



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ  
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ  
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ  
ΤΟΜΕΑΣ ΤΕΧΝΟΛΟΓΙΑΣ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ  
ΥΠΟΛΟΓΙΣΤΩΝ

## Ταίριασμα διδιάστατων εικόνων με ανάλυση σχήματος

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Γιαννέκου Βασιλική

**Επιβλέπων:** Στέφανος Κόλλιας  
Καθηγητής

Αθήνα , Οκτώβριος 2006





ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ  
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ  
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ  
ΤΟΜΕΑΣ ΤΕΧΝΟΛΟΓΙΑΣ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ  
ΥΠΟΛΟΓΙΣΤΩΝ

## Ταίριασμα διδιάστατων εικόνων με ανάλυση σχήματος

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Γιαννέκου Βασιλική

**Επιβλέπων:** Στέφανος Κόλλιας  
Καθηγητής

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 2006

Κόλλιας Στέφανος  
Καθηγητής

Αθήνα , Οκτώβριος 2006

.....  
Γιαννέκου Βασιλική

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

Copyright © Γιαννέκου Βασιλική

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 το 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.

## ΠΕΡΙΛΗΨΗ

Τα τελευταία χρόνια έχει παρατηρηθεί μια ραγδαία αύξηση του μεγέθους των συλλογών ψηφιακών εικόνων. Κατά καιρούς, πολλές μέθοδοι έχουν προταθεί με σκοπό την εύκολη πρόσβαση και χρήση των πληροφοριών που αυτές οι ψηφιακές εικόνες μεταφέρουν. Πρόσφατες έρευνες έχουν στραφεί στην αποδοτική επεξεργασία, αναζήτηση και ανάκτηση ομοίων εικόνων από μια βάση δεδομένων.

Σε αυτήν την διπλωματική εργασία μελετάμε την απεικόνιση και ανάκτηση ομοίων εικόνων με την μέθοδο της *curvature scale space* (CSS) υπό την παρουσία affine μετασχηματισμών. Πιο συγκεκριμένα, εξετάζουμε η ανθεκτικότητα της μεθόδου στους affine μετασχηματισμούς και συγκρίνουμε την απόδοση της σε σχέση με άλλες εναλλακτικές μεθόδους.

Για να επιτευχθεί η ανθεκτικότητα, χρησιμοποιήσαμε κανονικοποίηση της καμπύλης με βάση το affine μήκος (affine length) και εκτιμήσαμε την αποτελεσματικότητα αυτής της εφαρμογής.

Σε πειραματικό επίπεδο δημιουργήθηκε μια βάση δεδομένων με τις καμπύλες (περιγράμματα) διαφορετικών κατηγοριών σχημάτων. Αρχικά, με την εφαρμογή τυχαίων affine μετασχηματισμών, δημιουργήσαμε έναν αριθμό affine μετασχηματισμένων εκδοχών των παραπάνω καμπυλών. Στη συνέχεια, με την βοήθεια της CSS μεθόδου έγινε αναπαράσταση και ανάκτηση ομοίων καμπυλών των σχημάτων από την βάση δεδομένων. Το κόστος ταιριάσματος που προέκυψε, αποτελεί τον μέτρο σύγκρισης ομοιότητας καμπυλών και δείχνει την αποτελεσματικότητα της μεθόδου.

Τέλος, μελετήθηκε η εφαρμογή μιας εναλλακτικής μεθόδου κανονικοποίησης στην θέση της κανονικοποίησης με βάση το affine μήκος, η οποία φάνηκε να βελτιώνει αισθητά την απόδοση της CSS μεθόδου.

**ΛΕΞΕΙΣ-ΚΛΕΙΔΙΑ:** curvature scale space, affine μήκος, affine μετασχηματισμός, κανονικοποίηση, ανάκτηση ομοίων εικόνων.

## ABSTRACT

Recent years have seen a rapid increase of the size of digital image collections. Occasionally, several methods have been proposed aiming at the easy access and use of the information that these digital images carry. Recent researches have focused on the efficient processing, searching and retrieval of similar images from a database.

In this diploma thesis we study the representation and retrieval of similar images based on the curvature scale space (CSS) method in the presence of affine transformations. More specifically, we examine the robustness of the method under affine transformations and compare its performance with the performance of other alternative methods.

In order to achieve invariance, we used curve normalization based on affine length parametrisation and evaluated the effectiveness of this application.

In experimental level, a database of curves (contours) of different categories of shapes has been constructed. Initially, by applying random affine transforms, we created a number of affine transformed versions of the above curves. Afterwards, we used the CSS method in order to represent and retrieve similar curves of shapes from the database. The resulted matching cost constitutes a measure of comparison of curve similarity and indicates the effectiveness of the method.

Finally, we study the application of an alternative normalization method instead of the affine length based normalization, which appear to improve perceptibly the effectiveness of the CSS method.

**KEYWORDS:** curvature scale space, affine length, affine transformation, normalization, shape similarity retrieval

# Περιεχόμενα

Περιεχόμενα.....	7
1.Εισαγωγή .....	8
2. Ανάκτηση εικόνων.....	10
2.1. Εισαγωγικά .....	10
2.2. Ανάκτηση εικόνας με βάση το κείμενο .....	11
2.3. Ανάκτηση εικόνας με βάση το περιεχόμενο .....	12
2.3.1 Οπτικά χαρακτηριστικά εικόνων .....	12
2.4 Απεικόνιση εικόνων με βάση το σχήμα.....	16
2.4.1 Σταθερές Moment .....	17
2.4.2 Περιγραφείς Fourier.....	19
2.5. Σύστημα Περιγραφής του Περιεχομένου Δεδομένων Πολυμέσων (MPEG – 7) ..	20
3. Affine μετασχηματισμοί .....	24
3.1.Εισαγωγή .....	24
3.2 Θεωρητικά στοιχεία.....	24
3.3. Μήκος τόξου – Affine μήκος.....	26
4.Μέθοδος της curvature scale space (CSS).....	29
4.1.Εισαγωγικά .....	29
4.2.Θεωρητική μελέτη της μεθόδου .....	30
4.2.1.Καμπυλότητα .....	30
4.2.2. Κανονικοποίηση .....	31
4.2.3. Ομαλοποίηση Καμπύλης .....	32
4.3. Κατασκευή της CSS εικόνας .....	35
4.3.1 Σημεία μηδενισμού .....	35
4.4. Εξαγωγή μεγίστων από την CSS εικόνα.....	37
4.5. Αλγόριθμος ταιριάσματος.....	39
5. Πειράματα, αποτελέσματα, παρατηρήσεις .....	42
5.1 Υλοποίηση της εφαρμογής .....	42
5.2 Πειράματα και αποτελέσματα.....	42
5.2.1 Παρουσίαση ενδεικτικών αποτελεσμάτων για την εφαρμογή της μεθόδου ....	42
5.2.2 Πειράματα με affine μετασχηματισμούς της ίδιας αρχικής καμπύλης της ίδιας κατηγορίας σχήματος.....	48
5.2.3 Πειράματα με affine μετασχηματισμούς διαφορετικής αρχικής καμπύλης της ίδιας κατηγορίας σχήματος .....	52
5.2.4 Πειράματα με affine μετασχηματισμούς διαφορετικής αρχικής καμπύλης διαφορετικής κατηγορίας σχήματος .....	53
5.3 Συμπεράσματα .....	54
6. Μελέτη της CSS μεθόδου σε σχέση με άλλες μεθόδους.....	56
6.1 Εισαγωγικά .....	56
6.2 Παρουσίαση εναλλακτικής μεθόδου κανονικοποίησης.....	56
6.3 Προσαρμογή της εναλλακτικής μεθόδου κανονικοποίησης στην CSS μέθοδο – πειράματα, αποτελέσματα και παρατηρήσεις.....	57
6.4 Συμπεράσματα σύγκρισης των δύο μεθόδων – μελλοντικές επεκτάσεις .....	68
Βιβλιογραφία .....	71

# 1.Εισαγωγή

Η μεγάλη ανάπτυξη που σημειώνει η τεχνολογία στην διάρκεια των τελευταίων ετών, έχει καταστήσει τα επιτεύγματα της απαραίτητα στοιχεία της καθημερινότητας στους περισσότερους τομείς της ζωής μας. Κάθε πληροφορία του περιβάλλοντος κωδικοποιείται, αναλύεται και επεξεργάζεται με σκοπό να χρησιμοποιηθεί όσο το δυνατόν καλύτερα από τον άνθρωπο. Ιδιαίτερα στον τομέα της επικοινωνίας, η χρήση του διαδικτύου (internet) έχει αλλάξει ριζικά τον τρόπο διακίνησης και διαχείρισης της υπάρχουσας πληροφορίας, δίνοντας ταυτόχρονα απεριόριστες δυνατότητες εφαρμογών στον εκάστοτε χρήστη.

Ένα πολύ μεγάλο μέρος της πληροφορίας αυτής μεταφέρεται μέσω ψηφιακών εικόνων. Απαραίτητη επομένως κρίνεται η ανάγκη για εύρεση μεθόδων κατάλληλων να αναλύσουν, να επεξεργαστούν και τελικά να αποθηκεύσουν την πληροφορία των ψηφιακών εικόνων. Πολλές τέτοιες προσπάθειες έχουν γίνει κατά καιρούς με αποτέλεσμα την δημιουργία ενός ερευνητικού πεδίου, που ασχολείται αποκλειστικά με την επεξεργασία των εικόνων και τις εφαρμογές τους. Η όραση των υπολογιστών και η διαχείριση βάσεων δεδομένων αποτελούν συστατικά μέρη του ερευνητικού αυτού πεδίου, που προσεγγίζουν η κάθε μια διαφορετικά την επεξεργασία των ψηφιακών εικόνων.

Στη παρούσα διπλωματική εργασία παρουσιάζεται μια μέθοδος ανάλυσης και επεξεργασίας εικόνας βασισμένη στο σχήμα της και πιο συγκεκριμένα στο περίγραμμα της εικόνας. Η μέθοδος αυτή μπορεί να βρει εφαρμογή σε περιπτώσεις αναζήτησης και ανάκτησης ομοίων εικόνων από μια βάση δεδομένων, διαδικασία που συχνά απαιτείται για την εύρεση πληροφοριών στο διαδίκτυο. Η μέθοδος της *curvature scale space* (CSS) απεικόνισης αποτελεί μια από τις πιο αξιόπιστες μεθόδους ανάκτησης ομοίων εικόνων και μάλιστα είναι αυτή που χρησιμοποιείται από το σύστημα MPEG-7 για την πρότυπη περιγραφή του περιγράμματος του σχήματος μιας εικόνας.

Η δομή της εργασίας περιλαμβάνει 6 κεφάλαια στα οποία παρουσιάζεται και μελετάται η ανάκτηση εικόνων, η επίδραση των affine μετασχηματισμών σε αυτές, καθώς και η συμπεριφορά της CSS μεθόδου στην επεξεργασία τους.

Πιο συγκεκριμένα, στο κεφάλαιο 2 γίνεται μια αναφορά στην θεωρία της ανάκτησης εικόνων αλλά και στις ήδη υπάρχουσες μεθόδους που την υλοποιούν. Επιπλέον, όσον αφορά την ανάκτηση εικόνας με βάση το περιεχόμενο, γίνεται μια εκτενέστερη αναφορά στα οπτικά χαρακτηριστικά της εικόνας που είναι το χρώμα, η υφή, η κίνηση και το σχήμα. Ιδιαίτερα για το χαρακτηριστικό του σχήματος, στο οποίο βασίζεται και η CSS μέθοδος, παρουσιάζονται οι βασικότερες μέθοδοι επεξεργασίας του. Στο τέλος του κεφαλαίου, γίνεται μια αναφορά στο σύστημα MPEG-7, το οποίο πρακτικά υλοποιεί την παραπάνω θεωρητική ανάλυση.

Στη συνέχεια, στο κεφάλαιο 3 παρουσιάζεται μια θεωρητική μελέτη των affine μετασχηματισμών, καθώς επίσης και μια συνοπτική ανάλυση του μήκους τόξου και του affine μήκους.

Στο κεφάλαιο 4 παρουσιάζεται το θεωρητικό υπόβαθρο της CSS μεθόδου. Ειδικότερα, γίνεται μια ανάλυση ορισμένων βασικών θεωρητικών παραμέτρων, όπως η



καμπυλότητα, η κανονικοποίηση και η ομαλοποίηση της καμπύλης. Κατόπιν, παρουσιάζεται ο τρόπος κατασκευής της CSS εικόνας και ο αλγόριθμος ταιριάσματος.

Έπειτα, στο κεφάλαιο 5 παρουσιάζεται η υλοποίηση της μεθόδου καθώς και τα αποτελέσματα των πειραμάτων που εκτελέστηκαν με την εφαρμογή της. Τα πειράματα αφορούν σε καμπύλες και affine μετασχηματισμούς αυτών, που ανήκουν στην ίδια ή και διαφορετική κατηγορία σχήματος, ώστε να εξεταστεί κατά πόσο η μέθοδος είναι σε θέση να ταιριάζει όμοιες εικόνες από μια βάση δεδομένων.

Τέλος, στο κεφάλαιο 6 προτείνουμε την αντικατάσταση της affine length κανονικοποίησης με μια εναλλακτική μέθοδο κανονικοποίησης και παρουσιάζουμε αντίστοιχα πειράματα και τα αποτελέσματά τους. Δημιουργούμε με αυτόν τον τρόπο μια τροποποιημένη εκδοχή της CSS μεθόδου, η οποία φαίνεται να βελτιώνει την απόδοση της ήδη υπάρχουσας.

Η διπλωματική εργασία βασίστηκε σε μια μελέτη των F. Mokhtarian και S. Abbasi [1], οι οποίοι χρησιμοποίησαν την μέθοδο της curvature scale space αναπαράστασης για την ανάκτηση ομοίων εικόνων από μια βάση δεδομένων. Πάνω σε αυτήν την θεωρητική μελέτη, υλοποιήσαμε σε γλώσσα προγραμματισμού Matlab την CSS μέθοδο και πειραματιστήκαμε πάνω στην ανάκτηση και το ταίριασμα ομοίων εικόνων από μια ήδη υπάρχουσα βάση δεδομένων. Ως μέτρο αυτού του ταιριάσματος ορίσαμε το κόστος ταιριάσματος που προέκυψε εφαρμόζοντας τον προτεινόμενο αλγόριθμο ταιριάσματος. Καταλήξαμε έτσι σε συμπεράσματα για την αποτελεσματικότητα της CSS μεθόδου και την αξιοπιστία της σε πειραματικό επίπεδο. Αξιολογώντας τα παραπάνω, προσαρμόσαμε μια προτεινόμενη μέθοδο κανονικοποίησης στην ήδη υπάρχουσα CSS μέθοδο και εκτελέσαμε τα αντίστοιχα πειράματα. Μπορέσαμε επομένως να συγκρίνουμε τις δύο μεθόδους ως προς την απόδοσή τους και να διαπιστώσουμε την βελτίωση της μεθόδου με την εφαρμογή της εναλλακτικής μεθόδου κανονικοποίησης.

## 2. Ανάκτηση εικόνων

### 2.1. Εισαγωγικά

Τα τελευταία χρόνια, όλο και συχνότερα, οι ψηφιακές εικόνες βρίσκουν εφαρμογές σε συστήματα πολυμέσων, όπως οι προσωπικοί ηλεκτρονικοί υπολογιστές αλλά και ο χώρος του διαδικτύου (Internet). Οι χρήστες συστημάτων πολυμέσων έχουν πολύ συχνά μια αφηρημένη έννοια του τύπου των εικόνων που αναζητούν, για αυτό το λόγο παρουσιάστηκε η επιτακτική ανάγκη για ανάκτηση και αναζήτηση των εικόνων με βάση τα σημασιολογικά τους χαρακτηριστικά.

Επιπλέον, ο συνεχώς αυξανόμενος αριθμός ψηφιακών εικόνων, οδήγησε σε μια ανάλογη απαίτηση για αποτελεσματική διαχείριση και αποθήκευση τους σε βάσεις δεδομένων, ώστε να είναι πιο εύκολη η πρόσβαση σε αυτές. Για το λόγο αυτό αναπτύχθηκαν κατάλληλες μέθοδοι, που επιτρέπουν την ανάκτηση και επεξεργασία πληροφοριών από τις ψηφιακές εικόνες, με τέτοιο τρόπο ώστε η πληροφορία να μεταφέρεται όσο το δυνατόν αυτούσια στον χρήστη.

Η ανάκτηση εικόνων από μια βάση δεδομένων έχει υπάρξει αντικείμενο έρευνας ήδη από το 1970, με την βοήθεια δύο μεγάλων ερευνητικών κοινοτήτων, της Διαχείρισης Βάσης Δεδομένων (Database Management) και της Όρασης Υπολογιστών (Computer Vision). Οι δύο αυτοί τομείς μελετούν το πρόβλημα της ανάκτησης εικόνων από διαφορετική οπτική γωνία, μία βασισμένη στο κείμενο (*text-based*) και μία βασισμένη στα οπτικά χαρακτηριστικά της εικόνας (*visual-based*).[2]

Αρχικά μελετήθηκε η *βασισμένη στο κείμενο* ανάκτηση εικόνων, της οποίας τη θέση πήρε στις αρχές του 1990, η ανάκτηση εικόνων *με βάση το περιεχόμενο* – οπτικά χαρακτηριστικά – της εικόνας.

