.2 Ανακατασκευή της εικόνας Peppers, για διάφορα p_{max} , με χρήση της μεθόδου που παρουσιάζεται στη παρούσα εργασία

Λόγω των χαρακτηριστικών των Zernike coefficients που αναφέρθηκαν παραπάνω, οι χαμηλής τάξης Zernike moments χρησιμοποιούνται κυρίως σε εφαρμογές που έχουν να κάνουν με αναγνώριση οπτικών χαρακτηριστικών, βέλτιστη μοντελοποίηση του κερατοειδούς, αναγνώριση αντικειμένων [1], [2] και ομαδοποίηση χαρακτηριστικών [3]. Επίσης χρησιμοποιούνται ευρέως για την αφαίρεση θορύβου από τις εικόνες, για την εκτίμηση της στροφής μεταξύ των εικόνων καθώς και του καθορισμού των συμμετριών εντός του περιεχομένου αυτών. Από την άλλη πλευρά, οι υψηλής τάξης Zernike moments βρίσκουν εφαρμογή σε βιομετρικές μελέτες [4], όπου είναι πολύπλοκη η αναγνώριση και κατάταξη ανθρώπινων χαρακτηριστικών [5]. Κάτι τέτοιο ισχύει στην αναγνώριση χαρακτηριστικών προσώπου και χεριού [6]. Τέλος οι συντελεστές Zernike αξιοποιούνται από εφαρμογές σχετικά με watermarking [7], [8] και κατάτμηση των εικόνων, καθώς και στον εντοπισμό ακμών που υπάρχουν σε μία εικόνα [9].

3 Υπάρχουσες τεχνικές υπολογισμού των Zernike moments.

Όπως αναφέρθηκε προηγουμένως, τα Zernike moments ορίζονται ως ένα σύνολο ορθογώνιων πολυωνύμων μέσα στο μοναδιαίο κύκλο και ορίζονται επί των πολικών συντεταγμένων (ρ, θ) όπου $\rho \leq 1$ και $\theta \in [0, 2\pi)$. Για αυτό το λόγο ο υπολογισμός των συντελεστών Zernike απαιτεί μια μετατροπή του τυχόντος εικονοστοιχείου (i, j) , $i, j = 0, 1, 2, \dots, N - 1$, όπου N η διάσταση της τετραγωνικής εικόνας. Βάσει των ανωτέρω, ο κάτωθι τύπος αποτελεί το διακριτό μετασχηματισμό της εξίσωσης (1.1) :

$$Z_p^q = \lambda(p, N) \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} R_p^q(\rho_i^j) e^{-iq\theta_i^j} f(i, j) \quad \rho_i^j \leq 1 \quad (3.1)$$

3.1 Τεχνικές μετασχηματισμού μιας εικόνας, εγγεγραμμένης ή περιγεγραμμένης, στον μοναδιαίο κύκλο.

Ο γενικός τύπος μετασχηματισμού της εικόνας στο εσωτερικό ενός κύκλου δίνεται από τους τύπους :

$$\rho = \sqrt{(c_1 i + c_2)^2 + (c_1 j + c_2)^2} \quad (3.2)$$

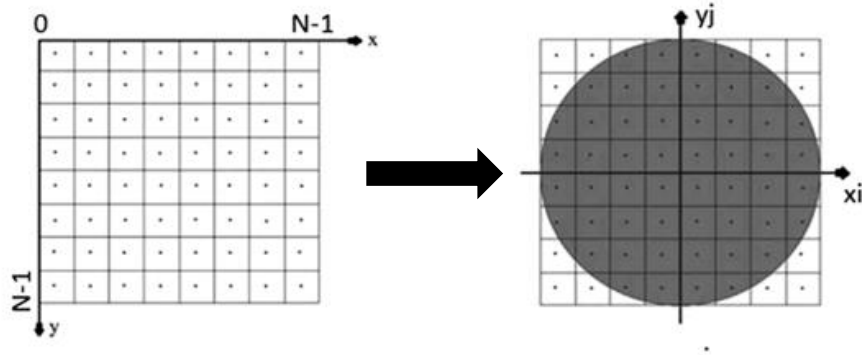
$$\theta = \tan^{-1} \left(\frac{c_1 j + c_2}{c_1 i + c_2} \right) \quad (3.3)$$

όπου (i, j) είναι οι δείκτες του τρέχοντος εικονοστοιχείου και τα c_1, c_2 ορίζονται όπως περιγράφεται στις επόμενες παραγράφους.

3.1.1 Ο μοναδιαίος κύκλος εγγράφεται εντός της τετραγωνικής εικόνας.

Σε αυτήν τη περίπτωση για τους (3.1), (3.2) και (3.3) ισχύει ότι

$$\lambda(p, N) = \frac{p + 1}{(N - 1)^2}, \quad c_1 = \frac{2}{N - 1}, \quad c_2 = -1$$

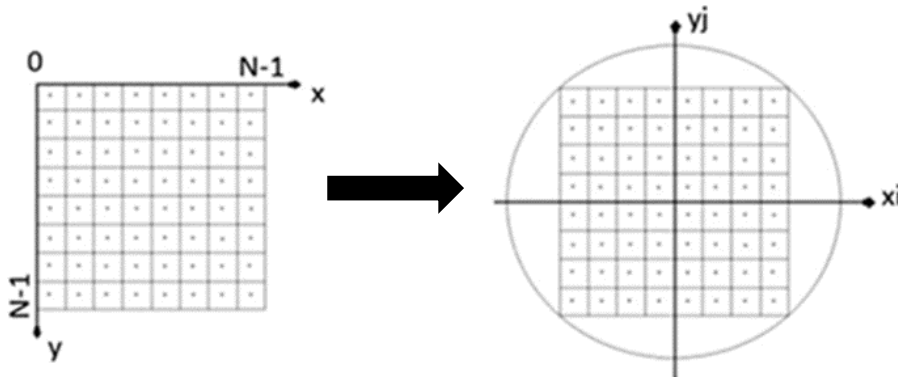


Εικόνα 3.1 Ο μοναδιαίος κύκλος εγγράφεται εντός της τετραγωνικής εικόνας

3.1.2 Όλη η εικόνα εντάσσεται εντός του περιγεγραμμένου κύκλου

Σε αυτήν τη περίπτωση για τους (3.1), (3.2) και (3.3) ισχύει ότι

$$\lambda(p, N) = \frac{2(p + 1)}{\pi(N - 1)^2}, \quad c_1 = \frac{\sqrt{2}}{N - 1}, \quad c_2 = -\frac{1}{\sqrt{2}}$$



Εικόνα 3.2 Όλη η εικόνα εντάσσεται εντός του περιγεγραμμένου κύκλου

3.2 Τεχνικές υπολογισμού των συντελεστών Zernike Z_p^q

Στο παρόν κεφάλαιο παρουσιάζονται κάποιες από τις πιο διαδεδομένες τεχνικές υπολογισμού των συντελεστών Zernike. Οι τεχνικές αυτές είναι οι Direct, Modified Kintner, Modified Prata, Q-recursive, Belkasim, Coefficient και Hybrid, οι οποίες εφαρμόζονται πάνω στα Cartesian/Square pixels (καρτεσιανά/ τετραγωνικά εικονοστοιχεία). Εξαιρώντας την Direct μέθοδο, οι υπόλοιπες τεχνικές έχουν ως σκοπό την ελάττωση της πολυπλοκότητας και χρόνου υπολογισμού των Zernike moments. Τέλος παρουσιάζεται και μια μέθοδος των Polar Zernike moments, η οποία εφαρμόζεται πάνω σε polar pixels (πολικά εικονοστοιχεία) [10],[11]. Σημειώνεται ότι άλλες τεχνικές υπολογισμού των συντελεστών Zernike, οι οποίες εφαρμόζονται πάνω στα Cartesian/Square pixels, παρουσιάζονται στα [12], [13], [14], [15] και [16], ενώ μια άλλη τεχνική υπολογισμού των Polar Zernike moments αναφέρεται στο [17].

3.2.1 Τεχνικές υπολογισμού των ποσοτήτων Zernike πάνω στα Cartesian pixels

Οι μέθοδοι υπολογισμού των ποσοτήτων Zernike χωρίζονται σε αυτές του ευθέως υπολογισμού και του αναδρομικού υπολογισμού.

3.2.1.1 Μέθοδοι ευθέως υπολογισμού

Δυο είναι οι πλέον διαδεδομένοι μέθοδοι ευθέως υπολογισμού του ακτινικού πολυωνύμου $R_p^q(\rho)$, η Direct και η Belkasim.

3.2.1.1.1 Ανάλυση της Direct μεθόδου

Direct μέθοδος ονομάζεται ο υπολογισμός των ακτινικών πολυωνύμων $R_p^q(\rho)$, σύμφωνα με τον ορισμό τους, όπως φαίνεται στην (1.4). Όπως μπορεί να παρατηρηθεί, ο υπολογισμός των ακτινικών πολυωνύμων εξαρτάται από παραγοντικά. Σύμφωνα με το [18] το πλήθος των πολλαπλασιασμών που χρειάζονται για τον υπολογισμό όλων των ακτινικών πολυωνύμων μιας τάξης p ισούται με

$$T(p) \approx \sum_{q=0}^p \frac{(p-q+2)(2p+q)}{2}, \quad p-q \text{ άρτιο} \quad (3.4)$$

Όπου το πλήθος των πολλαπλασιασμών για τον υπολογισμό των παραγοντικών ισούται με

$$\sum_{q=0}^p \frac{(p-q+2)(3p+q)}{4} \quad (3.5)$$

Όπως παρατηρείται από την (3.4) η πολυπλοκότητα υπολογισμού ενός ακτινικού πολυωνύμου είναι $O(p^3)$. Για να υπολογιστούν όλοι οι συντελεστές μιας τάξης p απαιτείται $\frac{p+1}{2}$ φορές η εκτέλεση της (1.4), αν p περιττό και $\frac{p+2}{2}$, αν p άρτιο. Επομένως όλα τα προαναφερθέντα καθιστούν την Direct μέθοδο ακατάλληλη για τον αποτελεσματικό/αποδοτικό υπολογισμό των Zernike moments. Επιπροσθέτως, ο υπολογισμός των συντελεστών μεγάλης τάξης είναι αδύνατη, διότι η εκτίμηση των αντίστοιχων παραγοντικών οδηγεί σε υπερχείλιση (overflow) και συνεπώς σε λάθος αποτελέσματα. Παράδειγμα η IEEE 754 double precision μπορεί να υπολογίσει μέχρι το $170!$, κρατώντας μόνο τα 16 πρώτα ψηφία του αποτελέσματος.

3.2.1.1.2 Ανάλυση της Belkasim μεθόδου [18]

Η μέθοδος αυτή προσπαθεί να ελαττώσει τη πολυπλοκότητα της Direct μεθόδου, καθώς και να επιταχύνει τον υπολογισμό των ακτινικών πολυωνύμων παραγοντοποιώντας κάποιους όρους αυτών. Συγκεκριμένα το πολυώνυμο $R_p^q(\rho)$ υπολογίζεται σύμφωνα με τον κάτωθι τύπο:

$$R_p^q(\rho) = B_{pp}^q R_p^p(\rho) + B_{p(p-2)}^q R_{p-2}^{p-2}(\rho) + \dots + B_{pq}^q R_q^q(\rho) \quad (3.6)$$

όπου

$$B_{pk}^q = \frac{(-1)^{\frac{p-k}{2}} \left(\frac{p+k}{2}\right)!}{\left(\frac{p-k}{2}\right)! \left(\frac{q+k}{2}\right)! \left(\frac{k-q}{2}\right)!} \quad (3.7)$$

Κάνοντας χρήση της (3.6) η (1.1) γράφεται ως

$$Z_p^q = \frac{p+1}{N^2} \left(B_{pp}^q \sum_{\theta} Z_p^p(\theta) e^{-iq\theta} + B_{p(p-2)}^q \sum_{\theta} Z_{p-2}^{p-2}(\theta) e^{-iq\theta} + \dots + B_{pq}^q \sum_{\theta} Z_q^q(\theta) e^{-iq\theta} \right) \quad (3.8)$$

Ο χρόνος υπολογισμού μπορεί να ελαττωθεί περισσότερο εάν

- Τα $Z_k^k, k = 1, 2, \dots, p$ υπολογιστούν μία φορά και μετά πολλαπλασιαστούν με το αντίστοιχο $e^{-iq\theta}$ κάθε φορά.
- Δεν εκτελούνται πολλαπλασιασμοί με τα $e^{-iq\theta}$ όταν το q είναι μηδέν.
- Για μεγάλα p και q το άθροισμα $\sum_{\theta} Z_p^p(\theta)e^{-iq\theta}$ προσεγγίζεται μέσω του γρήγορου μετασχηματισμού Fourier.

Όπως και στη περίπτωση της Direct μεθόδου, ο υπολογισμός των παραγοντικών καθώς και ο υπολογισμός ενός μόνο $R_p^q(\rho)$ σε κάθε αναδρομή, καθιστά απαγορευτική τη χρήση αυτής της μεθόδου στον υπολογισμό των συντελεστών Zernike.

Σύμφωνα με τα παραπάνω γίνεται επιτακτική η χρήση αναδρομικών τύπων, για τον γρήγορο υπολογισμό των συντελεστών Zernike, οι οποίοι και παρουσιάζονται στην επόμενη παράγραφο.

3.2.1.2 Αναδρομικές μέθοδοι υπολογισμού των Zernike moments

Στη παρούσα εργασία αναλύονται οι πιο διαδεδομένες μέθοδοι αναδρομικού υπολογισμού των συντελεστών Zernike. Ακολουθεί η ανάλυση των Coefficient, Modified Prata, Q-recursive, Modified Kintner και Hybrid μεθόδων.

3.2.1.2.1 Ανάλυση της μεθόδου “Coefficient” [18]

Επεκτείνοντας τη μέθοδο του κεφαλαίου 3.2.1.1.2 και εφαρμόζοντας τον μετασχηματισμό $k \leftarrow (p - 2k)$ στην εξίσωση (1.4), έχουμε ότι

$$R_p^q(\rho) = \sum_{k=q}^p B_{pk}^p \rho^k \quad (3.9)$$

Όπου,

$$B_{pp}^p(\rho) = 1 \quad (3.10)$$

$$B_{pp}^{q-2}(\rho) = \frac{B_{pp}^q(\rho)(p+q)}{p-q+2} \quad (3.11)$$

$$B_{p(k-2)}^q(\rho) = \frac{-B_{pk}^q(\rho)(k+q)(k-q)}{(p-k+2)(k+p)} \quad (3.12)$$

Το πρόβλημα αυτής της μεθόδου είναι ότι θα πρέπει σε κάθε αναδρομή να υπολογίζεται ο κάθε συντελεστής και η δύναμη ενός ακτινικού πολυωνύμου. Η πολυπλοκότητα αυτής της μεθόδου είναι $O(p^3)$, παρόλο που δεν υπάρχουν τα παραγοντικά. Επιπλέον είναι εμφανές ότι το $|B_{pk}^q|$ αυξάνεται με ρυθμό p^2 , όσο αυξάνεται η τάξη p . Άρα αυξάνεται δραματικά η πολυπλοκότητα σε μεγάλες τάξεις p .

3.2.1.2.2 Ανάλυση της μεθόδου “Modified Prata” [19]

Η μέθοδος Modified Prata υπολογίζει αναδρομικά τα ακτινικά πολυώνυμα $R_p^q(\rho)$ παρακάμπτοντας τα παραγοντικά. Συγκεκριμένα, τα $R_p^q(\rho)$ υπολογίζονται σύμφωνα με τους κάτωθι τύπους:

$$R_p^p(\rho) = \rho^p, \quad p = 0, 1, \dots, p_{max} \quad (3.13)$$

$$R_p^q(\rho) = pR_p^p(\rho) - (p-1)R_{p-2}^{p-2}(\rho), \quad q = 0, 1, \dots, p_{max}, \quad p = q + 2 \quad (3.14)$$

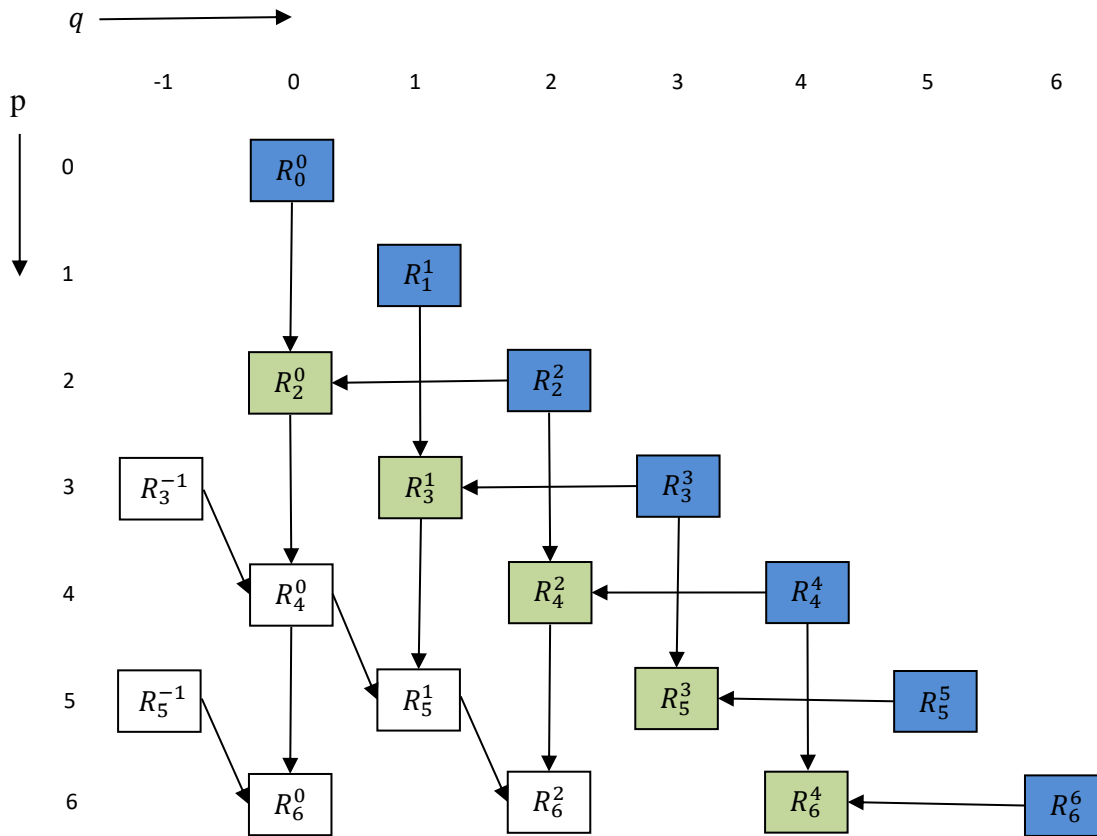
$$R_{p+q}^q(\rho) = L_1 \rho R_{p+q-1}^{|q-1|}(\rho) + L_2 R_{p+q-2}^q(\rho), \quad p = 4, 6, \dots, p_{max}, \quad q = 0, 1, \dots, p_{max} - p \quad (3.15)$$

Τα L_1 και L_2 δίνονται από τους τύπους :

$$L_1 = \frac{2p}{p+q} \quad (3.16)$$

$$L_2 = -\frac{p-q}{p+q} \quad (3.17)$$

Παράδειγμα ενός διαγράμματος ροής (flow chart) υπολογισμού των ακτινικών πολυωνύμων παρουσιάζεται στην Εικόνα 3.3.



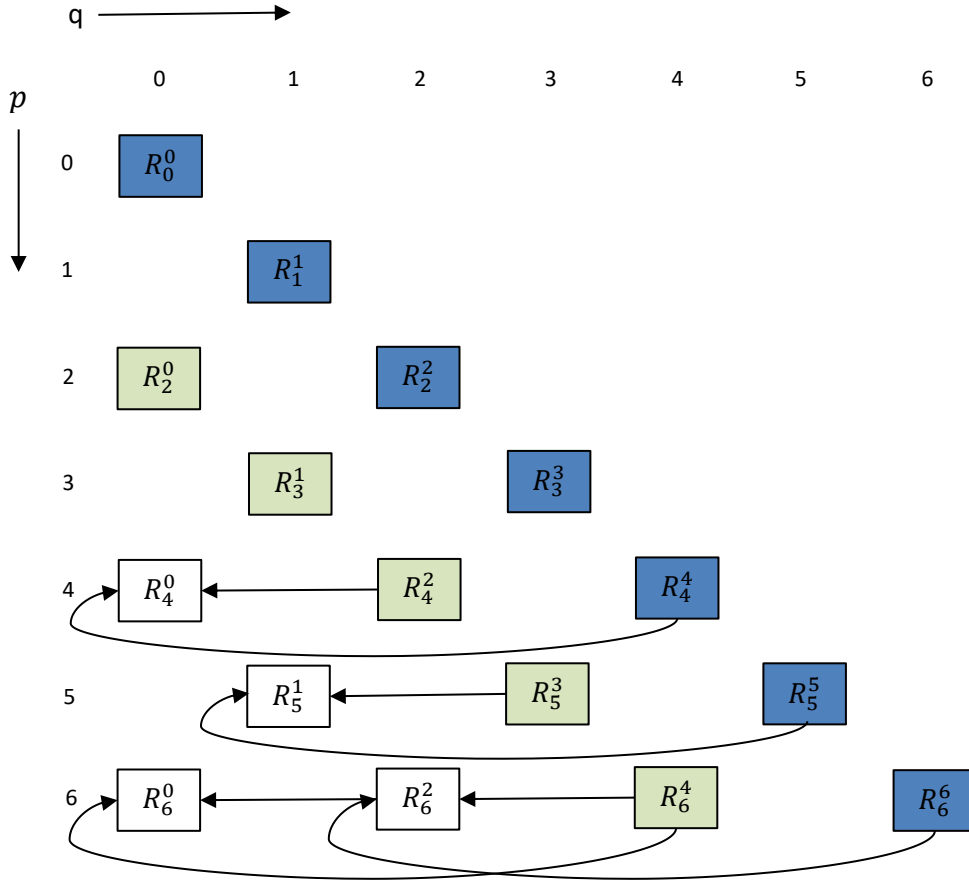
Εικόνα 3.3 Παράδειγμα flow υπολογισμού των ακτινικών πολυωνύμων μέσω των αναδρομικών σχέσεων του modified prata

Παρατηρείται ότι ο υπολογισμός των R_p^0 εμπεριέχει τον υπολογισμό των R_{p-1}^{-1} πολυωνύμων. Φαινομενικά για τον υπολογισμό των R_{p-1}^{-1} απαιτείται να γίνει η χρήση της Direct μεθόδου. Όμως αξιοποιώντας την ιδιότητα $R_p^q = R_p^{-q}$, αποφεύγεται η χρήση της Direct μεθόδου, δεδομένου ότι το R_p^q υπολογίζεται από τη συγκεκριμένη αναδρομική μέθοδο.

Η πολυπλοκότητα του αλγορίθμου αυτού είναι $O(p^2)$. Τέλος για τον υπολογισμό των συντελεστών Zernike γίνεται χρήση της (3.1) εξίσωσης.

3.2.1.2.3 Ανάλυση της μεθόδου “Q-Recursive”

Η Q- recursive μέθοδος που παρουσιάστηκε από τους Chong et al [18] βασίζεται στο γεγονός ότι ο υπολογισμός όλων των ακτινικών πολυωνύμων με δείκτη επαναληψιμότητας $q \neq p$ και $q \neq p - 2$ δεν εξαρτάται από ακτινικά πολυώνυμα χαμηλότερης τάξης, όπως ισχύει στις άλλες αναδρομικές μεθόδους. Οι τύποι υπολογισμού των ακτινικών πολυωνύμων σύμφωνα με αυτή τη μέθοδο είναι:



Εικόνα 3.4 Flow υπολογισμού των ακτινικών πολυωνύμων με τη μεθοδολογία του q-recursive για $p_{max} = 6$

$$R_p^p(\rho) = \rho^p, \quad p = 0, 1, \dots, p_{max} \quad (3.18)$$

$$R_p^q(\rho) = pR_p^p(\rho) - (p-1)R_{p-2}^{p-2}(\rho), \quad q = 2, 3, \dots, p_{max}, \quad q = p-2 \quad (3.19)$$

$$R_p^q(\rho) = H_1 \cdot R_p^{q+4}(\rho) + \left(H_2 + \frac{H_3}{\rho^2}\right) \cdot R_p^{q+2}(\rho) \quad (3.20)$$

Όπου,

$$H_1 = (q+4) \cdot \frac{q+3}{2} - (q+4) \cdot H_2 + H_3 \cdot \frac{(p+q+6) \cdot (p-q-4)}{8} \quad (3.21)$$

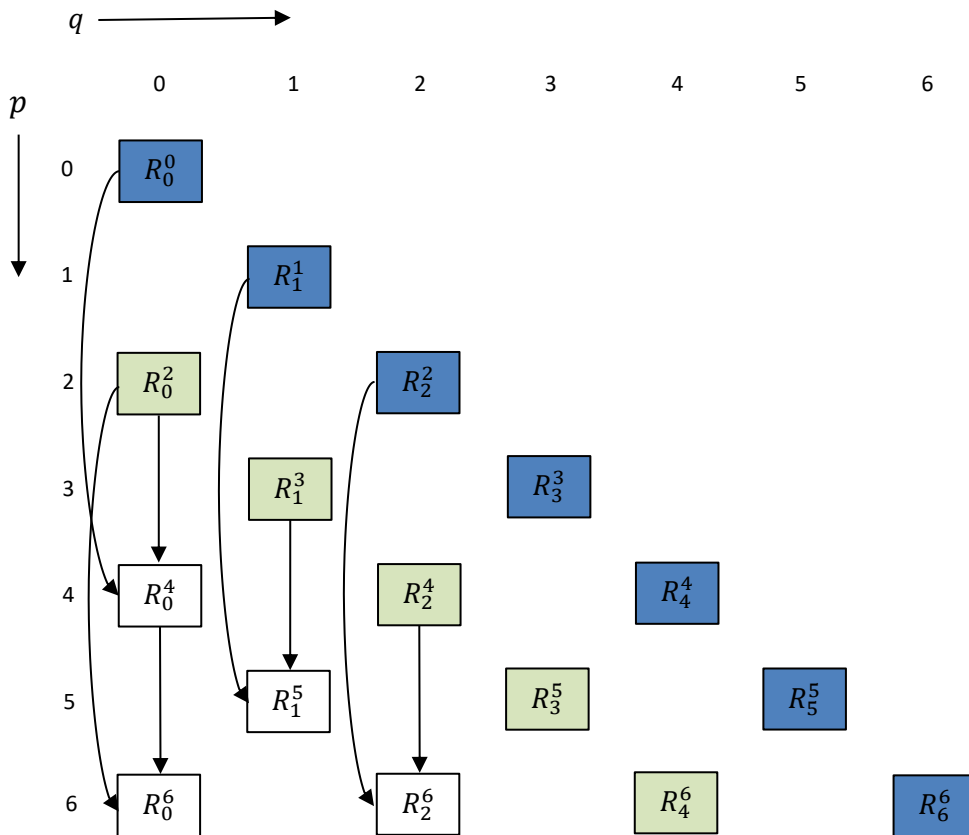
$$H_2 = H_3 \cdot (p+q+4) \cdot \frac{(p-q-2)}{4 \cdot (q+3)} + (q+2) \quad (3.22)$$

$$H_3 = -4 \cdot (q+2) \cdot \frac{q+1}{((p+q+2) \cdot (p-q))} \quad (3.23)$$

Η πολυπλοκότητα του αλγορίθμου αυτού είναι $O(p)$. Τέλος για τον υπολογισμό των συντελεστών Zernike γίνεται χρήση της (3.1) εξίσωσης.

3.2.1.2.4 Ανάλυση της μεθόδου “Modified Kintner”.

Η modified Kintner μέθοδος που παρουσιάστηκε από τους Chong et [18] επεκτείνει την Kintner μέθοδο, η οποία απαιτούσε τον υπολογισμό των ακτινικών πολυωνύμων μέσω της Direct μεθόδου, όταν $p \neq q$. Συνεπώς οι εξισώσεις που υπολογίζουν τα ακτινικά πολυώνυμα είναι οι κάτωθι:



Εικόνα 3.5 Flow υπολογισμού των ακτινικών πολυωνύμων με τη μεθοδολογία του Modified Kintner για $p_{max} = 6$

$$R_p^p(\rho) = \rho^p, \quad p = 0, 1, \dots, p_{max} \quad (3.24)$$

$$R_p^q(\rho) = pR_p^p(\rho) - (p-1)R_{p-2}^{p-2}(\rho), \quad q = 2, 3, \dots, p_{max}, \quad q = p-2 \quad (3.25)$$

$$R_q^p(\rho) = \frac{\left((K_2 \cdot \rho^2 + K_3) \cdot R_{p-2,q}(\rho) + K_4 \cdot R_{p-4,q}(\rho) \right)}{K_1} \quad (3.26)$$

Όπου,

$$K_1 = \frac{(p+q) \cdot (p-q) \cdot (p-2)}{2} \quad (3.27)$$

$$K_2 = 2 \cdot p \cdot (p-1) \cdot (p-2) \quad (3.28)$$

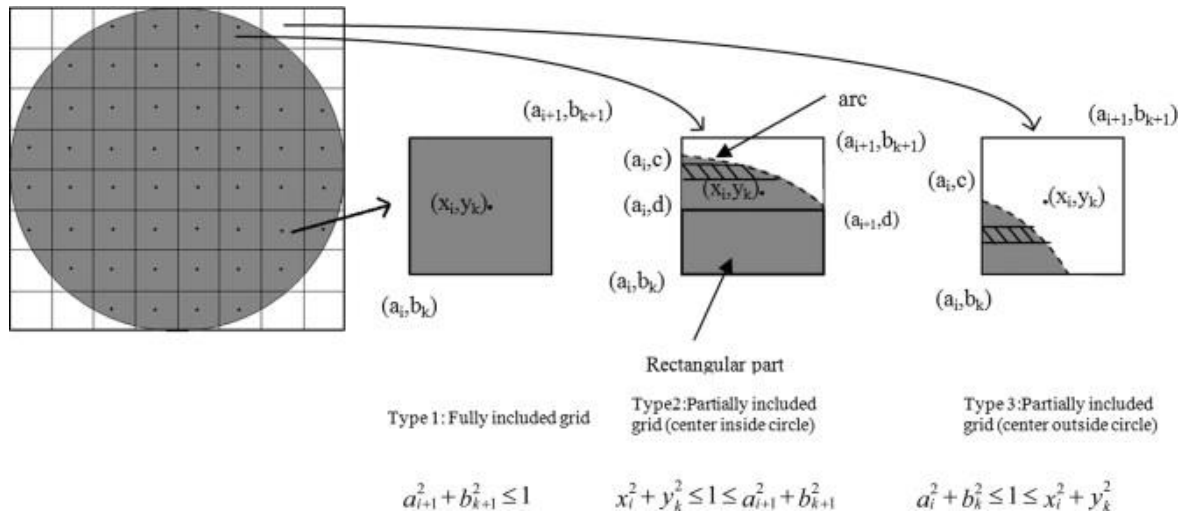
$$K_3 = -q^2 \cdot (p-1) - p \cdot (p-1) \cdot (p-2) \quad (3.29)$$

$$K_4 = -\frac{(p+q-2) \cdot (p-q-2) \cdot p}{2} \quad (3.30)$$

Η πολυπλοκότητα του αλγορίθμου είναι $O(p^2)$, ενώ ένα παράδειγμα flow chart για τον υπολογισμό των ακτινικών πολυωνύμων με $p_{max} = 6$ φαίνεται στην Εικόνα 3.5. Τέλος για τον υπολογισμό των συντελεστών Zernike γίνεται εκ νέου χρήση της (3.1) εξίσωσης.