Η αναζήτηση και ανάκτηση των εικόνων *με βάση το κείμενο* στηρίζεται στην ιδέα της απόδοσης των χαρακτηριστικών τους με λέξεις κλειδιά (*δεικτοδότηση*). Οι λέξεις αυτές δίνονται από τον ειδικό που τις επεξεργάζεται, και αποθηκεύονται σαν δεδομένα ώστε να είναι δυνατή η ανάκτησή τους κάθε φορά από το χρήστη. Η δεικτοδότηση, παρά το γεγονός ότι γίνεται από κάποιον ειδικό σχολιαστή, έχει το στοιχείο της υποκειμενικότητας. Επιπλέον, η οπτική του σχολιαστή μπορεί να διαφέρει από αυτή του χρήστη. Από έρευνες που έχουν διεξαχθεί προέκυψαν μεγάλες διαφορές στις λέξεις κλειδιά που διαφορετικά άτομα προσέδωσαν στην ίδια εικόνα.

Εξαιτίας της όξυνσης αυτών των προβλημάτων, υπήρξε από τις αρχές της δεκαετίας του 1990 μια έντονη ερευνητική δραστηριότητα για την εύρεση και εφαρμογή πιο αποτελεσματικών μεθόδων και τεχνικών ανάκτησης εικόνας από την παραδοσιακή, δηλαδή με βάση τις λέξεις κλειδιά. Η ερευνητική αυτή προσπάθεια είχε ως αποτέλεσμα την ανάκτηση εικόνας με βάση το περιεχόμενο.

Ανάκτηση εικόνας με *βάση το περιεχόμενο* (*content-based image retrieval*) ονομάζεται η τεχνική για την ανάκτηση εικόνας από μια ψηφιακή βιβλιοθήκη ή τον παγκόσμιο ιστό με τη χρήση χαρακτηριστικών στοιχείων που εξάγονται αυτόματα με βάση το οπτικό της περιεχόμενο. Πιο συγκεκριμένα γίνεται εξαγωγή και ανάλυση χαρακτηριστικών - γνωρισμάτων από την ίδια την εικόνα, δηλαδή οι εικόνες δεικτοδοτούνται αυτόματα (*image indexing*) με τη χρήση αλγόριθμων επεξεργασίας

εικόνας και μετατρέπονται σε μαθηματικά δεδομένα και στατιστικά στοιχεία ώστε να είναι μετρήσιμα και συγκρίσιμα μεγέθη. Αυτά, με την σειρά τους, χρησιμοποιούνται με τη βοήθεια του υπολογιστή για την ανάκτηση αλλά και την αναζήτηση ομοίων εικόνων.

Ακολουθεί μια παρουσίαση των μεθόδων ανάκτησης εικόνας με βάση το κείμενο και με βάση το περιεχόμενο της, καθώς και μια σύντομη αξιολόγηση αυτών.

## 2.2. Ανάκτηση εικόνας με βάση το κείμενο

Η ανάκτηση εικόνας με βάση το κείμενο (*text – based image retrieval TBIR*) εξαρτάται από την ανθρώπινη ικανότητα να προσθέσει σχόλια σε μια εικόνα, και να τα χρησιμοποιήσει σε μια ανάλογη έρευνα.[3, 4]

Τα σχόλια αυτά μπορούν να περιγράψουν το περιεχόμενο μιας εικόνας σε διαφορετικά επίπεδα, περιλαμβάνοντας τόσο τα αντικείμενα στην εικόνα, όσο και την σημασία που αυτά μπορούν να έχουν μέσα σε αυτή. Γι αυτό το λόγο, η μέθοδος TBIR βρίσκει εφαρμογή στην ανάκτηση εικόνων βασισμένη στην σημασιολογία της.

Παρόλα αυτά, η TBIR παρουσιάζει δύο βασικά μειονεκτήματα:

- Τα σχόλια που αποδίδονται σε μια εικόνα μπορεί να είναι προκατειλημμένα, καθώς αυτά αντιπροσωπεύουν την προσωπική άποψη του σχολιαστή για την εικόνα, βασισμένα στο προσωπικό του λεξιλόγιο.
- Ο σχολιασμός μιας εικόνας είναι μια κοπιαστική και χρονοβόρα εργασία, και, ως επί το πλείστον, περιλαμβάνει μόνο ένα μικρό υποσύνολο των πιθανών σημασιολογικών ερμηνειών μιας εικόνας.

Έτσι, ο σχολιαστής και ο χρήστης που αναζητά μια εικόνα, μπορεί να χρησιμοποιούν κάθε φορά διαφορετικούς όρους για να περιγράψουν το ίδιο αντικείμενο ή κατάσταση μίας εικόνας, ή μπορεί να ερμηνεύσουν διαφορετικά το περιεχόμενο της.

Με σκοπό την μείωση του χρόνου και της προσπάθειας σχολιασμού, έγιναν απόπειρες αυτοματοποίησης της διαδικασίας του σχολιασμού των εικόνων. [5, 6] Παρόλα αυτά, οι απόπειρες αυτές κατάφεραν μόνο να μειώσουν το βάρος του σχολιασμού, χωρίς να καταφέρουν όμως να αυτοματοποιήσουν πλήρως την διαδικασία σχολιασμού.

Για την επίλυση του παραπάνω προβλήματος, εισήχθη η έννοια της ανάκτησης εικόνων με βάση το περιεχόμενο (*content – based image retrieval, CBIR*), που υποστήριζε την ανάκτηση εικόνων σε περιπτώσεις όπου η ανάκτηση τους με βάση το κείμενο δεν ήταν δυνατή ή δεν είχε πρακτική εφαρμογή. Η CBIR μέθοδος υποστηρίζει αυτόματη καταγραφή των βασικών οπτικών χαρακτηριστικών μιας εικόνας, όπως το χρώμα, το σχήμα και η υφή, αλλά υστερεί στην καταγραφή της σημασιολογίας της, όπως έκανε η TBIR μέθοδος.

Έτσι, πρόσφατα, έγιναν προσπάθειες για συνδυασμό των δύο αυτών μεθόδων ανάκτησης, που οδήγησαν στην ανάπτυξη νέων συστημάτων ανάκτησης εικόνων, όπως το “Query by image” (IBM’s QBIC) σύστημα, σύμφωνα με το οποίο έγχρωμα σχέδια και λέξεις – κλειδιά χρησιμοποιούνται για αναζήτηση και ανάκτηση ομοίων εικόνων. [7].

Στην επόμενη ενότητα θα ασχοληθούμε αναλυτικά με την μελέτη της μεθόδου ανάκτησης με βάση το περιεχόμενο, καθώς πάνω σε αυτή έχει στραφεί μεγάλο μέρος των ερευνών για την αναζήτηση και ανάκτηση εικόνων.

## 2.3. Ανάκτηση εικόνας με βάση το περιεχόμενο

### 2.3.1 Οπτικά χαρακτηριστικά εικόνων

Όπως προαναφέρθηκε, η ανάκτηση εικόνων με βάση το περιεχόμενο εφαρμόζει την εξαγωγή οπτικών γνωρισμάτων για να χαρακτηρίσει μια εικόνα.

Τα χαρακτηριστικά αυτά κατηγοριοποιούνται γενικά σε *ολικά* (*global*) και *τοπικά* (*local*) χαρακτηριστικά. Τα μεν είναι αυτά που αφορούν και υπολογίζονται για το σύνολο της εικόνας, τα δε είναι αυτά που εξάγονται από κάποιο μέρος ή τμήμα της εικόνας (π.χ. εξαγωγή χαρακτηριστικών σχετικά με το χρώμα για ένα σχήμα ή ένα τμήμα της εικόνας).

Πιο συγκεκριμένα, οι κυριότεροι τύποι χαρακτηριστικών στοιχείων που εξάγονται από μια εικόνα είναι οι εξής:

#### ➤ Χρώμα (colour)

Η μέτρηση του χρώματος γίνεται με την αναπαράσταση του ως τρισδιάστατο διάνυσμα σε κατάλληλους διανυσματικούς χώρους. Σύμφωνα με την θεωρία της *τριχρωμίας* (trichromacy) [8], κάθε άγνωστο χρώμα μπορεί να παρασταθεί ως γραμμικός συνδυασμός τριών βασικών χρωμάτων που είναι γραμμικά ανεξάρτητα μεταξύ τους. Στην τεχνολογία ψηφιακών εικόνων αρκετά διαδεδομένος είναι ο χώρος *RGB*, που προκύπτει χρησιμοποιώντας μονοχρωματικά πρωταρχικά χρώματα, συγκεκριμένα κόκκινο *R*, πράσινο *G* και μπλε *B*.

Ο χρωματικός χώρος *RGB* καθώς και διάφορες παραλλαγές του (*XYZ*, *UVW*) που προκύπτουν μέσω γραμμικών μετασχηματισμών, παρουσιάζουν αρκετά μαθηματικά πλεονεκτήματα αλλά δεν αντιστοιχούν καλά στην ανθρώπινη αντίληψη χρώματος. Για αυτό το λόγο, προτάθηκε η ιδέα μη γραμμικών χώρων που προσεγγίζουν καλύτερα την ανθρώπινη αντίληψη. Από ένα μη γραμμικό μετασχηματισμό του χώρου *RGB*, προκύπτει ο μη γραμμικός χώρος *HSI*. Σύμφωνα με αυτόν, οι συνιστώσες της οπτικής αντίληψης *έγχρωμων* εικόνων είναι η *χροιά* (*Hue*), ο *κορεσμός* (*Saturation*) και η *φωτεινότητα* – *ένταση* (*brightness – Intensity*). Οι αντίστοιχες εξισώσεις μετατροπής από συνιστώσες *RGB* σε συνιστώσες *HSI* είναι:

$$I = (R + G + B) / 3$$

$$S = 1 - \min(R, G, B) / I$$

$$\theta = \cos^{-1} \left\{ \frac{0.5[(R - G) + (R - B)]}{[(R - G)^2 + (R - B)(G - B)]^{1/2}} \right\}$$

$$H = \begin{cases} \theta / 360, & G \geq B \\ 1 - \theta / 360, & G \leq B \end{cases}$$

Σε αρκετές εφαρμογές ανάκτησης εικόνων με βάση το χρώμα, τα χρωματικά χαρακτηριστικά μιας εικόνας μπορούν να αναπαρασταθούν από μια πιθανοτική κατανομή των χρωμάτων της εικόνας. Έτσι, το *χρωματικό ιστόγραμμα* (*histogram*) παραμένει η πιο δημοφιλής αναπαράσταση των χρωματικών κατανομών, καθώς είναι αμετάβλητο σε μικρές αλλαγές του αντικειμένου και είναι εύκολο να υπολογισθεί. Παρόλα αυτά, δεν είναι πολύ αποδοτικό εξαιτίας της μεγάλης του απαίτησης χώρου στην μνήμη.

Το παραπάνω πρόβλημα οδήγησε στην δημιουργία νέων μεθόδων ανάκτησης με βάση το χρώμα, όπως η χρήση των *coarser ιστογραμμάτων* [9], των *κυρίαρχων χρωμάτων* (*dominant colours*) ή των *signature colours*, [10, 11] καθώς και εφαρμογές μεθόδων συμπίεσης επεξεργασίας σήματος όπως ο *Karhunen – Loeve μετασχηματισμός*, ο *διακριτός μετασχηματισμός συνημίτονου*, οι *μετασχηματισμοί Hadamard και Haar* και άλλοι. [12, 13, 14]

Μια άλλη προσέγγιση στην ανάκτηση εικόνων με βάση το χρώμα, είναι ο συνδυασμός του με χωρικά χαρακτηριστικά γνωρίσματα. Προκύπτει έτσι μια συνδυαστική μέθοδος, χρήσιμη σε περιπτώσεις όπου αναζητάμε ενός συγκεκριμένου χρώματος αντικείμενο, που επιπλέον βρίσκεται σε μια συγκεκριμένη θέση μέσα στην εικόνα. Η μέθοδος αυτή αποδίδει ικανοποιητικά στην ανάκτηση ομοίων εικόνων, βασισμένη σε μία ήδη γνωστή εικόνα. Έχουν υπάρξει δύο σημαντικές προσπάθειες να συνδυαστούν χρωματικές και χωρικές πληροφορίες, πρώτα από τον Gong [15], και κατόπιν από τον Hsu [16], που όμως παρουσιάζουν μειονεκτήματα καθώς περιορίζονται από την μέθοδο των ιστογραμμάτων που χρησιμοποιούν.

### ➤ Υφή (*texture*)

Η υφή αναφέρεται στις περιοχές μιας εικόνας που παρουσιάζουν ομοιογένεια ή ανομοιογένεια αντίστοιχα, κάτι που προσδίδει στην εικόνα πολυχρωμία ή περιοχές «έντονου» χρωματισμού. Περιέχει σημαντικές πληροφορίες δομής των επιφανειών, καθώς και πληροφορίες σε σχέση με το περιβάλλον τους.

Αρχικά, αναπτύχθηκαν *γεωμετρικές και δομικές μέθοδοι* για την ανάλυση της υφής. Οι Julesz και Bergen [17], έκαναν μια έρευνα σχετικά με τους μηχανισμούς αντίληψης της υφής από τον ανθρώπινο οργανισμό. Πιο συγκεκριμένα, ασχολήθηκαν με το πώς το ανθρώπινο οπτικό σύστημα αντιλαμβάνεται την υφή ύστερα από στιγμιαία ή συνεχή παρατήρηση, με το κατά πόσο τα γεωμετρικά χαρακτηριστικά του αντικειμένου την επηρεάζουν, και ποια είναι τα βασικά χαρακτηριστικά της αντίληψης της υφής. Οδηγήθηκαν έτσι στο συμπέρασμα, ότι η απαραίτητη πληροφορία για τον διαχωρισμό της αντίληψης της υφής μεταφέρεται μέσω τοπικών διακεκριμένων χαρακτηριστικών, τα οποία και ονόμασαν *textons*. Για αυτό το λόγο όρισαν μια σειρά κανόνων, τους *heuristic κανόνες* [18], που σχετίζονται με τα *textons* και είναι πολύ σημαντικοί για την διάκριση της υφής με βάση τα *textons*.

Στην συνέχεια η έρευνα στράφηκε στον χώρο των *στατιστικών μεθόδων*. Έτσι, η *διδιάστατη γραμμική πρόβλεψη* (*2D linear prediction*) έχει ως στόχο να προβλέψει τοπικά τις τιμές μιας εικόνας, βρίσκοντας τον καλύτερο γραμμικό συνδυασμό τους με γειτονικές τιμές. Το μοντέλο αυτό πρόβλεψης βρίσκει εφαρμογή στην κωδικοποίηση, κατάταξη και κατάτμηση εικόνων. [19]

Αργότερα, εξετάστηκαν *ενεργειακές μέθοδοι* με τους Bovic και Havlicek [20] να προτείνουν το μοντέλο *AM – FM*, σύμφωνα με το οποίο η υφή μιας εικόνας μπορεί να εκφραστεί ως ένα διδιάστατο *AM – FM* σήμα της μορφής:

$$f(x, y) = a(x, y) \cos[\varphi(x, y)]$$

όπου το διδιάστατο συνημίτονο περιέχει και την κατά πλάτος (*AM*) και την κατά συχνότητα (*FM*) διαμόρφωση.

Η διαφοροποίηση στην χρήση του *AM – FM* μοντέλου σε σχέση με άλλες μεθόδους, είναι ότι δεν εφαρμόζεται στην εικόνα, αλλά σε μια εκδοχή της, φιλτραρισμένη από 2 *-D Gabor φίλτρα* [21]. Το πλεονέκτημα τους είναι ότι προσεγγίζουν ικανοποιητικά τα πρώτα στάδια φιλτραρίσματος που χρησιμοποιεί η ανθρώπινη όραση.

### ➤ **Κίνηση (motion)**

Το στοιχείο της κίνησης είναι πολύ σημαντικό όσον αφορά τις ακολουθίες κινούμενων εικόνων (*video*), αφού μπορεί να μεταφέρει πληροφορίες απαραίτητες για την ανάκτηση τους.

Η ανάλυση της κίνησης αποτελεί ένα σημαντικό στοιχείο έρευνας στην όραση υπολογιστών, καθώς αποσκοπεί στην εξαγωγή πληροφοριών για μια κινούμενη εικόνα.

Κάθε φορά που η κάμερα βρίσκεται σε σχετική κίνηση με το οπτικό της περιβάλλον, ένα κινούμενο πεδίο φωτεινότητας επιδρά στην εικόνα της κάμερας. Προκύπτει έτσι ένα διδιάστατο διανυσματικό πεδίο – χώρος, της ταχύτητας της προβολής των αντικειμένων πάνω στην εικόνα που ονομάζουμε *οπτική ροή (optical flow)*. [22]

Η οπτική ροή αναπαρίσταται από ένα 2D χωροχρονικό διανυσματικό πεδίο ( $u, v$ ) όπου  $u, v$  οι ταχύτητες στην  $x, y$  διεύθυνση αντίστοιχα. Για διακριτές χρονικές ακολουθίες εικόνων, μια μέση τιμή των ταχυτήτων μπορεί να δοθεί από το *διανυσματικό πεδίο μετατόπισης* (με περίοδο δειγματοληψίας  $\Delta t$ ) του οποίου τα διανύσματα καθορίζουν την αντιστοιχία των σημείων των διαδοχικών πλαισίων της εικόνας [23]

Σε κάποιες περιπτώσεις, το διανυσματικό πεδίο ταχυτήτων μιας εικόνας μπορεί να μην αντιστοιχεί με ακρίβεια στο πραγματικό 2D πεδίο κίνησης, που έχει προκύψει από την προβολή της 3D κίνησης πάνω στην εικόνα. Παρόλα αυτά, εξαιτίας των πληροφοριών που περιέχει και της εύκολης πρόσβασης σε αυτές, η μέτρηση του διανυσματικού πεδίου ταχυτήτων μιας εικόνας είναι πολύ σημαντική στην ανάλυση της κίνησης.

Η ανάλυση της οπτικής κίνησης βρίσκει εφαρμογή σε αρκετούς τεχνολογικούς τομείς όπως:

- *Βιομηχανική αυτοματοποίηση*: dynamic monitoring of industrial processes
- *Ρομποτική*: αυτόνομη πλοήγηση, έλεγχος κίνησης
- *Βίντεο*: συμπίεση, ψηφιακή τηλεόραση, τηλεδιάσκεψη, τηλεϊατρική, videophone
- *Ιατρική*: μελέτη της κίνησης της καρδιάς και των κυττάρων

- *Δορυφορικές επικοινωνίες και Μετεωρολογία*: εντοπισμός κίνησης νεφελωμάτων, remote sensing
- *Μεταφορά*: καταγραφή της κίνησης των οχημάτων σε έναν δρόμο

Γενικά, η μέτρηση και ανάλυση της οπτικής κίνησης γίνεται σε δύο στάδια: 1) Εκτίμηση της οπτικής ροής, π.χ. της διδιάστατης κίνησης πάνω στη εικόνα, και 2) εκτίμηση της 3D κίνησης και πιθανόν των 3D σχημάτων των πραγματικών αντικειμένων.

Η εκτίμηση της οπτικής ροής με την βοήθεια του διανυσματικού πεδίου ταχυτήτων μπορεί να χωριστεί σε δύο μεγάλες κατηγορίες, ανάλογα με τον τρόπο προσέγγισης, σε *συνεχή (continuous)* και *διακριτή (discrete)*, χρησιμοποιώντας *μοντέλα κλίσεως (gradient models)* ή *αντιστοιχία (correspondence)* των σημείων κίνησης αντίστοιχα. [24]

Μερικά παραδείγματα των μοντέλων παραγώγων είναι ο περιορισμός της οπτικής ροής [26], μια προσέγγιση βασισμένη στην μέθοδο των ελαχίστων τετραγώνων της οπτικής ροής σε μετασχηματισμένα κατά affine διανυσματικά πεδία χρησιμοποιώντας shape gramians [25] και κάποιοι περιορισμοί ομαλότητας που χρησιμοποιούνται για να παραγάγουν την οπτική ροή κατά μήκος των περιγραμμάτων. [27]

Βασικά μειονεκτήματα των παραπάνω μεθόδων είναι η μεγάλη τους απαίτηση σε υπολογισμούς, η εφαρμογή τους μόνο σε μικρές περιοχές της κίνησης και η ευαισθησία τους στην παρουσία θορύβου. Από την άλλη πλευρά, οι μέθοδοι αντιστοιχίας είναι πιο σταθερές στον θόρυβο και έχουν την δυνατότητα εφαρμογής και σε περιοχές κίνησης μεγαλύτερου μεγέθους. Για την μέτρηση του 2D πεδίου ταχυτήτων εφαρμόζονται οι μέθοδοι του *ταιριάσματος block (block matching)* [28], της *ανάλυσης Fourier* και του *ταιριάσματος χαρακτηριστικών*.

### ➤ Σχήμα (shape)

Τέλος, το σχήμα αποτελεί σημαντικό χαρακτηριστικό γνώρισμα, απαραίτητο στην διαδικασία ταιριάσματος εικόνων. Μπορεί να μελετηθεί με βάση το περίγραμμα του σχήματος ή με βάση την περιοχή που αυτό περικλείει. Έτσι, η ανάλυση του σχήματος χωρίζεται σε δύο βασικές κατηγορίες, αυτή που ασχολείται με τα *όρια του σχήματος (boundary-based)* και σ' αυτή που ασχολείται με την *περιοχή του (region-based)*.

Η ανάλυση με βάση τα όρια του σχήματος (περίγραμμα) χρησιμοποιεί μεθόδους βασισμένες στην καμπυλότητα και την διαφορική γεωμετρία. Από την άλλη πλευρά, η ανάλυση με βάση την περιοχή του σχήματος χρησιμοποιεί μεθόδους βασισμένες στα moments και στις προβολές, στους μετασχηματισμούς απόστασης (distance transforms) στο skeletonization και στις κατανομές μεγέθους (granulometries) [22].

Όσον αφορά τους μετασχηματισμούς απόστασης, χρησιμοποιείται ο *Ευκλείδειος μετασχηματισμός απόστασης (Euclidean distance transform)*, [29] ο οποίος όμως παρουσιάζει πολυπλοκότητα στον υπολογισμό του. Γι αυτό το λόγο, στην πράξη χρησιμοποιούνται κάποιες προσεγγίσεις του, που εμπεριέχουν λιγότερους υπολογισμούς. [30]

Η μέθοδος του *skeletonization* των ψηφιακών εικόνων βασίστηκε στον *μετασχηματισμό σκελετού (skeleton transform)*, τον οποίο πρότεινε ο Blum το 1967. [31]

Σύμφωνα με αυτόν, ένα ψηφιακό σχήμα μπορεί να αναπαρασταθεί από τους λεπτούς άξονες του (thin medial axis) και αντίστοιχα να ανακτηθεί από την πληροφορία που μεταφέρουν οι άξονες αυτοί. [32]

Η μέθοδος των *granulometries* βασίζεται σε μη γραμμικά φίλτρα ομαλοποίησης (μορφολογικά φίλτρα). Τα φίλτρα αυτά, μπορούν να αναπαρασταθούν από ένα μετασχηματισμό σκελετού, και να επιτελέσουν ομαλοποίηση πολλαπλής κλίμακας (multiscale smoothing) του υπό εξέταση σχήματος. Είναι μια χρήσιμη μέθοδος σε περιπτώσεις προσδιορισμού του μεγέθους, αλλά και στην πολλαπλής κλίμακας περιγραφή του σχήματος ή της υφής ενός αντικειμένου. [33]

Όσον αφορά τις μεθόδους των moments και της καμπυλότητας (curvature), γίνεται εκτενής ανάλυση στις επόμενες ενότητες, όπου και παρουσιάζεται αναλυτικά η μέθοδος της Curvature Scale Space (CSS).

## 2.4 Απεικόνιση εικόνων με βάση το σχήμα

Το πλεονέκτημα του χαρακτηριστικού του *σχήματος* (shape) μιας εικόνας, έχει να κάνει με το γεγονός ότι η ανθρώπινη αντίληψη έχει την δυνατότητα να αναγνωρίσει ένα αντικείμενο απλά και μόνο παρατηρώντας το σχήμα του. Το σχήμα, επομένως, μπορεί να μεταφέρει σημαντική πληροφορία στην διαδικασία της αναζήτησης και ανάκτησης μιας εικόνας. Για αυτό το λόγο, το χαρακτηριστικό του σχήματος ξεχωρίζει από τα άλλα οπτικά χαρακτηριστικά μιας εικόνας που προαναφέραμε, γιατί παρόλο που και αυτά είναι εξίσου σημαντικά, δεν έχουν την δυνατότητα να καθορίσουν την ταυτότητα ενός αντικειμένου.

Στην πράξη πολλές εφαρμογές χρησιμοποιούν το χαρακτηριστικό του σχήματος. Παραδείγματος χάριν, ένα σύστημα ασφαλείας που διαθέτει video, μπορεί με την βοήθεια μεθόδων με βάση το σχήμα, να διαγνώσει την είσοδο ενός εισβολέα. Επιπλέον, το σχήμα βρίσκει ευρεία εφαρμογή στον χώρο του διαδικτύου, όπου η αναζήτηση και ανάκτηση ενός αντικειμένου (image retrieval) γίνεται ευκολότερα και πιο γρήγορα μέσω των shape descriptor.[34]

Γενικά οι σχηματικές αναπαραστάσεις διαιρούνται σε δύο κατηγορίες, σ' αυτή που βασίζεται στα *όρια του σχήματος* (boundary-based ή contour-based) και σ' αυτή που βασίζεται στην *περιοχή* (region-based). Η πρώτη χρησιμοποιεί μόνο τα εξωτερικά όρια του σχήματος, δηλαδή το περίγραμμά του, ενώ η δεύτερη τη συνολική περιοχή του σχήματος. Όπως είναι φανερό, η βασισμένη στα όρια του σχήματος αναπαράσταση είναι κατάλληλη για αντικείμενα με παρόμοια χαρακτηριστικά περιγράμματος, ενώ η βασισμένη στην περιοχή είναι αποτελεσματικότερη σε σχήματα με παραπλήσια χωρική κατανομή των pixel τους. Οι πιο δημοφιλείς μέθοδοι αναπαράστασης αυτών, είναι οι *περιγραφείς Fourier* (Fourier Descriptors) και οι *σταθερές Moment* (Moment Invariants) αντίστοιχα.[2]



## 2.4.1 Σταθερές Moment

Όπως αναφέρθηκε παραπάνω, η αναπαράσταση εικόνων με βάση την περιοχή χρησιμοποιεί την μέθοδο των σταθερών moment (moment invariants). Οι σταθερές moment και οι μέθοδοι που υλοποιήθηκαν με την βοήθεια τους, χρησιμοποιήθηκαν σε ένα πλήθος εφαρμογών αναγνώρισης διδιάστατων εικόνων, και αποδείχθηκε ότι παραμένουν αμετάβλητα σε τυχαίους γεωμετρικούς μετασχηματισμούς.

Πρώτος ο Hu [35] εισήγαγε το 1961 την έννοια των σταθερών moment, βασισμένος σε μεθόδους των αλγεβρικών σταθερών. Χρησιμοποιώντας μη γραμμικούς συνδυασμούς των γεωμετρικών (αλγεβρικών) moment, κατάφερε να δημιουργήσει μια ομάδα σταθερών moment, που παρέμενε αμετάβλητη σε γεωμετρικούς μετασχηματισμούς, όπως η περιστροφή, μετάθεση και αλλαγή κλίμακας.

Για να γίνει πιο εύκολη η ανάκληση των εικόνων με την βοήθεια των σταθερών moment, ο Teague πρότεινε την ιδέα των *Zernike moments*. [37] Τα moments αυτά ήταν βασισμένα στην θεωρία των ορθογώνιων πολυώνυμων, που επέτρεπαν την δημιουργία ακόμα και υψηλότερης τάξης moment. Άλλα ορθογώνια moments είναι τα *Legendre moments*, που βασίζονται στα αντίστοιχα πολυώνυμα Legendre. Πιο πρόσφατα, εισήχθη η έννοια των *complex moments*, με σκοπό την πιο απλή και εύκολη κατασκευή των σταθερών moment. [38]

Οι σταθερές moment ορίζονται συνοπτικά ως εξής:

1. **Γεωμετρικά Moments** τάξεως  $(p + q)$  μιας συνάρτησης  $f(x, y)$

$$M_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p y^q f(x, y) dx dy$$

όπου  $p, q = 0, 1, 2, \dots, \infty$ .

2. **Legendre Moments** τάξεως  $(m + n)$

$$\lambda_{mn} = \frac{(2m+1)(2n+1)}{4} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} P_m(x) P_n(y) f(x, y) dx dy$$

όπου  $m, n = 0, 1, 2, \dots, \infty$ , και  $P_n, P_m$  τα πολυώνυμα Legendre.

3. **Zernike Moments** τάξεως  $n$  με  $l$  επαναλήψεις

$$A_{nl} = \frac{n+1}{\pi} \int_0^{2\pi} \int_0^{\infty} [V_{nl}(r, \theta)]^* \cdot f(r \cos \theta, r \sin \theta) r dr d\theta$$

όπου  $n = 0, 1, 2, \dots, \infty$ , και το  $l$  παίρνει θετικές και αρνητικές τιμές σύμφωνα με την συνθήκη:

$$n - |l| = \text{even}, \quad \text{όπου} \quad |l| \leq n$$

#### 4. *Complex Moments* τάξεως $(p + q)$

$$C_{pq} = \int_{-\infty-\infty}^{\infty} \int_{-\infty-\infty}^{\infty} (x+iy)^p (x-iy)^q f(x,y) dx dy$$

όπου  $p, q = 0, 1, 2, \dots, \infty$ , και  $i = \sqrt{-1}$

Οι παραπάνω σταθερές moment παραμένουν γενικά αμετάβλητες στην επίδραση affine μετασχηματισμών, όπως η αλλαγή κλίμακας, η μετάθεση και η περιστροφή. Στην περίπτωση όμως του affine μετασχηματισμού της κλίσης (shear) τα σύνορα του σχήματος ενός αντικειμένου αλλάζουν. Για το λόγο αυτό, ο Flusser [41,42], πρότεινε την ιδέα των *affine area moment invariants* που παρέμεναν αμετάβλητα σε γενικούς affine μετασχηματισμούς.

Τα Hu moments, όπως προαναφέραμε, αναφέρονται στην περιοχή, δηλαδή μπορούν να υπολογιστούν για όλα τα pixel ενός αντικειμένου, συμπεριλαμβανομένου και αυτά των συνόρων του. Ο υπολογισμός τους όμως απαιτεί πολύ χρόνο στις περιπτώσεις μεγάλων διαστάσεων αντικειμένων. Για το λόγο αυτό, ο Chen [43] παρουσίασε μια διαφορετική προσέγγιση των σταθερών moment, βασισμένη αυτή τη φορά στα όρια του σχήματος, την κατηγορία των *curve moment invariants*. Τα curve moment invariants έχουν πολλές ομοιότητες με τα Hu moments, με την διαφορά ότι απαιτούν υπολογισμό τους μόνο στα όρια του αντικειμένου. Επιπρόσθετα, διατηρούνται σταθερά σε περίπτωση αλλαγής κλίμακας, μετάθεσης ή περιστροφής του αντικειμένου. Για να εξασφαλιστεί όμως και εδώ η αμεταβλητότητα σε περίπτωση ενός “shear” μετασχηματισμού, ορίστηκε ένας νέος τύπος σταθερών moment, τα *affine weighted curve moments*, [44] πάνω σε μια παραμετροποίηση του συνόρου του σχήματος. Από αυτά κατόπιν προέκυψε μια νέα κατηγορία moment invariants που παρέμενε αμετάβλητη σε τυχαίους affine μετασχηματισμούς, τα *affine curve moment invariants*. [45] Ένα επιπλέον πλεονέκτημα αποτελεί η μεγαλύτερη ανοχή τους στον θόρυβο, καθώς γίνεται ομαλοποίηση της καμπύλης του συνόρου κατά την παραμετροποίηση του.

Σημαντικό μειονέκτημα στην μέθοδο των σταθερών moment αποτελεί η επίδραση του θορύβου που μπορεί να προέρχεται από την δειγματοληψία, την ψηφιοποίηση και την κβαντοποίηση που εφαρμόζεται στην εικόνα. Πιο συγκεκριμένα, λόγω της υποδειγματοληψίας (undersampling) μπορεί να εμπεριέχεται σφάλμα κατά την αλλαγή κλίμακας, η ψηφιοποίηση να επηρεάζει σε περίπτωση περιστροφής του αντικειμένου και τέλος η το σφάλμα στην κβαντοποίηση να υπεισέρχεται και αυτό στους υπολογισμούς των moments. [36]

Για αυτό το λόγο, επιδιώκουμε κάθε φορά οι σταθερές moment να είναι όσο το δυνατόν λιγότερο ευαίσθητες στην παρουσία θορύβου.

## 2.4.2 Περιγραφείς Fourier

Οι μετασχηματισμοί Fourier είναι οι πλέον διαδεδομένοι στην διαδικασία ανάλυσης ενός σχήματος, και τελευταία βρίσκουν εφαρμογή και στην διαδικασία ανάκτησης του. Οι *περιγραφείς Fourier* (Fourier descriptors) που προκύπτουν από τους μετασχηματισμούς Fourier, ουσιαστικά αναπαριστούν το σχήμα ενός αντικειμένου στο πεδίο της συχνότητας. Ο διακριτός 1-D μετασχηματισμός Fourier για μία συνάρτηση  $s(t)$  δίνεται από τον τύπο:

$$u_n = \frac{1}{N} \sum_{t=0}^{N-1} s(t) \exp\left(\frac{-j2\pi nt}{N}\right)$$

όπου  $t = 0, 1, \dots, L$  και  $n = 0, 1, \dots, N-1$  τα  $N$  σημεία στα οποία θεωρούμε ότι έγινε κανονικοποίηση της συνάρτησης κατά την δειγματοληψία. Η κανονικοποίηση που εφαρμόζεται αρχικά στην συνάρτηση, αποσκοπεί στην μελλοντική διαδικασία ταιριάσματος των σχημάτων, όπου η σύγκριση απαιτεί ίδιο αριθμό σημείων.