3.2.1.2.5 Ανάλυση της υβριδικής μεθόδου (Hybrid Method) [20]

Σύμφωνα με αυτή τη προσέγγιση, τα pixels μιας εικόνας εγγεγραμμένης στο μοναδιαίο κύκλο, διαιρούνται σε τρεις κατηγορίες που παρουσιάζονται στην εικόνα 3.6 .



Εικόνα 3.6 Οι τρεις τύποι των pixels, βάσει της Hybrid μεθόδου

Αναλυτικότερα,

- Τα pixels τύπου 1 κείνται εξ' ολοκλήρου εντός του μοναδιαίου κύκλου.
- Τα pixels τύπου 2 τέμνουν το μοναδιαίο κύκλο, αλλά το κέντρο τους βρίσκεται εντός αυτού.
- Τα pixels τύπου 3 επίσης τέμνουν το μοναδιαίο κύκλο, αλλά το κέντρο τους βρίσκεται εκτός αυτού.

Οι συντελεστές Zernike επί των pixels τύπου 1 υπολογίζονται με τις κλασικές προαναφερθείσες μεθόδους. Αντιθέτως, για να υπολογιστούν οι συντελεστές Zernike επί των pixels των τύπων 2 και 3, κάθε ένα από αυτά τα εικονοστοιχεία διαιρείται σε ένα ορθογώνιο πλέγμα $n \times n$ σημείων και εν συνεχεία επί του πλέγματος αυτού εφαρμόζεται η αριθμητική μέθοδος ολοκλήρωσης που είναι διεθνώς γνωστή με τον όρο $n \times n$ point Gaussian quadrature numerical integration method. Κατ' αυτόν τον τρόπο υπολογίζονται οι συντελεστές Zernike επί των τμημάτων των συνοριακών pixels που κείνται εντός του μοναδιαίου κύκλου και μόνο αυτού.

4 Σημαντικά αριθμητικά προβλήματα που παρουσιάζουν οι προαναφερθείσες μέθοδοι υπολογισμού των συντελεστών Zernike.

Οι παραπάνω μέθοδοι εισάγουν τρία είδη σφαλμάτων στον υπολογισμό των Zernike moments, το γεωμετρικό (geometric error), το σφάλμα ολοκλήρωσης (numerical integration error) και το σφάλμα πεπερασμένης ακρίβειας (finite precision error). Θα επακολουθήσει ανάλυση των τριών αυτών τύπων λάθους ευθύς παρακάτω [21].

4.1 Γεωμετρικό σφάλμα στον υπολογισμό των συντελεστών Zernike

Το γεωμετρικό σφάλμα οφείλεται στον μετασχηματισμό των συντεταγμένων της εικόνας στον μοναδιαίο κύκλο, καθώς και στην συνεπακόλουθη μεταβολή του σχήματος των τετραγωνικών εικονοστοιχείων (square pixels). Συγκεκριμένα,

- Στη περίπτωση που ο κύκλος εγγράφεται στην τετραγωνική εικόνα, τα τετραγωνικά pixels που το κέντρο τους βρίσκεται εκτός του κύκλου δε συμμετέχουν στον υπολογισμό των συντελεστών Zernike. Η πληροφορία που εμπεριέχεται σε αυτά τα pixels χάνεται. Συνεπώς υπάρχει μία υποβάθμιση στη ποιότητα της ανακατασκευασμένης εικόνας. Το πρόβλημα αυτό εντείνεται όταν αυτά τα pixels περιέχουν σημαντική πληροφορία.
- Στη περίπτωση που η εικόνα εντάσσεται εντός του κύκλου το γεωμετρικό σφάλμα που εμφανίζεται οφείλεται στο γεγονός ότι κάθε μετασχηματισμένο pixel εντός του κύκλου αντιστοιχεί σε παραπάνω από ένα τετραγωνικά pixels της εικόνας 3.2 . Επίσης, για να εντεθεί η εικόνα στον κύκλο, συχνά γίνεται σημαντική μείωση του μεγέθους της αρχικής εικόνας. Επιπροσθέτως, απαιτείται και μεγαλύτερο πλήθος συντελεστών Zernike για την ανακατασκευή της εικόνας, σε σύγκριση με τη μέθοδο που ο κύκλος εγγράφεται μέσα στην εικόνα.

4.2 Αριθμητικό σφάλμα ολοκλήρωσης στον υπολογισμό των συντελεστών Zernike.

Το σφάλμα ολοκλήρωσης προκαλείται από τη αναγκαιότητα υπολογισμού του διπλού ολοκληρώματος που εμφανίζεται στην (1.1). Ο υπολογισμός του διπλού ολοκληρώματος γίνεται με προσέγγιση μηδενικής τάξης, στην οποία θεωρείται ότι οι τιμές των Zernike πολυωνύμων είναι σταθερές εντός του pixel, καθώς και ότι στο κέντρο του pixel βρίσκεται η τιμή του. Σημειώνουμε ότι αυτό το λάθος μεγαλώνει όσο μικραίνει ο αριθμός των εικονοστοιχείων ή ισοδυνάμως η ανάλυση της εικόνας, δηλαδή το πλήθος των δειγμάτων που λαμβάνονται από μια συγκεκριμένη εικόνα. Αυτό το αριθμητικό λάθος οφείλεται στο γεγονός ότι τα σημεία με καρτεσιανές συντεταγμένες (x, y) εντός ενός τετραγωνικού pixel, μέσω του πολικού μετασχηματισμού, μετατρέπονται σε συναρτήσεις της μορφή $(\rho \cos\theta, \rho \sin\theta)$. Αυτή η τελευταία συναρτησιακή μορφή των συντεταγμένων κάθε σημείου της εικόνας, υποδηλώνει ότι οι τιμές των συντεταγμένων αυτών ταλαντώνονται στο διάστημα $[-\rho, \rho]$. Συνεπώς, και τα ακτινικά πολυώνυμα Zernike ταλαντώνονται στο διάστημα $[-1, 1]$, γεγονός που εισάγει τυχαίο λάθος προσέγγισης κατά τον υπολογισμό του διπλού ολοκληρώματος (1.1) και άρα των τελικών τιμών των Zernike moments, επίσης. Προφανώς αν ο αριθμός των εικονοστοιχείων κάθε πλευράς της εκάστοτε τετραγωνικής εικόνας είναι N ,

σε αυτή τη περίπτωση εάν χρησιμοποιηθεί μέγιστη τάξη p_{max} για τα Zernike moments, με $p_{max} < N$, τότε οπωσδήποτε θα υπάρχει απώλεια πληροφορίας, τόσο μεγαλύτερη όσο μεγαλύτερη είναι η διαφορά $N - p_{max}$.

5 Γένεση και διάδοση του σφάλματος πεπερασμένης ακρίβειας στον υπολογισμό των Zernike moments.

Στο τρέχον κεφάλαιο, με βάση τη παράγραφο 5 και τον ορισμό 1 της παραγράφου 2 του κεφαλαίου Α, θα γίνει ο εντοπισμός των βασικών πηγών f.p.e. λάθους σε διάσημους αλγορίθμους που υπολογίζουν τα πολυώνυμα Zernike καθώς και των σχετικών συντελεστών (Zernike moments).

5.1 Γένεση και διάδοση του σφάλματος πεπερασμένης ακρίβειας στον ευθύ τύπο υπολογισμού των ακτινικών πολυωνύμων Zernike $R_p^q(\rho)$.

Όπως προαναφέρθηκε, το βασικό πρόβλημα στους αλγορίθμους που υπολογίζουν τα Zernike πολυώνυμα με τον ευθύ τύπο, είναι η υπερχειλίση που επισυμβαίνει κατά τον υπολογισμό των παραγοντικών. Αυτό το πρόβλημα είναι ανεξάρτητο από το είδος των pixels που χρησιμοποιείται, δηλαδή είτε αυτά είναι τετράγωνα είτε πολικά pixels. Πολλοί αναδρομικοί αλγόριθμοι δημιουργήθηκαν με σκοπό τη λύση αυτού του προβλήματος. Η γένεση, διάδοση και συσσώρευση του λάθους πεπερασμένης ακρίβειας στους αναδρομικούς αυτούς αλγορίθμους ουδεμία σχέση έχει με υπερχειλίση. Αντιθέτως, το λάθος αυτό οφείλεται στους λόγους που αναφέρθηκαν στη προηγούμενη ενότητα 4 και οι οποίοι θα παρουσιαστούν εκτενώς στο παρόν κεφάλαιο, ευθύς παρακάτω.

Επιπλέον, στη παρούσα εργασία, παρουσιάζεται ένας πρωτότυπος αναδρομικός αλγόριθμος υπολογισμού των πολικών συντελεστών Zernike, ο οποίος μεταξύ άλλων παρακάμπτει και το πρόβλημα της υπερχειλίσης.

5.2 Γένεση, συσσώρευση και διάδοση του σφάλματος πεπερασμένης ακρίβειας στους επαναληπτικούς αλγορίθμους υπολογισμού των ακτινικών πολυωνύμων Zernike.

5.2.1 Ο υπολογισμός του ακτινικού πολυωνύμου $R_p^p(\rho)$: Η βασική πηγή λάθους κοινή για όλους τους σχετικούς αναδρομικούς αλγορίθμους.

Σε όλους τους αλγορίθμους που μελετήθηκαν στην παρούσα έρευνα, παρατηρείται μια κοινή πηγή σφάλματος, όταν γίνεται ο υπολογισμός των διαγώνιων ακτινικών στοιχείων $R_p^p(\rho)$. Συγκεκριμένα, το σφάλμα αυτό οφείλεται στον πολλαπλασιασμό, αφού $R_p^p(\rho) = \rho^p$. Κάθε φορά που το γινόμενο των mantissae δίνει αποτέλεσμα μικρότερο του δέκα, υπάρχει μεγάλη πιθανότητα να παράγεται ένα επιπλέον σφάλμα. Δεδομένου ότι το ρ^p μπορεί να εσωκλείει αρκετούς διαδοχικούς πολλαπλασιασμούς, είναι πιθανόν τα επιμέρους γινόμενα των mantissae να δίνουν συχνά αποτέλεσμα μικρότερο του δέκα. Άρα όσο μεγαλύτερος είναι ο εκθέτης p τόσο περισσότερα δεκαδικά ψηφία λάθος είναι πιθανόν να εμφανιστούν. Συνεπώς, υπάρχει άμεση σχέση μεταξύ του σφάλματος πεπερασμένης ακρίβειας που συσσωρεύεται στο $R_p^p(\rho)$ και της τιμής του p .

Επομένως, η συσσώρευση του λάθους πεπερασμένης ακρίβειας στα ακτινικά πολυώνυμα $R_p^q(\rho)$, όπου $|q| \leq p$, οφείλεται και στην ύπαρξη f.p.e. κατά τον υπολογισμό των $R_p^p(\rho)$.

5.2.2 Βασικές πηγές f.p. λάθους στον υπολογισμό των ακτινικών πολυωνύμων Zernike $R_p^q(\rho)$, $|q| < p$ με τη μέθοδο “Q-recursive”.

Ακολουθώντας την ανάλυση που παρουσιάστηκε σε σχέση με τα [22], [23], την ορολογία 4.1 και την ενότητα 5 του κεφαλαίου A, εκτελέστηκε ο Q-recursive αλγόριθμος με $n = 16$ και $m = 3n = 48$ δεκαδικά ψηφία στη mantissa, έχοντας την ίδια εικόνα ως είσοδο. Συγκεκριμένα, σε κάθε στάδιο του αλγορίθμου, οι υπολογισμοί γινόντουσαν με n και m ακρίβεια, πρακτικά ταυτόχρονα. Έπειτα, εφαρμόσαμε τα όσα περιγράφονται στον ορισμό 2 του κεφαλαίου A, ώστε να γίνει ο εντοπισμός και υπολογισμός του λάθους της n – αναπαράστασης όλων των ποσοτήτων του αλγορίθμου Q - recursive. Σύμφωνα με τους τύπους (3.20), (3.21), (3.22) και (3.23), για τον υπολογισμό $R_p^p(\rho)$ με $|q| \leq p - 4$, στην αρχή έγινε ο υπολογισμός των $R_{p,n}^q(\rho)$ και $R_{p,3n}^q(\rho)$, όπου $R_{p,n}^q(\rho)$ είναι η τιμή των ακτινικών πολυωνύμων $R_p^q(\rho)$ σε μήκος λέξης n και $R_{p,3n}^q(\rho)$ σε μήκος λέξης $m = 3n$. Εν συνεχεία, προβλήθηκε η ποσότητα $R_{p,3n}^q(\rho)$ σε n δεκαδικά ψηφία, παράγοντας τον αριθμό $\widetilde{R}_{p,n}^q(\rho)$. Εν συνεχεία, υπολογίστηκε

ο αριθμός λ των λανθασμένων δεκαδικών ψηφίων με το οποίο παράχθηκε ο $R_{p,n}^q(\rho)$, σύμφωνα με τον ορισμό 1 του κεφαλαίου Β. Αφού υπολογίστηκε το πλήθος λ των λανθασμένων δεκαδικών ψηφίων των $R_{p,n}^q(\rho)$ σε κάθε στάδιο του αλγορίθμου έγινε και ο υπολογισμός τη μέσης τιμής αυτού $\bar{\lambda}$. Η ποσότητα αυτή ήταν μεγαλύτερη του μηδενός ενώ παρέμεινε θετική και για όλες τις εικόνες που εδόθησαν ως είσοδοι στον αλγόριθμο.

Παρατηρήθηκε ότι, αφού ο τύπος (3.26) εκτελείται $[(p-4)/2]$ φορές, τότε στατιστικά οι δύο όροι $H_1 \cdot R_p^{q+4}(\rho)$ και $(H_2 + \frac{H_3}{\rho^2}) \cdot R_p^{q+2}(\rho)$ θα έχουν και αντίθετα πρόσημα, γεγονός που αναγκάζει τον εκθέτη του αποτελέσματος $R_q^p(\rho)$ να “βυθιστεί”. Αυτή η βύθιση του εκθέτη του $R_{p,n}^q(\rho)$ οδηγεί στην εμφάνιση τουλάχιστον του νομοτελειακού (deterministic) λάθους, με μεγάλη πιθανότητα εμφάνισης ενός επιπλέον τυχαίου / στοχαστικού (erratic) λάθους.

ΠΑΡΑΔΕΙΓΜΑΤΑ

Παράδειγμα 5.1

Έστω ότι

- 1) $\rho = 0.955031111111114$ η ακτίνα
- 2) $p = 1241$
- 3) $q = 1069$
- 4) $H_3 = -1.152701778385773 \cdot 10^1$
- 5) $H_2 = 1.351439929711910 \cdot 10^1$
- 6) $H_1 = -9.873815133469179 \cdot 10^{-1}$
- 7) Η $R_p^{q+2}(\rho)$ ποσότητα έχει υπολογιστεί με 5 δεκαδικά ψηφία λάθος
- 8) $R_p^{q+2}(\rho) = -5.809025928993933 \cdot 10^{-2}$
- 9) Η $R_p^{q+4}(\rho)$ ποσότητα έχει υπολογιστεί με 7 δεκαδικά ψηφία λάθος
- 10) $R_p^{q+4}(\rho) = -5.161439980542552 \cdot 10^{-2}$

Σύμφωνα με τα παραπάνω έχουμε

$$H_1 \cdot R_p^{q+4}(\rho) = 5.096310419037392 \cdot 10^{-02}$$

$$\left(H_2 + \frac{H_3}{\rho^2}\right) \cdot R_p^{q+2}(\rho) = -5.090422186484990 \cdot 10^{-02}$$

Αθροίζοντας τους παραπάνω δύο όρους έχουμε

$$R_{1241}^{1069}(\rho) = 5.888232552402062 \cdot 10^{-05}$$

Όπου και σημειώνονται τα λάθος δεκαδικά ψηφία με bold χαρακτήρες. Συνεπώς η βύθιση / μείωση του εκθέτη ισούται με $(-2) - (-5) = 3$, το οποίο σημαίνει ότι τρία επιπλέον λανθασμένα δεκαδικά ψηφία γεννήθηκαν κατά τη πράξη της αφαίρεσης, λόγω του αιτιοκρατικού λάθους πεπερασμένης ακρίβειας. Επαναλαμβάνοντας την προαναφερθείσα διαδικασία για μήκος λέξης $n = 32$ αφ' ενός και $m = 96$ αφετέρου, επιβεβαιώνεται ότι όντως η ποσότητα $R_q^p(\rho)$ υπολογίζεται με δέκα λανθασμένα δεκαδικά ψηφία στο τέλος της mantissa (εφτά που προϋπήρχαν και τρία επιπλέον ψηφία που γεννήθηκαν από το αιτιοκρατικό λάθος). Με αυτόν τον τρόπο, επιβεβαιώσαμε επίσης την ισχύ ενός θεωρήματος που θα διατυπωθεί στα πλαίσια αυτής της εργασίας, σύμφωνα με το οποίο, η "μόλυνση" μιας ποσότητας από λάθος πεπερασμένης ακρίβειας είναι κατ' ουσία ανεξάρτητη από το χρησιμοποιούμενο μήκος λέξης, ενώ προφανώς συσσωρεύεται στα τελευταία δεκαδικά ψηφία της mantissa.

Η επαναλαμβανόμενη εμφάνιση αυτού του φαινομένου οδηγεί στη συσσώρευση λάθους πεπερασμένης ακρίβειας σε μεγαλύτερης τάξης ακτινικών πολυωνύμων Zernike.

Επομένως μια βασική πηγή λάθους πεπερασμένης ακρίβειας είναι η αφαίρεση που γίνεται στον υπολογισμό του $R_q^p(\rho)$ με $|q| \leq p - 4$, μέσω των σχέσεων (3.20), (3.21), (3.22) και (3.23) όταν οι δύο όροι που συμμετέχουν στην πράξη έχουν αντίθετα πρόσημα.

Δεν υπάρχει άλλος τύπος στον αλγόριθμο του Q-recursive που να παράγει συστηματικά λάθος πεπερασμένης ακρίβειας $\bar{\lambda}$. Άρα μόνο ο υπολογισμός του $R_q^p(\rho)$ με $|q| \leq p - 4$ μπορεί να θεωρηθεί ως η κύρια πηγή λάθους στον αλγόριθμο αυτό.

5.2.3 Βασικές πηγές λάθους στον υπολογισμό των ακτινικών πολυωνύμων Zernike

$R_p^q(\rho)$, $|q| < p$ με τη μέθοδο "Modified Kintner".

Ακολουθώντας τα ίδια βήματα όπως και στον Q-recursive αλγόριθμο, αποδείχτηκε ότι η βασική πηγή λάθους πεπερασμένης ακρίβειας είναι η αφαίρεση κατά τον υπολογισμό του $R_q^p(\rho)$ με $|q| \leq p - 4$, όπου σύμφωνα με τους τύπους (3.26), (3.27), (3.28), (3.29) και (3.30) παρατηρήθηκε ότι, αφού ο τύπος (3.26) εκτελείται $\lfloor (p - 4)/2 \rfloor$ φορές, τότε στατιστικά οι δύο όροι $(K_2 \cdot \rho^2 + K_3) \cdot R_{p-2,q}(\rho)$ και $K_4 \cdot R_{p-4,q}(\rho)$ θα έχουν αρκετές φορές αντίθετα πρόσημα, αναγκάζοντας τον εκθέτη του αποτελέσματος

$R_q^p(\rho)$ να βυθιστεί. Αυτή η βύθιση του εκθέτη του ακτινικού πολυωνύμου $R_p^q(\rho)$ οδηγεί στην εμφάνιση τουλάχιστον του νομοτελειακού (deterministic) λάθους πεπερασμένης ακρίβειας.

ΠΑΡΑΔΕΙΓΜΑΤΑ

Παράδειγμα 5.2

Έστω ότι

1. $\rho = 0.955031111111114$ η ακτίνα
2. $p = 433$
3. $q = 327$
4. $K_4 = -17067128$
5. $K_3 = -126814464$
6. $K_2 = 161242272$
7. $K_1 = 17360680$
8. Η $R_{p-2}^q(\rho)$ ποσότητα έχει εκτιμηθεί με 1 δεκαδικό ψηφίο λάθος
9. $R_{p-2}^q(\rho) = 7.114767418907587 \cdot 10^{-2}$
10. Η $R_{p-4}^q(\rho)$ ποσότητα έχει εκτιμηθεί με 3 δεκαδικά ψηφία λάθος
11. $R_{p-4}^q(\rho) = 8.999657507691747 \cdot 10^{-2}$

Σύμφωνα με τα παραπάνω έχουμε

$$\frac{(K_2 \cdot \rho^2 + K_3) \cdot R_{p-2,q}(\rho)}{K_1} = 8.299731685398935 \cdot 10^{-02}$$

$$\frac{K_4 \cdot R_{p-4,q}(\rho)}{K_1} = -8.847482163137392 \cdot 10^{-02}$$

Αθροίζοντας τους παραπάνω δύο όρους έχουμε

$$R_{327}^{433}(\rho) = -5.477504777384568 \cdot 10^{-03}$$

Δεν υπάρχει άλλος τύπος στον αλγόριθμο του modified Kintner που να παράγει συστηματικά λάθος πεπερασμένης ακρίβειας $\bar{\lambda}$. Άρα μόνο ο τύπος υπολογισμού του $R_q^p(\rho)$ μπορεί να θεωρηθεί ως η κύρια πηγή λάθους.

5.2.4 Βασικές πηγές λάθους στον υπολογισμό των ακτινικών πολυωνύμων Zernike

$R_p^q(\rho)$, $|q| < p$ με τη μέθοδο “Modified Prata”.

Η βασική πηγή λάθους f.p.e. στον συγκεκριμένο αλγόριθμο είναι ο τύπος (3.15). Όπως και στις προηγούμενες περιπτώσεις, η αιτία της γένεσης και συσσώρευσης επιπλέον λανθασμένων δεκαδικών ψηφίων στη mantissa των πολυωνύμων $R_p^q(\rho)$ είναι η αφαίρεση και η ενδεχόμενη βύθιση του εκθέτη του αποτελέσματος.

Συνοπτικά, σε όλους τους αναδρομικούς αλγόριθμους που χρησιμοποιούνται για τον υπολογισμό των ακτινικών πολυωνύμων Zernike, η βασική πηγή λάθους πεπερασμένης ακρίβειας είναι η ύπαρξη ενός μαθηματικού τύπου που υπολογίζει τα σχετικά πολυώνυμα, ο οποίος είναι της μορφής $R_p^q(\rho) = \text{Term1} + \text{Term2}$, όπου Term1 και Term2 είναι αριθμοί που έχουν αντίθετο πρόσημο. Επομένως εάν ισχύουν τα παρακάτω

$$\tau = \max\{E(\text{Term1}), E(\text{Term2})\}$$

$$\delta = E(R_p^q(\rho)) - \tau$$

$$\varepsilon_{bef} = \max\{[edd(\text{Term1}), edd(\text{Term2})]\}$$

$$\varepsilon_{aft} = \{[edd(R_p^q(\rho))]\}$$

με $\varepsilon_{bef}, \varepsilon_{aft}$ τα νομοτελειακά λάθη ακριβώς πριν και μετά την εκτέλεση της βασικής πηγής λάθους πεπερασμένης ακρίβειας, τότε ισχύει πάντα

$$\varepsilon_{aft} = \varepsilon_{bef} + \delta, \quad \delta > 0 \quad (5.1)$$

Για να υπολογιστεί το συνολικό λάθος πεπερασμένης ακρίβειας που γεννάται και συσσωρεύεται στις ποσότητες R_p^q , είναι απολύτως απαραίτητο στο προαναφερθέν νομοτελειακό λάθος να προστεθεί αλγεβρικά και το τυχαίο / στοχαστικό f.p. error.

Στον πίνακα 5.1 παρουσιάζονται διάφορα ενδεικτικά αποτελέσματα του λάθους πεπερασμένης ακρίβειας στα ακτινικά πολυώνυμα Zernike R_p^q με $p_{max} = 1200$.

Πλήθος λανθασμένων δεκαδικών ψηφίων	Ποσοστό του λάθους δεκαδικών ψηφίων
0	15.27
1	20.58
2	51.51
3	11.85
4	0.62
5	0.08
6	0.03
7	0.02
8	0.02
9	0.008
10	0.002
11	0.00007

Πίνακας 5.1 Σφάλμα πεπερασμένης ακρίβειας στα ακτινικά πολυώνυμα Zernike R_p^q με $p_{max} = 1200$

5.3 Το λάθος πεπερασμένης ακρίβειας στους συντελεστές Zernike

Σύμφωνα με την ανάλυση που παρουσιάστηκε ως τώρα, για να υπολογιστούν οι συντελεστές Zernike $Z(p, q)$ πρέπει να εκτελεστούν αλληπάλληλες προσθαφαιρέσεις, όπου κάθε μία θα αντιστοιχεί σε ένα επιμέρους άθροισμα. Εάν S_i είναι το μερικό άθροισμα στη παραπάνω διαδικασία, τότε για να πάρουμε το επόμενο S_{i+1} , πρέπει

- Για τετραγωνικά pixels, ο παράγοντας $f(x, y)(R_p^q(\rho)e^{jq\theta})^*$ να προστεθεί στο S_i .
- Ενώ αναφορικά με τα pixels κυκλικού τομέα, στο S_i να προστεθεί ο παράγοντας

$$\hat{f}(\rho_k, \theta_l) \iint_{\Omega_{k,l}} R_p^q(\rho) e^{-jq\theta} \rho \, d\rho d\theta.$$

Από τη διαδικασία αυτή προκύπτουν δυο ενδεχόμενα:

- Να γίνει βύθιση του εκθέτη του S_{i+1} , σε σχέση με τον εκθέτη του μερικού αθροίσματος S_i , κατά μία θετική ποσότητα δ_{i+1} . Τότε σε αυτή τη περίπτωση το S_{i+1} έχει παραχθεί με τα γνωστά επιπλέον λάθος δεκαδικά ψηφία, λόγω του νομοτελειακού λάθους. Αυτά τα λανθασμένα ψηφία μπορούν να μεταβληθούν από την επίδραση του στοχαστικού λάθους. Όσον αφορά αυτό το τυχαίο λάθος, αν $\lambda_i = edd(S_i)$ και $\lambda_{i+1} = edd(S_{i+1})$, τότε πρακτικά το λ_{i+1} παίρνει τιμές από $\lambda_i + 1$ μέχρι $\lambda_i - 4$ · με πιθανότερες τιμές να είναι μηδέν ή ± 1 .
- Να γίνει αύξηση του εκθέτη του S_{i+1} κατά μία ποσότητα ζ_{i+1} . Σε αυτή τη περίπτωση γίνεται δεξιά ολίσθηση ώστε ο S_{i+1} να διατηρήσει τη IEEE 754 ή γενικότερα την επιστημονική μορφή του. Κατά τη πρόσθεση του όρου $f(x,y)(R_p^q(\rho)e^{jq\theta})^*$ για τα τετραγωνικά pixels ή $\hat{f}(\rho_k, \theta_l) \iint_{\Omega_{k,l}} R_p^q(\rho)e^{-jq\theta} \rho \, d\rho d\theta$ για τα πολικά στο S_i , μπορεί το S_{i+1} να υπολογιστεί με λιγότερα δεκαδικά ψηφία, σύμφωνα με την ανάλυση της παραγράφου 4 του κεφαλαίου Α.

Παρακάτω ακολουθούν δυο παραδείγματα αύξησης του λάθους σε ένα στάδιο υπολογισμού των συντελεστών Zernike.

ΠΑΡΑΔΕΙΓΜΑΤΑ

Παράδειγμα 5.3

- Τάξη $p = 311$ και $q = 39$
- Ακτίνα $\rho = 9.294362229410000 \cdot 10^{-1}$
- Γωνία $\theta = -2.246981658962470$
- Τιμή του pixel $f(x,y) = 0.9$

$$\text{τρέχων προσθεταίος } f(x,y)(R_p^q(\rho)e^{jq\theta})^* = -4.193433704404736 \cdot 10^{-5}$$

$$\text{προηγούμενο άθροισμα } S_i = 4.193493539312988 \cdot 10^{-5}$$

$$\text{νέο άθροισμα } S_{i+1} = S_i + f(x,y)(R_p^q(\rho)e^{jq\theta})^* = 5.983490825207938 \cdot 10^{-10}$$

Ως συνήθως, τα λανθασμένα δεκαδικά ψηφία παρατίθεται με bold.

Παράδειγμα 5.4

- Τάξη $p = 611$ και $q = 129$
- Ακτίνα $\rho = 8.619499842727100 \cdot 10^{-1}$
- Γωνία $\theta = 2.400870624913580$
- Τιμή του pixel $f(x, y) = 0.5$

$$\text{τρέχων προσθεταίος } f(x, y)(R_p^q(\rho)e^{jq\theta})^* = 1.231101585568280 \cdot 10^{-5}$$

$$\text{προηγούμενο άθροισμα } S_i = -1.230745216775373 \cdot 10^{-5}$$

$$\text{νέο άθροισμα } S_{i+1} = S_i + f(x, y)(R_p^q(\rho)e^{jq\theta})^* = 3.563687929071841 \cdot 10^{-9}$$

Τα λανθασμένα δεκαδικά ψηφία παρατίθενται με bold.