Η  $s(t)$  (*shape signature*) είναι μία συνάρτηση που προκύπτει από τις συντεταγμένες του συνόρου του σχήματος. Υπάρχουν διάφορων ειδών τέτοιες συναρτήσεις, αλλά κατά κανόνα χρησιμοποιείται αυτή με βάση την απόσταση των σημείων του ορίου του σχήματος από το κέντρο μάζας του (*centroid distance function*), καθώς φαίνεται να έχει την καλύτερη απόδοση [46].

Οι συντελεστές  $u_n, n = 0, 1, \dots, N-1$  ονομάζονται περιγραφείς Fourier (Fourier descriptors, FD) του σχήματος και συμβολίζονται ως  $FD_n, n = 0, 1, \dots, N-1$

Παρόλο που οι μονοδιάστατοι περιγραφείς Fourier είναι αμετάβλητοι εν γένει σε γεωμετρικούς μετασχηματισμούς, υπάρχουν περιπτώσεις που αυτό δεν συμβαίνει. Για αυτό το λόγο, εισήχθη η έννοια των *generic* Fourier περιγραφέων, που προκύπτουν από τον διακριτό 2-D Fourier μετασχηματισμό.

Ο διακριτός 2-D μετασχηματισμός Fourier για μία συνάρτηση  $f(x, y), (0 \leq x < M, 0 \leq y < N)$  δίνεται από τον τύπο:

$$F(u, v) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) \exp[-j2\pi(ux/M + vy/N)]$$

όπου  $u$  και  $v$  συχνότητες στην οριζόντια και κάθετη κατεύθυνση αντίστοιχα.

Επειδή όμως αυτός μεταβάλλεται κατά την περιστροφή ενός αντικειμένου, το σχήμα «μεταφέρθηκε» στο χώρο των πολικών συντεταγμένων και ορίστηκε ο διακριτός *πολικός* μετασχηματισμός Fourier ως εξής:

$$PF_1(\rho_l, \psi_m) = \sum_p \sum_i f(r_p, \theta_i) r_p \exp[-j2\pi r_p \rho_l \sin(\theta_i + \psi_m)]$$

όπου  $r_p = p/R$ ,  $\theta_i = i(2\pi/T), (0 \leq i < T)$ ,  $\rho_l = l, (0 \leq l < R)$  και  $\psi_m = m\theta_i$ .

Οι μεταβλητές R και T αναφέρονται στην ακτινική και γωνιακή συχνότητα αντίστοιχα. Οι πολικοί Fourier συντελεστές που προκύπτουν, χρησιμοποιούνται για τον καθορισμό των αντίστοιχων Fourier περιγραφών (FD).[39]

Οι χαμηλής τάξεως Fourier περιγραφείς παρέχουν πληροφορίες για τα γενικά χαρακτηριστικά του σχήματος, ενώ οι υψηλότερης τάξεως περιέχουν πληροφορίες για τις λεπτομέρειες που εμφανίζει. Όπως φαίνεται και από τον παραπάνω τύπο, προκύπτει ένας αρκετά μεγάλος αριθμός Fourier περιγραφών, εκ των οποίων σε πολλές περιπτώσεις αγνοούνται αυτοί των υψηλών συχνοτήτων, καθώς αφορούν σε λεπτομέρειες του σχήματος που καθιστούν δύσκολή την σχηματική του αναπαράσταση.

Σημαντικό πλεονέκτημα των περιγραφών Fourier αποτελεί το γεγονός ότι παρουσιάζουν ανθεκτικότητα στον θόρυβο, σε αντίθεση με τις σταθερές moment. Επίσης, η μέθοδος των περιγραφών Fourier εφαρμόζει εύκολα την κανονικοποίηση, και φαίνεται να διατηρεί ικανοποιητικά τις πληροφορίες που μεταφέρει.[40]. Οι μέθοδοι Fourier χρησιμοποιούνται, όπως προαναφέραμε, στην σχηματική ανάλυση και αναπαράσταση, στην ανάκτηση σχήματος αντικειμένων, στην κατηγοριοποίηση αλλά και στην κωδικοποίηση σχημάτων.

## **2.5. Σύστημα Περιγραφής του Περιεχομένου Δεδομένων Πολυμέσων (MPEG – 7)**

Αξίζει να αναφέρουμε ότι με σκοπό την εξαγωγή των οπτικών χαρακτηριστικών μιας εικόνας, αναπτύχθηκε μία βάση δεδομένων πολυμέσων ,το πρότυπο *Σύστημα Περιγραφής του Περιεχομένου Δεδομένων Πολυμέσων*, γνωστό και ως *MPEG-7*. Στόχος του είναι να καθορίσει ένα πρότυπο (standard) τρόπο περιγραφής των στοιχείων μιας εικόνας, ώστε να είναι ευκολότερη και ακριβέστερη η ανάκτηση της από τον χρήστη. Έτσι εισήχθη η έννοια του *περιγραφέα* (descriptor) για κάθε ένα από τα οπτικά χαρακτηριστικά μιας εικόνας ,και μάλιστα αποδόθηκαν παραπάνω από ένας περιγραφείς σε κάθε χαρακτηριστικό ώστε η αναφορά στην εικόνα να είναι ακόμα πιο λεπτομερής.

Στην συνέχεια παραθέτουμε μερικούς από τους πιο ενδεικτικούς περιγραφείς για κάθε οπτικό χαρακτηριστικό. [48]

### **➤ Χρώμα**

Για το χαρακτηριστικό του χρώματος ο MPEG-7 διαθέτει πρότυπους περιγραφείς, που επιλέχθηκαν για το μικρή τους απαίτηση σε μνήμη και υπολογισμούς, καθώς και για την αποδοτικότητα τους σε περιπτώσεις ταιριάσματος εικόνων βασισμένες στα χρωματικά χαρακτηριστικά τους. Συνοπτικά είναι οι εξής:

- *color space descriptor*

Επιτρέπει την επιλογή του χρωματικού χώρου που θα χρησιμοποιηθεί στην περιγραφή. Με αυτόν σχετίζεται και ο *color quantization descriptor*, ο οποίος καθορίζει την κατανομή του χρωματικού χώρου στο επίπεδο των διακριτών bin της εικόνας. Και οι δύο αυτοί περιγραφείς χρησιμοποιούνται σε συνδυασμό με άλλους περιγραφείς χρώματος.

- *dominant color descriptor*

Επιτρέπει τον καθορισμό των τιμών του κυρίαρχου χρώματος μια εικόνας, όπως επίσης και τις στατιστικές τους ιδιότητες σε σχέση με την κατανομή και την ποικιλομορφία του. Σκοπός του είναι η αποτελεσματική και ρεαλιστική αναπαράσταση των χρωμάτων της περιοχής της εικόνας που εξετάζουμε κάθε φορά.

- *scalable color descriptor*

Προκύπτει από το χρωματικό ιστόγραμμα που καθορίζεται από τον χρωματικό χώρο HSV. Χρησιμοποιεί κωδικοποίηση συντελεστή μετασχηματισμένου κατά Haar, δίνοντας μια κλιμακωτή αναπαράσταση της περιγραφής.

- *group of frames or group of pictures descriptor*

Είναι μια επέκταση του *scalable color descriptor* για μια ομάδα εικόνων ή πλαισίων σε μια ακολουθία video. Βασίζεται στην συνάθροιση των χρωματικών ιδιοτήτων αυτόνομων κάθε φορά των εικόνων ή video πλαισίων.

- *colour structure descriptor*

Βασίζεται στα χρωματικά ιστογράμματα, αλλά έχει ως στόχο να αναγνωρίζει τοπικές χρωματικές κατανομές, με την βοήθεια ενός «παραθύρου» δόμησης (*structuring window*).

- *colour layout descriptor*

Αναπαριστά την χρωματική διάταξη των αντιπροσωπευτικών χρωμάτων σε ένα πλέγμα που «επιβάλλεται» σε μια περιοχή ή εικόνα. Η αναπαράσταση βασίζεται στους συντελεστές του *διακριτού μετασχηματισμού του συνημίτονου*. Είναι πολύ αποδοτικός σε εφαρμογές αναζήτησης, τόσο σε ακίνητες όσο και σε κινούμενες εικόνες.

## ➤ Υφή

Οι περιγραφείς υφής στο MPEG-7 σύστημα έχουν την δυνατότητα της αναζήτησης και ανάκτησης ομοίων εικόνων από βάσεις δεδομένων αλλά και από ακολουθίες video. Μια εικόνα μπορεί να θεωρηθεί ως «μωσαϊκό» ομογενών περιοχών υφής, συνδεδεμένες κατά τέτοιο τρόπο ώστε να μπορούν να χαρακτηρίσουν μια εικόνα. Για να υποστηρίξει μια τέτοιου είδους αναζήτηση ο MPEG-7 έχει ορίσει τρεις πρότυπους περιγραφείς, τους εξής:

- *homogeneous texture descriptor*

Παρέχει μια επαρκή αναπαράσταση της υφής – χρησιμοποιώντας 62 αριθμούς – που αποτελείται από την μέση ενέργεια και την ενεργειακή απόκλιση με βάση ένα καθορισμένο αριθμό καναλιών συχνοτήτων. Είναι χρήσιμος στην ανάκτηση ομοίων εικόνων και πολύ αποτελεσματικός στον χαρακτηρισμό ομογενών περιοχών μίας εικόνας.

- *texture browsing descriptor*

Είναι αρκετά συμπίεσμένος (compact) και απαιτεί μέγιστο αριθμό 12 bit. Παρέχει χαρακτηρισμό της υφής, που προσεγγίζει την ανθρώπινη αντίληψη, όσον αφορά την κανονικότητα (regularity), την τραχύτητα (coarseness) και την κατευθυντικότητα (directionality) του σχεδίου υφής.

- *edge histogram descriptor*

Αναπαριστά την χωρική κατανομή των ακμών μιας εικόνας. Είναι ιδιαίτερα χρήσιμος στο ταίριασμα περιοχών με ανομοιογένεια, όπως οι φυσικές, sketch και clip art εικόνες.]

## ➤ **Κίνηση**

Οι περιγραφείς της κίνησης πρέπει να είναι εύκολο να εξαχθούν από το περιεχόμενο ενός video. Επιπλέον, πρέπει να μην περιορίζεται όσο το δυνατόν η εφαρμογή τους, και να απαιτούν καθορισμένο χώρο μνήμης. Σύμφωνα με τα παραπάνω, ο MPEG-7 όρισε τέσσερις πρότυπους περιγραφείς κίνησης, τους εξής:

- *Motion activity*

Αναπαριστά την αφηρημένη έννοια της «έντασης της κίνησης» ή του «ρυθμού της δράσης» σε μια ακολουθία video. Μας επιτρέπει να εκφράσουμε με ακρίβεια την δραστηριότητα σε μία δοσμένη ακολουθία video, και χρησιμοποιείται στις περιπτώσεις δημιουργίας αλλά και γρήγορης αναζήτησης video.

- *Camera motion*

Περιγράφει τις λειτουργίες μιας κάμερας, όπως η μετάθεση, η περιστροφή και η αλλαγές του μήκους εστίασης, όπως και συνδυασμούς τους. Μπορούμε με αυτό τον τρόπο να γνωρίζουμε τις τυχόν διαδοχικές αλλαγές της κάμερας, και το πως αυτές επηρεάζουν την αναπαράσταση μιας εικόνας.

- *Motion trajectory*

Περιγράφει τις μετατοπίσεις ενός αντικειμένου στο χρόνο, καθώς αυτό ορίζεται χωροχρονικά σε μια περιοχή της εικόνας. Οι διαδοχικές χωροχρονικές «θέσεις» του αντικειμένου αναπαριστώνται με βάση ένα χαρακτηριστικό του σημείο, όπως το κέντρο μάζας του.

- *Parametric motion*

Αναπαριστά την κίνηση και/ή την παραμόρφωση μιας περιοχής της εικόνας, εξαιτίας διαφόρων μοντέλων παραμετρικής κίνησης (*translational, rotation/scaling, affine, perspective, quadratic*).

## ➤ Σχήμα

Όπως προαναφέραμε, υπάρχουν δύο μέθοδοι περιγραφής του σχήματος μιας εικόνας, η βασισμένη στην περιοχή (*region – based*) και στο περίγραμμα (*contour – based*) του σχήματος. Ορίστηκαν επομένως και οι αντίστοιχοι περιγραφείς:

- *region – based shape descriptor*

Αναπαριστά την κατανομή των pixels μιας διδιάστατης εικόνας. Η τεχνική ανάλυσης σχήματος που χρησιμοποιεί βασίζεται στην μέθοδο των moments.

- *contour – based shape descriptor*

Βασίζεται στην Curvature Scale – Space (CSS) αναπαράσταση του περιγράμματος μιας εικόνας. Το πλεονέκτημα του είναι ότι δεν επηρεάζεται από παραμορφώσεις, και τυχαίους μετασχηματισμούς (π.χ. affine μετασχηματισμούς).

- *3 – D shape descriptor*

Χρησιμοποιείται στην περίπτωση των 3 – D αντικειμένων, τα οποία αναπαριστώνται ως διακριτά πολυγωνικά 3 – D δικτυωτά πλέγματα (meshes).

### 3. Affine μετασχηματισμοί

#### 3.1.Εισαγωγή

Όλες οι μέθοδοι ανάλυσης και επεξεργασίας σχήματος αντιμετωπίζουν ένα σημαντικό πρόβλημα. Ως γνωστόν, ένα σχήμα καθορίζεται από τις  $x$  και  $y$  συντεταγμένες των σημείων του συνόρου του (περίγραμμα), που όμως αλλάζει ανάλογα με την άποψη του αντικειμένου στο χώρο. Η αλλαγή αυτή αποτελεί μετασχηματισμό προοπτικής προβολής του αντικειμένου (perspective transformation) και διαφοροποιεί σε μεγάλο βαθμό την αρχική μορφή του αντικειμένου.

Η μεταβολή αυτή μπορεί να προσεγγιστεί με έναν γενικό *affine μετασχηματισμό* (affine transformation). Το βασικό πλεονέκτημα του είναι ότι αποτελεί γραμμικό μετασχηματισμό, γεγονός που διευκολύνει τους υπολογισμούς κατά την διαδικασία ανάλυσης και επεξεργασίας μιας εικόνας.

Πρακτικά, affine μετασχηματισμό θεωρούμε ότι έχουμε όταν η κάμερα αλλάξει σημείο παρατήρησης σε σχέση με το αντικείμενο, με αποτέλεσμα το σύνορο του αντικειμένου να έχει παραμορφωθεί. Η παραμόρφωση αυτή αντιστοιχεί σε τυχαίους affine μετασχηματισμούς, όπως η *αλλαγή στην κλίμακα* (scaling), η *περιστροφή* (rotation) και η *μετάθεση* (translation) και *κλίση* (shear).

Οι παραπάνω μετασχηματισμοί αποτελούν *μετασχηματισμούς συντεταγμένων*, καθώς σχετίζονται άμεσα με τις τιμές των συντεταγμένων του συνόρου του αντικειμένου. Υπάρχουν όμως και οι *μετασχηματισμοί παραμέτρου* που μπορεί να οφείλονται στο γεγονός ότι οι καμπύλες των συνόρων προέρχονται από κατάτμηση εικόνων. Σαν αποτέλεσμα, εμφανίζεται θόρυβος κατάτμησης και δειγματοληψία μη ομοιόμορφη ως προς το μήκος τόξου. Επίσης, σημαντική παράμετρο αποτελεί και το *σημείο αναφοράς* της καμπύλης, το σημείο δηλαδή που επιλέγεται αυθαίρετα για να περιγράψει την κλειστή καμπύλη. Τέλος, η φορά διαγραφής της καμπύλης μπορεί να είναι αριστερόστροφη ή δεξιόστροφη.

#### 3.2 Θεωρητικά στοιχεία

Στην περίπτωση ενός διδιάστατου αντικειμένου, ένας *affine* μετασχηματισμός μπορεί να δοθεί από τις παρακάτω μαθηματικές εξισώσεις:

$$\begin{cases} x_a(u) = ax(u) + by(u) + e \\ y_a(u) = cx(u) + dy(u) + f \end{cases}$$



όπου  $x_a(u), y_a(u)$ , αντιπροσωπεύουν τις κατά affine μετασχηματισμένες συντεταγμένες του αντικειμένου συναρτήσει μιας τυχαίας παραμέτρου  $u$ .

Υπάρχουν έξι βαθμοί ελευθερίας στον παραπάνω μετασχηματισμό. Η μετάθεση και η αλλαγή κλίμακας αναπαριστώνται από δύο βαθμούς ελευθερίας η καθεμία, ενώ η αλλαγή στον προσανατολισμό χρειάζεται ένα βαθμό. Η παράμετρος που υπολείπεται αναφέρεται στην κλίση (shear).