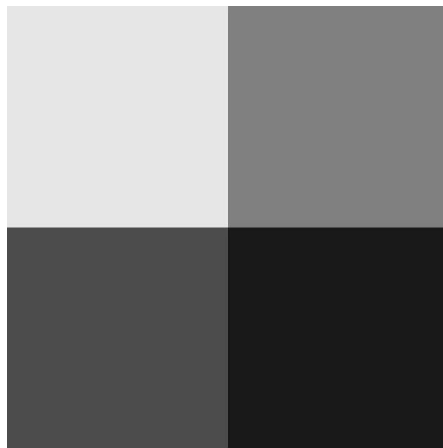
Όπως παρατηρείται στα παραπάνω, η γένεση και διάδοση του σφάλματος πεπερασμένης ακρίβειας είναι αρκετά σύνθετη, αφού επηρεάζεται από πολλούς διαφορετικούς παράγοντες, οι εν πολλοίς είναι τυχαίοι. Συνεπώς η προσπάθεια μοντελοποίησης της βρίσκει πολύ μεγάλες δυσκολίες. Όμως, στη παρούσα εργασία αποδεικνύεται ότι, κατά τους υπολογισμό των Zernike polynomials και moments, το παραγόμενο και συσσωρευόμενο λάθος πεπερασμένης ακρίβειας μπορεί να ανιχνευθεί πλήρως, καθώς και ότι αυτό εξαρτάται από δύο κυρίως παράγοντες:

1. Το είδος των pixels που χρησιμοποιείται, δηλαδή το σφάλμα εξαρτάται από την επιλογή χρήσης μεταξύ πολικών, τετράγωνων ή υβριδικών pixels.
2. Από την επιλογή της αναδρομικής μεθόδου για τον υπολογισμό των ακτινικών πολυωνύμων Zernike (δηλαδή Kintner, Q – recursive, Prata ή οποιονδήποτε άλλο). Αυτοί οι αλγόριθμοι επηρεάζουν τόσο το πλήθος του σφάλματος πεπερασμένης ακρίβειας που συσσωρεύεται στα ακτινικά πολυώνυμα, όσο και κατά συνέπεια το λάθος που εμφανίζεται στο διπλό άθροισμα που υπολογίζει τους συντελεστές Zernike.

Με σκοπό να επαληθευθούν πειραματικά τα παραπάνω, χρησιμοποιήθηκαν δύο ειδών τεχνητές εικόνες καθώς και ένα πλήθος πραγματικών εικόνων, εκ των οποίων οι περισσότερες χρησιμοποιούνται διεθνώς στην επεξεργασία εικόνας.



Εικόνα 5.1 Μονόχρωμες τεχνητές εικόνες, μιας απόχρωσης, που χρησιμοποιήθηκαν για να αποδειχθεί ότι το σφάλμα πεπερασμένης ακρίβειας εξαρτάται από τον αλγόριθμο υπολογισμού των Zernike polynomials.



Εικόνα 5.2 Η τεχνητή εικόνα που χρησιμοποιήθηκε για να αποδειχθεί ότι η γένεση και διάδοση του σφάλματος πεπερασμένης ακρίβειας εξαρτάται μόνο από τον αλγόριθμο που υπολογίζει τους συντελεστές Zernike.

Αρχικά, χρησιμοποιήθηκαν μονόχρωμες (grayscale) τεχνητές εικόνες, διάφορων μεγεθών και μίας απόχρωσης. Με αυτόν τον τρόπο βγήκε κοινός παράγον ο όρος $f(x, y)$ από το άθροισμα Z_p^q , ως άμεση συνέπεια του επιμεριστικού και έπειτα του προσεταιριστικού νόμου. Με αυτόν τον τρόπο εξασφαλίστηκε ότι ο υπολογισμός των συντελεστών Zernike επηρεάζεται μόνο από το είδος των pixels, από το f.p.e. των ακτινικών πολυωνύμων, καθώς και από τις συμμετρίες. Επαναλαμβάνεται με έμφαση ότι το λάθος των ακτινικών πολυωνύμων εξαρτάται πάντα από τη τάξη p του εκάστοτε συντελεστή Zernike.

Το δεύτερο σετ που χρησιμοποιήθηκε ως είσοδος, αποτελείται από εικόνες που διαιρούνται σε τέσσερις (grayscale) τομείς, όπου ο κάθε ένας έχει μία μόνο απόχρωση του γκρι, διαφορετική από τους άλλους τρεις. Οι διαστάσεις των εικόνων αυτού του σετ ήταν ίδιες με αυτές του πρώτου, όπως ήταν ίδιες και η μέγιστη τάξη p_{max} των συντελεστών Zernike.

Οι παραπάνω τεχνητές εικόνες δώσανε πρακτικά, για όλους τους συνδυασμούς p και q , την ίδια συσσωρευμένη ποσότητα σφάλματος πεπερασμένης ακρίβειας στους διάφορους συντελεστές Zernike Z_p^q .

Έπειτα εφαρμόστηκε η ίδια μέθοδος για ένα μεγάλο πλήθος από πραγματικές εικόνες. Τα σχετικά αποτελέσματα παρουσιάζονται παρακάτω. Σημειώνεται εκ νέου ότι οι διαστάσεις των πραγματικών εικόνων ήταν ίδιες με αυτές των τεχνητών, καθώς και η μέγιστη τάξη p_{max} των συντελεστών Zernike.

Τα πειράματα που πραγματοποιήθηκαν απέδειξαν ότι για όλα τα είδη εικόνας που αναφέρονται παραπάνω, το σφάλμα πεπερασμένης ακρίβειας που γεννάται κατά τον υπολογισμό των συντελεστών Zernike είναι σχεδόν ίδιο, όπως φαίνεται στον πίνακα 5.2 και 5.3.

Μέγεθος εικόνας	p_{max}	Μέση τιμή του e.d.d. Πραγματικό μέρος / φανταστικό μέρος	Τυπική απόκλιση του e.d.d. Πραγματικό μέρος / φανταστικό μέρος
512 x 512	500	2.85 / 2.62	0.91 / 0.89
512 x 512	600	2.92 / 2.71	0.90 / 0.87
512 x 512	700	2.98 / 2.76	0.89 / 0.84
700 x 700	800	3.47 / 3.26	0.82 / 0.86
1200 x 1200	1000	3.71 / 3.62	0.86 / 0.90
1200 x 1200	1200	3.75 / 3.70	0.84 / 0.89
1200 x 1200	1600	3.82 / 3.79	0.81 / 0.81
1600 x 1600	1600	3.87 / 3.80	0.83 / 0.93

Πίνακας 5.2 Η μέση τιμή και τυπική απόκλιση του f.p.e. στα πολυώνυμα Zernike, σύμφωνα με τη μέθοδο Modified Kintner. Οι συντελεστές Zernike υπολογίστηκαν με τον μοναδιαίο κύκλο να εγγράφεται εντός της τετραγωνικής εικόνας

Μέγεθος εικόνας	p_{max}	Μέση τιμή του e.d.d. Πραγματικό μέρος / φανταστικό μέρος	Τυπική απόκλιση του e.d.d. Πραγματικό μέρος / φανταστικό μέρος
512 x 512	500	5.98 / 5.89	1.16 / 1.24
512 x 512	600	6.22 / 6.13	1.19 / 1.26
512 x 512	700	6.43 / 6.35	1.19 / 1.25
700 x 700	800	6.89 / 6.85	1.21 / 1.26
1200 x 1200	1000	7.11 / 7.03	1.23 / 1.29
1200 x 1200	1200	7.34 / 7.28	1.17 / 1.23
1200 x 1200	1600	7.61 / 7.56	1.17 / 1.24
1600 x 1600	1600	7.67 / 7.68	1.23 / 1.36

Πίνακας 5.3 Η μέση τιμή και τυπική απόκλιση του f.p.e. στα πολυώνυμα Zernike, σύμφωνα με τη μέθοδο Q-Recursive.

Οι συντελεστές Zernike υπολογίστηκαν με τον μοναδιαίο κύκλο να εγγράφεται εντός της τετραγωνικής εικόνας

Συνεπώς, σύμφωνα με όλη την ανάλυση που προηγήθηκε, το f.p. λάθος εξαρτάται από τους παρακάτω παράγοντες:

- I. Από τη τιμή που έχει ο δείκτης p . Όση μεγαλύτερη είναι η τιμή, τόσα περισσότερα λανθασμένα ψηφία συσσωρεύονται στους συντελεστές Z_p^q .
- II. Από το γεωμετρικό σχήμα των pixels, τετράγωνα ή κυκλικούς τομείς. Έχει αποδειχθεί από πολλούς ερευνητές ότι η χρήση των τετράγωνων pixels είναι άμεσα συνδεδεμένη με τη γένεση σημαντικού γεωμετρικού και αριθμητικού λάθους, που αλλοιώνει την αξιοπιστία των αποτελεσμάτων.
- III. Από το μέγεθος της εικόνας. Κατά κανόνα, όσο μεγαλύτερες είναι οι διαστάσεις της εικόνας, τόσο μεγαλύτερο λάθος πεπερασμένης ακρίβειας συσσωρεύεται στους συντελεστές Zernike.
- IV. Από το είδος του αναδρομικού αλγορίθμου που χρησιμοποιείται για τον υπολογισμό των ακτινικών πολυωνύμων Zernike. Από τους αλγορίθμους που μελετήθηκαν ο modified Kintner υποφέρει λιγότερο από λάθος πεπερασμένης ακρίβειας, όπως και ο αλγόριθμος που παρουσιάζεται στη παρούσα εργασία.

Επαναλαμβάνεται emphaticά ότι η γένεση και διάδοση του λάθους πεπερασμένης ακρίβειας στους συντελεστές Zernike εξαρτάται μόνο από τους παραπάνω παράγοντες και είναι ανεξάρτητο από το

περιεχόμενο της εικόνας. Επιπροσθέτως, οι όποιες συμμετρίες υπάρχουν δεν επηρεάζουν το πλήθος των e.d.d. που παράγεται.

Όντως, υπολογίζοντας τους ίδιους συντελεστές Zernike με χρήση των συμμετριών τη μία φορά και την άλλη χωρίς, παρατηρήθηκε ότι τα αποτελέσματα ήταν παρόμοια όσον αφορά τη γένεση, διάδοση και συσσώρευση του σφάλματος πεπερασμένης ακρίβειας. Συνεπώς, η σημαντικότητα των συμμετριών έγκειται στη ελάττωση του χρόνου εκτέλεσης του εκάστοτε αλγορίθμου και στη μείωση της υπολογιστικής πολυπλοκότητας αυτού.

5.4 Παραδείγματα ανακατασκευής εικόνων με χρήση των τετραγωνικών pixels

Από τις εικόνες 5.3 - 5.11 παρατηρούμε ότι όσο το p_{max} μεγαλώνει, τόσο μεγαλώνουν και τα σφάλματα ανακατασκευής της εικόνας. Από την άλλη πλευρά όμως, η αύξηση του p_{max} είναι απολύτως αναγκαία, δεδομένου ότι για μικρότερα p_{max} , η ανακατασκευασμένη εικόνα είναι “θολή”, εμφανίζει απώλεια πληροφορίας ιδιαίτερα στις λεπτομέρειες και γενικά δεν είναι κατάλληλη για περεταίρω επεξεργασία. Συνεπώς, η ανεύρεση αλγορίθμων που επιτρέπουν την ανακατασκευή εικόνων αναδεικνύοντας όλες τις επιθυμητές λεπτομέρειες αυτών, χωρίς ταυτόχρονα όμως να εμφανίζονται τα προαναφερθέντα σοβαρά αριθμητικά σφάλματα, είναι απολύτως απαραίτητη.



Αρχική εικόνα



$$p_{max} = 100$$



$$p_{max} = 200$$



$$p_{max} = 250$$

Εικόνα 5.3 Ανακατασκευή της εικόνας living room πάνω στα Cartesian pixels, μεγέθους 128×128 και για p_{max} , 100, 200 και 250. Η εικόνα έχει εγγραφεί εντός του μοναδιαίου κύκλου.



Αρχική εικόνα



$$p_{max} = 50$$



$$p_{max} = 100$$



$$p_{max} = 150$$

Εικόνα 5.4 Ανακατασκευή της εικόνας living room πάνω στα Cartesian pixels, μεγέθους 128×128 και για p_{max} , 50, 100 και 150. Ο μοναδιαίος κύκλος έχει εγγραφεί εντός της εικόνας.



Αρχική εικόνα



$$p_{max} = 100$$



$$p_{max} = 200$$



$$p_{max} = 250$$

Εικόνα 5.5 Ανακατασκευή της εικόνας cameraman πάνω στα Cartesian pixels, μεγέθους 128×128 και για p_{max} , 100, 200 και 250. Η εικόνα έχει εγγραφεί εντός του μοναδιαίου κύκλου.



Αρχική εικόνα



$$p_{max} = 50$$

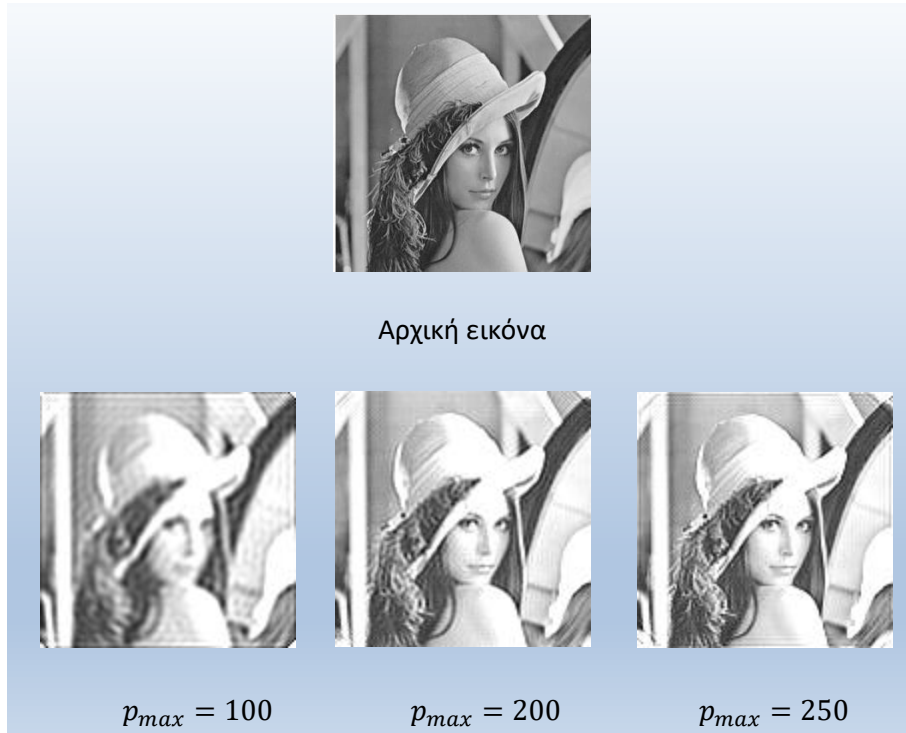


$$p_{max} = 100$$

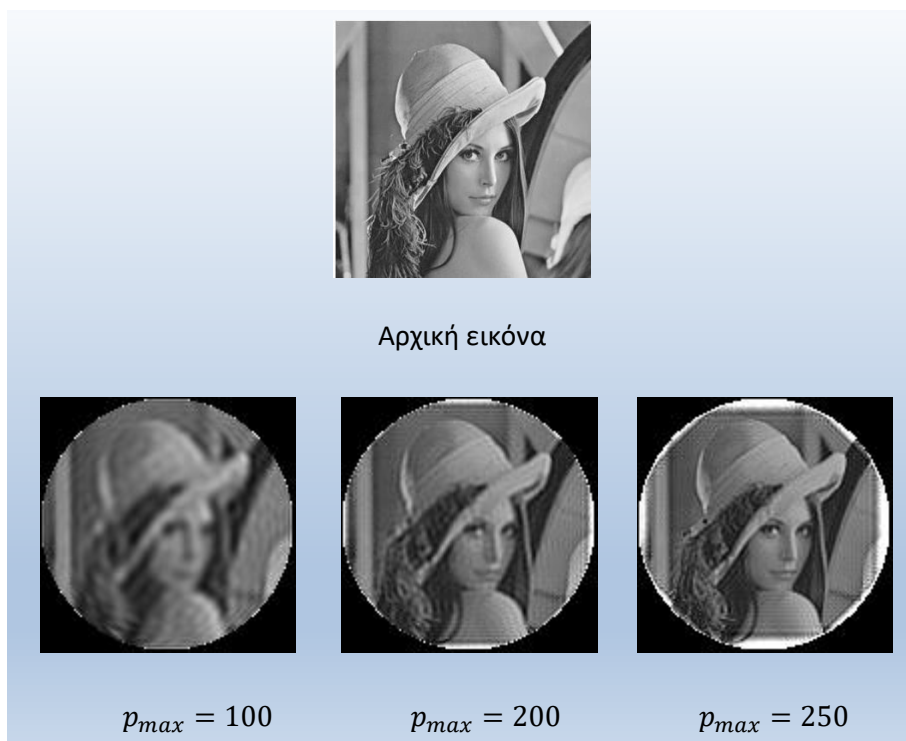


$$p_{max} = 150$$

Εικόνα 5.6 Ανακατασκευή της εικόνας cameraman πάνω στα Cartesian pixels, μεγέθους 128×128 και για p_{max} , 50, 100 και 150. Ο μοναδιαίος κύκλος έχει εγγραφεί εντός της εικόνας



Εικόνα 5.7 Ανακατασκευή της εικόνας lena πάνω στα Cartesian pixels, μεγέθους 128x128 και για p_{max} 100, 200 και 250. Η εικόνα έχει εγγραφεί εντός του μοναδιαίου κύκλου.



Εικόνα 5.8 Ανακατασκευή της εικόνας lena πάνω στα Cartesian pixels, μεγέθους 128x128 και για p_{max} 50, 100 και 150. Ο μοναδιαίος κύκλος έχει εγγραφεί εντός της εικόνας



Εικόνα 5.9 Η εικόνας Barbara, μεγέθους 512x512



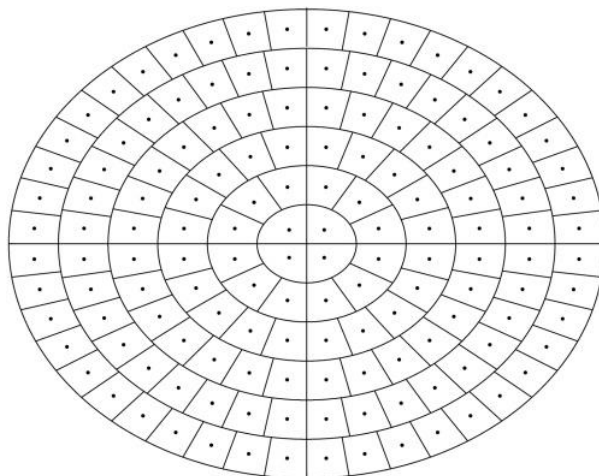
Εικόνα 5.10 Ανακατασκευή της εικόνας barbara πάνω στα Cartesian pixels, μεγέθους 512x512 και για r_{max} 100, 200 και 250. Η εικόνα έχει εγγραφεί εντός του μοναδιαίου κύκλου



Εικόνα 5.11 Ανακατασκευή της εικόνας barbara πάνω στα Cartesian pixels, μεγέθους 512×512 και για p_{max} 50, 100 και 150. Ο μοναδιαίος κύκλος έχει εγγραφεί εντός της εικόνας

6 Ευθεία Μέθοδος υπολογισμού των ποσοτήτων Zernike που χρησιμοποιεί Polar Pixels

Η μέθοδος που χρησιμοποιεί pixels σε σχήμα κυκλικού τομέα, όπως φαίνεται στην Εικόνα 6.1, αναπτύχθηκε από τους Xin Y, Pawlak M., Liao S., η δουλειά των οποίων παρουσιάζεται στα [10] και [11].



Εικόνα 6.1 Η εικόνα με pixels σε σχήμα κυκλικού τομέα

Για αυτά τα pixels, όπως αναφέρονται και στο [24], θα χρησιμοποιηθούν είτε ο αρχικός αγγλικός όρος polar pixels είτε οι ελληνικοί όροι “pixels κυκλικού τομέα” ή / και “πολικά pixels”. Για τα polar pixels ο τύπος του υπολογισμού των συντελεστών Zernike είναι ο κάτωθι:

$$Z_p^q = \frac{p+1}{\pi} \sum_{k=1}^U \sum_{l=1}^{(2k-1)V} \hat{f}(\rho_k, \theta_l) \iint_{\Omega_{k,l}} R_p^q(\rho) e^{-jq\theta} \rho \, d\rho d\theta \quad (6.1)$$

Όπου,

$$\iint_{\Omega_{k,l}} R_p^q(\rho) e^{-jq\theta} \rho \, d\rho d\theta = \int_{\rho_{k-1}}^{\rho_k} R_p^q(r) r \, dr \int_{\theta_{in}}^{\theta_{out}} e^{-jq\theta} \, d\theta \quad (6.2)$$

$$\hat{f}(\rho_k, \theta_l) = \sum_{i=k-1}^{k+2} \sum_{j=l-1}^{l+2} f(x_i, y_j) \times h\left(\frac{\rho_k \cos(\theta_l) - x_i}{\Delta}\right) h\left(\frac{\rho_k \sin(\theta_l) - y_j}{\Delta}\right) \quad (6.3)$$

Οι μεταβλητές που εμφανίζονται στον ανωτέρω τύπο

(6.3) έχουν ως εξής:

- $x_i = \frac{2*(i+\frac{1}{2})-N-1}{N}$ και $y_j = \frac{2*(j+\frac{1}{2})-N-1}{N}$, οι συντεταγμένες του (i, j) pixel
- Το $\Delta = \frac{2}{N}$
- $k = \left\lfloor \frac{\rho_k \cos(\theta_l)}{\Delta} \right\rfloor + \frac{N}{2}$
- $l = \left\lfloor \frac{\rho_k \sin(\theta_l)}{\Delta} \right\rfloor + \frac{N}{2}$
- h είναι η bi-cubic παρεμβολή, με τύπο

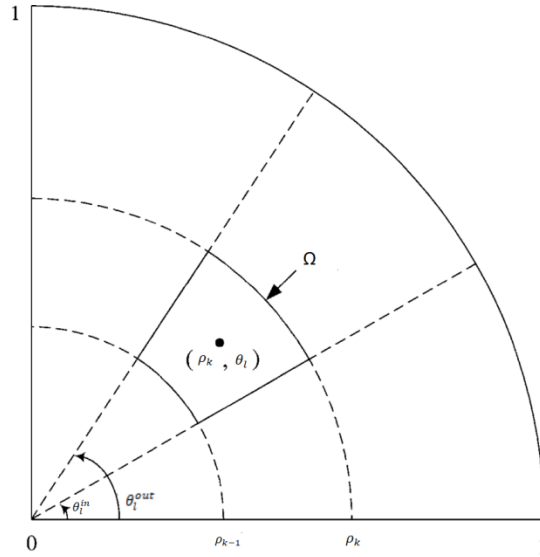
$$h(x) = \begin{cases} \frac{3}{2}|x|^3 - \frac{5}{2}|x|^2 + 1 & , \quad |x| \leq 1 \\ -\frac{1}{2}|x|^3 + \frac{5}{2}|x|^2 - 4|x| + 2, & 1 < |x| \leq 1 \\ 0 & , \quad \text{αλλιώς} \end{cases}$$

Η εξίσωση

(6.2) αναλύεται περαιτέρω ως εξής :

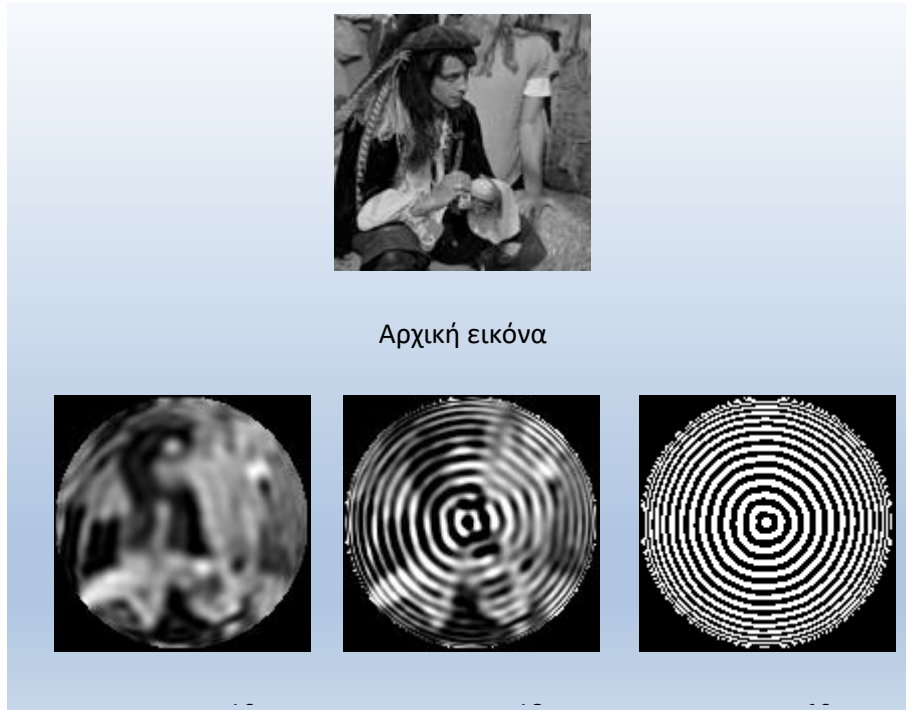
$$\int_{\rho_{k-1}}^{\rho_k} R_p^q(r) r dr = \sum_{s=0}^{(p-|q|)/2} \frac{(-1)^s (n-s)! [(\rho_k)^{p-2s+2} - (\rho_{k-1})^{p-2s+2}]}{(p-2s+2)s! \left(\frac{p+|q|}{2} - s\right)! \left(\frac{p-|q|}{2} - s\right)!} \quad (6.4)$$

$$\int_{\theta_l^{in}}^{\theta_l^{out}} e^{-jq\theta} d\theta = \begin{cases} \frac{j}{q} (e^{-jq\theta_l^{out}} - e^{-jq\theta_l^{in}}) & q \neq 0 \\ \theta_l^{out} - \theta_l^{in} & q = 0 \end{cases} \quad (6.5)$$



Εικόνα 6.2 Ρixel σχήματος κυκλικού τομέα

Το τυχόν polar pixel είναι ένας κυκλικός τομέας με εσωτερική ακτίνα ρ_{k-1} , εξωτερική ακτίνα ρ_k , ο οποίος εκκινεί από τη πολική γωνία $\theta_l^{in} = \theta_l - \frac{\Delta\theta}{2}$ και περατούται στη πολική γωνία $\theta_l^{out} = \theta_l + \frac{\Delta\theta}{2}$ (Εικόνα 6.2), όπου θ_l η πολική γωνία του άξονα συμμετρίας του πολικού εικονοστοιχείου υπ' αριθμό l , $l = 1, 2, \dots, (2k - 1)V$. Η γωνία $\Delta\theta$ είναι το βήμα με το οποίο παράγεται ο κάθε κυκλικός τομέας. Επιπλέον, $\hat{f}(\rho_k, \theta_l)$ είναι η εκτίμηση της τιμής της εικόνας επί του polar pixel, η οποία επιτυγχάνεται μέσω της bi-cubic παρεμβολής, όπως περιγράφεται στα [10] και [11]. Επιπροσθέτως, το U ορίζεται ως το πλήθος των ομόκεντρων κυκλικών δακτυλίων στους οποίους διαιρείται η εικόνα, ενώ ο k -οστός ομόκεντρος κυκλικός δακτύλιος περιέχει $(2k - 1)V$ pixels, όπου V κατάλληλα επιλεγμένος ακέραιος αριθμός που διαιρείται με το τέσσερα και $k = 1, 2, \dots, U$. Σύμφωνα με τα [10] και [11] πρέπει να γίνει κατάλληλη επιλογή των V και U , διότι ο μοναδιαίος κύκλος χωρίζεται σε VU^2 pixels σχήματος κυκλικού τομέα. Συνεπώς, όσο μικρότερο είναι το πλήθος VU^2 των polar pixels στο οποίο υποδιαιρείται η τυχούσα εικόνα, τόσο φτωχότερη / ανεπαρκή πληροφορία εμπεριέχουν οι συντελεστές Zernike. Αντιστρόφως, όσο το πλήθος VU^2 των polar pixels αυξάνει, τόσο καλύτερα οι συντελεστές Zernike μπορούν να περιγράψουν την εικόνα, εις βάρος όμως της ταχύτητας υπολογισμού αυτών. Σύμφωνα με τα παραπάνω προτείνεται το V να ισούται με τον αριθμό τέσσερα (4), ενώ το U να είναι μεταξύ $\frac{N}{2}$ και N , όπου N η διάσταση της εικόνας.



Εικόνα 6.3 Προσπάθεια ανακατασκευής της εικόνας map, μεγέθους 128x128 με χρήση της υπάρχουσας μεθόδου υπολογισμού των polar Zernike moments, με double precision ακρίβεια.

Όμως ο αλγόριθμος που παρουσιάζεται στα παραπάνω papers δεν είναι δεόντως γρήγορος και αποδοτικός, ενώ έχει πρόβλημα σε μεγάλες τάξεις p_{max} πολυωνύμων Zernike, αφού ο υπολογισμός τους γίνεται μέσω ενός ευθέως τύπου που εμπεριέχει παραγοντικά. Όπως αναφέρθηκε και στις παραπάνω μεθόδους που υπολογίζουν τα πολυώνυμα Zernike με τον ευθύ τρόπο, ο υπολογισμός των παραγοντικών περιορίζεται από τις υπολογιστικές δυνατότητες της εκάστοτε μηχανής που χρησιμοποιείται για την εκτέλεση του αλγορίθμου. Στην εικόνα 6.3 φαίνεται η αδυναμία ανακατασκευής της εικόνας, μεγέθους 128x128 pixels, διότι η IEEE 754 double ακρίβεια δεν επιτρέπει τον υπολογισμό υψηλότερης τάξης Zernike moments. Επίσης απαιτούνται πολύπλοκες πράξεις υπολογισμού του κάθε συντελεστή ξεχωριστά. Συνεπώς η πολυπλοκότητα του αλγορίθμου είναι μεγάλη και ο χρόνος που απαιτείται για τον αποτελεσματικό υπολογισμό όλων των Zernike moments είναι τεράστιος. Παρακάτω παρουσιάζεται αναλυτικά μια νέα μεθοδολογία υπολογισμού των polar Zernike moments, με αναδρομικό τρόπο, με σκοπό την άρση όλων των παραπάνω περιορισμών.