Οι παραπάνω εξισώσεις δίνουν την δυνατότητα να συσχετιστούν δύο διαφορετικές οπτικές γωνίες παρατήρησης ενός αντικειμένου. Αν υποθέσουμε ότι οι συντεταγμένες  $x(u), y(u)$  είναι γνωστές, αναφέρονται δηλαδή σε συγκεκριμένη γνωστή γωνία παρατήρησης, τότε οι συντεταγμένες μιας άγνωστης γωνίας παρατήρησης  $x_a(u), y_a(u)$ , μπορούν να εκφραστούν σαν γραμμικός συνδυασμός των γνωστών  $x(u), y(u)$ . Κατά τον ίδιο τρόπο μπορούν να προκύψουν τυχαίοι affine μετασχηματισμοί της οπτικής γωνίας ενός αντικειμένου, επιλέγοντας κάθε φορά διαφορετικές affine παραμέτρους. Παρακάτω, ορίζονται και εξετάζονται ξεχωριστά κάθε μια από αυτές τις παραμέτρους.

Αν θεωρήσουμε ότι ο πίνακας  $X$  περιέχει τις συντεταγμένες  $x(u), y(u)$  ενός αντικειμένου, δηλαδή είναι της μορφής  $X = \begin{bmatrix} x(u) \\ y(u) \end{bmatrix}$ , τότε ένας γενικός affine μετασχηματισμός του,  $Y$ , δίνεται από τον τύπο:

$$Y = AX + B$$

Συνδέοντας την εξίσωση πινάκων με τις παραπάνω εξισώσεις για τα  $x(u), y(u)$ , μπορούμε να πούμε ότι οι πίνακες  $A$  και  $B$  αντιστοιχίζονται ως εξής:

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

$$B = \begin{bmatrix} e \\ f \end{bmatrix}$$

Ο πίνακας  $A$  μπορεί να αποτελείται από επιμέρους πίνακες ή από το γινόμενο τους όπου ο καθένας αντιπροσωπεύει και έναν γεωμετρικό μετασχηματισμό [1]. Έτσι, η αλλαγή στην κλίμακα (scaling), η περιστροφή (rotation), και η κλίση (shear) ορίζονται ως εξής:

1) **Αλλαγή κλίμακας (scaling)**

$$A_{scaling} = \begin{pmatrix} S_x & 0 \\ 0 & S_y \end{pmatrix}$$

2) **Περιστροφή (rotation)**

$$A_{rotation} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$$

### 3) Κλίση (shear)

Υπάρχουν τρεις επιμέρους περιπτώσεις shearing [47]

- Shearing του σχήματος κατά  $k$  στον άξονα των  $x$  οπότε και προκύπτει ο ακόλουθος επιμέρους πίνακας:

$$A_{shear} = \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}$$

- Shearing του σχήματος κατά  $k$  στον άξονα των  $y$  οπότε και προκύπτει ο ακόλουθος επιμέρους πίνακας:

$$A_{shear} = \begin{pmatrix} 1 & 0 \\ k & 1 \end{pmatrix}$$

- Shearing του σχήματος κατά  $k_x$  στον άξονα των  $x$  και κατά  $k_y$  στον άξονα των  $y$  οπότε και προκύπτει ο ακόλουθος επιμέρους πίνακας:

$$A_{shear} = \begin{pmatrix} 1 & k_x \\ k_y & 1 \end{pmatrix}$$

Αντίστοιχα ο πίνακας  $B$  αναφέρεται στην *μετάθεση* (translation) που προκύπτει όταν το σχήμα μετατοπιστεί κατά ένα παράγοντα στον  $x$  και  $y$  άξονα. Έτσι, αν το σχήμα έχει μετατοπιστεί κατά  $h$  στον άξονα των  $x$  και κατά  $v$  στον άξονα των  $y$ , ο μετασχηματισμός προκύπτει ως εξής:

### 4) Μετάθεση (translation)

$$Y = A \begin{bmatrix} x-h \\ y-v \end{bmatrix} + B + \begin{bmatrix} h \\ v \end{bmatrix} = A \begin{bmatrix} x \\ y \end{bmatrix} + (I-A) \begin{bmatrix} h \\ v \end{bmatrix} + B = A \begin{bmatrix} x \\ y \end{bmatrix} + B'$$

όπου ο πίνακας  $I$  είναι ο μοναδιαίος

Παρατηρούμε πως η μετάθεση δεν μεταβάλλει καθόλου τον πίνακα  $A$ . [47]

## 3.3. Μήκος τόξου – Affine μήκος

Ο υπολογισμός του μήκους τόξου μία παραμετρικής καμπύλης αποτελεί βασικό πρόβλημα στην γεωμετρική ανάλυση των εικόνων και κατ' επέκταση και στην γραφική των υπολογιστών. Αρχικά, οι Guenter και Parent [49] χρησιμοποίησαν το ολοκλήρωμα

της παραγώγου της συνάρτησης της καμπύλης για τον υπολογισμό του μήκους τόξου, αργότερα οι Vincent και Forsey [50] πρότειναν μία μέθοδο βασισμένη εξ ολοκλήρου στον υπολογισμό σημείων.

Εστω  $f : [a, b] \rightarrow \mathbb{R}^d, d \geq 2$  μια κανονική παραμετρική καμπύλη με συνεχή παράγωγο  $f'(t) \neq 0, t \in [a, b]$ . Για κάθε  $t \in [a, b]$  υπάρχει μια ευθεία γραμμή, η οποία περιέχει τόσο το σημείο  $f(t)$  όσο και το διάνυσμα  $f'(t)$ . Το μήκος τόξου ορίζεται ως εξής:

$$L(f) = \int_a^b |f'(t)| dt$$

Στην περίπτωση όπου  $d = 2$ , δηλαδή το σχήμα είναι δύο διαστάσεων  $x, y$  τότε το μήκος τόξου προκύπτει από τον τύπο:

$$L(x, y) = \int_a^b (\dot{x}^2 + \dot{y}^2)^{1/2}$$

Στην περίπτωση μιας διδιάστατης εικόνας με συντεταγμένες  $x(u), y(u)$ , όπου  $u$  μια τυχαία παράμετρος, μας ενδιαφέρει ο υπολογισμός του κανονικοποιημένου (normalized) μήκους τόξου σύμφωνα με τον τύπο[51]:

$$s(u) = \frac{\int_0^u (\dot{x}^2 + \dot{y}^2)^{1/2}}{\int_0^1 (\dot{x}^2 + \dot{y}^2)^{1/2}}$$

όπου  $(\dot{x}, \dot{y})$  οι παράγωγοι ως προς  $u$  πρώτης τάξης

Η παραμετροποίηση μιας καμπύλης με βάση το μήκος τόξου ονομάζεται *κανονική απεικόνιση* (natural representation), και παραμένει αμετάβλητη σε ευκλείδειους μετασχηματισμούς του αντικειμένου. Παρόλα αυτά, το μήκος τόξου δεν διατηρείται σταθερό στην παρουσία affine μετασχηματισμών, γεγονός που οδήγησε τους Cyganski και Orr [52], να προτείνουν μια νέα παράμετρο που παραμένει αμετάβλητη σε τυχαίους affine μετασχηματισμούς, το *affine μήκος τόξου* (affine length).

Στην περίπτωση μιας διδιάστατης εικόνας με συντεταγμένες  $x(u), y(u)$ , όπου  $u$  μια τυχαία παράμετρος, μας ενδιαφέρει ο υπολογισμός του κανονικοποιημένου (normalized) affine length, που ορίζεται ως εξής [51]:

$$\tau(u) = \frac{\int_0^u (\dot{x}\ddot{y} - \ddot{x}y)^{1/3}}{\int_0^1 (\dot{x}\ddot{y} - \ddot{x}y)^{1/3}}$$

όπου  $(\dot{x}, \dot{y}), (\ddot{x}, \ddot{y})$  οι παράγωγοι ως προς  $u$  πρώτης και δεύτερης τάξης αντίστοιχα.

Σημαντικό μειονέκτημα στον υπολογισμό του affine length αποτελεί το γεγονός ότι απαιτεί τον υπολογισμό παραγώγων μεγαλύτερης τάξεως, με αποτέλεσμα η διαδικασία να γίνεται χρονοβόρα και να εμπεριέχει μεγαλύτερη πιθανότητα σφάλματος. Παρόλα αυτά, όπως θα φανεί και στη συνέχεια από τα πειραματικά αποτελέσματα, η χρήση του affine length παρουσιάζει ικανοποιητική απόδοση στην εφαρμογή της CSS μεθόδου.

## 4. Μέθοδος της curvature scale space (CSS)

### 4.1. Εισαγωγικά

Η ιδέα της *curvature scale space (CSS)* αναπαράστασης βασίζεται στην θεωρία της αναπαράστασης ενός αντικειμένου σε *πολλαπλή κλίμακα* (multi-scale representation). Η θεωρία της πολλαπλής κλίμακας στηρίζεται το γεγονός ότι κάθε αντικείμενο μπορεί να παρουσιάζει διαφορετικές δομές σε διαφορετικές κλίμακες. Με άλλα λόγια, το υπό εξέταση αντικείμενο, εμφανίζεται με διαφορετικές ιδιότητες ανάλογα με την κλίμακα της παρατήρησης. Για ένα σύστημα όρασης υπολογιστών που αναλύει μια άγνωστη εικόνα, δεν υπάρχει κανένας τρόπος να είναι γνωστό εκ των προτέρων ποιες κλίμακες είναι κατάλληλες για την περιγραφή των στοιχείων του. Ως εκ τούτου, η μόνη λογική προσέγγιση είναι να εξεταστούν όλες οι κλίμακες περιγραφών ταυτόχρονα.

Η έννοια της αναπαράστασης σε πολλαπλή κλίμακα εισήχθη για πρώτη φορά στην όραση των υπολογιστών και στην ανάλυση εικόνων από τους Marr και Witkin [54, 55]. Η βασική ιδέα ήταν να ενσωματώσει ένα αρχικό σήμα της μορφής  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  σε μία σειρά σημάτων, φιλτραρισμένα σε διαφορετικές κλίμακες με αύξοντα αριθμό, οι οποίες περιείχαν πληροφορίες για τις λεπτομέρειες του σχήματος. Το φιλτράρισμα κάθε φορά του σήματος, σε κλίμακα  $\sigma \in \mathbb{R}$  επιτυγχάνεται με μια συνάρτηση  $f : \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}$  και ορίζεται ως εξής:

$$F(x, \sigma) = (Q_\sigma(f))(x)$$

όπου  $Q_\sigma$  είναι μια συνάρτηση φίλτρου συναρτήσεως του  $\sigma$ , που μπορεί να είναι μία γραμμική συνάρτηση (Gaussian smoothing) ή μια μη γραμμική συνάρτηση (morphological filter).[56]

Πολύ σύντομα, η μέθοδος της scale – space αναπαράστασης προκάλεσε το ενδιαφέρον στην όραση υπολογιστών, καθώς τα αποτελέσματα της βρίσκουν ευρεία εφαρμογή σε πολλούς τομείς, όπως η εξαγωγή οπτικών χαρακτηριστικών και η κατηγοριοποίηση τους στην ανάλυση ενός σχήματος. Έτσι, κατά καιρούς, αναπτύχθηκαν πολλές τεχνικές βασισμένες στην scale – space αναπαράσταση, όπως οι κατανομές μεγέθους (size distributions) η μέθοδος των granulometries [57, 58] και μία μέθοδος που πρότεινε ο Bangham [59] βασισμένη στα διαδοχικά εναλλακτικά φίλτρα. Μια επίσης διαφορετική προσέγγιση πρότειναν οι Leymarie και Levine βασισμένη στα top-hat or bottom-hat φίλτρα. [60].

Κάνοντας μια βιβλιογραφική ανασκόπηση, διαπιστώνουμε ότι η CSS αναπαράσταση βρίσκει τις ρίζες της στην εξέλιξη της καμπύλης (curve evolution) και στην παραμόρφωση κυρτότητας (curve deformation). Πιο συγκεκριμένα, η δειγματοληπτημένη (resampled) CSS καμπύλη υλοποιεί την παραμόρφωση κυρτότητας[53]. Αυτό επιτυγχάνεται με την συνέλιξη κάθε συντεταγμένης της κλειστής επίπεδης καμπύλης με την *Γκαουσιανή συνάρτηση (Gaussian function)* σε διαφορετικά διαδοχικά επίπεδα κλίμακας. Σαν αποτέλεσμα, η διαδικασία δεν είναι πλέον μια παραμόρφωση κυρτότητας.

Η CSS μέθοδος χρησιμοποιεί την παραμετροποίηση της καμπύλης με βάση το μήκος τόξου, η οποία όμως δεν μένει αμετάβλητη στην παρουσία affine μετασχηματισμών. Επομένως, αναμένουμε μία διαφοροποίηση στην εικόνα της CSS αναπαράστασης όταν αυτή έχει υποστεί affine μετασχηματισμούς. Έχει αποδειχθεί ότι η σταθερότητα της μεθόδου υπό την επίδραση τυχαίων affine μετασχηματισμών επιτυγχάνεται μόνο με την παραμετροποίηση της καμπύλης με βάση το *affine μήκος τόξου* (affine length)[1]. Σε αυτόν ακριβώς τον παράγοντα έγκειται και η αποτελεσματικότητα της CSS μεθόδου, η οποία επιδιώκει να διατηρεί το περίγραμμα ενός σχήματος σταθερό σε affine μετασχηματισμούς. Σε επόμενη ενότητα γίνεται αναφορά σε αυτήν την κανονικοποίηση με βάση το affine μήκος τόξου και στο πώς αυτή επιδρά στην αποτελεσματικότερη εφαρμογή της μεθόδου.

## 4.2. Θεωρητική μελέτη της μεθόδου

### 4.2.1. Καμπυλότητα

Θεωρούμε πρακτικά ότι μία γραμμή παρουσιάζει καμπυλότητα, σε κάθε περίπτωση όπου αυτή παύει να είναι ευθεία. Κατά την ανάλυση μιας καμπύλης, σημαντικό ρόλο παίζει η ανεύρεση των σημείων καμπυλότητας (corner points) κατά μήκος αυτής. Αυτά αποτελούν ιδιομορφίες της καμπύλης – όπως π.χ. άπειρες τιμές καμπυλότητας – και αναγνωρίζονται σαν *extrema points*. Γνωρίζοντας τις θέσεις αυτών των σημείων κανείς μπορεί να λάβει σημαντικές πληροφορίες για το περίγραμμα ενός σχήματος, καθώς και για τιμές άλλων σημείων επί της καμπύλης.

Ο υπολογισμός της καμπυλότητας ακολουθεί τα εξής βήματα:

Αρχικά, θεωρούμε την παραμετρική διανυσματική εξίσωση  $\Gamma$  μίας κλειστής καμπύλης ως εξής:

$$\bar{\Gamma}(u) = (x(u), y(u))$$

όπου  $u$  είναι μια αυθαίρετη παράμετρος.

Το διάνυσμα *unit tangent* για κάθε σημείο της, μπορεί να εκφραστεί ως εξής:

$$\bar{T}(u) = \frac{\bar{\Gamma}'(u)}{\|\bar{\Gamma}'(u)\|}$$

όπου  $\bar{\Gamma}'(u)$  η παράγωγος της καμπύλης ως προς  $u$ .

Επιπλέον, η *ταχύτητα* με την οποία το διάνυσμα θέσης διατρέχει την καμπύλη δίνεται από τον τύπο:

$$\|\bar{\Gamma}'(u)\| = \sqrt{[x'(u)]^2 + [y'(u)]^2}$$

Σε κάθε σημείο της καμπύλης, η γωνία που σχηματίζει το διάνυσμα  $\vec{T}(u)$  και ο άξονας των  $x$ , ορίζεται ως εξής:

$$\theta(u) = \arctan[y'(u)/x'(u)]$$

Για δεδομένο ρυθμό αλλαγής της γωνίας  $d\theta$ , η καμπυλότητα ορίζεται ως εξής:

$$k(u) = \frac{x'(u)y''(u) - x''(u)y'(u)}{[(x'(u))^2 + (y'(u))^2]^{3/2}}$$

όπου εδώ η πρώτη και δεύτερη παράγωγος ορίζονται ως  $x'(u), x''(u)$  και  $y'(u), y''(u)$  για τα  $x(u), y(u)$  αντίστοιχα.

Από τον τύπο φαίνεται ότι η συνάρτηση καμπυλότητας υπολογίζεται μόνο με την βοήθεια των παραμετρικών παραγώγων και για αυτό το λόγο, παραμένει αμετάβλητη σε περίπτωση περιστροφής και μετάθεσης. Παρόλα αυτά, η καμπυλότητα εξαρτάται από μια τυχόν αλλαγή στην κλίμακα του σχήματος, γεγονός που απαιτεί την κανονικοποίηση (normalization) της καμπύλης με τέτοιο τρόπο, ώστε αυτή να παραμένει αμετάβλητη σε έναν τυχαίο affine μετασχηματισμό.[61]

#### 4.2.2. Κανονικοποίηση

Η μέθοδος της κανονικοποίησης αποτελεί έναν ικανοποιητικό τρόπο αντιμετώπισης του προβλήματος των γεωμετρικών affine μετασχηματισμών. Σκοπός της είναι η περιγραφή καμπυλών με τέτοιο τρόπο ώστε αυτές να παραμένουν αναλλοίωτες σε affine μετασχηματισμούς. Επιπλέον διατηρεί ακέραια την υπάρχουσα πληροφορία, καθώς το σχήμα της αρχικής καμπύλης παραμένει ουσιαστικά αμετάβλητο.