7 Ένας νέος, εύρωστος και αποδοτικός αναδρομικός αλγόριθμος για τον υπολογισμό των συντελεστών Zernike, κάνοντας χρήση των πολικών (polar) pixels.

Στα [10] και [11] παρουσιάζεται μία πρωτότυπη μέθοδος, η οποία μειώνει σημαντικά το γεωμετρικό λάθος και το σφάλμα ολοκλήρωσης που υπάρχει στις διάφορες συμβατικές μεθόδους υπολογισμού των συντελεστών Zernike και της συνεπακόλουθης ανακατασκευής μίας εικόνας. Σε αυτή τη μέθοδο, τα pixels δε θεωρούνται τετράγωνα αλλά κυκλικοί τομείς.

Επαναλαμβάνουμε ότι ο ευθύς (direct) υπολογισμός των Zernike moments πάνω στα polar pixels γίνεται από τον κάτωθι τύπο

$$Z_p^q = \frac{p+1}{\pi} \sum_{k=1}^U \sum_{l=1}^{(2k-1)V} \hat{f}(\rho_k, \theta_l) \int_{\rho_{k-1}}^{\rho_k} R_p^q(r) r dr \int_{\theta_l^{in}}^{\theta_l^{out}} e^{-jq\theta} d\theta \quad (7.1)$$

με

$$\int_{\rho_{k-1}}^{\rho_k} R_p^q(r) r dr = \sum_{s=0}^{(p-|q|)/2} \frac{(-1)^s (n-s)! [(\rho_k)^{p-2s+2} - (\rho_{k-1})^{p-2s+2}]}{(p-2s+2)s! \left(\frac{p+|q|}{2} - s\right)! \left(\frac{p-|q|}{2} - s\right)!} \quad (7.2)$$

Όπου το τυχόν polar pixel είναι ένας κυκλικός τομέας με εσωτερική ακτίνα ρ_{k-1} , εξωτερική ακτίνα ρ_k , ο οποίος εκκινεί από τη πολική γωνία $\theta_l^{in} = \theta_l - \frac{\Delta\theta}{2}$ και περατούται στη πολική γωνία $\theta_l^{out} = \theta_l + \frac{\Delta\theta}{2}$ (Εικόνα 6.2). Η γωνία $\Delta\theta$ είναι το βήμα με το οποίο παράγεται ο κάθε κυκλικός τομέας. Επιπλέον, $\hat{f}(\rho_k, \theta_l)$ είναι η εκτίμηση της τιμής της εικόνας επί του polar pixel, η οποία επιτυγχάνεται μέσω της bi-cubic παρεμβολής, όπως περιγράφεται στα [10] και [11].

Όμως ο αλγόριθμος που παρουσιάζεται στα παραπάνω papers δεν είναι δεόντως γρήγορος και αποδοτικός, ενώ έχει πρόβλημα σε μεγάλες τάξεις p_{max} πολυωνύμων Zernike, αφού ο υπολογισμός τους γίνεται μέσω ενός ευθέως τύπου που εμπεριέχει παραγοντικά και επομένως προκαλεί νομοτελειακά υπερχειλίση (overflow). Στο παρόν κεφάλαιο θα γίνει η παρουσίαση μιας νέας αναδρομικής σχέσης όπου υπολογίζει το παρακάτω ολοκλήρωμα, που είναι κομμάτι του υπολογισμού που εκτελείται στο δεξί μέλος της σχέσεως (7.1):

$$A_p^q(\rho) = \int_{\rho_{k-1}}^{\rho_k} R_p^q(r) r dr \quad (7.3)$$

Ο πρωτότυπος αναδρομικός αλγόριθμος που προτείνεται στην παρούσα εργασία παρήχθη με έναυσμα τη φιλοσοφία του αλγορίθμου modified Kintner που αναφέρονται στο κεφάλαιο 3.2.1.2.4. Οι νέες αυτές αναδρομικές σχέσεις υπολογίζουν το ακτινικό μέρος των συντελεστών Zernike γρήγορα και ιδιαίτερως αποδοτικά, παρακάμπτοντας πρακτικά πλήρως, την υπερχειλίση. Όσον αφορά το κομμάτι του αλγορίθμου που σχετίζεται με τις γωνίες, συγκεκριμένες συμμετρίες παρουσιάζονται που ελαττώνουν τη ολική πολυπλοκότητα των αντίστοιχων υπολογισμών τουλάχιστον κατά ένα παράγοντα της τάξης του τέσσερα (4). Επιπροσθέτως, οι νέοι αυτοί τύποι έχουν τις εξής εξαιρετικές ιδιότητες:

- Διατηρούν πλήρως τα πλεονεκτήματα του υπολογισμού των Zernike moments, επί των πολικών pixels [10], [11] και επομένως δεν παράγουν πρακτικά γεωμετρικό λάθος και λάθος ολοκλήρωσης, όπως θα φανεί σαφέστατα σε επόμενα παραδείγματα.
- Δεν προκαλεί καθόλου υπερχειλίση και γενικά
- Δεν εμφανίζουν μεγάλο σφάλμα πεπερασμένης ακρίβειας, όπως θα φανεί στη συνέχεια.

7.1 Ο υπολογισμός των ακτινικών πολυωνύμων Zernike σύμφωνα με τη νέα αναδρομική μέθοδο.

Τονίζεται ότι ως A_p^q συμβολίζεται το ακτινικό πολυώνυμο των συντελεστών Zernike που υπολογίζεται πάνω σε ένα πολικό pixel, σαν αυτό που παρουσιάστηκε στην εισαγωγή του παρόντος κεφαλαίου 7 και Εικόνα 6.2. Σκοπός του παρόντος κεφαλαίου είναι να παρουσιαστούν οι αναδρομικές σχέσεις που υπολογίζουν το $A_p^q(\rho)$ εύρωστα και αποδοτικά. Ακολουθεί η περιγραφή των βημάτων του προτεινόμενου αλγορίθμου:

ΒΗΜΑ 1

Εφαρμόζεται η ιδιότητα :

$$\int_{\rho_{k-1}}^{\rho_k} \{R_p^q(r) + R_p^{q+2}(r)\} dr = \frac{1}{p+1} [R_{p+1}^{q+1}(r) - R_{p-1}^{q+1}(r)]_{r=\rho_{k-1}}^{r=\rho_k} \quad (7.4)$$

Η οποία έχει αποδειχθεί από τον Prata και Rusch στο [25].

ΒΗΜΑ 2

Εκτελείται ολοκλήρωση κατά μέρη στην εξίσωση που παρουσιάστηκε στο βήμα ένα, ώστε να εκφραστεί το άθροισμα $A_p^q(\rho) + A_p^{q+2}(\rho)$ στο τυχόν k ομόκεντρο κυκλικό δακτύλιο (ενότητα 6, κεφάλαιο Β) ως ακολούθως:

$$A_p^q(\rho) + A_p^{q+2}(\rho) = \frac{1}{p+1} \left(\left[r(R_{p+1}^{q+1}(r) - R_{p-1}^{q+1}(r)) \right]_{r=\rho_{k-1}}^{r=\rho_k} - \int_{\rho_{k-1}}^{\rho_k} (R_{p+1}^{q+1}(r) - R_{p-1}^{q+1}(r)) \right) \quad (7.5)$$

ΒΗΜΑ 3

Με στόχο τη παράκαμψη του υπολογισμού του ολοκληρώματος του τύπου (7.5) του προηγούμενου βήματος 2, διατυπώνουμε τη σχέση αυτή ώστε ο δεύτερος δείκτης του A να είναι τη πρώτη φορά q και τη δεύτερη φορά $q = q - 2$, ούτως ώστε τελικά να λάβουμε τη σχέση:

$$2A_p^q(\rho) + A_p^{q+2}(\rho) + A_p^{q-2}(\rho) = \frac{1}{p+1} \left(\left[\frac{1}{p} [R_p^q(r) - R_{p-2}^q(r)]_{r=\rho_{k-1}}^{r=\rho_k} - \frac{1}{p+2} [R_{p+2}^q(r) - R_p^q(r)]_{r=\rho_{k-1}}^{r=\rho_k} + r (R_{p+1}^{q+1}(r) + R_{p+1}^{q-1}(r) - R_{p-1}^{q+1}(r) - R_{p-1}^{q-1}(r)) \right]_{r=\rho_{k-1}}^{r=\rho_k} \right). \quad (7.6)$$

Το δεξί μέλος στη ανωτέρω σχέσης (7.6) είναι μία συνάρτηση των $R_p^q(\rho)$.

ΒΗΜΑ 4

Εφαρμόζοντας τις ήδη γνωστές αναδρομικές σχέσεις για τα $R_p^q(\rho)$ υπολογίζεται η αντίστοιχη αναδρομική σχέση για τα $A_p^q(\rho)$

ΒΗΜΑ 5

Υπολογίζονται όλα τα $A_q^q(\rho)$ και $A_q^{q-2}(\rho)$ για όλα τα q τέτοια ώστε $q \leq p_{max}$. Συνδυάζοντας αυτές τις τιμές με τις αναδρομικές σχέσεις του βήματος 4, υπολογίζεται το $A_p^q(\rho)$ που είναι το ζητούμενο.

ΑΠΟΔΕΙΞΗ

ΑΝΑΛΥΣΗ ΤΟΥ ΒΗΜΑΤΟΣ 1

Διαφορίζοντας τη σχέση (7.4) του βήματος 1 λαμβάνουμε:

$$R_p^q(\rho) + R_p^{q+2}(\rho) = \frac{1}{p+1} \frac{d[R_{p+1}^{q+1}(\rho) - R_{p-1}^{q+1}(\rho)]}{d\rho}. \quad (7.7)$$

Πράγματι, ολοκληρώνοντας τη παραπάνω σχέση για μία ακτίνα που αποτελεί τη πλευρά ενός πολικού (polar) pixel, από $\rho = \rho_{k-1}$ έως ρ_k , εξάγεται ο τύπος του βήματος 1.

ΑΝΑΛΥΣΗ ΤΟΥ ΒΗΜΑΤΟΣ 2

Ο ορισμός του ακτινικού πολυωνύμου Zernike, που κάνει χρήση των polar pixels, είναι:

$$A_p^q(\rho) = \int_{\rho_{k-1}}^{\rho_k} R_p^q(r) r dr$$

Ξαναγράφοντας της ίδια σχέση για $q = q + 2$ και αθροίζοντας κατά μέλη δίνεται η παρακάτω σχέση

$$A_p^q(\rho) + A_p^{q+2}(\rho) = \int_{\rho_{k-1}}^{\rho_k} \{R_p^q(r) + R_p^{q+2}(r)\} r dr$$

Χρησιμοποιώντας την διαφορική σχέση της ανάλυσης του βήματος 1 (7.7)

$$A_p^q(\rho) + A_p^{q+2}(\rho) = \int_{\rho_{k-1}}^{\rho_k} \left\{ \frac{1}{p+1} \frac{d[R_{p+1}^{q+1}(r) - R_{p-1}^{q+1}(r)]}{dr} \right\} r dr$$

Ολοκληρώνοντας στο δεξί μέρος κατά μέρη λαμβάνουμε:

$$A_p^q(\rho) + A_p^{q+2} = \frac{1}{p+1} \left([r(R_{p+1}^{q+1}(r) - R_{p-1}^{q+1}(r))]_{r=\rho_{k-1}}^{r=\rho_k} - \int_{\rho_{k-1}}^{\rho_k} (R_{p+1}^{q+1}(r) - R_{p-1}^{q+1}(r)) dr \right) \quad (7.8)$$

ΑΝΑΛΥΣΗ ΤΟΥ ΒΗΜΑΤΟΣ 3

Η σχέση (7.8) για τους δείκτες (p, q) και $(p, q - 2)$ γίνεται αντίστοιχα:

$$\begin{aligned}
 & \bullet A_p^q(\rho) + A_p^{q+2}(\rho) = \frac{1}{p+1} \left([r(R_{p+1}^{q+1}(r) - R_{p-1}^{q+1}(r))]_{r=\rho_{k-1}}^{r=\rho_k} - \right. \\
 & \quad \left. \int_{\rho_{k-1}}^{\rho_k} (R_{p+1}^{q+1}(r) - R_{p-1}^{q+1}(r)) dr \right) \\
 & \bullet A_p^{q-2}(\rho) + A_p^q(\rho) = \frac{1}{p+1} \left([r(R_{p+1}^{q-1}(r) - R_{p-1}^{q-1}(r))]_{r=\rho_{k-1}}^{r=\rho_k} - \right. \\
 & \quad \left. \int_{\rho_{k-1}}^{\rho_k} (R_{p+1}^{q-1}(r) - R_{p-1}^{q-1}(r)) dr \right)
 \end{aligned} \tag{7.9}$$

Αθροίζοντας κατά μέλη τις δυο παραπάνω ισότητες συνάγεται ότι:

$$\begin{aligned}
 & A_p^{q+2} + A_p^{q-2} + 2A_p^q = \\
 & \frac{1}{p+1} \left([r(R_{p+1}^{q+1}(r) + R_{p+1}^{q-1}(r) - R_{p-1}^{q+1}(r) - R_{p-1}^{q-1}(r))]_{r=\rho_{k-1}}^{r=\rho_k} - \int (R_{p+1}^{q+1}(r) + R_{p+1}^{q-1}(r)) dr \right. \\
 & \quad \left. + \int (R_{p-1}^{q+1}(r) + R_{p-1}^{q-1}(r)) dr \right)
 \end{aligned}$$

Ακολουθως, με χρήση της σχέση (7.7) του βήματος 1, ο τελευταίος τύπος γίνεται

$$\begin{aligned}
 & A_p^{q+2} + A_p^{q-2} + 2A_p^q = \\
 & \frac{1}{p+1} \left([r(R_{p+1}^{q+1}(r) + R_{p+1}^{q-1}(r) - R_{p-1}^{q+1}(r) - R_{p-1}^{q-1}(r))]_{r=\rho_{k-1}}^{r=\rho_k} \right. \\
 & \quad \left. - \frac{1}{p+2} [R_{p+2}^q(r) - R_p^q(r)]_{r=\rho_{k-1}}^{r=\rho_k} + \frac{1}{p} [R_p^q(r) - R_{p-2}^q(r)]_{r=\rho_{k-1}}^{r=\rho_k} \right)
 \end{aligned} \tag{7.10}$$

ΑΝΑΛΥΣΗ ΤΟΥ ΒΗΜΑΤΟΣ 4

Στην αρχή του βήματος αυτού υπολογίζονται τα $R_p^q(\rho)$ κάνοντας χρήση των ήδη γνωστών αναδρομικών αλγορίθμων που παρουσιάστηκαν στο κεφάλαιο 3.2.1.2. Έπειτα υπολογίζονται τα $A_q^q(\rho)$ για $0 \leq q \leq p_{max}$ και $A_q^{q-2}(\rho)$ για $2 \leq q \leq p_{max}$. Επίσης υπολογίζεται το $A_q^{q-4}(\rho)$ για $4 \leq q \leq p_{max}$. Εν συνεχεία,

υπολογίζονται τα A_q^{q-6}, A_q^{q-8} , κτλ, μέχρι να υπολογιστούν με αναδρομικό τρόπο όλοι οι συντελεστές Zernike που αφορούν τα pixels σχήματος κυκλικού τομέα.

ΑΝΑΛΥΣΗ ΤΟΥ ΒΗΜΑΤΟΣ 5

Τα αποτελέσματα των προηγούμενων βημάτων συγκεντρώνονται στην ακόλουθη πρόταση:

ΠΡΟΤΑΣΗ Β.7.1

Τα ακτινικά πολυώνυμα $A_p^q(\rho)$ (δηλαδή το ακτινικό μέρος των συντελεστών Zernike), όταν υπολογίζονται πάνω σε πολικά pixels, ικανοποιούν τις παρακάτω σχέσεις:

$$\bullet A_p^q(\rho) = \frac{[r^{p+2}]_{r=\rho_{k-1}}^{r=\rho_k}}{p+2}, \text{ όταν } p = q \quad (7.11)$$

$$\bullet A_p^q(\rho) = \left(1 - \frac{2}{p+2}\right) [r^{p+2}]_{r=\rho_{k-1}}^{r=\rho_k} - \left(1 - \frac{1}{p}\right) [r^p]_{r=\rho_{k-1}}^{r=\rho_k}, \text{ όταν ισχύει } q = p - 2 \quad (7.12)$$

$$\bullet A_p^q(\rho) = \frac{1}{p+1} \left(\frac{1}{p} [R_p^{q+2}(r) - R_{p-2}^{q+2}(r)]_{r=\rho_{k-1}}^{r=\rho_k} - \frac{1}{p+2} [R_{p+2}^{q+2}(r) - R_p^{q+2}(r)]_{r=\rho_{k-1}}^{r=\rho_k} + [r (R_{p+1}^{q+3}(r) + R_{p+1}^{q+1}(r) - R_{p-1}^{q+3}(r) - R_{p-1}^{q+1}(r))]_{r=\rho_{k-1}}^{r=\rho_k} \right) - (2A_p^{q+2}(\rho) + A_p^{q+4}(\rho)), \text{ για } q = p - 4, q = p - 6 \text{ κτλ} \quad (7.13)$$

Επομένως, εφαρμόζοντας τις σχέσεις (7.11), (7.12) και (7.13), είναι δυνατόν να υπολογιστούν αποδοτικά όλα τα $A_p^q(\rho)$. □

7.2 Εκμετάλλευση των συμμετριών που εμφανίζονται στο γωνιακό σκέλος των $Z_p^q(\rho_k, \theta_l)$, με σκοπό τη μείωση της υπολογιστικής πολυπλοκότητας.

Σκοπός αυτού του κεφαλαίου είναι η εύρεση συμμετριών στους συντελεστές Zernike που χρησιμοποιούν polar pixels, ώστε να μειωθεί η υπολογιστική πολυπλοκότητα και ο χρόνος εκτέλεσης του αλγορίθμου. Συγκεκριμένα, στις δημοσιεύσεις των [10] και [11] αποδεικνύεται ότι

$$\int_{\theta_l^{in}}^{\theta_l^{out}} e^{-jq\theta} d\theta = \begin{cases} \frac{j}{q} (e^{-jq\theta_l^{out}} - e^{-jq\theta_l^{in}}) & q \neq 0 \\ \theta_l^{out} - \theta_l^{in} & q = 0 \end{cases} \quad (7.14)$$

Θα εξεταστεί η περίπτωση όπου το q είναι διάφορο του μηδενός, αφού για $q = 0$ το γωνιακό ολοκλήρωμα ισούται με $\theta_l^{out} - \theta_l^{in}$. Συνεπώς, αν $q \neq 0$ και $\Delta\theta \triangleq \theta_l^{out} - \theta_l^{in}$ τότε η προαναφερθείσα σχέση γίνεται

$$\frac{j}{q} \left(e^{-jq\theta_l^{out}} - e^{-jq\theta_l^{in}} \right) = \frac{j}{q} e^{-jq\theta_l^{in}} \left(e^{-jq(\Delta\theta)} - 1 \right) \quad (7.15)$$

Το $\Delta\theta$ είναι σταθερό και αμετάβλητο για όλα τα πολικά pixels ενός κυκλικού τομέα, επομένως και ο παράγοντας $e^{-jq(\Delta\theta)}$ αρκεί να υπολογιστεί μία φορά για κάθε q .

Στη συνέχεια, όσον αφορά το θ_l^{in} , παρατηρείται ότι $\theta_l^{in} = \mu \cdot \Delta\theta$, όπου $\mu = 0, 1, \dots, (2k - 2)V$. Επομένως, αφού ισχύει ότι $\Delta\theta = \frac{2\pi}{(2k-1)V}$, το V θα πρέπει να είναι ακέραιο πολλαπλάσιο του αριθμού τέσσερα, ώστε να υπάρξουν συμμετρίες. Πράγματι, ο αύξων αριθμός μ ενός pixel στον k -οστό πλήρη κυκλικό δακτύλιο ισούται με ([10] και [11]):

$$\mu = \frac{(2k-1)V}{4}x + m \text{ pixels}, \quad x = 0, 1, \dots, 3, \quad m = 1, 2, \dots, 2k - 1 \quad (7.16)$$

Το x εκφράζει τον αύξοντα αριθμό που χαρακτηρίζει το τεταρτημόριο στο οποίο βρίσκεται το εκάστοτε l -οστό pixel σχήματος κυκλικού τομέα. Δηλαδή, για το πρώτο τεταρτημόριο το x είναι 0, για το 2^ο το x είναι 1 κτλ.

Αξιοποιώντας το τύπο του Euler

$$e^{jx} = \cos x + j \sin x \quad (7.17)$$

Το $e^{-jq\theta_l^{in}}$ γίνεται

$$e^{-jq\theta_l^{in}} = \cos(q\theta_l^{in}) - j \sin(q\theta_l^{in})$$

Αναλύοντας τον παραπάνω τύπο

- Για το $\cos(q\theta_l^{in})$ προκύπτει ότι

$$\cos(q\theta_l^{in}) = \cos(q \cdot \mu \cdot \Delta\theta) = \cos\left(q \cdot \left(x \cdot \frac{(2k-1)V}{4} + m\right) \cdot \Delta\theta\right)$$

- Για το $\sin(q\theta_l^{in})$ ισχύει κατά αναλογία

$$\sin(q\theta_l^{in}) = \sin(q \cdot \mu \cdot \Delta\theta) = \sin\left(q \cdot \left(x \cdot \frac{(2k-1)V}{4} + m\right) \cdot \Delta\theta\right)$$

Για τη προτεινόμενη τιμή του V , $V = 4$, το q γράφεται ως $q = 4\tau + v$, $\tau \in \mathbb{N}$, $v \in \{0,1,2,3\}$. Επομένως οι παραπάνω σχέσεις γίνονται

$$\cos\left(q \left(x \frac{(2k-1)V}{4} + m\right) \Delta\theta\right) = \cos\left((4\tau + v)x \frac{\pi}{2} + qm\Delta\theta\right) \quad (7.18)$$

$$\sin\left(q \left(x \frac{(2k-1)V}{4} + m\right) \Delta\theta\right) = \sin\left((4\tau + v)x \frac{\pi}{2} + qm\Delta\theta\right) \quad (7.19)$$

Επομένως για $x = 1$ προκύπτει:

- Για $q = 4\tau \Rightarrow$

- $\cos\left((4\tau) \frac{\pi}{2} + qm\Delta\theta\right) = \cos\left(4\tau \frac{\pi}{2}\right) \cos(qm\Delta\theta) - \sin\left(4\tau \frac{\pi}{2}\right) \sin(qm\Delta\theta) = \cos(qm\Delta\theta)$

- $\sin\left((4\tau) \frac{\pi}{2} + qm\Delta\theta\right) = \sin\left(4\tau \frac{\pi}{2}\right) \cos(qm\Delta\theta) + \cos\left(4\tau \frac{\pi}{2}\right) \sin(qm\Delta\theta) = \sin(qm\Delta\theta)$

- Για $q = 4\tau + 1 \Rightarrow$

- $\cos\left((4\tau + 1) \frac{\pi}{2} + qm\Delta\theta\right) =$

$$\cos\left((4\tau + 1)\frac{\pi}{2}\right)\cos(qm\Delta\theta) - \sin\left((4\tau + 1)\frac{\pi}{2}\right)\sin(qm\Delta\theta) = -\sin(qm\Delta\theta)$$

$$\circ \sin\left((4\tau + 1)\frac{\pi}{2} + qm\Delta\theta\right) =$$

$$\sin\left((4\tau + 1)\frac{\pi}{2}\right)\cos(qm\Delta\theta) + \cos\left((4\tau + 1)\frac{\pi}{2}\right)\sin(qm\Delta\theta) = \cos(qm\Delta\theta)$$

• Για $q = 4\tau + 2 \Rightarrow$

$$\circ \cos\left((4\tau + 2)\frac{\pi}{2} + qm\Delta\theta\right) =$$

$$\cos\left((4\tau + 2)\frac{\pi}{2}\right)\cos(qm\Delta\theta) - \sin\left((4\tau + 2)\frac{\pi}{2}\right)\sin(qm\Delta\theta) = -\cos(qm\Delta\theta)$$

$$\circ \sin\left((4\tau + 2)\frac{\pi}{2} + qm\Delta\theta\right) =$$

$$\cos\left((4\tau + 2)\frac{\pi}{2}\right)\sin(qm\Delta\theta) + \sin\left((4\tau + 2)\frac{\pi}{2}\right)\cos(qm\Delta\theta) = -\sin(qm\Delta\theta)$$

• Για $q = 4\tau + 3 \Rightarrow$

$$\circ \cos\left((4\tau + 3)\frac{\pi}{2} + qm\Delta\theta\right) =$$

$$\cos\left((4\tau + 3)\frac{\pi}{2}\right)\cos(qm\Delta\theta) - \sin\left((4\tau + 3)\frac{\pi}{2}\right)\sin(qm\Delta\theta) = \sin(qm\Delta\theta)$$

$$\circ \sin\left((4\tau + 3)\frac{\pi}{2} + qm\Delta\theta\right) =$$

$$\cos\left((4\tau + 3)\frac{\pi}{2}\right)\sin(qm\Delta\theta) + \sin\left((4\tau + 3)\frac{\pi}{2}\right)\cos(qm\Delta\theta) = -\cos(qm\Delta\theta)$$

Εν συνεχεία για $x = 2$ προκύπτει:

• Για $q = 4\tau \Rightarrow$

$$\circ \cos\left((4\tau)2\frac{\pi}{2} + qm\Delta\theta\right) = \cos\left(4\tau 2\frac{\pi}{2}\right)\cos(qm\Delta\theta) - \sin\left(4\tau 2\frac{\pi}{2}\right)\sin(qm\Delta\theta) = \cos(qm\Delta\theta)$$

- $\sin\left((4\tau)2\frac{\pi}{2} + qm\Delta\theta\right) = \sin\left(4\tau 2\frac{\pi}{2}\right) \cos(qm\Delta\theta) + \cos\left(4\tau 2\frac{\pi}{2}\right) \sin(qm\Delta\theta) = \sin(qm\Delta\theta)$

- Για $q = 4\tau + 1 \Rightarrow$
 - $\cos\left((4\tau + 1)2\frac{\pi}{2} + qm\Delta\theta\right) = \cos\left((4\tau + 1)2\frac{\pi}{2}\right) \cos(qm\Delta\theta) - \sin\left((4\tau + 1)2\frac{\pi}{2}\right) \sin(qm\Delta\theta) = -\cos(qm\Delta\theta)$
 - $\sin\left((4\tau + 1)2\frac{\pi}{2} + qm\Delta\theta\right) = \sin\left((4\tau + 1)2\frac{\pi}{2}\right) \cos(qm\Delta\theta) + \cos\left((4\tau + 1)2\frac{\pi}{2}\right) \sin(qm\Delta\theta) = -\sin(qm\Delta\theta)$

- Για $q = 4\tau + 2 \Rightarrow$
 - $\cos\left((4\tau + 2)2\frac{\pi}{2} + qm\Delta\theta\right) = \cos\left((4\tau + 2)2\frac{\pi}{2}\right) \cos(qm\Delta\theta) - \sin\left((4\tau + 2)2\frac{\pi}{2}\right) \sin(qm\Delta\theta) = \cos(qm\Delta\theta)$
 - $\sin\left((4\tau + 2)2\frac{\pi}{2} + qm\Delta\theta\right) = \cos\left((4\tau + 2)2\frac{\pi}{2}\right) \sin(qm\Delta\theta) + \sin\left((4\tau + 2)2\frac{\pi}{2}\right) \cos(qm\Delta\theta) = \sin(qm\Delta\theta)$

- Για $q = 4\tau + 3 \Rightarrow$
 - $\cos\left((4\tau + 3)2\frac{\pi}{2} + qm\Delta\theta\right) = \cos\left((4\tau + 3)2\frac{\pi}{2}\right) \cos(qm\Delta\theta) - \sin\left((4\tau + 3)2\frac{\pi}{2}\right) \sin(qm\Delta\theta) = -\cos(qm\Delta\theta)$
 - $\sin\left((4\tau + 3)2\frac{\pi}{2} + qm\Delta\theta\right) = \cos\left((4\tau + 3)2\frac{\pi}{2}\right) \sin(qm\Delta\theta) + \sin\left((4\tau + 3)2\frac{\pi}{2}\right) \cos(qm\Delta\theta) = -\sin(qm\Delta\theta)$

Για $x = 3$ προκύπτει:

- Για $q = 4\tau \Rightarrow$
 - $\cos\left((4\tau)3\frac{\pi}{2} + qm\Delta\theta\right) = \cos\left(4\tau 3\frac{\pi}{2}\right)\cos(qm\Delta\theta) - \sin\left(4\tau 3\frac{\pi}{2}\right)\sin(qm\Delta\theta) = \cos(qm\Delta\theta)$
 - $\sin\left((4\tau)3\frac{\pi}{2} + qm\Delta\theta\right) = \sin\left(4\tau 3\frac{\pi}{2}\right)\cos(qm\Delta\theta) + \cos\left(4\tau 3\frac{\pi}{2}\right)\sin(qm\Delta\theta) = \sin(qm\Delta\theta)$

- Για $q = 4\tau + 1 \Rightarrow$
 - $\cos\left((4\tau + 1)3\frac{\pi}{2} + qm\Delta\theta\right) = \cos\left((4\tau + 1)3\frac{\pi}{2}\right)\cos(qm\Delta\theta) - \sin\left((4\tau + 1)3\frac{\pi}{2}\right)\sin(qm\Delta\theta) = \sin(qm\Delta\theta)$
 - $\sin\left((4\tau + 1)3\frac{\pi}{2} + qm\Delta\theta\right) = \sin\left((4\tau + 1)3\frac{\pi}{2}\right)\cos(qm\Delta\theta) + \cos\left((4\tau + 1)3\frac{\pi}{2}\right)\sin(qm\Delta\theta) = -\cos(qm\Delta\theta)$

- Για $q = 4\tau + 2 \Rightarrow$
 - $\cos\left((4\tau + 2)3\frac{\pi}{2} + qm\Delta\theta\right) = \cos\left((4\tau + 2)3\frac{\pi}{2}\right)\cos(qm\Delta\theta) - \sin\left((4\tau + 2)3\frac{\pi}{2}\right)\sin(qm\Delta\theta) = -\cos(qm\Delta\theta)$
 - $\sin\left((4\tau + 2)3\frac{\pi}{2} + qm\Delta\theta\right) = \sin\left((4\tau + 2)3\frac{\pi}{2}\right)\cos(qm\Delta\theta) + \cos\left((4\tau + 2)3\frac{\pi}{2}\right)\sin(qm\Delta\theta) = -\sin(qm\Delta\theta)$

- Για $q = 4\tau + 3 \Rightarrow$
 - $\cos\left((4\tau + 3)3\frac{\pi}{2} + qm\Delta\theta\right) = \cos\left((4\tau + 3)3\frac{\pi}{2}\right)\cos(qm\Delta\theta) - \sin\left((4\tau + 3)3\frac{\pi}{2}\right)\sin(qm\Delta\theta) = -\sin(qm\Delta\theta)$
 - $\sin\left((4\tau + 3)3\frac{\pi}{2} + qm\Delta\theta\right) = \sin\left((4\tau + 3)3\frac{\pi}{2}\right)\cos(qm\Delta\theta) + \cos\left((4\tau + 3)3\frac{\pi}{2}\right)\sin(qm\Delta\theta) = \cos(qm\Delta\theta)$

$$\cos\left((4\tau + 3)3\frac{\pi}{2}\right) \sin(qm\Delta\theta) + \sin\left((4\tau + 3)3\frac{\pi}{2}\right) \cos(qm\Delta\theta) = \cos(qm\Delta\theta)$$

Αξιοποιώντας τα παραπάνω αποτελέσματα, διατυπώνεται το κάτωθι λήμμα, που εκφράζει τις διάφορες συμμετρίες του ολοκληρώματος (7.14) στον υπολογισμό των συντελεστών Zernike Z_p^q .

Λήμμα Β.7.1 Στο γωνιακό σκέλος της σχέσης (7.1), εφαρμόζονται συγκεκριμένες συμμετρίες, οι οποίες και παρουσιάζονται στον πίνακα 7.1, για την προτεινόμενη τιμή $V = 4$ ([10] και [11]). Στη γενική περίπτωση που $V = 4V_1$, $V_1 \in N$ ένας αντίστοιχος πίνακας μπορεί να παραχθεί με ευθύγραμμο τρόπο. Συνεπώς, ο υπολογισμός του ολοκληρώματος $\int_{\theta_{in}}^{\theta_{out}} e^{-jq\theta} d\theta$ πρέπει και αρκεί να πραγματοποιηθεί μόνο για τη περίπτωση που το polar pixel βρίσκεται στο πρώτο τεταρτημόριο μέσω των σχέσεων (7.18) και (7.19). Με αυτόν τον τρόπο η υπολογιστική πολυπλοκότητα και ο χρόνος εκτέλεσης του αλγορίθμου, για τον υπολογισμό των συντελεστών Zernike, μειώνεται κατά ένα παράγοντα της τάξης του τέσσερα (4).

□

$q \setminus$ Τεταρτημόριο	1	2	3	4
4τ	$\cos(qm\Delta\theta)$	$\cos(qm\Delta\theta)$	$\cos(qm\Delta\theta)$	$\cos(qm\Delta\theta)$
$4\tau + 1$	$\cos(qm\Delta\theta)$	$-\sin(qm\Delta\theta)$	$-\cos(qm\Delta\theta)$	$\sin(qm\Delta\theta)$
$4\tau + 2$	$\cos(qm\Delta\theta)$	$-\cos(qm\Delta\theta)$	$\cos(qm\Delta\theta)$	$-\cos(qm\Delta\theta)$
$4\tau + 3$	$\cos(qm\Delta\theta)$	$\sin(qm\Delta\theta)$	$-\cos(qm\Delta\theta)$	$-\sin(qm\Delta\theta)$

$q \setminus$ Τεταρτημόριο	1	2	3	4
4τ	$\sin(qm\Delta\theta)$	$\sin(qm\Delta\theta)$	$\sin(qm\Delta\theta)$	$\sin(qm\Delta\theta)$
$4\tau + 1$	$\sin(qm\Delta\theta)$	$\cos(qm\Delta\theta)$	$-\sin(qm\Delta\theta)$	$-\cos(qm\Delta\theta)$
$4\tau + 2$	$\sin(qm\Delta\theta)$	$-\sin(qm\Delta\theta)$	$\sin(qm\Delta\theta)$	$-\sin(qm\Delta\theta)$
$4\tau + 3$	$\sin(qm\Delta\theta)$	$-\cos(qm\Delta\theta)$	$-\sin(qm\Delta\theta)$	$\cos(qm\Delta\theta)$

Πίνακας 7.1 Πίνακες που παρουσιάζουν τις τελικές συμμετρίες των πολικών pixels, για την προτεινόμενη τιμή $V = 4$. Ο i) παρουσιάζει τις συμμετρίες για το κομμάτι του συνημίτονου στον τύπο του Euler, ενώ ο ii) παρουσιάζει τις συμμετρίες που αφορούν το ημίτονο.

Μέγεθος εικόνας	p_{max}	Βαθμός ταχύτητας αρχικού αλγορίθμου [11] / προτεινόμενου αλγορίθμου
128 x 128	25	1.1
128 x 128	100	1.7
128 x 128	150	1.7
512 x 512	80	2.9
300 x 300	120	3.8

Πίνακας 7.2 Πίνακας που δείχνει πόσες φορές ταχύτερος είναι ο προτεινόμενος αλγόριθμος (αναδρομικός) από τον

Σημειώνεται ότι ο παραπάνω πίνακας περιέχει p_{max} μέχρι 150 και διαστάσεις μέχρι **512x512**, διότι για μεγαλύτερες διαστάσεις και τιμές του p_{max} , ο αρχικός αλγόριθμος καταστρέφεται λόγω υπερχειλίσης.

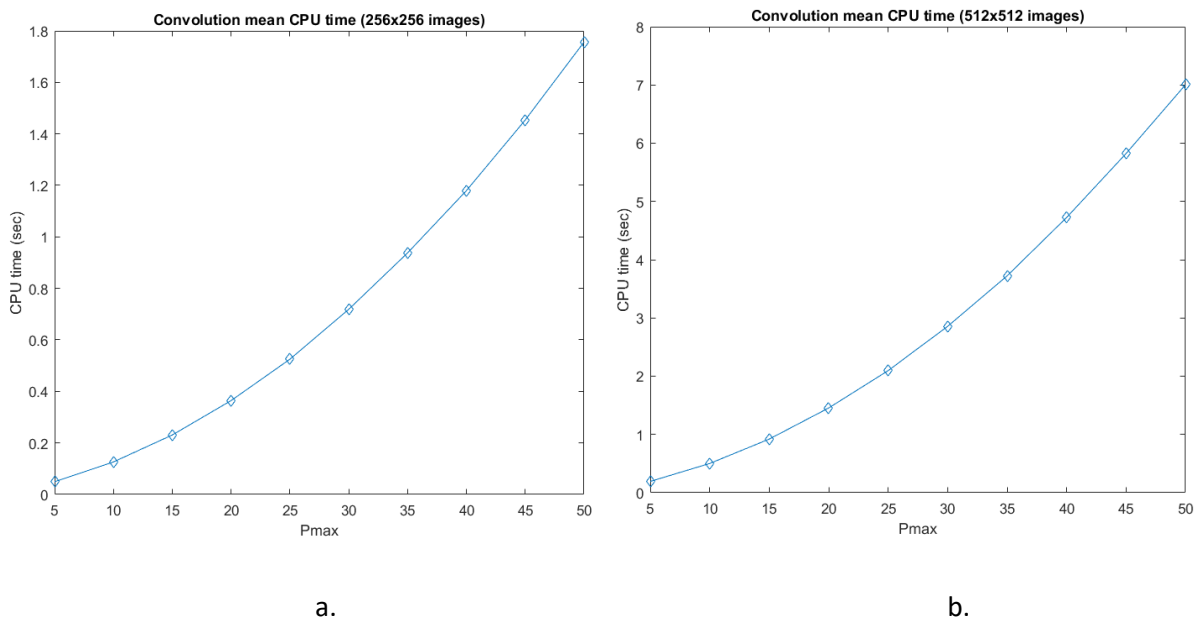
8 Ανάλυση των επιδόσεων του νέου αλγορίθμου.

8.1 Ο Προτεινόμενος αλγόριθμος είναι γρήγορος

Όπως αναμενόταν, ο αλγόριθμος που προτείνεται στη παρούσα εργασία, είναι αρκετά πιο γρήγορος από τον κλασικό αλγόριθμο που χρησιμοποιεί τα πολικά pixels [10]. Αυτό οφείλεται στο γεγονός ότι στον νέο αλγόριθμο που προτείνεται εδώ εφαρμόζεται μια αναδρομική μέθοδος υπολογισμού των ακτινικών πολωνύμων Zernike $A_p^q(\rho)$. Για να αποδειχθεί αυτό στη πράξη, πραγματοποιήθηκαν τα εξής πειράματα. Χρησιμοποιώντας το ίδιο υλικό και λογισμικό υπολογίστηκαν οι συντελεστές Zernike για εικόνες διάφορων διαστάσεων και για διάφορα p_{max} , τόσο με τη κλασική προσέγγιση, όσο και με την αναδρομική προσέγγιση που εισήχθη στην παρούσα εργασία. Τα σχετικά αποτελέσματα παρουσιάζονται στον πίνακα 7.2. Επισημαίνεται ότι ο προτεινόμενος αλγόριθμος μπορεί να υπολογίσει τα πολυώνυμα Zernike για εικόνες μεγάλου μεγέθους, κάνοντας χρήση της διπλής (double) ακρίβειας της υπολογιστικής μηχανής, για λόγους τους οποίους παρουσιάζονται στο κεφάλαιο 7. Για παράδειγμα, ο αναδρομικός αλγόριθμος μπορεί να υπολογίσει τους συντελεστές Zernike για εικόνες μεγέθους 512x512, 1200x1200, 2048x2048 κτλ, κάνοντας χρήση της IEEE 754 double ακρίβειας. Όμως, στις περιπτώσεις που ο κλασικός υπολογισμός των συντελεστών Zernike που χρησιμοποιούν pixels κυκλικού τομέα απαιτούσε να γίνουν οι υπολογισμοί με μεγαλύτερη ακρίβεια, η ίδια ακρίβεια εφαρμόζονταν ταυτόχρονα και στους υπολογισμούς του αναδρομικού αλγορίθμου. Σκοπός αυτής της προσέγγισης είναι να εκτελούνται και οι

δύο αλγόριθμοι κάτω από τις ίδιες συνθήκες. Σε αυτές τις περιπτώσεις απαίτησης αυξημένης ακρίβειας, έγινε χρήση της MPFR βιβλιοθήκης (multiple-precision floating-point computations with *correct rounding*).

Επιπροσθέτως, η αξιολόγηση της ταχύτητας εκτέλεσης της (7.1) σχέσης που υπολογίζει τους συντελεστές Zernike, βασίστηκε στη μέθοδο που προτείνεται στο [9]. Επιπλέον εφαρμόστηκαν και οι συμμετρίες που παρουσιάστηκαν στο κεφάλαιο 7.2, με στόχο να ελαχιστοποιηθεί ο χρόνος εκτέλεσης. Για τα συγκεκριμένα πειράματα, χρησιμοποιήθηκε ένας υπολογιστής με επεξεργαστή συχνότητας 3.2Ghz και μνήμης 4GB RAM. Επίσης μόνο ένα CPU thread χρησιμοποιήθηκε με σκοπό την ισότιμη σύγκριση με τις άλλες μεθόδους. Τα σχετικά αποτελέσματα παρουσιάζονται στις εικόνες 8.1a και 8.1b και αφορούν τον υπολογισμό των συντελεστών Zernike για εικόνες 256x256 και 512x512, αντιστοίχως.



Εικόνα 8.1 Μέσος χρόνος που απαιτείται για τον υπολογισμό των πολικών συντελεστών Zernike μέσω της (7.1). Η α) εικόνα αφορά των υπολογισμό των πολωνύμων για διάσταση εικόνας 256x256 ενώ η β) για εικόνα διαστάσεων 512x512.

Οι προαναφερθείσες εικόνες 8.1a και 8.1b αφορούν το χρόνο υπολογισμού των Zernike moments σε εικονοστοιχεία σχήματος κυκλικού τομέα. Επομένως, ο προτεινόμενος αλγόριθμος διατηρεί όλα τα πλεονεκτήματα του αντίστοιχου αλγορίθμου που παρουσιάζεται στα [10] και [11]. Συγκεκριμένα, ο νέος αλγόριθμος που εισάγεται εδώ παρουσιάζει μικρό γεωμετρικό και αριθμητικό λάθος και επιπλέον μικρό

λάθος πεπερασμένης ακρίβειας, ενώ είναι πολύ γρήγορος και αποδοτικός. Για παράδειγμα, για μια εικόνα 1600x1600 τα ακτινικά πολυώνυμα Zernike A_q^p , με $p_{max} = 50$, είχαν υπολογιστεί σε 0.06 sec, κάνοντας χρήση του νέου αλγορίθμου. Επαναλαμβάνεται ότι ο εισαχθείς αλγόριθμος μπορεί να εφαρμοστεί και σε εικόνες μεγάλου μεγέθους, με μεγάλα p_{max} , με πολύ καλές επιδόσεις, όπως αναλύεται παρακάτω.

8.2 Ο αλγόριθμος που προτείνεται στη παρούσα εργασία εμφανίζει αξιοσημείωτα μικρό γεωμετρικό λάθος, μικρό σφάλμα ολοκλήρωσης και μικρό σφάλμα πεπερασμένης ακρίβειας.

Πράγματι, ο νέος αυτός αλγόριθμος έχει όλα τα προτερήματα και πλεονεκτήματα της προσέγγισης που παρουσιάζεται στο [10], που οφείλονται στο γεγονός ότι γίνεται χρήση των πολικών pixels, αντί των τετραγωνικών. Ο αλγόριθμος αυτός εφαρμόστηκε σε μια σειρά από εικόνες που χρησιμοποιούνται ευρέως για δοκιμές στην αναγνώριση προτύπων και στην επεξεργασία εικόνας (Εικόνες 8.2a, 8.3a, 8.4a και 8.5). Τα αποτελέσματα της ανακατασκευής των εικόνων, κάνοντας χρήση της προτεινόμενης μεθόδου, παρουσιάζονται στις εικόνες 8.2c, 8.3b, 8.4b και 8.5 b. Τονίζεται, όπως φαίνεται παρακάτω, ότι ο προτεινόμενος αλγόριθμος συμπεριφέρεται πολύ ικανοποιητικά και για εικόνες διαστάσεων 300x300, 512x512 και 1200x1200, όπου ο αρχικός αλγόριθμος των [10] και [11] αποτυγχάνει.

Τα λάθη ανακατασκευής κάθε επεξεργασθείσας εικόνας παρουσιάζονται στους πίνακες 8.1, 8.2, 8.3 και 8.4. Σε αυτούς τους πίνακες έχουν συμπεριληφθεί δύο εκτιμήσεις του λάθους ανακατασκευής. Συγκεκριμένα:

- Το μέσο τετραγωνικό σφάλμα (The Mean Square Error - MSE), το οποίο υπολογίζεται από τη σχέση

$$MSE = \frac{\sum_{(x,y) \in K} (RCI(x,y) - OI(x,y))^2}{\sum_{(x,y) \in K} OI(x,y)^2} \quad (8.1)$$

Όπου με K συμβολίζεται ο εγγεγραμμένος κύκλος στον οποίο έχει εντεθεί η εικόνα. RCI είναι η ανακατασκευασθείσα εικόνα και OI είναι η αρχική εικόνα.

- Το σχετικό σφάλμα απόλυτης τιμής (The Relative Absolute Error - RABSE), το οποίο υπολογίζεται μέσω της σχέσης

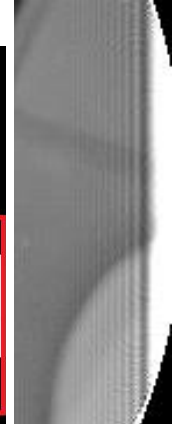
$$\text{RABSE} = \frac{\sum_{(x,y) \in K} |\text{RCI}(x,y) - \text{OI}(x,y)|}{\sum_{(x,y) \in K} |\text{OI}(x,y)|} \quad (8.2)$$

Από τη ποιότητα των ανακατασκευασμένων εικόνων και από τα αντίστοιχα λάθη ανακατασκευής, θεμελιώνεται ότι δεν υπάρχει σημαντικό γεωμετρικό λάθος ούτε ιδιαίτερο λάθος ολοκλήρωσης, στις εικόνες που χρησιμοποιήθηκαν ως είσοδοι στον αναδρομικό αλγόριθμο.



a. αρχική (1200 x 1200)

αρχική (512 x 512)



b. ανακατασκευασμένες εικόνες από τους συντελεστές Zernike που παρήχθησαν μέσω της modified Kintner's μεθόδου ($p_{max}=900$ (αριστερά), $p_{max}=450$ (δεξιά))



c. ανακατασκευασμένες εικόνες από τους συντελεστές Zernike που παρήχθησαν μέσω της προτεινόμενης μεθόδου ($p_{max}=900$ (αριστερά), $p_{max}=450$ (δεξιά))

Εικόνα 8.2 Σχήματα που καταδεικνύουν τη ποιότητα των ανακατασκευασμένων εικόνων, όπως αυτές παρήχθησαν με χρήση του αλγορίθμου της παρούσας εργασίας. Ο αναδρομικός αλγόριθμος που υπολογίζει τους συντελεστές Zernike με χρήση των πολικών ρικελς, καθιστά εφικτό τον υπολογισμό υψηλής τάξης πολυωνύμων, τόσο από πλευράς ακρίβειας όσο και από πλευράς χρόνου. Τέλος, ο προτεινόμενος αλγόριθμος διατηρεί χαμηλά το γεωμετρικό λάθος.



a. αρχικές εικόνες - 300X300



b. ανακατασκευασμένες εικόνες - 300X300

Εικόνα 8.3 Αποτελέσματα ανακατασκευής των εικόνων "Lena", "cameraman", "peppers", μεγέθους 300x300



a. αρχικές εικόνες - 300X300



b. ανακατασκευασμένες εικόνες - 300X300

Εικόνα 8.4 Αποτελέσματα ανακατασκευής των εικόνων "house", "jet-plane", "living room", μεγέθους 300x300



α. αρχικές εικόνες - 300X300



β. ανακατασκευασμένες εικόνες - 300X300

Εικόνα 8.5 Αποτελέσματα ανακατασκευής των εικόνων "man", "ship", "walk bridge", μεγέθους 300x300

Όνομα Εικόνας	P_{max}													
	300		400		500		600		700		800		900	
	MSE	RABSE	MSE	RABSE	MSE	RABSE	MSE	RABSE	MSE	RABSE	MSE	RABSE	MSE	RABSE
Lena	0.0177	0.1358	0.0062	0.0787	0.0085	0.0944	0.0104	0.1052	0.0103	0.1072	0.0096	0.0926	0.0107	0.1012
Peppers	0.0022	0.0386	0.0009	0.0232	0.0006	0.0194	0.0004	0.0163	0.0004	0.0155	0.0003	0.0155	0.0005	0.0189
Cameraman	0.0059	0.0754	0.0016	0.0372	0.0006	0.0218	0.0010	0.0300	0.0014	0.0379	0.0015	0.0393	0.0012	0.0347
House	0.0023	0.0434	0.0027	0.0493	0.0022	0.0458	0.0024	0.0514	0.0028	0.0521	0.0025	0.0457	0.0017	0.0369
Jet Plane	0.0043	0.0598	0.0057	0.0714	0.0004	0.0138	0.0007	0.0208	0.0005	0.0163	0.0007	0.0204	0.0020	0.0410
Living Room	0.0063	0.0679	0.0044	0.0600	0.0051	0.0693	0.0026	0.0491	0.0009	0.0244	0.0005	0.0179	0.0005	0.0182
Man	0.0099	0.0943	0.0060	0.0738	0.0048	0.0686	0.0039	0.0618	0.0033	0.0544	0.0011	0.0310	0.0014	0.0359
Ship	0.0080	0.0835	0.0021	0.0389	0.0010	0.0252	0.0019	0.0419	0.0024	0.0484	0.0036	0.0556	0.0028	0.0486
Walk Bridge	0.0070	0.0681	0.0040	0.0520	0.0019	0.0340	0.0011	0.0257	0.0009	0.0232	0.0007	0.0212	0.0007	0.0234

Πίνακας 8.1 Το λάθος ανακατασκευής που σχετίζεται με τις 9 πρότυπες εικόνες, διαστάσεων 512x512, για τις οποίες οι συντελεστές Zernike παρήχθησαν κάνοντας χρήση της προτεινόμενης μεθόδου. Το λάθος υπολογίστηκε μέσω της (8.1) MSE και (8.2) RABSE σχέσης.

Όνομα Εικόνας	P_{max}													
	100		200		300		400		500		600		700	
	MSE	RABSE	MSE	RABSE	MSE	RABSE	MSE	RABSE	MSE	RABSE	MSE	RABSE	MSE	RABSE
Lena	0.0091	0.0729	0.0077	0.0757	0.0078	0.0898	0.0064	0.0722	0.0068	0.0847	0.0098	0.0910	0.0069	0.0756
Peppers	0.0092	0.0755	0.0056	0.0699	0.0023	0.0422	0.0012	0.0299	0.0012	0.0300	0.0016	0.0370	0.0012	0.0308
Cameraman	0.0133	0.1144	0.0117	0.1057	0.0019	0.0406	0.0011	0.0308	0.0012	0.0314	0.0009	0.0292	0.0007	0.0244
House	0.0029	0.0544	0.0022	0.0422	0.0027	0.0538	0.0039	0.0591	0.0031	0.0554	0.0029	0.0488	0.0038	0.0535
Jet Plane	0.0154	0.1187	0.0138	0.1115	0.0009	0.0209	0.0011	0.0292	0.0007	0.0201	0.0014	0.0334	0.0013	0.0312
Living Room	0.0049	0.0536	0.0043	0.0487	0.0012	0.0237	0.0006	0.0176	0.0005	0.0171	0.0005	0.0187	0.0006	0.0195
Man	0.0211	0.1105	0.0103	0.0952	0.0052	0.0700	0.0039	0.0616	0.0040	0.0626	0.0048	0.0727	0.0024	0.0503
Ship	0.0099	0.0902	0.0085	0.0871	0.0030	0.0503	0.0048	0.0648	0.0036	0.0600	0.0061	0.0731	0.0058	0.0700
Walk Bridge	0.0085	0.0811	0.0068	0.0683	0.0020	0.0354	0.0013	0.0310	0.0011	0.0286	0.0009	0.0256	0.0010	0.0256

Πίνακας 8.2 Το λάθος ανακατασκευής που σχετίζεται με τις 9 πρότυπες εικόνες, διαστάσεων 300x300, για τις οποίες οι συντελεστές Zernike παρήχθησαν κάνοντας χρήση της προτεινόμενης μεθόδου. Το λάθος υπολογίστηκε μέσω της (8.1) MSE και (8.2) RABSE σχέσης.

Όνομα Εικόνας	P_{max}					
	1000		1100		1200	
	MSE	RABSE	MSE	RABSE	MSE	RABSE
Lena	0.0109	0.1180	0.0071	0.0980	0.0071	0.0890
Peppers	0.0005	0.0170	0.0003	0.0140	0.0002	0.0150
Car-Mountain	0.0011	0.0338	0.0010	0.0343	0.0011	0.0370
Cameraman	0.0014	0.0354	0.0021	0.0381	0.0023	0.0410
Man	0.0024	0.0458	0.0019	0.0405	0.0026	0.0524

Πίνακας 8.3 Το λάθος ανακατασκευής που σχετίζεται με τις 9 πρότυπες εικόνες, διαστάσεων 1200x1200, για τις οποίες οι συντελεστές Zernike παρήχθησαν κάνοντας χρήση της προτεινόμενης μεθόδου. Το λάθος υπολογίστηκε μέσω της (8.1) MSE και (8.2) RABSE σχέσης.

Επιπλέον, σύμφωνα με τα πειράματα που εκτελέστηκαν παρατηρήθηκε ότι ο προτεινόμενος αλγόριθμος είναι ανεξάρτητος από τη στροφή της εικόνας. Για να αποδειχθεί αυτό ένα πλήθος από στραμμένες εκδοχές των εικόνων των 8.2a, 8.3a, 8.4a και 8.5 έχουν παραχθεί, αφού πρώτα εντέθηκαν μέσα στον μοναδιαίο εγγεγραμμένο κύκλο. Εν συνεχεία έλαβε χώρα η ανακατασκευή αυτών. Παρακάτω παρουσιάζονται τα αποτελέσματα για την εικόνα “peppers”, για τέσσερις (4) στραμμένες εκδοχές της (Εικόνα 8.6). Τα σχετικά λάθη ανακατασκευής παρουσιάζονται στον Πίνακα 8.4. Σύμφωνα με όσα παρουσιάζονται, επιβεβαιώνεται ότι η μεθοδολογία που προτείνεται είναι ανεξάρτητη από τη στροφή, όσον αφορά την ποιότητα ανακατασκευής της εικόνας.



Γωνία στροφής: 45°



Γωνία στροφής : 120°



Γωνία στροφής : 210°



Γωνία στροφής : 315°

Εικόνα 8.6 Ανακατασκευασμένες εικόνες των στραμμένων εκδοχών της εικόνας "Peppers", με γωνίες 45°, 120°, 215°, 315°.

Γωνία στροφής	$P_{max} = 500$	
	MSE	RABSE
0°	0.0012	0.0300
45°	0.0019	0.0384
120°	0.0013	0.0309
215°	0.0014	0.0337
315°	0.0024	0.0416

Πίνακας 8.4 Λάθος ανακατασκευής των στραμμένων εικόνων "Peppers" με γωνίες στροφής 45°, 120°, 215°, 315°. Στη πρώτη γραμμή παρουσιάζεται το λάθος ανακατασκευής της μη

8.3 Η συμπεριφορά του προτεινόμενου αλγορίθμου όσον αφορά το σφάλμα πεπερασμένης ακρίβειας.

Για να εξεταστεί η συμπεριφορά του προτεινόμενου αλγορίθμου, αναφορικά με το σφάλμα πεπερασμένης ακρίβειας, έγιναν τα εξής πειράματα. Ο αλγόριθμος που παρουσιάζεται στη παρούσα εργασία εκτελέστηκε τρεις φορές. Τη πρώτη φορά με δεκάξι δεκαδικά (53 bits) ψηφία ακρίβεια στη mantissa, τη δεύτερη φορά με 32 (107 bits), και τη τρίτη φορά με 64 δεκαδικά ψηφία (214 bits) ακρίβεια, έχοντας την ίδια εικόνα ως είσοδο. Οι τιμές της εικόνας ήταν εκφρασμένες στη μικρότερη ακρίβεια των 16 δεκαδικών ψηφίων. Εν συνεχεία, πραγματοποιήθηκε η σύγκριση των αποτελεσμάτων που παράχθηκαν από τις τρεις αυτές εκτελέσεις, σύμφωνα με τις μεθόδους της παραγράφου 6 του κεφαλαίου Α. Το ίδιο πείραμα εφαρμόστηκε τόσο στις πρότυπες εικόνες που παρουσιάζονται στις Εικόνες 8.2, 8.3, 8.4 και 8.5 όσο και στις τεχνητές εικόνες με χρήση των οποίων πραγματοποιήθηκαν τα πειράματα του κεφαλαίου 5. Τα σχετικά αποτελέσματα παρουσιάζονται στο πίνακα 8.5. Ακολουθεί η παράθεση των συμπερασμάτων που συνάγονται από όλη τη προηγούμενη ανάλυση, καθώς και από τα πραγματοποιηθέντα πειράματα:

- Ο προτεινόμενος αλγόριθμος δεν εμφάνισε υπερχειλίση στον υπολογισμό των ακτινικών πολυωνύμων Zernike, όπως ήταν αναμενόμενο, λόγω της πλήρους εξάλειψης των παραγοντικών κατά τη διαδικασία υπολογισμού τους. Δηλαδή, το αλγοριθμικό σχήμα που προτείνεται παρακάμπτει τα παραγοντικά που είναι η κύρια αιτία υπερχειλίσης στον αρχικό αλγόριθμο που χρησιμοποιεί pixels κυκλικού τομέα (polar pixels).
- Η βασική πηγή λάθους πεπερασμένης ακρίβειας στον υπολογισμό των πολυωνύμων Zernike Z_p^q με την προτεινόμενη μέθοδο, είναι ίδιου τύπου με αυτή που παρουσιάζεται στο κεφάλαιο 5.
- Δεν υπάρχει συστηματική πηγή λάθους πεπερασμένης ακρίβειας στον αλγόριθμο που παρουσιάζεται στη παρούσα μελέτη. Για αυτό το λόγο ο αλγόριθμος μπορεί να εφαρμοστεί σε εικόνες μεγάλου μεγέθους και με υψηλόβαθμο μέγιστο βαθμό p_{max} των πολυωνύμων Zernike, όπως φαίνεται στις εικόνες 8.2c, 8.3b, 8.4b και 8.5 b και στα αντίστοιχα λάθη ανακατασκευής στους πίνακες 8.2, 8.1, 8.3 και 8.4. Από όλη την ανάλυση και τα πειράματα που έγιναν, θεμελιώνεται ότι το σφάλμα πεπερασμένης ακρίβειας διατηρείται σε χαμηλά επίπεδα, κατά την ανακατασκευή μίας εικόνας, για διάφορα μεγέθη της εικόνας και διάφορα p_{max} .

Εν κατακλείδι, ο αλγόριθμος που παρουσιάζεται στη παρούσα εργασία είναι γρήγορος, και εμφανίζει αξιοσημείωτα μικρό γεωμετρικό σφάλμα, καθώς και σφάλματα ολοκλήρωσης και πεπερασμένης ακρίβειας.