Με την εφαρμογή της κανονικοποίησης, το κατά affine μετασχηματισμένο σχήμα έρχεται σε μια «πρότυπη μορφή», απαλλαγμένη από οποιοδήποτε μετασχηματισμό που αυτό μπορεί να έχει υποστεί. Με τον όρο «πρότυπη» θέση εννοούμε τη θέση μιας καμπύλης ή μιας εικόνας, η οποία ορίζεται με τέτοιο τρόπο ώστε όλοι οι affine μετασχηματισμοί του ιδίου αντικειμένου να κανονικοποιούνται πάντα στην ίδια θέση.[62]

Σαν αποτέλεσμα, το κανονικοποιημένο σχήμα και όλα τα χαρακτηριστικά του γνωρίσματα είναι πλέον σταθερά, ενώ παράλληλα οι παράμετροι της κανονικοποίησης δίνουν ανά πάσα στιγμή την θέση του αρχικού σχήματος.

Έστω ότι επιλέγουμε ένα συγκεκριμένο σημείο  $P(x, y)$  της αρχικής μας καμπύλης, του οποίου ένα γειτονικό σημείο θα χρησιμοποιηθεί για την κανονικοποίηση του. Αρχικά, χρησιμοποιούμε την παράμετρο του μήκους τόξου (arc length) για την κανονικοποίηση η οποία ορίζεται ως εξής:

$$d\tau = |\dot{x}^2 + \dot{y}^2|^{1/2} dt$$

Όμως, όπως ήδη αναφέραμε στην ενότητα των affine μετασχηματισμών, το μήκος τόξου μεταβάλλεται στην περίπτωση τυχαίων affine μετασχηματισμών, με αποτέλεσμα να μην δίνει πλέον την αναπαράσταση της αρχικής καμπύλης. Για αυτό το λόγο χρησιμοποιήσαμε το *affine μήκος τόξου* (affine length) που ορίζεται αντίστοιχα ως εξής:

$$d\tau = |\dot{x}\ddot{y} - \ddot{x}\dot{y}|^{1/3} dt$$

το οποίο παραμένει σταθερό στην παρουσία τυχαίων affine μετασχηματισμών.

Έτσι, αν ο πίνακας του affine μετασχηματισμού είναι ο  $A$  της μορφής:

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

τότε, υπό την επίδραση του  $A$ , η παράμετρος  $d\tau$  μετασχηματίζεται στην μορφή [63]:

$$d\tau_A = (\det A)^{1/3} d\tau = (ad)^{1/3} d\tau$$

Με αυτόν τον τρόπο το affine μήκος τόξου μετασχηματίζεται γραμμικά και αποτελεί ένα ικανοποιητικό μέτρο ολοκλήρωσης της καμπύλης.

Όπως ήδη αναφέραμε στην εισαγωγή του κεφαλαίου (βλ. Ενότητα 4.1), η κανονικοποίηση της καμπύλης με βάση το affine μήκος τόξου αποτελεί καθοριστικό παράγοντα στην καλή απόδοση της CSS μεθόδου. Πιο συγκεκριμένα, η κανονικοποίηση επιτυγχάνει να αντιμετωπίσει ικανοποιητικά το πρόβλημα της ύπαρξης affine μετασχηματισμών κατά την διαδικασία εφαρμογής της μεθόδου. Αποτελεί επομένως παράγοντα αξιοπιστίας της μεθόδου, εφόσον βασικό μέλημα στην ανάλυση και επεξεργασία μιας εικόνας είναι η διατήρηση της αρχικής μορφής της, ανεξάρτητα των μετασχηματισμών που μπορεί αυτή να έχει υποστεί.

### 4.2.3. Ομαλοποίηση Καμπύλης

Ως γνωστόν, η ποιότητα μιας εικόνας μπορεί να αλλοιωθεί με την παρουσία τυχαίου θορύβου. Ο θόρυβος αυτός μπορεί να προκύψει κατά την λήψη, μετάδοση ή επεξεργασία της εικόνας, και μπορεί να εξαρτάται ή όχι από το περιεχόμενο της. Συνήθως, ο θόρυβος περιγράφεται με στοιχεία πιθανοτήτων και μπορεί να είναι:

- **Λευκός θόρυβος** (white noise), ο οποίος ουσιαστικά είναι ένα σταθερό φάσμα, δηλαδή η ένταση του δεν μειώνεται με αύξηση της συχνότητας. Παρόλα αυτά, αποτελεί μια πρώτη και όχι ρεαλιστική προσέγγιση του θορύβου μιας εικόνας.



- **Γκαουσιανός θόρυβος** (Gaussian noise), ο οποίος αποτελεί μια ικανοποιητική προσέγγιση του πραγματικού θορύβου που σε πολλές περιπτώσεις προκύπτει κατά την επεξεργασία μιας εικόνας.

Επιπλέον, κατά την επεξεργασία του σχήματος μιας εικόνας, πολλές φορές απαιτείται η εφαρμογή μαθηματικών πράξεων πάνω στην καμπύλη του συνόρου του, με αποτέλεσμα, η επίδραση του θορύβου να πολλαπλασιάζεται και να παραμορφώνει την αρχική καμπύλη, και ιδιαίτερα στα σημεία όπου παρατηρείται μεγάλη καμπυλότητα.

Η *ομαλοποίηση της καμπύλης* (curve smoothing) έχει σκοπό να μειώσει την επίδραση του θορύβου, διατηρώντας ουσιαστικά τα καμπύλα τμήματα ανεπηρέαστα στη μορφή τους, ανεξάρτητα της όποιας επεξεργασίας μπορεί η καμπύλη να έχει υποστεί.

Σε θεωρητικό επίπεδο, αν  $f(t)$  είναι μια συνάρτηση ως προς το χρόνο (διακριτές τιμές), τότε η ομαλοποίηση είναι η εφαρμογή μιας συνέλιξης ενός πυρήνα (kernel) με την συνάρτηση  $f(t)$  με διακριτό παράθυρο  $NW$ . Το διακριτό παράθυρο  $NW$  στην πράξη δηλώνει ότι σε κάθε σημείο της καμπύλης χρησιμοποιούνται για την ομαλοποίηση  $NW$  προηγούμενα και  $NW$  επόμενα σημεία, δηλαδή ο πυρήνας της συνέλιξης έχει συνολικά μέγεθος  $(2*NW+1)$  σημεία.

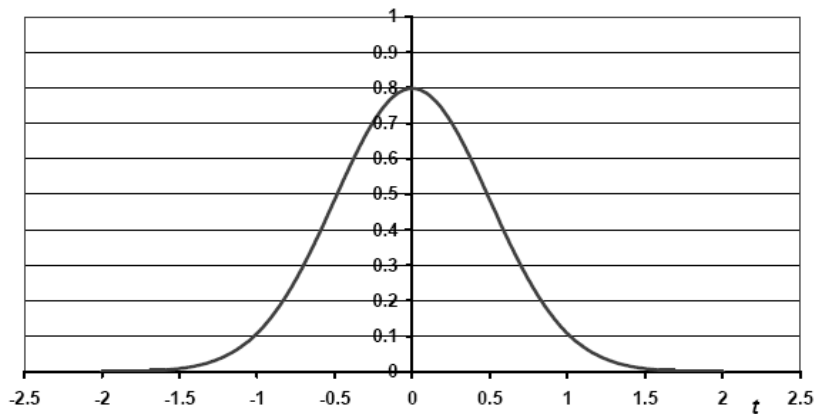
Για την ομαλοποίηση της καμπύλης ενός σχήματος μπορούν να χρησιμοποιηθούν διάφορα είδη συναρτήσεων – φίλτρων, όπως η ορθογώνια και η τριγωνική συνάρτηση, το φίλτρο ενδιάμεσης τιμής και άλλες. Η πλέον διαδεδομένη συνάρτηση που χρησιμοποιείται όμως στην ομαλοποίηση είναι η *Γκαουσιανή συνάρτηση* (Gaussian function). Έχει παρατηρηθεί ότι η Γκαουσιανή συνάρτηση μειώνει αρκετά ικανοποιητικά την επίδραση του θορύβου σε σύγκριση με άλλα φίλτρα, και για αυτό το λόγο θα την χρησιμοποιήσουμε στην CSS αναπαράσταση.

Θεωρητικά, η Γκαουσιανή συνάρτηση ορίζεται ως εξής:

$$g(t) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(t-\mu)^2}{2\sigma^2}\right]$$

όπου  $\mu$  η μέση τιμή και  $\sigma$  η τυπική απόκλιση.

Ακολουθεί μια σχηματική αναπαράσταση της Γκαουσιανής συνάρτησης:



**Σχήμα 1:** Γκαουσιανή συνάρτηση ομαλοποίησης ( $\mu=0$ ,  $\sigma=0.5$ )

Στην CSS αναπαράσταση, η ομαλοποίηση της καμπύλης προηγείται της μέτρησης καμπυλότητας με σκοπό να μειώσει όσο είναι δυνατόν τις επιδράσεις του θορύβου. Ουσιαστικά γίνεται συνέλιξη της καμπύλης με την Γκαουσιανή συνάρτηση για διαφορετικές τιμές του πλάτους  $\sigma$ , με αποτέλεσμα να προκύπτουν διαφορετικά επίπεδα ομαλοποίησης για την καμπύλη. Η διαδοχική αυτή ομαλοποίηση υλοποιεί την εξέλιξη της καμπύλης στον χρόνο (curve evolution).

Ο υπολογισμός ξεκινά εφαρμόζοντας την συνέλιξη κάθε συντεταγμένης της καμπύλης με την Γκαουσιανή συνάρτηση. Αν  $g(u, s)$  είναι η μονοδιάστατη Γκαουσιανή συνάρτηση πλάτους  $s$ , τότε οι μεταβλητές  $X(u, s)$  και  $Y(u, s)$  αναπαριστούν τις *ομαλοποιημένες* συντεταγμένες της καμπύλης και υπολογίζονται ως εξής:

$$X(u, s) = x(u) * g(u, s)$$

$$Y(u, s) = y(u) * g(u, s)$$

όπου  $*$  δηλώνει την συνέλιξη.

Σύμφωνα με τις ιδιότητες της συνέλιξης, οι παράγωγοι των  $X(u, s)$  και  $Y(u, s)$  μπορούν να υπολογιστούν εύκολα ως εξής:

$$\dot{X}(u, s) = x(u) * \dot{g}(u, s)$$

$$\ddot{X}(u, s) = x(u) * \ddot{g}(u, s)$$

Αντίστοιχα προκύπτουν οι εξισώσεις για τα  $\dot{Y}(u, s)$ ,  $\ddot{Y}(u, s)$

Με γνωστές λοιπόν την πρώτη και δεύτερη παράγωγο της Γκαουσιανής συνάρτησης προκύπτει η *καμπυλότητα* της ομαλοποιημένης καμπύλης ως εξής:

$$\kappa(u, s) = \frac{\dot{X}(u, s)\ddot{Y}(u, s) - \ddot{X}(u, s)\dot{Y}(u, s)}{(\dot{X}^2(u, s) + \dot{Y}^2(u, s))^{3/2}}$$

Όσο το  $s$  αυξάνει, το σχήμα της καμπύλης αλλάζει. Η διαδικασία αυτή καλείται *εξέλιξη της καμπύλης* (curve evolution).

Αν για την θέση της τυχαίας παραμέτρου  $u$  επιλέξουμε το μήκος τόξου, τότε η  $\Gamma(u)$  αποτελεί την *κανονική απεικόνιση* της καμπύλης (natural representation). Η απεικόνιση όμως αυτή της καμπύλης μεταβάλλεται στην περίπτωση affine μετασχηματισμών, με αποτέλεσμα η μορφή της αρχικής καμπύλης να διαφέρει από αυτή της κατά affine μετασχηματισμένης. Για το λόγο αυτό, το μήκος τόξου αντικαταστάθηκε από το affine μήκος τόξου, όπως αυτό ορίστηκε σε προηγούμενη ενότητα, και ουσιαστικά κανονικοποίησε την καμπύλη, ώστε αυτή να παραμένει αμετάβλητη σε affine μετασχηματισμούς.

### 4.3. Κατασκευή της CSS εικόνας

Η κατασκευή της CSS εικόνας βασίζεται στο θεωρητικό υπόβαθρο που παρουσιάστηκε στις παραπάνω ενότητες. Πιο συγκεκριμένα, κάθε αντικείμενο αναπαρίσταται από τις  $x$  και  $y$  συντεταγμένες των οριακών του σημείων. Στην περίπτωση μας, τα περιγράμματα (contours) των εικόνων που χρησιμοποιήσαμε βρίσκονται αποθηκευμένα σε μία βάση δεδομένων, ταξινομημένα σε κατηγορίες ανάλογα με το σχήμα τους.

Αρχικά, για να κανονικοποιήσουμε το μήκος τόξου, εφαρμόζουμε δειγματοληψία της καμπύλης του συνόρου σε  $L$  σημεία, και αναπαριστάμε εκ νέου το αντικείμενο. Στη συνέχεια, ομαλοποιούμε την καμπύλη με την βοήθεια της Γκαουσιανής συνάρτησης, και προκύπτει έτσι η ομαλοποιημένη καμπύλη της μορφής:

$$\Gamma_\sigma = (X(u, s), Y(u, s))$$

Η καμπύλη  $\Gamma_\sigma$  όπως φαίνεται παραπάνω, εκφράζεται συναρτήσει μιας παραμέτρου  $u$  και του πλάτους  $s$  της Γκαουσιανής συνάρτησης. Για να επιτύχουμε την αμεταβλητότητα της στην περίπτωση affine μετασχηματισμών, στην θέση της παραμέτρου  $u$  χρησιμοποιήσαμε το affine μήκος και μεταβάλλαμε το πλάτος  $s$  με σταθερό βήμα. Σαν αποτέλεσμα προέκυψε η εξέλιξη της καμπύλης  $\Gamma_\sigma$ , για διαφορετικές τιμές του  $s$ . Κατά την εξέλιξη της καμπύλης διαπιστώσαμε ότι με την αύξηση της τιμής του  $s$ , παρατηρείται μια διαδοχική ομαλοποίηση των καμπύλων τμημάτων του περιγράμματος του σχήματος, και μάλιστα για αρκετά υψηλές τιμές του  $s$ , η καμπύλη τείνει να γίνει τελείως κυρτή.

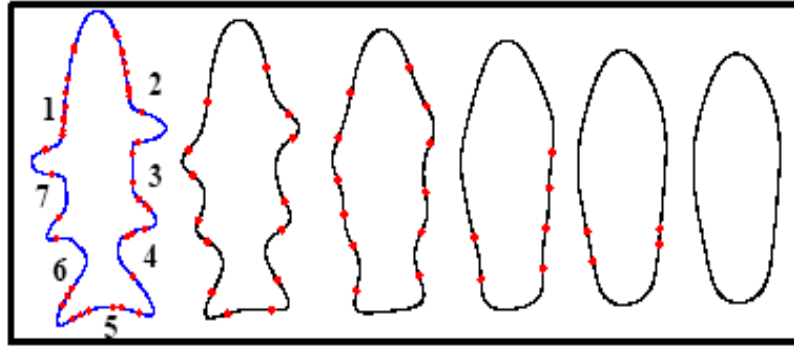
Στη συνέχεια, για κάθε στάδιο της εξέλιξης της καμπύλης, μηδενίζοντας την καμπυλότητα εντοπίσαμε τα σημεία μηδενισμού της καμπύλης, για τα οποία θα αναφερθούμε στην ενότητα που ακολουθεί.

#### 4.3.1 Σημεία μηδενισμού

Τα *σημεία μηδενισμού* (zero crossings) μιας καμπύλης προκύπτουν από το μηδενισμό της εξίσωσης καμπυλότητας σε κάθε στάδιο της εξέλιξης της καμπύλης. Πιο συγκεκριμένα, επιλύοντας την εξίσωση:

$$\kappa(u, s) = 0$$

προκύπτουν ζεύγη αριθμών  $(u, s)$ , για τις διαφορετικές τιμές του  $s$ . Τα ζεύγη αυτά μπορούμε να τα δούμε συνοπτικά στο παρακάτω σχήμα, όπου φαίνεται και η εξέλιξη της καμπύλης:

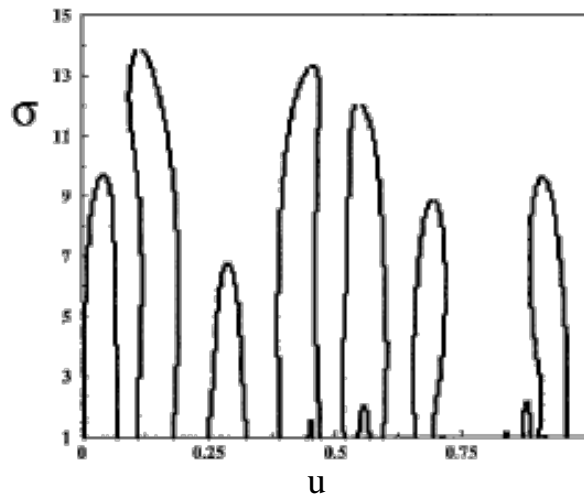


**Σχήμα 2:** Εξέλιξη της καμπύλης (curve evolution)

Παρατηρούμε ότι τα σημεία μηδενισμού βρίσκονται εκατέρωθεν ενός καμπύλου τμήματος και μάλιστα είναι περισσότερα στα σημεία όπου υπάρχει αλλαγή της καμπυλότητας. Κατά την εξέλιξη της διαδικασίας αυτά μειώνονται, κάτι που αναμένουμε άλλωστε, καθώς η καμπύλη συνεχώς ομαλοποιείται. Πιο συγκεκριμένα, στις περιοχές όπου ομαλοποιείται ένα καμπύλο τμήμα, τα σημεία μηδενισμού ενώνονται. Στο τέλος, και για μεγάλες τιμές του  $s$ , δεν συναντάμε πλέον σημεία μηδενισμού καθώς η καμπύλη έχει πλέον τελείως ομαλοποιηθεί.