Μέγεθος εικόνας	p_{max}	Μέση τιμή του e.d.d. Πραγματικό μέρος / φανταστικό μέρος	Τυπική απόκλιση του e.d.d. Πραγματικό μέρος / φανταστικό μέρος
512 x 512	500	4.38 / 4.22	1.11 / 1.13
512 x 512	600	4.55 / 4.36	1.12 / 1.13
512 x 512	700	4.67 / 4.51	1.12 / 1.14
700 x 700	800	4.68 / 4.53	1.09 / 1.07
1200 x 1200	1000	5.25 / 5.04	1.16 / 1.15
1200 x 1200	1200	5.38 / 5.20	1.16 / 1.16
1200 x 1200	1600	5.40 / 5.30	1.15 / 1.17
1600 x 1600	1600	5.45 / 5.34	1.13 / 1.13

Πίνακας 8.5 Η μέση τιμή και τυπική απόκλιση του σφάλματος πεπερασμένης ακρίβειας της προτεινόμενης μεθόδου για διάφορα μεγέθη εικόνων και διάφορα p_{max}

9 Βιβλιογραφία Β' Μέρους

- [1] Z. Iscan, Z. Dokur, and T. Ölmez, "Tumor detection by using Zernike moments on segmented magnetic resonance brain images," *Expert Syst. Appl.*, vol. 37, no. 3, pp. 2540–2549, 2010.
- [2] B. J. Chen *et al.*, "Quaternion Zernike moments and their invariants for color image analysis and object recognition," *Signal Processing*, vol. 92, no. 2, pp. 308–318, 2012.
- [3] Campbell C. E., "A new method for describing the aberrations aberrations of the eye using Zernike polynomials," *Optim. Vis. Sci.*, vol. 80, no. 1, pp. 77–83, 2003.
- [4] S. Sharma and P. Khanna, "Computer-Aided Diagnosis of Malignant Mammograms using Zernike Moments and SVM," *J. Digit. Imaging*, vol. 28, no. 1, pp. 77–90, 2014.
- [5] M. N. Singh C. , Walia E., "Rotation invariant complex Zernike moments features and their applications to human face and character recognition," *IET Comput. Vis.*, vol. 5, no. 5, pp. 255–265, 2011.
- [6] H. A. Qader, A. R. Ramli, and A.-H. Syed, "Fingerprint Recognition Using Zernike Moments.pdf," *Int. Arab J. Inf. Technol.*, vol. 4, no. 4, pp. 372–376, 2007.

- [7] Y. Xin, S. Liao, and M. Pawlak, "Circularly orthogonal moments for geometrically robust image watermarking," *Pattern Recognit.*, vol. 40, no. 12, pp. 3740–3752, 2007.
- [8] S. Xiang, J. Huang, R. Yang, C. Wang, and H. Liu, "Robust audio watermarking based on low-order Zernike moments," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 4283 LNCS, pp. 226–240, 2006.
- [9] C. Singh, A. Aggarwal, and S. K. Ranade, "A new convolution model for the fast computation of Zernike moments," *AEU - Int. J. Electron. Commun.*, vol. 72, pp. 104–113, 2017.
- [10] Y. Xin, M. Pawlak, and S. Liao, "Image Reconstruction with Polar Zernike Moments." pp. 394–403, 2005.
- [11] Y. Xin, M. Pawlak, and S. Liao, "Accurate computation of Zernike moments in polar coordinates," *IEEE Transactions on Image Processing*, vol. 16, no. 2. pp. 581–587, 2007.
- [12] L. Kotoulas and I. Andreadis, "Accurate calculation of image moments," *IEEE Transactions on Image Processing*, vol. 16, no. 8. pp. 2028–2037, 2007.
- [13] S. K. Hwang and W. Y. Kim, "A novel approach to the fast computation of Zernike moments," *Pattern Recognition*, vol. 39, no. 11. pp. 2065–2076, 2006.
- [14] C. Singh and R. Upneja, "Accurate calculation of high order pseudo-Zernike moments and their numerical stability," *Digit. Signal Process. A Rev. J.*, vol. 27, no. 1, pp. 95–106, 2014.
- [15] K. M. Hosny, "A systematic method for efficient computation of full and subsets Zernike moments," *Information Sciences*, vol. 180, no. 11. pp. 2299–2313, 2010.
- [16] C. Singh and E. Walia, "Algorithms for fast computation of Zernike moments and their numerical stability," *Image Vis. Comput.*, vol. 29, no. 4, pp. 251–259, 2011.
- [17] W. M. Liu C., Huang X.H., "Fast computation of Zernike moments in polar coordinates," *IET Image Process.*, vol. 6, no. 7, pp. 996–1004, 2012.
- [18] C. W. Chong, P. Raveendran, and R. Mukundan, "A comparative analysis of algorithms for fast computation of Zernike moments," *Pattern Recognition*, vol. 36, no. 3. pp. 731–742, 2003.
- [19] C. Singh and E. Walia, "Fast and numerically stable methods for the computation of Zernike

- moments," *Pattern Recognition*, vol. 43, no. 7. pp. 2497–2506, 2010.
- [20] C. Singh and R. Upneja, "Fast and accurate method for high order Zernike moments computation," *Appl. Math. Comput.*, vol. 218, no. 15, pp. 7759–7773, 2012.
- [21] Chong-Yaw WeeRaveendran Paramesran, "On the computational aspects of Zernike moments," *Image Vis. Comput.*, no. 25, pp. 967–980, 2007.
- [22] G. A. Papakostas, Y. S. Boutalis, C. N. Papaodysseus, and D. K. Fragoulis, "Numerical stability of fast computation algorithms of Zernike moments," *Applied Mathematics and Computation*, vol. 195, no. 1. pp. 326–345, 2008.
- [23] G. A. Papakostas, Y. S. Boutalis, C. N. Papaodysseus, and D. K. Fragoulis, "Numerical error analysis in Zernike moments computation," *Image and Vision Computing*, vol. 24, no. 9. pp. 960–969, 2006.
- [24] C. Singh and E. Walia, "Computation of Zernike moments in improved polar configuration," *IET Image Process.*, vol. 3, no. 4, pp. 217–227, 2009.
- [25] W. V. Prata, A., Rusch, "Algorithm for computation of Zernike polynomials expansion coefficients," *Appl. Opt.*, vol. 28, no. 4, pp. 749–754, 1989.

С. ΥΛΟΠΟΙΗΣΗ ΑΛΓΟΡΙΘΜΩΝ ΠΟΥ
ΕΠΙΤΕΛΟΥΝ ΑΥΤΟΜΑΤΗ
ΑΝΑΓΝΩΡΙΣΗ ΓΡΑΦΕΑ ΜΕ
ΕΦΑΡΜΟΓΗ ΣΕ ΣΗΜΑΝΤΙΚΑ ΠΑΛΑΙΑ
ΚΕΙΜΕΝΑ.

1 Εισαγωγή στην αυτόματη αναγνώριση γραφέα παλαιών (ιστορικών) κειμένων.

1.1 Η σημασία της Αυτόματης Αναγνώρισης Γραφέα παλαιών κειμένων.

Η διάσωση, συντήρηση, καθώς και η ορθή χρονολόγηση γραπτών κειμένων και επιγραφών αποτελεί ένα βασικό κομμάτι της Αρχαιολογίας και της ιστορικής μελέτης. Μέσω αυτών, οι επιστήμονες δύνανται:

1. να “ξεκλειδώσουν” διάφορες πτυχές ενός αρχαίου / μεσαιωνικού πολιτισμού
2. να αναδείξουν σημαντικά αρχαία κείμενα, που αφορούν είτε επιστημονικά είτε λογοτεχνικά έργα.
3. να αποδώσουν τα κείμενα αυτά στη σωστή εποχή και στο σωστό γραφέα.

Δυστυχώς όμως, αρκετοί γραφείς των αρχαίων και μεσαιωνικών κειμένων δεν υπέγραφαν τα κείμενά τους, ούτε ανέφεραν πουθενά την ημερομηνία γραφής τους. Επομένως, η χρονολόγηση αυτών και η απόδοσή τους στους αντίστοιχους γραφείς είναι εξαιρετικά δύσκολη.

Μέχρι σήμερα, η ταυτοποίηση, κατάταξη και χρονολόγηση ενός ιστορικού κειμένου γίνεται με συγκριτικές χρονολογήσεις με βάση το περιεχόμενο, το συντακτικό, τη γραμματική, τη μορφή των συμβόλων της αλφαβήτου, καθώς και την εμπειρία του εκάστοτε αξιολογητή / αρχαιολόγου. Αυτός ο τρόπος προσέγγισης αποδίδει, αλλά δεν είναι αδιαμφισβήτητος, ποσοτικός και αντικειμενικός και ως εκ τούτου αφήνει περιθώρια για αμφισβήτηση σε περιπτώσεις που δεν είναι ξεκάθαρες. Συνεπώς, υπάρχει αναγκαιότητα δημιουργίας εργαλείων και μεθοδολογιών που θα προσφέρουν αδιαμφισβήτητη και αντικειμενική ταυτοποίηση του γραφέα και μέσω αυτών χρονολογική κατάταξη παλαιών κειμένων και επιγραφών.

Στόχος της μεθοδολογίας που αναφέρεται στη παρούσα εργασία και πρωτοπαρουσιάστηκε στα [1], [2] είναι να ομαδοποιήσει ένα σύνολο κειμένων βάσει της ομοιότητας του γραφικού χαρακτήρα εκείνων που τα έγραψαν, με αυστηρά μαθηματικά κριτήρια. Με αυτή τη προσέγγιση γίνεται προσπάθεια προσδιορισμού των αριθμό των διαφορετικών γραφέων ενός συνόλου παλαιών κειμένων, και εν συνεχεία την απόδοση κάθε κειμένου ξεχωριστά στο γραφέα του. Μέσω αυτής της προσπάθειας είναι εφικτή η ιστορική κατάταξη των κειμένων που μελετώνται, είτε αφορά τα πρωτότυπα είτε τα χειρόγραφα αντίγραφα.

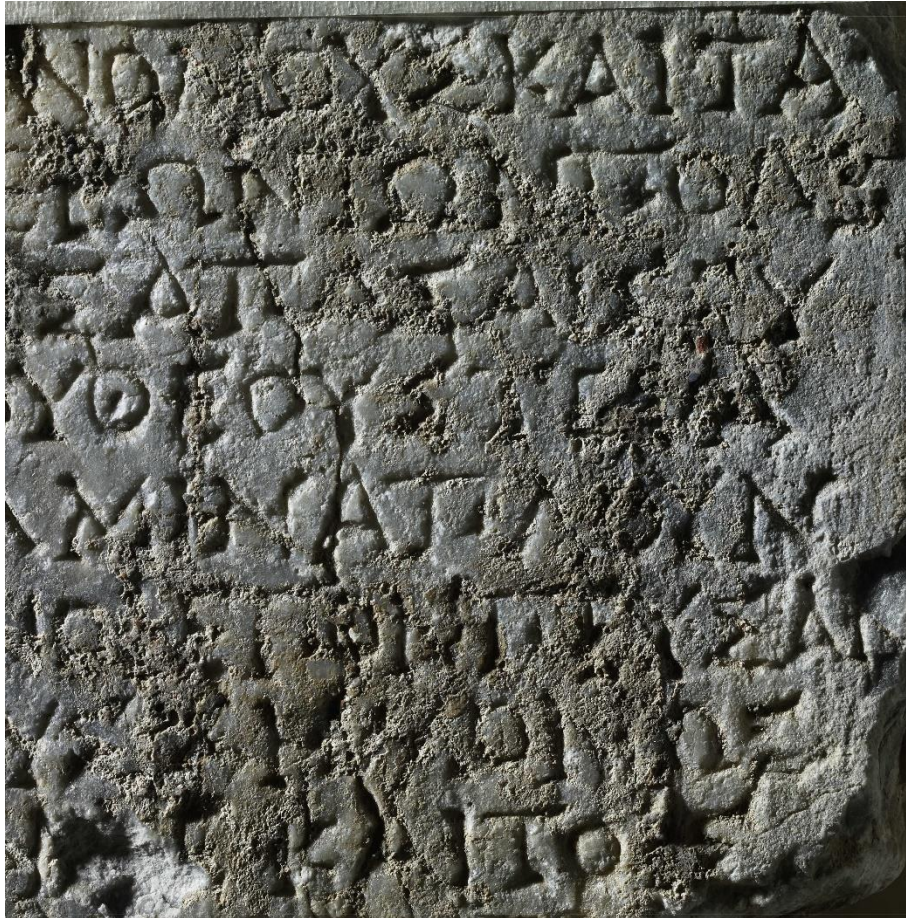
Χαρακτηριστικό παράδειγμα ανάγκης ταυτοποίησης του γραφέα αφορά τα κείμενα “Φυσικής απάνθισμα”, “Το Σαγανάκι της τρέλας” , “Δοκιμασμένη Φιλία” και “Το Σύνταγμα της Ελληνικής Δημοκρατίας”. Από αυτά, το περιεχόμενο του βιβλίου “Φυσικής απάνθισμα”, που εκδόθηκε στη Βιέννη το 1790, και το περιεχόμενου του έργου “Το Σύνταγμα της Ελληνικής Δημοκρατίας”, αποδίδονται τεκμηριωμένα στον Ρήγα Φεραίο Βελεστινλή. Ειδικότερα, θεωρείται ιστορικά τεκμηριωμένο ότι το έργο “Φυσικής απάνθισμα” εγράφη δια χειρός του ιδίου του Ρήγα Φεραίου, ενώ για το κείμενο “Το Σύνταγμα της Ελληνικής Δημοκρατίας” κάτι τέτοιο δε διαφαίνεται. Επιπλέον, δεν υπάρχει τελεσίδικη απόδοση των κειμένων “Το Σαγανάκι της τρέλας” και “Δοκιμασμένη Φιλία”, που σχετικά πρόσφατα ανακαλύφθηκαν στη Ρουμανία, στο γραφέα τους.

1.2 Οι εγγενείς δυσκολίες στη προσπάθεια ταυτοποίησης του γραφέα ενός παλαιού κειμένου.

Πολλές είναι οι δυσκολίες που έπρεπε να αντιμετωπιστούν στη προσπάθεια ταυτοποίησης ενός γραφέα παλαιού κειμένου, με τη μέθοδο που παρουσιάζεται στη παρούσα εργασία και πρωτοπαρουσιάστηκε στα [1], [2]. Μια από τις βασικότερες δυσκολίες είναι ότι δεν υπήρχαν σε κάποιο σύστημα βάσεων δεδομένων πληροφορίες σχετικά με τους γραφείς που επιχειρήθηκε να ταυτοποιηθούν. Επιπροσθέτως, σε αρκετές περιπτώσεις, ήταν άγνωστος μέχρι και ο αριθμός των γραφών που καλείτο η ερευνητική ομάδα να αναγνωρίσει. Ο βαθμός δυσκολίας αυξάνονταν όταν δεν υπήρχαν κείμενα που μπορούσαν να χρησιμοποιηθούν σαν σύνολο εκπαίδευσης (training set) στο σύστημα αναγνώρισης γραφέα. Παρακάτω αναφέρονται γενικές δυσκολίες που υπάρχουν στη προσπάθεια αναγνώρισης ενός γραφέα αρχαίων επιγραφών ή / και γενικότερα ενός παλαιού κειμένου.

- Πολλές επιγραφές, στις οποίες έχει εφαρμοστεί η μεθοδολογία που αναφέρεται στη παρούσα εργασία, έχουν υποστεί σημαντικές φθορές. Συνεπώς είναι δύσκολη και η εξαγωγή πολύτιμης πληροφορίας από αυτές, όπως μπορεί να φανεί και στην εικόνα 1.1 .
- Ένα από τα σημαντικά προβλήματα είναι ότι ένας γραφέας χρησιμοποιεί πάνω από ένα πρότυπο για την αποτύπωση ενός γράμματος. Χαρακτηριστικό παράδειγμα είναι το κείμενο “Δοκιμασμένη Φιλία”, στο οποίο το σύμβολο της αλφαβήτου σ έχει διαφορετικές υλοποιήσεις, ανάλογα με το αν ανήκει στα συμπλέγματα συμφώνων “στ” και “σκ” ή όχι. Επίσης, πολλά γράμματα αλλάζουν μορφή όταν ενώνονται με το επόμενο γράμμα της λέξης. Τέλος το γράμμα “ω” έχει αρκετές σχεδιαστικές ομοιότητες με το γράμμα “π” στα κείμενα αυτά.

- Ένα άλλο χαρακτηριστικό παράδειγμα δυσκολίας, το οποίο αναφέρεται στα [1], [2] είναι το εξής. Σε πολλά κείμενα του Μεσαίωνα και της Αρχαιότητας υπάρχουν δύο υλοποιήσεις του γράμματος “κ”, μία σαν “κ” και μία σαν το λατινικό γράμμα “u”. Επιπροσθέτως, σε κάποια από αυτά τα κείμενα το γράμμα “ω” έχει παραπάνω από μία υλοποίηση, όπως μπορεί να φανεί και από την εικόνα 1.2 .



Εικόνα 1.1 Αρχαία Επιγραφή που έχει υποστεί φθορές.



Εικόνα 1.2 Δυο διαφορετικές υλοποιήσεις του “ω” από ένα βυζαντινό κείμενο, στο οποίο εφαρμόστηκε η μέθοδος που αναφέρεται στη παρούσα εργασία [1], [2].

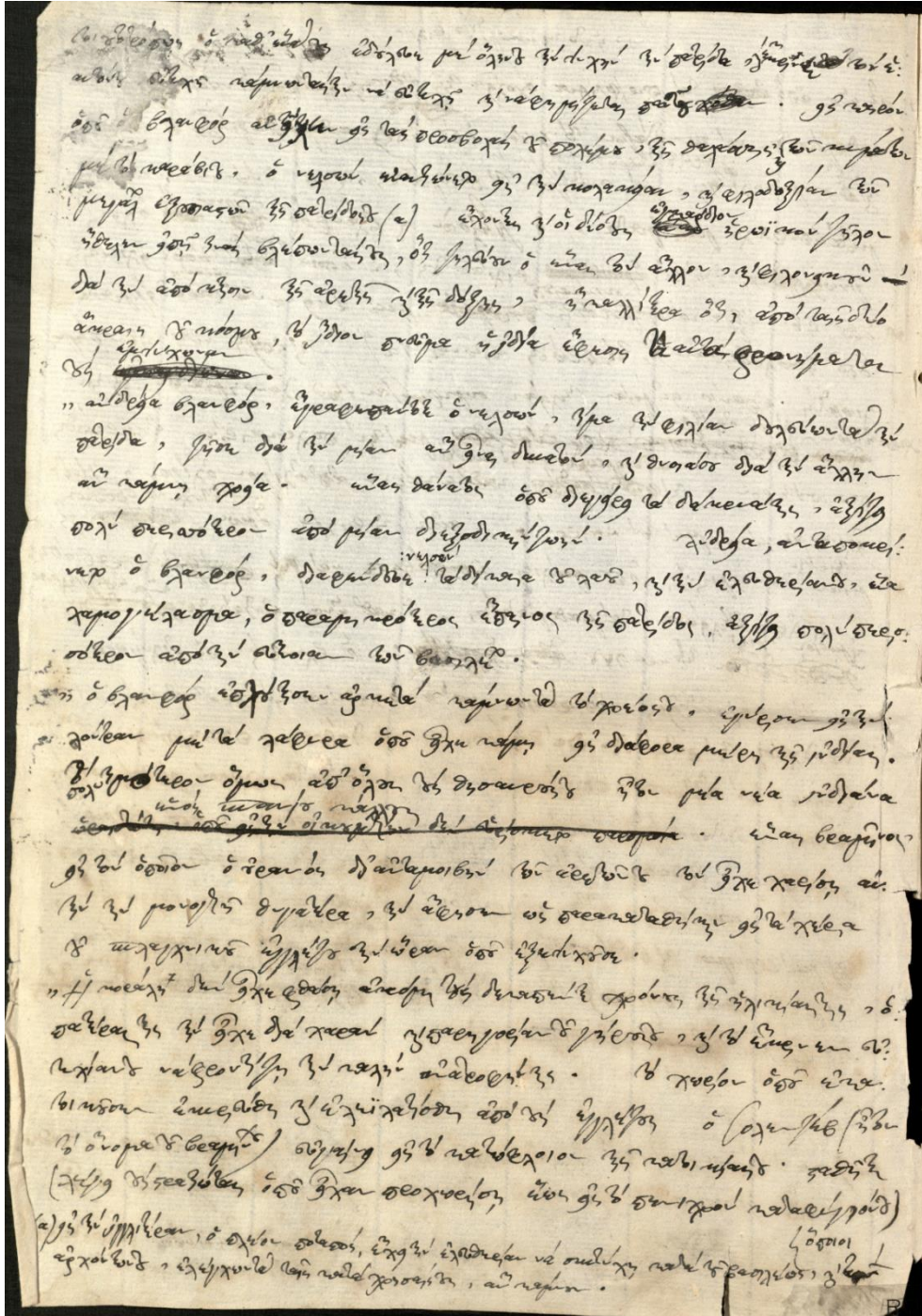
Μερικές άλλες δυσκολίες που εμφανίζονται είναι:

- Ο τρόπος γραφής ενός γράμματος από τον ίδιο γραφέα μπορεί να αλλάξει με τη πάροδο του χρόνου.
- Η συναισθηματική κατάσταση του γραφέα τη στιγμή της συγγραφής του κειμένου μπορεί να μεταβάλλεται και άρα να επηρεάζει το γραφικό του χαρακτήρα.
- Η ποιότητα του υλικού στο οποίο αποτυπώνονταν το κείμενο μπορεί να είναι κάθε φορά διαφορετική. Το ίδιο ισχύει και για το εργαλείο γραφής.
- Πολλές φορές η μη συστηματική συντήρηση των παλαιών κειμένων προκαλεί αρκετές αλλοιώσεις στη μορφή των γραμμάτων.

2 Συνοπτική παρουσίαση μίας νέας μεθόδου ταυτοποίησης του γραφέα παλαιών κειμένων που παρουσιάστηκε για πρώτη φορά από την ερευνητική μας ομάδα.

Στην παρούσα εργασία υλοποιείται υπολογιστικά με λίαν αποδοτικό τρόπο μια νέα μέθοδος ταυτοποίησης του γραφέα ενός κειμένου, η οποία παρουσιάστηκε για πρώτη φορά στα [1], [2]. Η μέθοδος αυτή χρησιμοποιεί την πρωτότυπη μαθηματική έννοια της επιπέδου καμπυλότητας ([1], [2]) και βασίζεται στην στατιστική ομοιότητα των υλοποιήσεων, αναφορικά με πολλά χαρακτηριστικά των περιγραμμάτων των ιδίων συμβόλων της αλφαβήτου σε όλη την έκταση των δυο κειμένων. Η συγκεκριμένη μέθοδος αποτελείται από τα βήματα που περιγράφονται στις αμέσως επόμενες παραγράφους 2.1, 2.2, 2.3, 2.4.

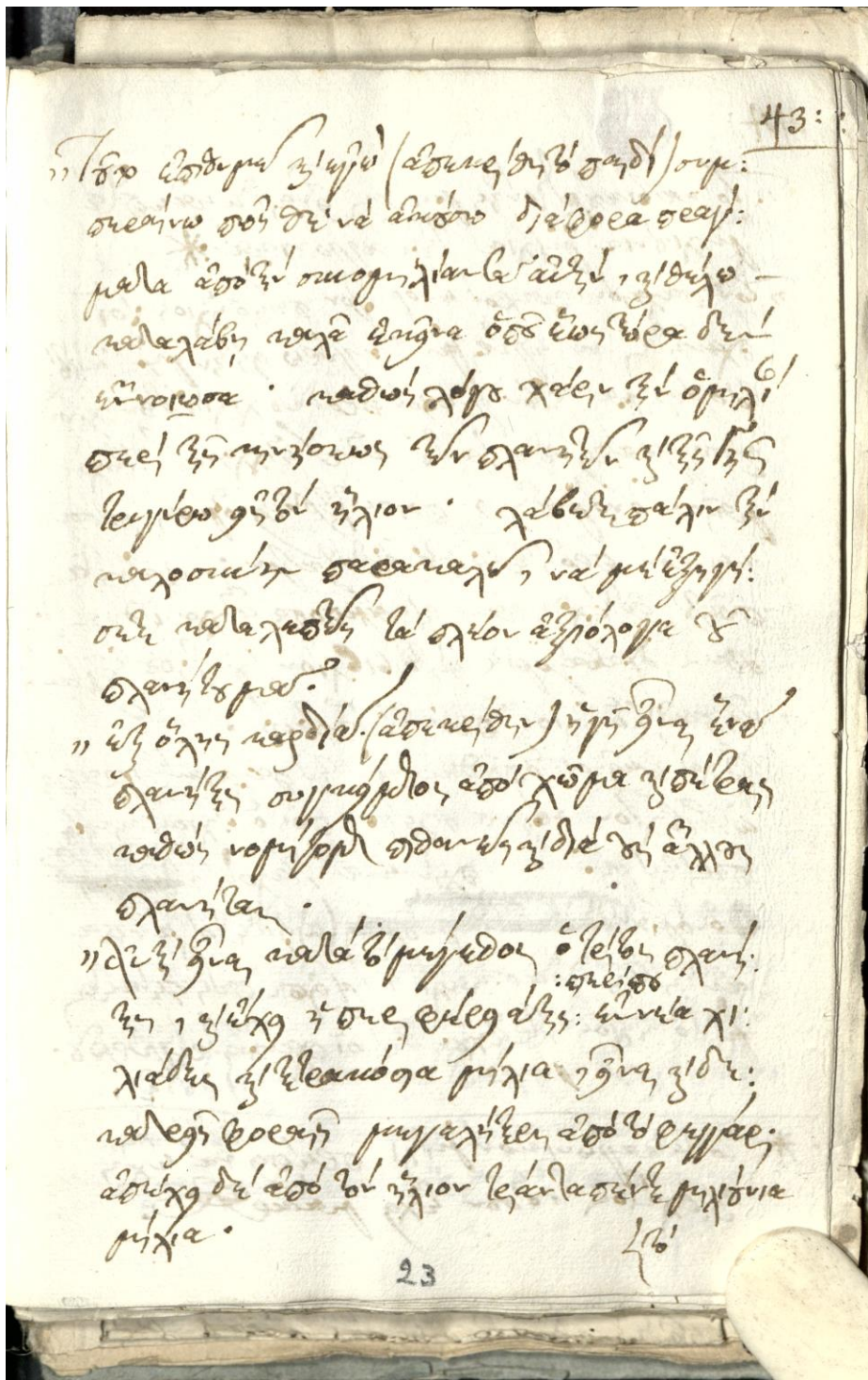
2.1 Εξαγωγή μεμονωμένων γραμμάτων από το κείμενο, ένθεση αυτών σε μεμονωμένες εικόνες και αυτόματη εξαγωγή των περιγραμμάτων τους.



Εικόνα 2.1 Σελίδα από το βιβλίο Δοκιμασμένη Φιλία που εικάζεται ότι ανήκει στον Ρήγα Φεραίο.



Εικόνα 2.2 Γράμματα που έχουν εξαχθεί από το κείμενο Δοκιμασμένη Φιλία, καθώς και τα αντίστοιχα περιγράμματά τους.



ἡ τὸ ἀποκρυφῶς (ἀποκρυφῶς) οὐκ
 οὐκ ἔστιν ἀποκρυφῶς ἀποκρυφῶς
 μαλα ἀποκρυφῶς ἀποκρυφῶς ἀποκρυφῶς
 ἀποκρυφῶς ἀποκρυφῶς ἀποκρυφῶς ἀποκρυφῶς
 ἀποκρυφῶς ἀποκρυφῶς ἀποκρυφῶς ἀποκρυφῶς
 ἀποκρυφῶς ἀποκρυφῶς ἀποκρυφῶς ἀποκρυφῶς
 ἀποκρυφῶς ἀποκρυφῶς ἀποκρυφῶς ἀποκρυφῶς
 ἀποκρυφῶς ἀποκρυφῶς ἀποκρυφῶς ἀποκρυφῶς
 ἀποκρυφῶς ἀποκρυφῶς ἀποκρυφῶς ἀποκρυφῶς

ἡ ἀποκρυφῶς ἀποκρυφῶς ἀποκρυφῶς ἀποκρυφῶς
 ἀποκρυφῶς ἀποκρυφῶς ἀποκρυφῶς ἀποκρυφῶς
 ἀποκρυφῶς ἀποκρυφῶς ἀποκρυφῶς ἀποκρυφῶς
 ἀποκρυφῶς ἀποκρυφῶς ἀποκρυφῶς ἀποκρυφῶς
 ἀποκρυφῶς ἀποκρυφῶς ἀποκρυφῶς ἀποκρυφῶς
 ἀποκρυφῶς ἀποκρυφῶς ἀποκρυφῶς ἀποκρυφῶς
 ἀποκρυφῶς ἀποκρυφῶς ἀποκρυφῶς ἀποκρυφῶς
 ἀποκρυφῶς ἀποκρυφῶς ἀποκρυφῶς ἀποκρυφῶς

ἡ ἀποκρυφῶς ἀποκρυφῶς ἀποκρυφῶς ἀποκρυφῶς
 ἀποκρυφῶς ἀποκρυφῶς ἀποκρυφῶς ἀποκρυφῶς
 ἀποκρυφῶς ἀποκρυφῶς ἀποκρυφῶς ἀποκρυφῶς
 ἀποκρυφῶς ἀποκρυφῶς ἀποκρυφῶς ἀποκρυφῶς
 ἀποκρυφῶς ἀποκρυφῶς ἀποκρυφῶς ἀποκρυφῶς
 ἀποκρυφῶς ἀποκρυφῶς ἀποκρυφῶς ἀποκρυφῶς
 ἀποκρυφῶς ἀποκρυφῶς ἀποκρυφῶς ἀποκρυφῶς
 ἀποκρυφῶς ἀποκρυφῶς ἀποκρυφῶς ἀποκρυφῶς

Εικόνα 2.3 Σελίδα από το βιβλίο "Φυσικής απάνθισμα".



Εικόνα 2.4 Γράμματα που έχουν εξαχθεί από το “Φυσικής Απάνθισμα”, καθώς και τα αντίστοιχα περιγράμματά τους

2.2 Μια πρώτη προσαρμογή ζευγών υλοποιήσεων του ίδιου συμβόλου της αλφαβήτου.

Έστωσαν δυο τυχούσες υλοποιήσεις ΣA_1 και ΣA_2 του ίδιου συμβόλου της αλφαβήτου ΣA , οι οποίες ενδέχεται να έχουν εντοπιστεί είτε στο ίδιο κείμενο είτε σε δύο διαφορετικά. Για να ελέγξουμε την ομοιότητα των ΣA_1 και ΣA_2 με όσον τον δυνατόν πιο αντικειμενικό και μετρήσιμο τρόπο, θα προσαρμόσουμε βέλτιστα την ΣA_2 στην ΣA_1 , εφαρμόζοντας στην πρώτη την τετράδα των αφινικών μετασχηματισμών:

- Ομοιοθεσία
- Στροφή
- Παράλληλη μετατόπιση κατά μήκος του άξονα των x
- Παράλληλη μετατόπιση κατά μήκος του άξονα των y

2.2.1 Υπολογισμός της βέλτιστης ομοιοθεσίας που προσαρμόζει την υλοποίηση ΣA_2 στην ΣA_1

Μια αρχική εκτίμηση του συντελεστή ομοιοθεσίας λ , ο οποίος θα μεγεθύνει ή θα σμικρύνει την υλοποίηση ΣA_2 προς την κατεύθυνση της βέλτιστης προσαρμογής της στη ΣA_1 , θα χρησιμοποιήσουμε την νέα έννοια της επιπέδου καμπυλότητας. Η οντότητα αυτή ορίστηκε για πρώτη φορά στα [1], [2] και στην ουσία εκφράζει τον βαθμό ομοιότητας της συνολικής καμπυλότητας των περιγραμμάτων των ΣA_2 και ΣA_1 , υπό την προϋπόθεση ότι η υλοποίηση ΣA_2 έχει υποστεί κατάλληλη ομοιοθεσία. Για τον προσδιορισμό μίας πολύ καλής εκτίμησης αυτού του συντελεστή ομοιοθεσίας, έστω λ , επιτελούνται τα εξής βήματα ([1], [2]):

Βήμα 1: Ενθέτουμε τις εικόνες αμφοτέρων των υλοποιήσεων ΣA_2 στην ΣA_1 στο ίδιο χωρίο (εικόνα), έστω I .

Επομένως υποθέτουμε ότι η καμπύλη Γ_2 , που συνιστά το περίγραμμα της υλοποίησης ΣA_2 , βρίσκεται στην αρχική της θέση. Για κάθε σημείο $M(x, y) \in \Gamma_2$ υπολογίζουμε την ισοϋψή της Γ_1 , η οποία είναι το περίγραμμα της ΣA_1 . Η απόσταση αυτή δίνεται από τον τύπο:

$$f(x, y) = d(M, \Gamma_1) = \sqrt{d^2(x, \Gamma_1) + d^2(y, \Gamma_1)} \quad (2.1)$$

Όπου $d(x, \Gamma_1)$ και $d(y, \Gamma_1)$ είναι οι αποστάσεις του σημείου M από την Γ_1 κατά μήκος του x και y άξονα αντίστοιχα.

Βήμα 2: Υπολογίζουμε τη ποσότητα $\vec{\nabla} f(x, y)$ σε κάθε σημείο $M(x, y)$ της Γ_2 και εν συνεχεία τις ποσότητες g_a, g_b, ρ^2 μέσω των σχέσεων:

$$\rho^2 = \left[\frac{\partial f}{\partial x} \frac{\partial f}{\partial y} \right] \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} = \left(\frac{\partial f}{\partial x} \right)^2 + \left(\frac{\partial f}{\partial y} \right)^2$$

$$g_a = 3 \left(\frac{\partial f}{\partial x} \right)^2 - 2\rho^2$$

$$g_b = 3 \left(\frac{\partial f}{\partial y} \right)^2 - 2\rho^2$$

Βήμα 3: Θεωρούμε ότι οι συντελεστές ομοιοθεσίας κατά μήκος των αξόνων x και y είναι a και b αντίστοιχα. Ο στόχος είναι να βρεθούν τα a και b που μετασχηματίζουν το μέγεθος της ΣA_2 ώστε να είναι πολύ παραπλήσιο με το μέγεθος της ΣA_1 . Αρχικά θέτουμε $a = 1, b = 1$, ώστε η διαδικασία να εκκινήσει από την κατάσταση στην οποία η καμπύλη Γ_2 είναι απαραμόρφωτη. Υπολογίζουμε το επικαμπύλιο ολοκλήρωμα πάνω στη Γ_1

$$\varepsilon_1 = \oint_{\Gamma_1} \vec{\nabla} \cdot \vec{n} \, dl$$

Το ολοκλήρωμα αυτό υπολογίζεται αριθμητικά μόνο μία φορά και για την κοντινότερη ισοϋψή της Γ_1 προς την κατεύθυνση της Γ_2 , εάν η κατεύθυνση αυτή είναι καλώς ορισμένη ή ισοδύναμα, προς την εξωτερική πλευρά της Γ_1 στο ψηφιακό χωρίο I της εικόνας που συμπεριλαμβάνει τις υλοποιήσεις του ΣA_2 και ΣA_1 του τυχόντος γράμματος ΣA .

Βήμα 4: Σε κάθε σημείο $M(x, y)$ της Γ_2 υπολογίζουμε την ευκλείδεια απόσταση αυτού από τη Γ_1 , καθώς και το μοναδιαίο διάνυσμα $\vec{n} = \frac{\vec{\nabla} f(x, y)}{\|\vec{\nabla} f(x, y)\|}$ και τέλος την ποσότητα $C(x, y) = \vec{\nabla} \cdot \vec{n}$. Υπενθυμίζεται ότι το $f(x, y)$ υπολογίζεται όπως στο Βήμα 1. Το σύνολο των τιμών αυτών $C(x, y)$ παίζει το ρόλο των αρχικών τιμών της επίπεδης καμπυλότητας Γ_2 σε σχέση με τη Γ_1 . Για αυτές τις τιμές της επίπεδης καμπυλότητας εφεξής θα χρησιμοποιείται το σύμβολο $C_0^{\Gamma_2}(x, y)$. Ο δείκτης μηδέν (0) υποδεικνύει ότι οι τιμές αυτές της καμπυλότητας υπολογίζονται στη θέση μηδέν, δηλαδή στην αρχική θέση, όπου η Γ_2 είναι απαραμόρφωτη, ενώ ο δείκτης Γ_2 υποδεικνύει ότι οι τιμές αυτές της καμπυλότητας αφορούν την καμπύλη Γ_2 .

Βήμα 5: Δεδομένων των τιμών $C_0^{\Gamma_2}(x, y)$ για κάθε σημείο της καμπύλης Γ_2 και δεδομένου ότι τα g_a και g_b έχουν ήδη υπολογιστεί, σε πλήρη αντιστοιχία με τις ποσότητες που ορίστηκαν στον Βήμα 2 οι προαναφερθείσες ποσότητες συμβολίζονται ως g_a^0, g_b^0, ρ_0^2 . Τότε, για κάθε σημείο της Γ_2 , υπολογίζεται η ποσότητα

$$\delta C_0^{\Gamma_2}(x, y) = C_0^{\Gamma_2}(x, y) \frac{g_a^0 \delta a + g_b^0 \delta b}{\rho_0^2} \quad (2.2)$$

Μέσω της (2.2), για κάθε σημείο της Γ_2 υπολογίζεται η ποσότητα:

$$C_1^{F_2}(x, y) = C_0^{F_2}(x, y) + \delta C_0^{F_2}(x, y)$$

Δηλαδή είναι η νέα τιμή της καμπυλότητας, όταν μεταβάλλαμε το συντελεστή ομοιοθεσίας κατά μήκος του x κατά $\delta\alpha$ και αυτόν κατά μήκος του y κατά $\delta\beta$.

Βήμα 6: Εάν οι συναρτήσεις $C_1^{F_2}(x, y)$ και $C_0^{F_2}(x, y)$ εντεθούν σε ένα αλγόριθμο ελαχιστοποίησης, τελικά ελαχιστοποιείται το σφάλμα:

$$|\zeta^C| = \left| 2 \left(\oint_{\Gamma_2} \vec{\nabla} \cdot \vec{n} dl - \oint_{\Gamma_1} \vec{\nabla} \cdot \vec{n} dl \right) \right| \quad (2.3)$$

Όπου τα ολοκληρώματα υπολογίζονται αριθμητικά.

Βήμα 7: Με δεδομένα τα $\delta\alpha_0, \delta b_0$ του προηγούμενου βήματος, προσδιορίζουμε τη νέα θέση της καμπύλης Γ_2 , η οποία συμβολίζεται $\tilde{\Gamma}_2(x_1, y_1)$, όπου κάθε σημείο (x_1, y_1) προκύπτει από κάθε αρχικό σημείο (x, y) της $\Gamma_2(x, y)$ από την ακόλουθη σχέση:

$$\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} \delta\alpha & 0 \\ 0 & \delta b \end{bmatrix} \right) \begin{bmatrix} x \\ y \end{bmatrix}$$

Βήμα 8: Θέτοντας τη τιμή της $C_1^{F_2}(x, y)$ στο σημείο (x_1, y_1) της $\tilde{\Gamma}_2$, έχουμε της συνάρτηση $C_1^{F_2}(x_1, y_1)$.

Βήμα 9: Υπολογίζουμε το συναρτησιοειδές $|\zeta^C|$ για την νέα καμπύλη $\tilde{\Gamma}_2(x_1, y_1)$ μέσω της σχέσης (2.3) που παρουσιάστηκε στο Βήμα 6. Συγκεκριμένα, υπολογίζεται η νέα ποσότητα $\varepsilon_2 = \oint_{\tilde{\Gamma}_2} \vec{\nabla} \cdot \vec{n} dl$ μέσω της αριθμητικής ολοκλήρωσης της ποσότητας $C_1^{F_2}(x_1, y_1)$ πάνω στη καμπύλη $\tilde{\Gamma}_2(x_1, y_1)$, αφού για κάθε σημείο $C_1^{F_2}(x_1, y_1)$ ισχύει $C_1^{F_2}(x_1, y_1) = \vec{\nabla} \cdot \vec{n}$. Στην περίπτωση κατά την οποία η ποσότητα $|\zeta^C|$ είναι μικρότερη από ένα κατάλληλα επιλεγμένο κατώφλι, το οποίο συνήθως σχετίζεται με τον αλγόριθμο ελαχιστοποίησης ο οποίος χρησιμοποιείται, η διαδικασία σταματά. Σε αντίθετη περίπτωση, η διαδικασία προχωρά στο επόμενο Βήμα.

Βήμα 10: Σε αυτό το βήμα υπολογίζεται η κάτωθι ποσότητα:

$$\tilde{\nabla}(x_1, y_1)\tilde{f}(x_1, y_1) = \mathbf{A}^T \tilde{\nabla}(x, y)f(x, y) \Leftrightarrow \begin{bmatrix} \frac{\partial \tilde{f}}{\partial x_1} \\ \frac{\partial \tilde{f}}{\partial y_1} \end{bmatrix} = \mathbf{A}^T \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} \quad (2.4)$$

Όπου η $\mathbf{A} = \begin{bmatrix} 1 + \delta a & 0 \\ 0 & 1 + \delta b \end{bmatrix}$. Εφόσον έχουν υπολογιστεί τα $\frac{\partial \tilde{f}}{\partial x_1}, \frac{\partial \tilde{f}}{\partial y_1}$, τότε μέσω των σχέσεων του βήματος 2, υπολογίζονται τα g_a^1, g_b^1, ρ_1^2 .

Βήμα 11: Επιστρέφοντας στο Βήμα 5, επαναλαμβάνονται τα Βήματα 5 έως 10 για τη καμπύλη $\tilde{\Gamma}_2(x_1, y_1)$. Συγκεκριμένα, χρησιμοποιώντας τις νέες τιμές των $C_1^{\Gamma_2}(x_1, y_1), g_a^1, g_b^1$, και ρ_1^2 στη θέση των $C_0^{\Gamma_2}(x, y), g_a^0, g_b^0$, και ρ_0^2 λαμβάνεται η νέα μετασχηματισμένη εκδοχή της $\Gamma_2(x, y)$, και ούτω καθεξής, έως ότου ικανοποιηθεί η συνθήκη τερματισμού του Βήματος 9. Δηλαδή η διαδικασία σταματά όταν η συνάρτηση σφάλματος $|\zeta^C|$ λάβει τιμή κάτω από το καθορισμένο / επιθυμητό κατώφλι.

Με το πέρας αυτής της διαδικασίας επιτυγχάνουμε τα εξής:

- i. Εντοπίζουμε τη βέλτιστη μετασχηματισμένη εκδοχή της καμπύλης Γ_2 , δηλαδή τη $\tilde{\Gamma}_2$, με την έννοια ότι η $\tilde{\Gamma}_2$ έχει συνολικά ελάχιστη διαφορά επιπέδου καμπυλότητας από τη Γ_1 .
- ii. Προσδιορίζουμε την αντιστοιχία κάθε σημείου M_1 της Γ_1 με το κατάλληλο σημείο M_2 της Γ_2 , ούτως ώστε αυτά να εμφανίζουν συνολικά την ελάχιστη διαφορά καμπυλότητας.
- iii. Κυρίως, δε, εντοπίζουμε τη βέλτιστη ομοιοθεσία, η οποία μπορεί να τροποποιήσει το μέγεθος της Γ_2 , ώστε αυτή η καμπύλη να έχει μέγεθος που είναι το πλησιέστερο δυνατό στο μέγεθος της Γ_1 .

2.2.2 Ενσωμάτωση της στροφής και της παράλληλης μετατόπισης για μια πρώτη καλή προσαρμογή της καμπύλης Γ_2 στη Γ_1 .

Όπως έχει ήδη αποδειχθεί στα [1], [2], η ελαχιστοποίηση της επίπεδης καμπυλότητας είναι ανεξάρτητη της στροφής και της παράλληλης μετατόπισης. Συνεπώς, ως υποθεθεί ότι μια πρώτη ικανοποιητική τιμή του συντελεστού ομοιοθεσίας λ της Γ_2 , ώστε αυτή να προσαρμοστεί ικανοποιητικά στη Γ_1 , έχει ήδη υπολογιστεί. Τότε θα ανέμενε κανείς ότι θα υπάρχουν άπειρες θέσεις της μετασχηματισμένης $\tilde{\Gamma}_2$ ως προς τη Γ_1 για τις οποίες η $|\zeta^C|$ έχει την ελάχιστη τιμή. Επομένως, ένας αλγόριθμος προσαρμογής θα πρέπει

να λαμβάνει υπόψη τόσο τη στροφή, όσο και την παράλληλη μετατόπιση. Με το σκεπτικό αυτό, επελέγη το σφάλμα προσαρμογής που περιγράφεται κατωτέρω:

Έστω P τυχόν σημείο της Γ_2 και έστω $d(P)$ η Ευκλείδεια απόσταση του P από τη Γ_1 . Το σφάλμα προσαρμογής της $\tilde{\Gamma}_2$ στη Γ_1 το οποίο έχει επιλεγεί είναι το επικαμπύλιο ολοκλήρωμα της απόστασης όλων των σημείων P της $\tilde{\Gamma}_2$ από τη Γ_1 , δηλαδή το

$$T(\tilde{\Gamma}_2, \Gamma_1) = \oint_{\tilde{\Gamma}_2} d(P) dl$$

Η ποσότητα $T(\tilde{\Gamma}_2, \Gamma_1)$ ελαχιστοποιείται με χρήση ενός οποιοδήποτε αλγορίθμου ελαχιστοποίησης, όπως ο Nelder – Mead αλγόριθμος, ο οποίος και χρησιμοποιήθηκε.

Σε κάθε σημείο του χρησιμοποιούμενου αλγορίθμου ελαχιστοποίησης, η νέα θέση της $\tilde{\Gamma}_2$ επανυπολογίζεται μέσω των αφινικών μετασχηματισμών:

$$\begin{bmatrix} x_{n+1} \\ y_{n+1} \end{bmatrix} = \begin{bmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{bmatrix} \begin{bmatrix} x_n - \bar{x}(\Gamma_1) \\ y_n - \bar{y}(\Gamma_1) \end{bmatrix} + \begin{bmatrix} \bar{x}(\Gamma_1) \\ \bar{y}(\Gamma_1) \end{bmatrix} + \begin{bmatrix} \Delta_x \\ \Delta_y \end{bmatrix} \quad (2.5)$$

Όπου

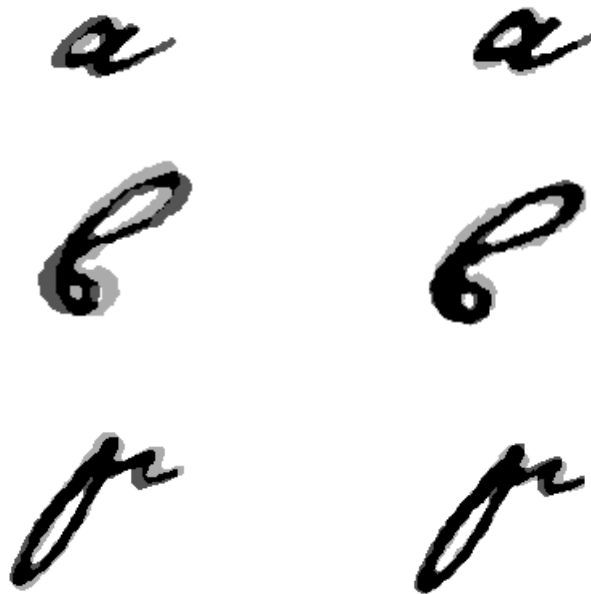
- $(\bar{x}(\Gamma_1), \bar{y}(\Gamma_1))$ είναι το κέντρο βάρους της καμπύλης Γ_1 .
- (x_n, y_n) οι συντεταγμένες του τυχόντος σημείου P της n -οστής εκδοχής της $\tilde{\Gamma}_2$ κατά την n -οστή αναδρομή του αλγορίθμου ελαχιστοποίησης.
- (x_{n+1}, y_{n+1}) οι συντεταγμένες του τυχόντος σημείου P της $n + 1$ -οστής εκδοχής της $\tilde{\Gamma}_2$ κατά την $n + 1$ -οστή αναδρομή του αλγορίθμου ελαχιστοποίησης.
- φ η γωνία στροφής που επιβάλλεται στην $\tilde{\Gamma}_2$, ώστε αυτή να μεταβεί από την n -οστή στη $n + 1$ -οστή εκδοχή αυτής.
- (Δ_x, Δ_y) οι παράλληλες μετατοπίσεις κατά μήκος των αξόνων x και y αντίστοιχα, ούτως ώστε η $\tilde{\Gamma}_2$ να μεταβεί από την n -οστή στη $n + 1$ -οστή εκδοχή αυτής.

2.3 Το χρησιμοποιηθέν βέλτιστο κριτήριο ομοιότητας μεταξύ δυο υλοποιήσεων της αλφαβήτα.

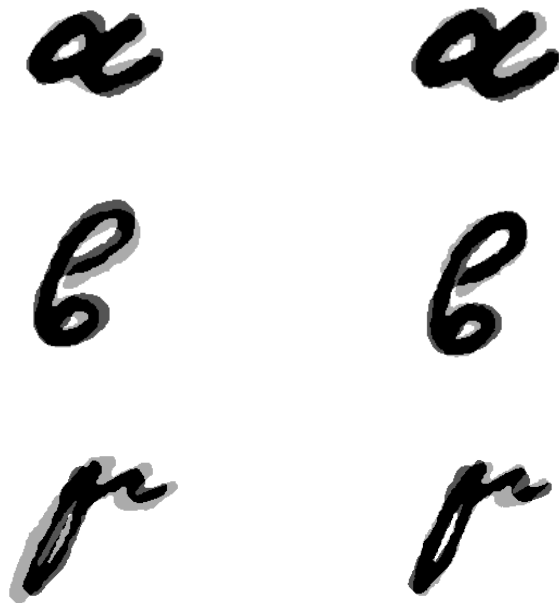
Ας υποθεθεί ότι δυο τυχαίες υλοποιήσεις του ίδιου γράμματος, οι οποίες μπορούν να ανήκουν είναι στο ίδιο κείμενο είτε σε διαφορετικά, έχουν προσαρμοστεί βέλτιστα μεταξύ τους με μία οποιαδήποτε μέθοδο, σε μία συγκεκριμένη θέση. Στη θέση αυτή, η ομοιότητα των δύο υλοποιήσεων ΣA_1 και ΣA_2 εκφράζεται αριθμητικά βέλτιστα με το κάτωθι κριτήριο.

$$\mu = \frac{EMB(\Sigma A_1 \cap \Sigma A_2(\lambda, \varphi, dx, dy))}{EMB(\Sigma A_1 \cup \Sigma A_2(\lambda, \varphi, dx, dy))}$$

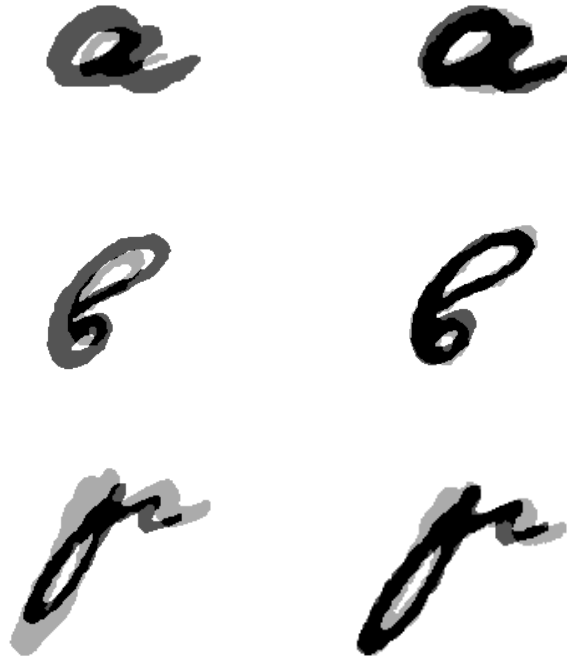
Στην ιδανική περίπτωση, όπου οι υλοποιήσεις ΣA_1 και ΣA_2 είναι ταυτοτικά ίδιες, η τιμή του κριτηρίου ισούται με 1. Αντιστρόφως όσο η ΣA_1 και ΣA_2 διαφοροποιούνται η τιμή του κριτηρίου όλο και μικραίνει, τείνοντας προς το μηδέν.



Εικόνα 2.5 Τρία διαφορετικά ζεύγη υλοποιήσεων τριών συμβόλων της αλφαβήτου, των α,β,μ κατά σειρά, τα οποία εξήχθησαν από το “Το Φυσικής Απάνθισμα”. Κάθε ζεύγος είναι τοποθετημένο στην ίδια εικόνα έχοντας μόνον κοινό κέντρο βάρους, ενώ κανένα μέλος του ζεύγους δεν έχει υποστεί στροφή ή ομοιοθεσία σε σχέση με την αρχική τους θέση στο κείμενο. Το αποτέλεσμα της βέλτιστης προσαρμογής των ζευγών της αριστερής στήλης μετά από τη κατάλληλη ομοιοθεσία, στροφή και παράλληλη μετατόπιση παρουσιάζεται στη δεξιά στήλη. Η βέλτιστη αυτή προσαρμογή γίνεται επί τη βάση του κριτηρίου 2.4



Εικόνα 2.6 Τρία διαφορετικά ζεύγη υλοποιήσεων τριών συμβόλων της αλφαβήτου, των α,β,μ κατά σειρά, τα οποία εξήχθησαν από το “Το Σαγανάκι της Τρέλας”. Κάθε ζεύγος είναι τοποθετημένο στην ίδια εικόνα έχοντας μόνον κοινό κέντρο βάρους, ενώ κανένα μέλος του ζεύγους δεν έχει υποστεί στροφή ή ομοιοθεσία σε σχέση με την αρχική τους θέση στο κείμενο. Το αποτέλεσμα της βέλτιστης προσαρμογής των ζευγών της αριστερής στήλης μετά από τη κατάλληλη ομοιοθεσία, στροφή και παράλληλη μετατόπιση παρουσιάζεται στη δεξιά στήλη. Η βέλτιστη αυτή προσαρμογή γίνεται επί τη βάση του κριτηρίου (2.6)



Εικόνα 2.7 Τρία διαφορετικά ζεύγη υλοποιήσεων τριών συμβόλων της αλφαβήτου, των α,β,μ κατά σειρά, από τα οποία ένα εξάχθηκε από το “Το Φυσικής Απάνθισμα” και ένα από το “Το Σαγανάκι της Τρέλας”. Κάθε ζεύγος είναι τοποθετημένο στην ίδια εικόνα έχοντας μόνον κοινό κέντρο βάρους, ενώ κανένα μέλος του ζεύγους δεν έχει υποστεί στροφή ή ομοιοθεσία σε σχέση με την αρχική τους θέση στο κείμενο. Το αποτέλεσμα της βέλτιστης προσαρμογής των ζευγών της αριστερής στήλης μετά από τη κατάλληλη ομοιοθεσία, στροφή και παράλληλη μετατόπιση παρουσιάζεται στη δεξιά στήλη. Η βέλτιστη αυτή προσαρμογή γίνεται επί τη βάσει του κριτηρίου (2.6)

2.4 Τελεσίδικη, βέλτιστη προσαρμογή δυο υλοποιήσεων του ίδιου γράμματος.

Ας υποθεθεί ότι η μέθοδος της παραγράφου 2.2 έδωσε μια πρώτη σχετικά ικανοποιητική προσαρμογή δυο συμβόλων της αλφαβήτου, των ΣA_1 και ΣA_2 . Χωρίς καμία βλάβη της γενικότητας, ας θεωρηθεί ότι η υλοποίηση ΣA_1 είναι η πρότυπη, ενώ η ΣA_2 προσαρμόστηκε στη ΣA_1 μετά από μία ομοιοθεσία με συντελεστή λ , μία στροφή κατά γωνία φ και μία παράλληλη μετατόπιση κατά το διάνυσμα $dx \cdot \vec{i} + dy \cdot \vec{j}$, μετά την εφαρμογή της μεθόδου 2.2. Έστω επίσης ότι σε αυτή τη θέση προσαρμογής, το κριτήριο που εισήχθη στη παράγραφο 2.3 έδωσε μία τιμή μ . Επαναλαμβάνεται ότι εάν $\Sigma A_2(\lambda, \varphi, dx, dy)$ είναι η μετασχηματισμένη υλοποίηση της ΣA_2 σύμφωνα με τις τιμές των ορισμάτων της, τότε ισχύει

$$\mu = \frac{EMB(\Sigma A_1 \cap \Sigma A_2(\lambda, \varphi, dx, dy))}{EMB(\Sigma A_1 \cup \Sigma A_2(\lambda, \varphi, dx, dy))} \quad (2.6)$$

Στο σημείο αυτό ορίζουμε τα εξής διαστήματα των τιμών των παραμέτρων:

- $[\lambda_A, \lambda_T]$ εκατέρωθεν του λ και συμμετρικά ως προς αυτό. Το διάστημα αυτό διακριτοποιείται με ένα βήμα $\delta\lambda$, οπότε λαμβάνεται το σύνολο των τιμών του συντελεστή ομοιοθεσίας $I_\lambda = \{\lambda_A, \lambda_A + \delta\lambda, \lambda_A + 2\delta\lambda, \dots, \lambda_A + N_\lambda \cdot \delta\lambda \equiv \lambda_T\}$.
- $[\varphi_A, \varphi_T]$ εκατέρωθεν του φ και συμμετρικά ως προς αυτό. Το διάστημα αυτό υφίσταται διαμέριση με ένα βήμα $\delta\varphi$ και τότε λαμβάνεται το σύνολο των τιμών της γωνίας στροφής $I_\varphi = \{\varphi_A, \varphi_A + \delta\varphi, \varphi_A + 2\delta\varphi, \dots, \varphi_A + N_\varphi \cdot \delta\varphi \equiv \varphi_T\}$.
- $[dx_A, dx_T]$ εκατέρωθεν του dx και συμμετρικά ως προς αυτό. Το διάστημα αυτό υφίσταται διαμέριση με ένα βήμα δx και τότε λαμβάνεται το σύνολο των τιμών της μετατόπισης που γίνεται παράλληλα προς τον άξονα των x $I_{dx} = \{dx_A, dx_A + \delta x, dx_A + 2\delta x, \dots, dx_A + N_{dx} \cdot \delta x \equiv dx_T\}$.
- $[dy_A, dy_T]$ εκατέρωθεν του dy και συμμετρικά ως προς αυτό. Το διάστημα αυτό υφίσταται διαμέριση με ένα βήμα δy και τότε λαμβάνεται το σύνολο των τιμών της μετατόπισης που γίνεται παράλληλα προς τον άξονα των y $I_{dy} = \{dy_A, dy_A + \delta y, dy_A + 2\delta y, \dots, dy_A + N_{dy} \cdot \delta y \equiv dy_T\}$.

Εν συνεχεία, ορίζεται ένα τετραδιάστατο χωρίο μέσω του καρτεσιανού γινομένου $I = I_\lambda \times I_\varphi \times I_{dx} \times I_{dy}$.