Στην συνέχεια μπορούμε να απεικονίσουμε τα σημεία μηδενισμού σαν μία δυαδική εικόνα (binary image), που ονομάζεται κανονική CSS εικόνα (regular CSS image) της καμπύλης.

Η κανονική CSS εικόνα της καμπύλης φαίνεται ενδεικτικά στο παρακάτω σχήμα:



**Σχήμα 3:** Κανονικοποιημένη CSS απεικόνιση της καμπύλης (regular CSS image)

Στον κατακόρυφο άξονα βρίσκονται οι τιμές του πλάτους της Γκαουσιανής συνάρτησης  $\sigma$ , το οποίο συνεχώς αυξάνεται κατά την εξέλιξη της καμπύλης, και στον

οριζόντιο άξονα βρίσκονται οι τιμές της θέσης των σημείων της καμπύλης όπου συναντάμε τα σημεία μηδενισμού.

Έτσι η διασταύρωση κάθε οριζόντιας γραμμής με τα καμπύλα τμήματα της CSS εικόνας δείχνει τις θέσεις των σημείων μηδενισμού της αντίστοιχης εξελισσόμενης καμπύλης για διαφορετικές τιμές του  $\sigma$ .

Συγκρίνοντας τα δύο σχήματα (σχήμα 2 και 3) παρατηρούμε ότι η πληροφορία που μεταφέρει η εικόνα μεταφέρεται αυτούσια. Ειδικότερα, διαπιστώνουμε πως οι καμπύλες του αρχικού μας σχήματος (σχήμα 2) απεικονίζονται όλες στην CSS αναπαράσταση (σχήμα 3) και μάλιστα σε αντιστοιχία όσον αφορά το μέγεθος τους. Αυτό αποτελεί ένα από τα πλεονεκτήματα της μεθόδου, καθώς φαίνεται να διατηρεί την πληροφορία για τα τοπικά χαρακτηριστικά του αρχικού σχήματος.

#### **4.4. Εξαγωγή μεγίστων από την CSS εικόνα**

Η εξαγωγή των *μεγίστων σημείων* της CSS εικόνας αποτελεί σημαντικό μέρος της εφαρμογής της μεθόδου, καθώς τα μέγιστα αυτά σημεία έχουν την ικανότητα να χαρακτηρίσουν επαρκώς την εικόνα. Όπως φαίνεται από το σχήμα 3, κάθε CSS εικόνα παρουσιάζει καμπύλες, στις κορυφές των οποίων βρίσκονται οι θέσεις των μεγίστων σημείων μηδενισμού. Συνήθως οι μικρές καμπύλες αναπαριστούν πληροφορία για τον υπάρχοντα θόρυβο σε μια εικόνα, για αυτό το λόγο κρατάμε μόνο τα μέγιστα αυτά που αντιστοιχούν σε τιμές του  $\sigma$  μεγαλύτερες από ένα κατώφλι  $\tau$ , το οποίο εδώ ορίζεται ως το 0.2 της μέγιστης τιμής του  $\sigma$ . [64]

Τα μέγιστα που προκύπτουν αποτελούν ζεύγη αριθμών των τιμών  $\sigma$  και  $u$  της CSS εικόνας και είναι ανάλογα του αριθμού των καμπυλών του αρχικού σχήματος. Παραδείγματος χάριν, στο σχήμα 2 υπάρχουν 7 καμπύλες στην αρχική εικόνα, αντίστοιχα στο σχήμα 3 υπάρχουν 7 μέγιστα και συνεπώς η εικόνα θα αναπαρασταθεί – χαρακτηριστεί από 7 ζεύγη ακεραίων αριθμών, της μορφής

$$\{\sigma_i \quad u_i\}$$

όπου  $i=1, \dots, 7$  ο αριθμός των σημείων στην περίπτωση αυτή.

Ακολουθεί η παρουσίαση της μεθόδου εύρεσης των μεγίστων της CSS εικόνας, και πιο συγκεκριμένα γίνεται αναφορά στην ιδέα πάνω στην οποία βασίζεται η μέθοδος αυτή.

##### **Μέθοδος εύρεσης μεγίστων**

Οι θέσεις των μεγίστων δεν είναι διαθέσιμες, και για αυτό και πρέπει να εξαχθούν από την CSS εικόνα. Όπως φαίνεται και από το παραπάνω σχήμα, η CSS εικόνα αποτελείται από καμπύλα τμήματα τα οποία ενώνονται μεταξύ τους με σημεία, εκτός από τις περιοχές όπου παρουσιάζονται μέγιστα. Σε εκείνες τις περιοχές, η κλίση τείνει να μηδενιστεί. Αυτό έχει ως αποτέλεσμα, ακόμα και με καλύτερη δειγματοληψία και κανονικοποίηση της αρχικής καμπύλης, να μην είναι εύκολο να «κλείσουν» τα καμπύλα

τμήματα της CSS εικόνας σε αυτές τις περιοχές. Για να θεωρηθεί επομένως ένα ζευγάρι σημείων ως ζευγάρι μεγίστων, πρέπει να ικανοποιεί τις εξής προϋποθέσεις:

- Το ένα σημείο να βρίσκεται πολύ κοντά στο άλλο
- Να μην υπάρχει άλλο σημείο μηδενισμού σε υψηλότερη κλίμακα, δηλαδή για μεγαλύτερη τιμή του  $\sigma$  να μην συναντάμε άλλο σημείο μηδενισμού στην γειτονική του περιοχή.

Βασισμένοι στην παραπάνω λογική υλοποιήσαμε μια μέθοδο εύρεσης των μεγίστων η οποία ικανοποιούσε τα παραπάνω κριτήρια. Αναλυτική αναφορά στην μέθοδο αυτή γίνεται στην ενότητα επεξήγησης της υλοποίησης της CSS μεθόδου (βλ. Ενότητα 5.1).

## 4.5. Αλγόριθμος ταιριάσματος

Μετά την εξαγωγή των θέσεων των μεγίστων κάθε καμπύλης, ακολουθεί η διαδικασία του ταιριάσματος (*matching*). Ο αλγόριθμος ταιριάσματος είναι ιδιαίτερα χρήσιμος κατά την σύγκριση δύο CSS εικόνων, καθώς συγκρίνει δύο ομάδες μεγίστων, της αρχικής και της κατά affine μετασχηματισμένης καμπύλης.

Τα βήματα που ακολουθούμε συνοψίζονται παρακάτω:

- 1) Δημιουργούμε ένα κόμβο ο οποίος αποτελείται από τα μέγιστα εκείνα της αρχικής και της μετασχηματισμένης καμπύλης που έχουν την μέγιστη τιμή στον άξονα των  $s$  (τυπική απόκλιση  $\sigma$ ). Έχουμε δηλαδή ένα ζευγάρι μεγίστων της μορφής:

*{1<sup>ο</sup> μέγιστο κατά  $s$  της αρχικής καμπύλης, 1<sup>ο</sup> μέγιστο κατά  $s$  της μετασχηματισμένης καμπύλης}*

Αν υπάρχουν παραπάνω από ένα μέγιστα στην αρχική καμπύλη που έχουν την  $s$  – συντεταγμένη τους πολύ κοντά – σε ποσοστό 80% – σε σχέση με το αντίστοιχο μέγιστο της μετασχηματισμένης καμπύλης, τότε δημιουργούμε επιπλέον κόμβους της μορφής:

*{1<sup>ο</sup> «κοντινό» μέγιστο κατά  $s$  της αρχικής καμπύλης, 1<sup>ο</sup> μέγιστο κατά  $s$  της μετασχηματισμένης καμπύλης}*

*{2<sup>ο</sup> «κοντινό» μέγιστο κατά  $s$  της αρχικής καμπύλης, 1<sup>ο</sup> μέγιστο κατά  $s$  της μετασχηματισμένης καμπύλης}*

- 2) Αρχικοποιούμε το κόστος κάθε κόμβου ως την απόλυτη διαφορά των  $s$  – συντεταγμένων της αρχικής και μετασχηματισμένης καμπύλης. Υπολογίζουμε έπειτα την CSS παράμετρο ολίσθησης  $a$  (CSS shift parameter) για κάθε κόμβο ως εξής:

$$a = U_{or} - U_{aff}$$

όπου  $U$  είναι η οριζόντια συντεταγμένη ενός μεγίστου, και οι δείκτες  $or$  και  $aff$  αναφέρονται στην αρχική και μετασχηματισμένη καμπύλη αντίστοιχα.

### **Παρατήρηση:**

Η παράμετρος κυκλικής ολίσθησης  $a$  χρησιμοποιείται για να εξαλείψει την επίδραση της αλλαγής στον προσανατολισμό (orientation) αλλά και της ύπαρξης διαφορετικών αρχικών σημείων (starting points) της καμπύλης. Από την στιγμή που εφαρμόζουμε μια τέτοια ολίσθηση στην μετασχηματισμένη καμπύλη, αυτή

πλέον μπορεί να συγκριθεί και στην συνέχεια να ταιριαστεί με την αντίστοιχη αρχική (βλ. Σχήματα 4 και 5).

Μια υπόθεση είναι ότι η αποτελεσματικότερη ολίσθηση είναι αυτή που εξισώνει την οριζόντια συντεταγμένη του 1<sup>ου</sup> μεγίστου της αρχικής καμπύλης – όπως αυτό ορίστηκε παραπάνω – με την οριζόντια συντεταγμένη του αντίστοιχου 1<sup>ου</sup> μεγίστου της μετασχηματισμένης καμπύλης. Παρόλα αυτά, στην πράξη, ο μόνος τρόπος να βρούμε την καλύτερη ολίσθηση είναι να εκτιμήσουμε όλα τα κόστη ταιριάσματος για κάθε πιθανή ολίσθηση, και να επιλέξουμε το μικρότερο εξ αυτών.

3) Εφόσον έχουμε ήδη ταιριάξει τα πρώτα μέγιστα της αρχικής και της μετασχηματισμένης καμπύλης συνεχίζουμε την διαδικασία ταιριάσματος για το επόμενο μέγιστο. Δημιουργούμε έτσι ανάλογους κόμβους για το 2<sup>ο</sup> μέγιστο κατά  $s$  της μετασχηματισμένης καμπύλης υπολογίζοντας την *Ευκλείδεια απόσταση* του από τα υπόλοιπα μέγιστα της αρχικής καμπύλης. Τα ήδη ταιριασμένα μέγιστα αγνοούνται.

4) Από τις Ευκλείδειες αποστάσεις επιλέγουμε την μικρότερη, και το ζευγάρι των μεγίστων της αρχικής καμπύλης που αντιστοιχεί σε αυτή την απόσταση, το ταιριάζουμε με το 2<sup>ο</sup> μέγιστο της μετασχηματισμένης καμπύλης που εξετάζουμε. Δυο μέγιστα θεωρούνται ταιριασμένα εφόσον βρίσκονται κοντά στον οριζόντιο άξονα, δηλαδή η οριζόντια απόσταση τους είναι μικρότερη από το 0.2 της μέγιστης οριζόντιας απόστασης.

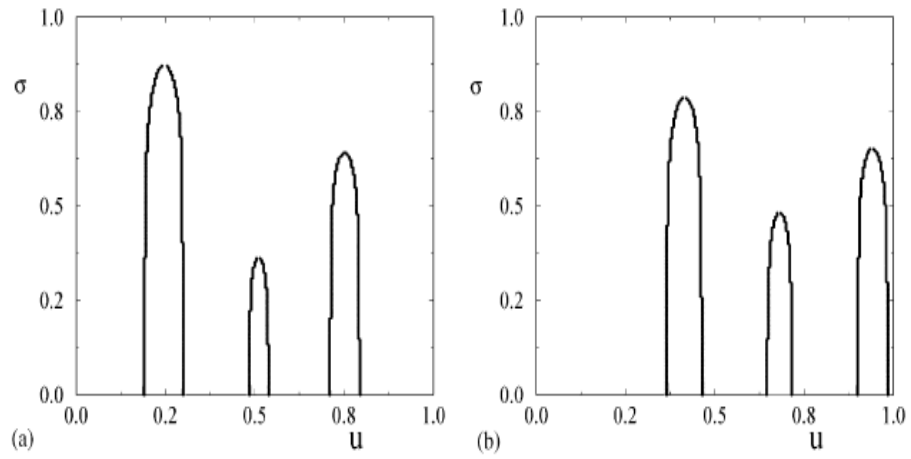
Το κόστος του ταιριάσματος ορίζεται ως την απόσταση των δύο μεγίστων.

Σε κάθε άλλη περίπτωση, όταν δηλαδή το μέγιστο της μετασχηματισμένης καμπύλης δεν θεωρείται ταιριασμένο, το κόστος ταιριάσματος ορίζεται ως το ύψος ( $s$  – συντεταγμένη) του μεγίστου αυτού.

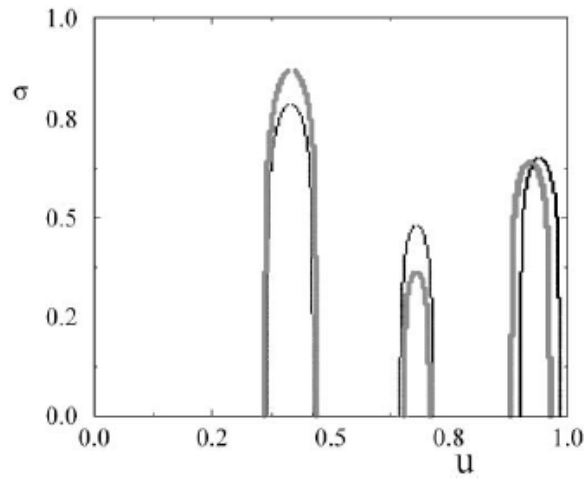
5) Η διαδικασία επαναλαμβάνεται μέχρι ότου όλα τα μέγιστα της μετασχηματισμένης καμπύλης εξεταστούν και ταιριαστούν με εκείνα της αρχικής.

6) Το ολικό κόστος του ταιριάσματος είναι το άθροισμα των αποστάσεων των ταιριασμένων μεγίστων. Σε αυτό προστίθεται το κόστος των μη ταιριασμένων μεγίστων της μετασχηματισμένης καμπύλης. Επιπλέον, στην περίπτωση όπου ο αριθμός των μεγίστων της μετασχηματισμένης καμπύλης είναι μεγαλύτερος από αυτόν της αρχικής, στο κόστος προστίθεται και το ύψος ( $s$  – συντεταγμένη) των υπολειπόμενων μεγίστων.





**Σχήμα 4:** Παράδειγμα CSS εικόνων που πρέπει να ταιριαστούν



**Σχήμα 5:** Εφαρμογή κυκλικής ολίσθησης για το ταίριασμα των εικόνων του σχ.4. Σε αυτό το παράδειγμα, η καλύτερη ολίσθηση επιτυγχάνεται με ταίριασμα των πρώτων μεγίστων των εικόνων.

## 5. Πειράματα, αποτελέσματα, παρατηρήσεις

### 5.1 Υλοποίηση της εφαρμογής

Η μέθοδος της CSS υλοποιήθηκε σε γλώσσα προγραμματισμού Matlab. Η κύρια συνάρτηση του προγράμματος αποτελεί μια εφαρμογή της παραπάνω θεωρητικής ανάλυσης και αποτελείται από επιμέρους συναρτήσεις όπου κάθε μια ξεχωριστά υλοποιεί τα αντίστοιχα επιμέρους θεωρητικά μέρη της.

Θα αναφερθούμε συγκεκριμένα στην υλοποίηση για την εύρεση των μεγίστων μιας CSS εικόνας. Ο τρόπος που υλοποιήσαμε την συνάρτηση για την εύρεση μεγίστων, βασίζεται στην ιδέα ότι η CSS εικόνα μπορεί να αναπαρασταθεί από ένα *αραιό πίνακα*  $H$ , που έχει μονάδες στις θέσεις των σημείων μηδενισμού, και μηδενικά σε όλες τις υπόλοιπες. Με αυτό τον τρόπο, μπορέσαμε έπειτα να εξετάσουμε κάθε στοιχείο του πίνακα, και με βάση κριτήρια για την εύρεση μεγίστων να εντοπίσουμε τα ζητούμενα μέγιστα.

Αξίζει επίσης να αναφερθεί ότι για την υλοποίηση της Γκαουσιανής συνάρτησης χρησιμοποιήσαμε την συνάρτηση της Matlab *fspecial*, η οποία δημιουργεί μια Γκαουσιανή συνάρτηση με παραμέτρους *hsize*, *s* που αποτελούν το μέγεθος παραθύρου και την τυπική απόκλιση αντίστοιχα. Οι παράμετροι αυτοί ορίζονται σαν εξωτερικές μεταβλητές, και μπορούν να δοθούν κάθε φορά από τον χρήστη.