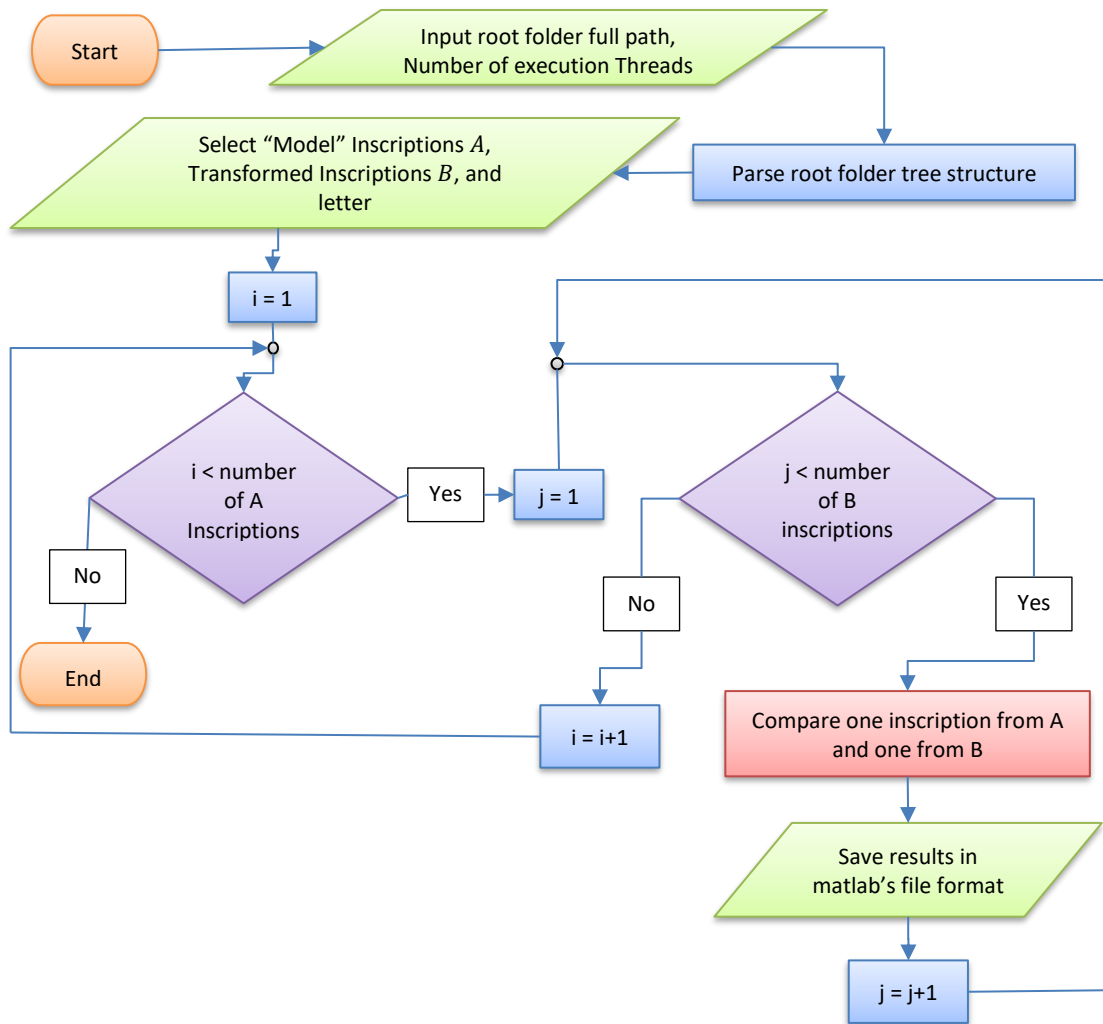
Ακολούθως, για κάθε σημείο $(\lambda_i, \varphi_j, dx_k, dy_m)$ μετατοπίζουμε αντίστοιχα την υλοποίηση $\Sigma A_2(\lambda, \varphi, dx, dy)$ και υπολογίζουμε το κριτήριο:

$$\mu(\lambda_i, \varphi_j, dx_k, dy_m) = \frac{EMB(\Sigma A_1 \cap \Sigma A_2(\lambda_i, \varphi_j, dx_k, dy_m))}{EMB(\Sigma A_1 \cup \Sigma A_2(\lambda_i, \varphi_j, dx_k, dy_m))} \quad (2.7)$$

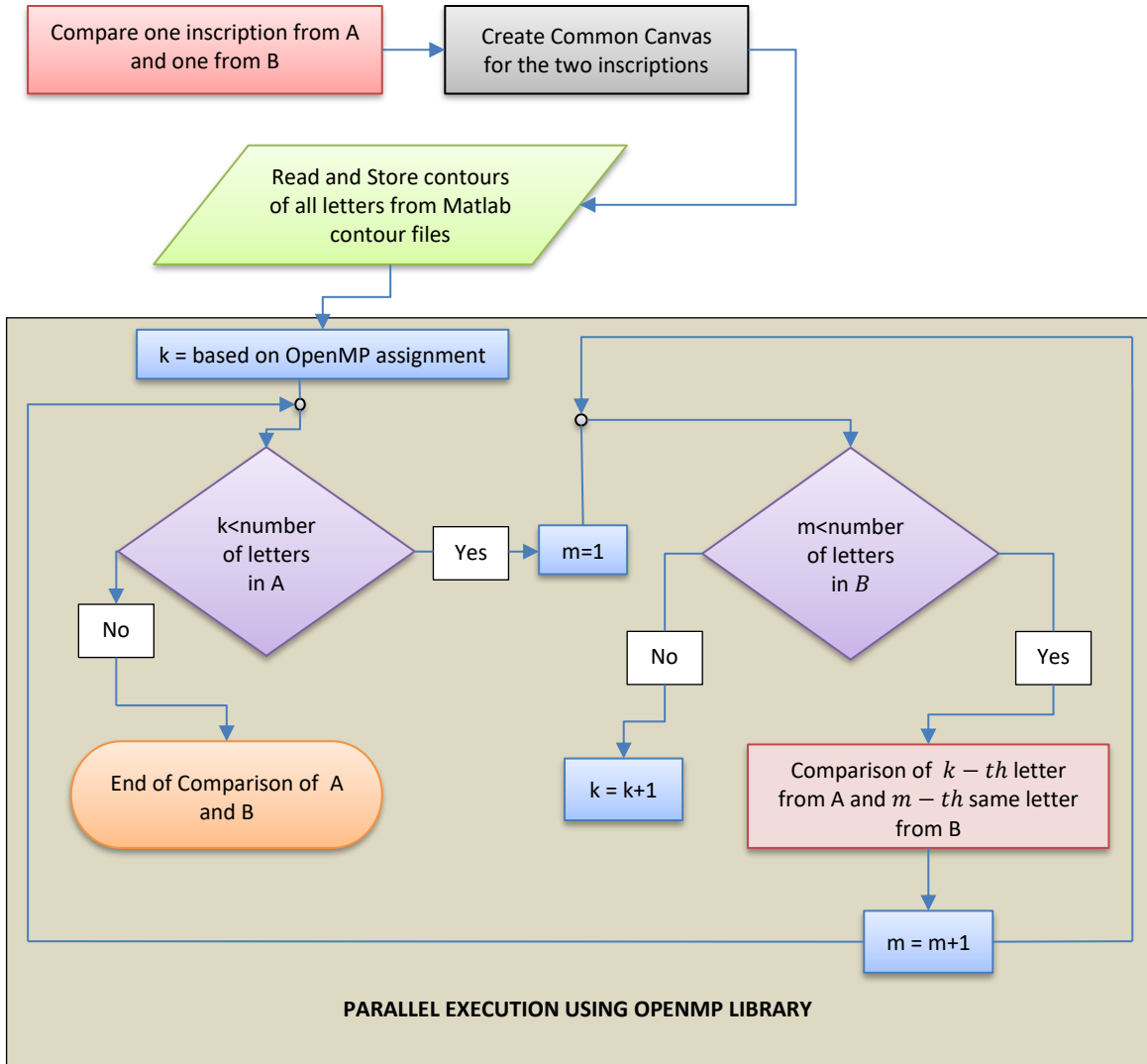
Εξ όλων αυτών των τιμών $\mu(\lambda_i, \varphi_j, dx_k, dy_m)$ επιλέγουμε τη μέγιστη μ^{max} και θεωρούμε απολύτως εύλογα και σε πλήρη συμφωνία με όλη τη προηγούμενη ανάλυση, ότι η τετράδα των τιμών $(\lambda_i^{max}, \varphi_j^{max}, dx_k^{max}, dy_m^{max})$ που παράγει αυτό το μέγιστο μ^{max} , είναι τελεσιδίκως η τετράδα των τιμών των αφινικών μετασχηματισμών που δίνει τη βέλτιστη προσαρμογή της υλοποίησης ΣA_2 στη ΣA_1 . Προφανώς, η ακριβής τιμή αυτού του μέγιστου μ^{max} φανερώνει και τον πραγματικό βαθμό ομοιότητας των δυο υλοποιήσεων των ΣA_2 και ΣA_1 του ίδιου τυχόντος συμβόλου της αλφαβήτου.

3 Αλγοριθμική υλοποίηση που επιτυγχάνει μεγάλη επιτάχυνση της διαδικασίας ταυτοποίησης γραφέα.

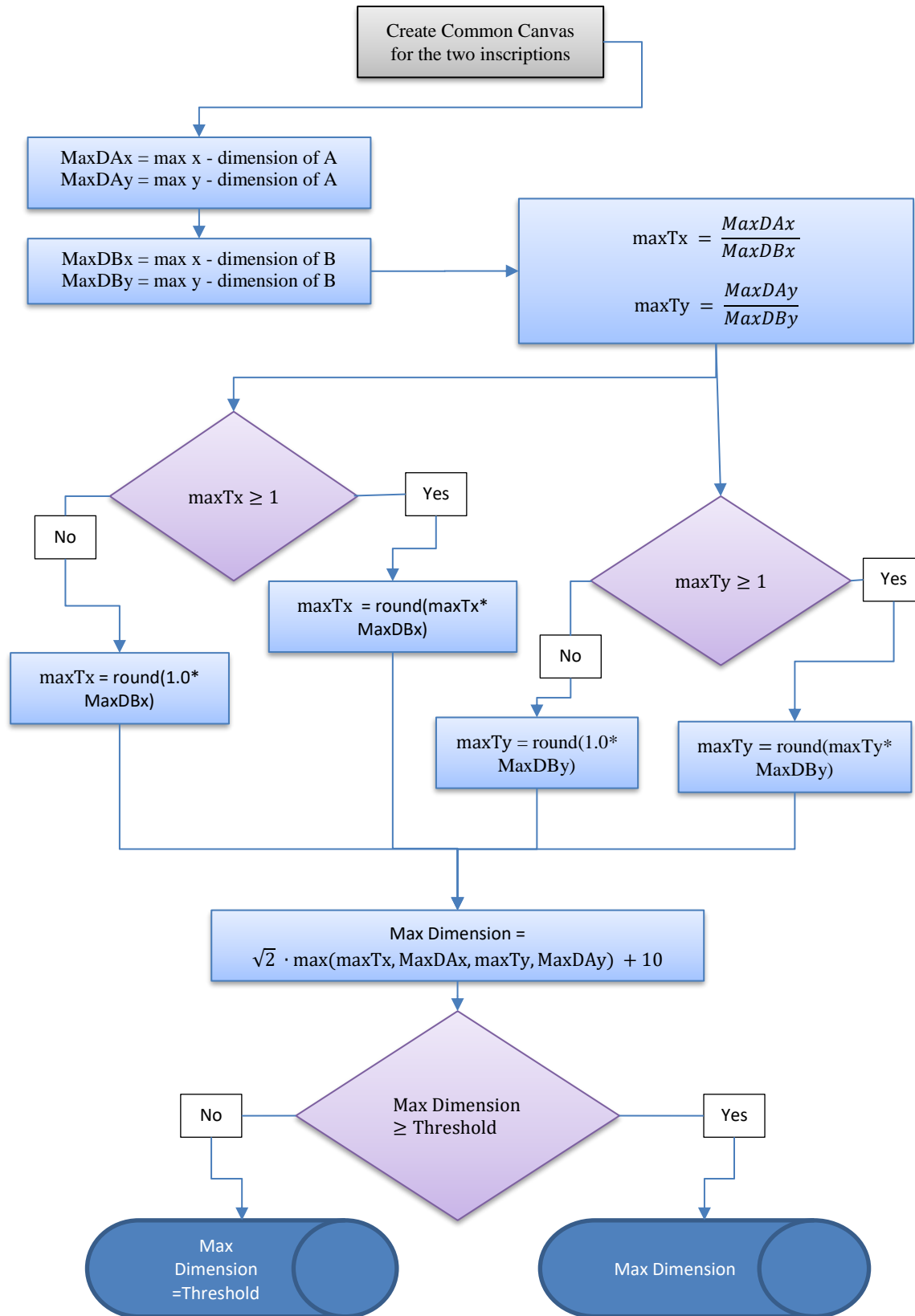
3.1 Διαγράμματα ροής της μεθόδου που παρουσιάζεται στη παρούσα εργασία.



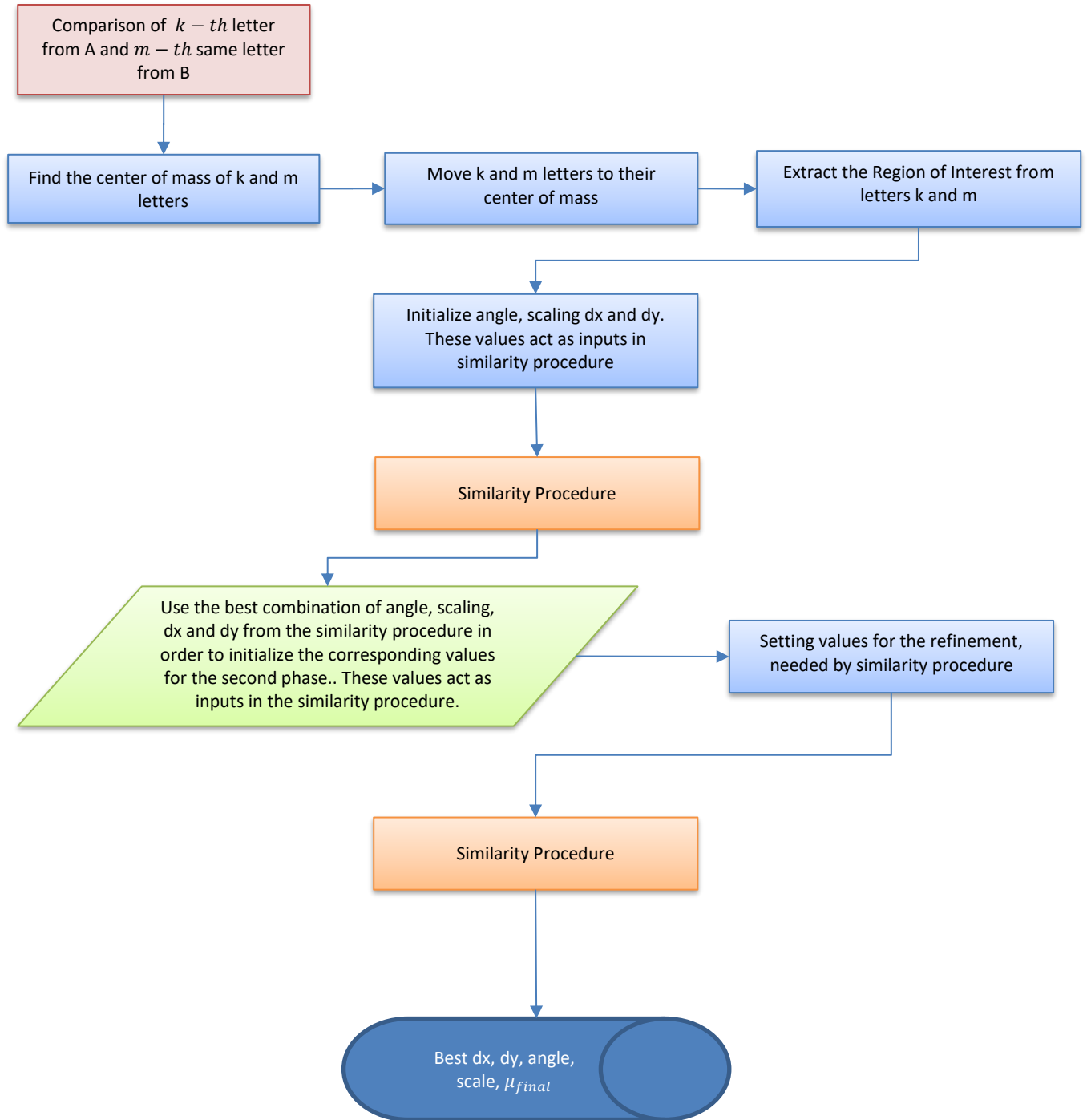
Εικόνα 3.1 Διάγραμμα ροής που αναπαριστά τον αλγόριθμο εκκίνησης της μεθόδου σύγκρισης δυο επιγραφών



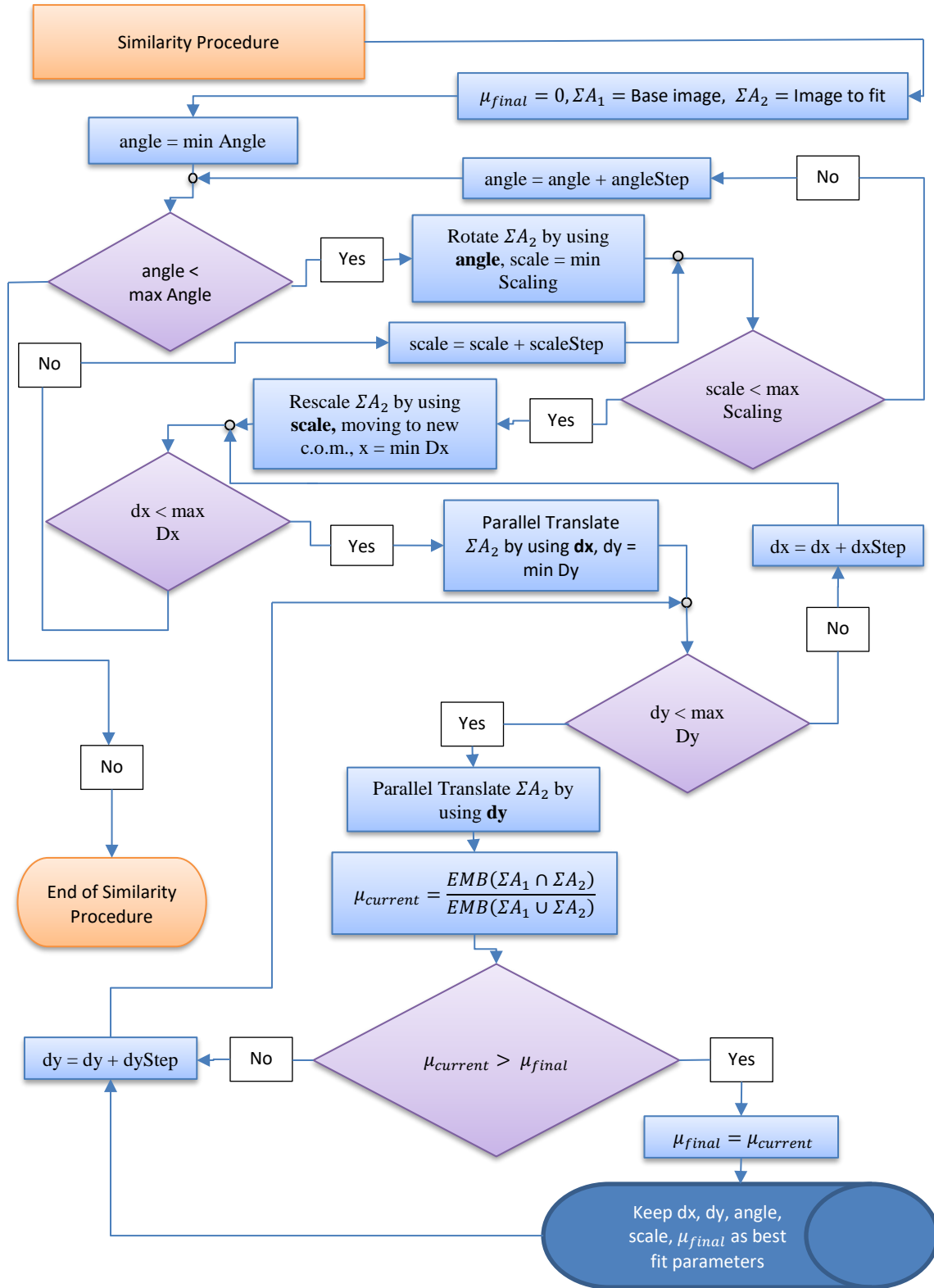
Εικόνα 3.2 Διάγραμμα ροής που αναπαριστά τον αλγόριθμο σύγκρισης της μεθόδου σύγκρισης δυο επιγραφών. Σε αυτό το σημείο γίνεται σύγκριση δυο γραμμάτων, του k -οστού από την A και του m -οστού ίδιου συμβόλου από τη B



Εικόνα 3.3 Καθορισμός των διαστάσεων του κοινού καμβά, στον οποίο θα τοποθετούνται τα γράμματα προς σύγκριση.



Εικόνα 3.4 Διάγραμμα ροής που δείχνει τη σύγκριση δυο γραμμάτων σε δύο φάσεις



Εικόνα 3.5 Διάγραμμα ροής του αλγορίθμου σύγκρισης δυο γραμμάτων.

3.2 Πλεονεκτήματα της επιλογής υλοποίησης του αλγορίθμου ταυτοποίησης γραφέα.

Η υλοποίηση του προαναφερθέντος διαγράμματος ροής αυτόματης αναγνώρισης γραφέα, πραγματοποιήθηκε με χρήση της γλώσσα προγραμματισμού C++ στο λειτουργικό σύστημα Linux (Linux Mint, Ubuntu, openSuse). Ως εκ τούτου ο αναπτυχθείς κώδικας παρουσίασε σημαντικά πλεονεκτήματα σε σχέση με ανάλογους κώδικες που είχαν αναπτυχθεί σε διερμηνείς (π.χ τον διερμηνέα γραμμής MATLAB). Συγκεκριμένα, ο νέος αυτός υλοποιημένος αλγόριθμος:

- a. Είναι πολύ ταχύς, δεκάδες φορές ταχύτερος από ένα αντίστοιχο σύνολο κωδίκων, αναπτυγμένο σε διερμηνέα. Ενδεικτικά αναφέρεται ότι η σύγκριση δυο κειμένων όταν εκτελείτο στο διερμηνέα MATLAB απαιτούσε χρόνο δυο περίπου εβδομάδων, μέχρι να προσφέρει ορθά αποτελέσματα, ενώ η υλοποίηση που παρουσιάστηκε εδώ προσφέρει πρακτικά τα ίδια αποτελέσματα το πολύ εντός 10 ωρών, για το ίδιο hardware. Για την επιτάχυνση της διαδικασίας αναγνώρισης γραφέα στη παρούσα προσέγγιση, έγινε εκμετάλλευση των δυνατοτήτων που προσφέρει η βιβλιοθήκη OpenMP.
- b. Κάνοντας χρήση της βιβλιοθήκης OpenCV, επετεύχθη βελτίωση της ακρίβειας των παραγομένων αποτελεσμάτων.
- c. Με εκμετάλλευση των αποτελεσμάτων της έρευνας που αφορά το λάθος πεπερασμένης ακρίβειας, τοποθετήθηκαν σημεία ελέγχου σε διάφορα σημεία του προγράμματος, ώστε να αποτραπεί κατά το δυνατόν η εμφάνιση αριθμητικά λανθασμένων αποτελεσμάτων. Αυτός ο έλεγχος και τα αντίστοιχα σημεία ελέγχου υλοποιήθηκαν με βάση τα αποτελέσματα του τμήματος Α της παρούσης εργασίας.

3.3 Πρώτα αποτελέσματα που ελήφθησαν.

Ο αλγόριθμος που υλοποιήθηκε εφαρμόστηκε πιλοτικά για την επιβεβαίωση ή μη των υλοποιήσεων του συμβόλου της αλφαβήτου “α” στα κείμενα “Φυσικής απάνθισμα” και “Το Σαγανάκι της τρέλας” με σκοπό να ελεγχθεί εάν αυτή η ομοιότητα είναι συμβατή ή όχι με την υπόθεση ότι το “α” έχει γραφεί από τον ίδιο γραφέα ή όχι με κάποια σημαντική πιθανοφάνεια. Τα σχετικά αποτελέσματα παρουσιάζονται στους παρακάτω πίνακες.

Μέση ομοιότητα των υλοποιήσεων του συμβόλου της αλφαβήτου “α” μεταξύ των σελίδων του κειμένου “Το Σαγανάκι της τρέλας”.				
		Σελίδες από το “Το Σαγανάκι της τρέλας”		
		Ι0026	Ι0034	Ι0035
Σελίδες από το “Το Σαγανάκι της τρέλας”	Ι0026	72.62%	65.50%	67.74%
	Ι0034	65.59%	66.87%	66,75%
	Ι0035	67.68%	66.54%	68.79%

Πίνακας 3.1

Μέση ομοιότητα των υλοποιήσεων του συμβόλου της αλφαβήτου “α” μεταξύ των σελίδων του κειμένου “Φυσικής Απάνθισμα”				
		Σελίδες από το “Φυσικής Απάνθισμα”		
		00031	00034	00091
Σελίδες από το “Φυσικής Απάνθισμα”	00031	71.31%	68.08%	68.09%
	00034	67.86%	68.60%	65.18%
	00091	68.44%	65.78%	68.92%

Πίνακας 3.2

Μέση ομοιότητα των υλοποιήσεων του συμβόλου της αλφαβήτου “α” έχοντας το κείμενο “Το Σαγανάκι της τρέλας” ως το πρότυπο κείμενο.				
		Σελίδες από το “Φυσικής Απάνθισμα”		
		00031	00034	00091
Σελίδες από το “Το Σαγανάκι της τρέλας”	Ι0026	67.64%	66.58%	67.41%
	Ι0034	65.17%	65.20%	63.49%
	Ι0035	66.99%	66.27%	65.47%

Πίνακας 3.3

Μέση ομοιότητα των υλοποιήσεων του συμβόλου της αλφαβήτου “α” έχοντας το κείμενο “Φυσικής Απάνθισμα” ως το πρότυπο κείμενο.				
		Σελίδες από το “Το Σαγανάκι της τρέλας”		
		Ι0026	Ι0034	Ι0035
Σελίδες από το “Φυσικής Απάνθισμα”	00031	68.39%	65.94%	68.09%
	00034	67.14%	65.75%	67.13%
	00091	68.45%	64.53%	66.81%

Πίνακας 3.4

Τα αποτελέσματα που περιλαμβάνονται στους ανωτέρω πίνακες υποδεικνύουν σημαντική πιθανοφάνεια τα μελετηθέντα κείμενα να έχουν γραφεί από το ίδιο χέρι, όσον αφορά το σύμβολο “α” της αλφαβήτου και μόνο. Η σχετική ολοκληρωμένη, επίσης πρωτότυπη, μελέτη, μαζί με την πλήρη στατιστική ανάλυση των σχετικών αποτελεσμάτων θα είναι αντικείμενο επομένων εργασιών.

4 Βιβλιογραφία C’ Μέρους

- [1] Φώτιος Γ. Γιαννόπουλος, “Πρωτότυπες Μέθοδοι Αυτόματης Αναγνώρισης Γραφέα και Αντιμετώπισης Υπολογιστικών Προβλημάτων Υλοποίησης,” 2018.
- [2] C. Papaodysseus *et al.*, “Identifying the writer of ancient inscriptions and Byzantine codices. A novel approach,” *Comput. Vis. Image Underst.*, vol. 121, pp. 57–73, 2014.

Γλωσσάριο Τεχνικών όρων

Ελληνικοί τεχνικοί όροι	English technical terms
αριθμητικό σφάλμα πεπερασμένης ακρίβειας	finite precision error
λάθος δεκαδικά ψηφία	erroneous decimal digits
σωστά δεκαδικά ψηφία	correct decimal digits
μήκος λέξης	finite word length
σφάλμα στρογγύλευσης	round-off error
νομοτελειακό λάθος	deterministic error
τυχαίο λάθος	stochastic / erratic error
διωνυμική κατανομή	binomial distribution
συσώρευση σφάλματος πεπερασμένης ακρίβειας	accumulation of finite precision error
διάδοση σφάλματος πεπερασμένης ακρίβειας	propagation of finite precision error
ευσταθειοποίηση	stabilization
εκθέτης	exponent
τάξης αριθμού	number order
αριστερή ολίσθηση αριθμού	left shift of a number
πολυώνυμα Zernike	Zernike polynomials
συντελεστές Zernike	Zernike coefficients / moments
τάξη Zernike	Zernike order
επεξεργασία εικόνας	image processing
εικονοστοιχείο	pixel
τετραγωνικό εικονοστοιχείο	square pixel
πολικό εικονοστοιχείο	polar pixel
Μέθοδοι ευθέως υπολογισμού των συντελεστών Zernike	direct methods of computation of Zernike moments
Αναδρομικοί μέθοδοι υπολογισμού των συντελεστών Zernike	recursive methods of Zernike moments computation
γεωμετρικό σφάλμα	geometric error
σφάλμα ολοκλήρωσης	numerical integration error
δικυβική παρεμβολή	bi-cubic interpolation

αναγνώριση γραφέα	writer identification
καμπυλότητα	curvature
ισούψεις	level sets

Συνοπτικό Βιογραφικό Σημείωμα

Ο Χαλάτσης Κωνσταντίνος γεννήθηκε το 1987 στην Αθήνα. Αποφοίτησε από το τμήμα Πληροφορικής και Τηλεπικοινωνιών του Εθνικού και Καποδιστριακού Πανεπιστημίου Αθηνών το 2010 και εκπόνησε τη διδακτορική του διατριβή στη Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών του Ε.Μ.Π. Τα ερευνητικά του ενδιαφέροντα συμπεριλαμβάνουν τους τομείς της Αναγνώρισης Προτύπων, τη Ψηφιακή επεξεργασία εικόνας, της μελέτης του Λάθους Πεπερασμένης Ακρίβειας, της Αυτόματης Αναγνώρισης Μουσικών Συνθέσεων και της Βιοϊατρικής Μηχανικής.

Δημοσιεύσεις

Διεθνή επιστημονικά περιοδικά με κριτές

1. C. Chalatsis, C. Papaodysseus, D. Arabadjis and F. Giannopoulos, "Finite Precision Error Analysis of Zernike Moments Computation Schemes and a New, Efficient, Robust Recursive Algorithm," *Digital Signal Processing*, vol. 79, pp. 75-90, August 2018.

Διεθνή επιστημονικά συνέδρια με κριτές

1. Papaodysseus, C. Chalatsis, D. Arabadjis, S. Zannos, F. Giannopoulos, P. Rousopoulos and E. Kalfa, "A General, Novel Methodology for Studying the Generation of Finite Precision Error in Any Algorithm," *Proceedings IEEE International Symposium on Signal Processing and Information Technology, ISSPIT*, 2013.

Υπό κρίση σε διεθνή επιστημονικά περιοδικά με κριτές

1. Chalatsis, C. Papaodysseus, D. Arabadjis, A. Mamatsis. "Exact analysis of the finite precision error generated in important chaotic maps and complete numerical remedy of these schemes." *Journal of Computational and Applied Mathematics Elsevier*.
2. Papaodysseus, C. Chalatsis, M.Exarchos, D. Arabadjis. "Exact tracking and the actual sources of the Finite Precision Error generated in any number of successive subtractions", *Mathematics and Computers in Simulation Elsevier*.