Τέλος, με σκοπό οι καμπύλες που εξετάζουμε να διαθέτουν κοινό πλήθος σημείων περιγράμματος, ώστε να είμαστε σε θέση αργότερα να τις ταιριάξουμε, εφαρμόσαμε μια πρώτη κανονικοποίηση στις συντεταγμένες των υπό εξέταση καμπυλών με βάση το μήκος τόξου (*arc length*). Στην συνέχεια εφαρμόσαμε κανονικά την κανονικοποίηση με βάση το *affine* μήκος, όπως φαίνεται και στα σχήματα της ενότητας που ακολουθεί.

### 5.2 Πειράματα και αποτελέσματα

#### 5.2.1 Παρουσίαση ενδεικτικών αποτελεσμάτων για την εφαρμογή της μεθόδου

Για την εκτέλεση των πειραμάτων χρησιμοποιήσαμε μια βάση δεδομένων με περιγράμματα ταξινομημένα σε κατηγορίες ανάλογα με το σχήμα τους, π. χ. Αυτοκίνητα, αεροπλάνα, ποτήρια κτλ.

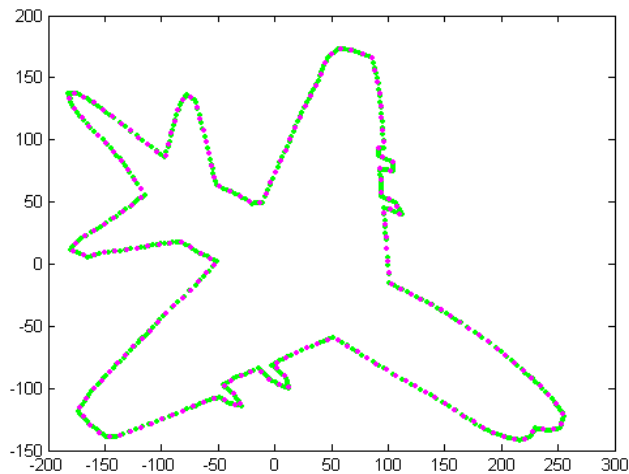
Αρχικά, εξετάσαμε την αποτελεσματικότητα της μεθόδου στην περίπτωση των *affine* μετασχηματισμών. Έτσι, δώσαμε ως είσοδο την αρχική καμπύλη, και μετά από την εφαρμογή τυχαίων *affine* παραμέτρων, πήραμε την κατά *affine* μετασχηματισμένη καμπύλη. Στην συνέχεια εφαρμόσαμε την μέθοδο της CSS για το ταίριασμα των δύο καμπυλών. Το πείραμα εφαρμόστηκε σε *affine* μετασχηματισμένες καμπύλες που ανήκουν στην ίδια αλλά και σε διαφορετικές κατηγορίες.

Σαν ενδεικτικό παράδειγμα αναφέρουμε την περίπτωση ενός μετασχηματισμένου κατά affine σχήματος και τα αποτελέσματα που λάβαμε κατά την εφαρμογή της CSS μεθόδου.

Έστω ότι επιλέγουμε από την κατηγορία των αεροπλάνων την εικόνα του A2.

Το πλήθος  $L$  των σημείων, επιλέγεται στην τιμή 200 [1], αν και κατά την διάρκεια των πειραμάτων έχουμε την δυνατότητα να μεταβάλλουμε την τιμή του.

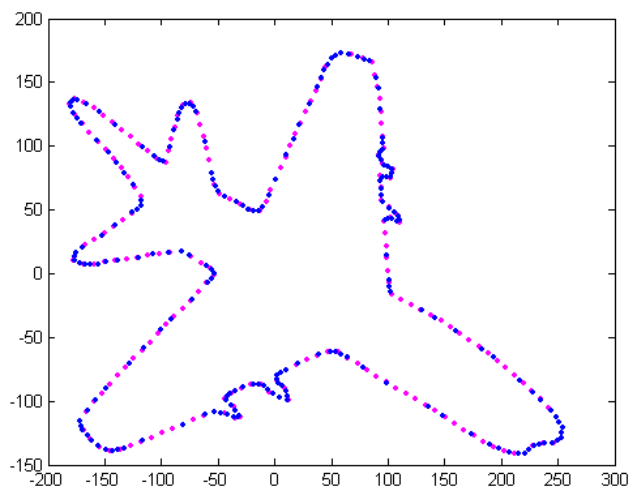
Μετά την αρχική δειγματοληψία με βάση το μήκος τόξου, λαμβάνουμε την εξής εικόνα:



**Σχήμα 5-1:** Κανονικοποίηση καμπύλης με βάση το μήκος τόξου.

Με το πράσινο χρώμα δείχνουμε την αρχική μας καμπύλη, ενώ με το μοβ την καμπύλη μετά από μια πρώτη δειγματοληψία με βάση το μήκος τόξου σε  $L=200$  σημεία.

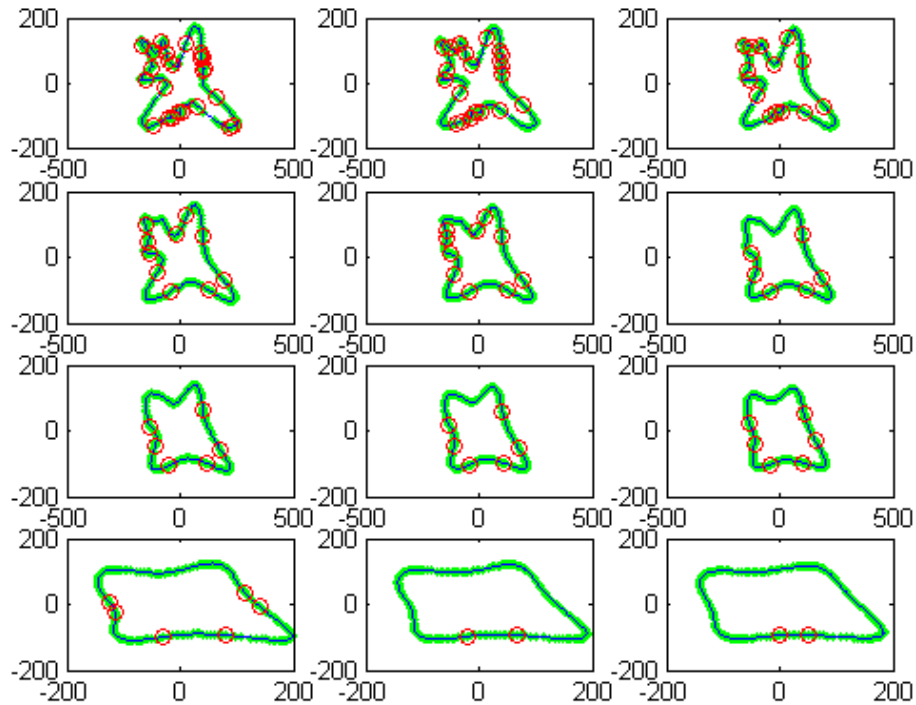
Ακολουθεί η αντίστοιχη εικόνα με την δειγματοληψία με βάση το affine length



**Σχήμα 5-2:** Κανονικοποίηση καμπύλης με βάση το affine μήκος.

Με μοβ χρώμα δείχνουμε την κανονικοποιημένη με βάση το μήκος τόξου καμπύλη, ενώ με το μπλε χρώμα την καμπύλη που προκύπτει μετά την κατά affine μήκος δειγματοληψία.

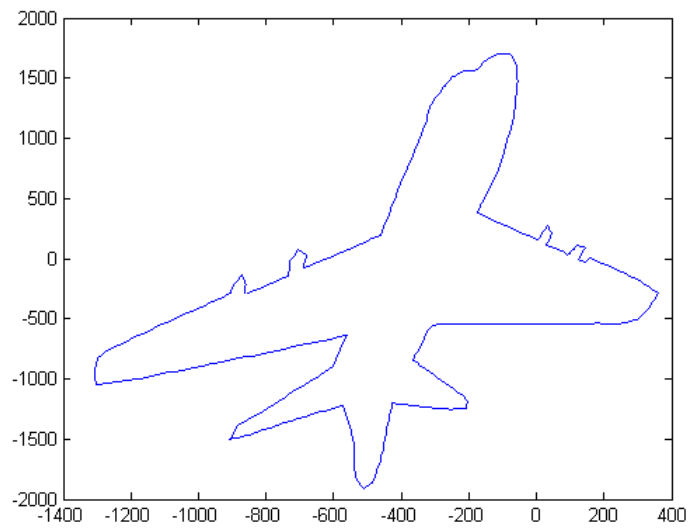
Στην συνέχεια παρουσιάζεται η εξέλιξη της καμπύλης κατά την ομαλοποίηση:



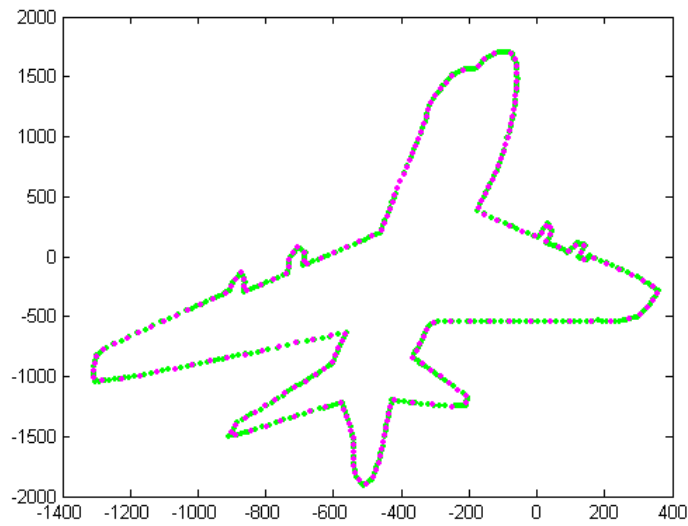
**Σχήμα 5-3:** Εξέλιξη της καμπύλης (curve evolution)

Ξεκινώντας από την τιμή  $s=0.1$  και με βήμα 1 παρατηρούμε ότι αρχικά υπάρχουν πολλά σημεία μηδενισμού (σημειώνονται με κόκκινα κυκλάκια) ενώ όσο η ομαλοποίηση εξελίσσεται αυτά μειώνονται και φτάνουν στο τελευταίο στάδιο να μένουν μόνο δύο, καθώς η ομαλοποίηση έχει σχεδόν ολοκληρωθεί (το επόμενο βήμα θα ήταν να μην εμφανίζεται κανένα σημείο μηδενισμού).

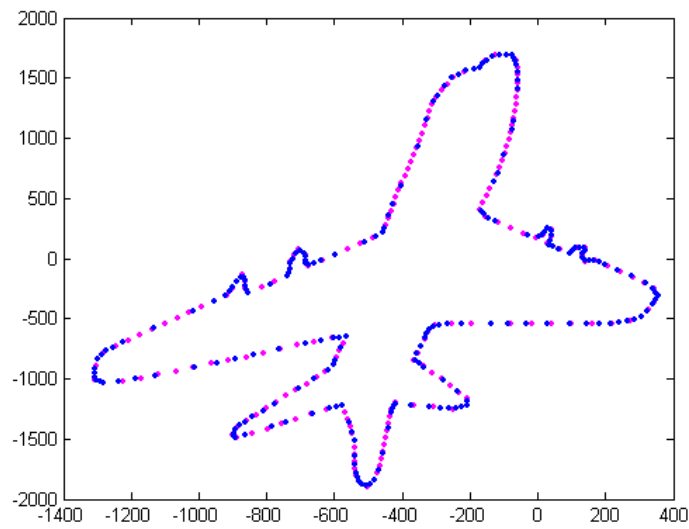
Ανάλογες εικόνες προκύπτουν και για την affine μετασχηματισμένη καμπύλη που παρουσιάζονται συνοπτικά ως εξής:



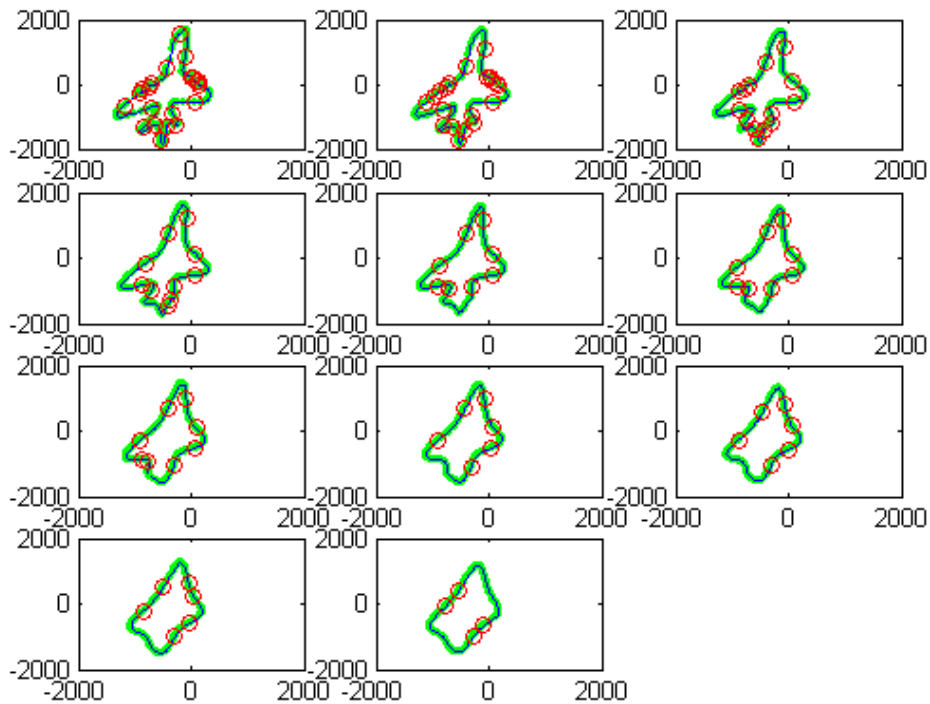
**Σχήμα 5-4:** Affine μετασχηματισμένη καμπύλη



**Σχήμα 5-5:** Κανονικοποίηση καμπύλης με βάση το μήκος τόξου.

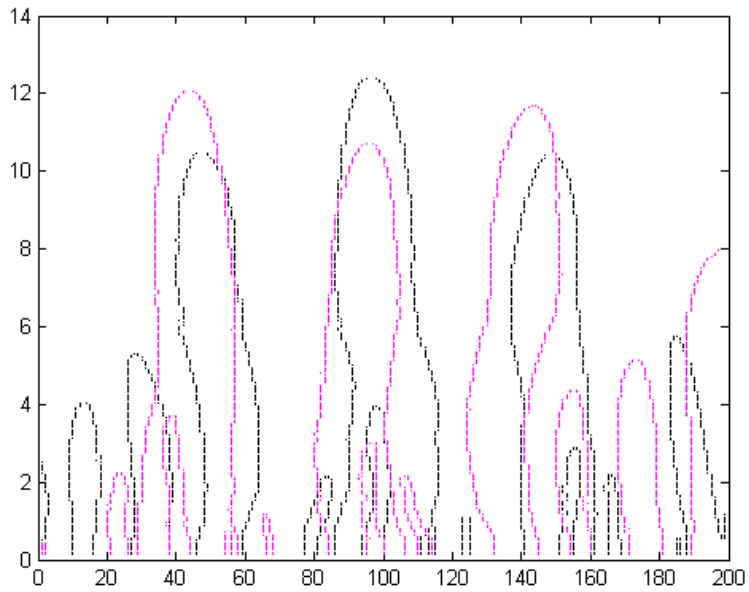


**Σχήμα 5-6:** Κανονικοποίηση καμπύλης με βάση το affine μήκος.



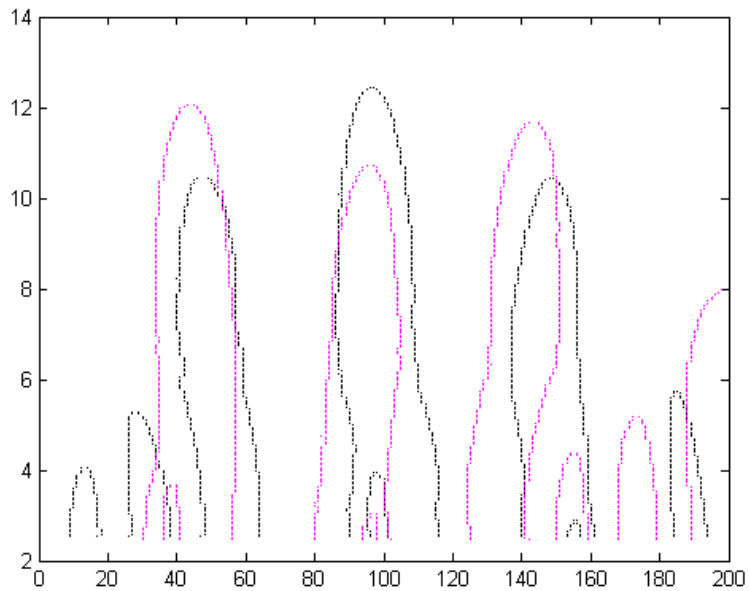
**Σχήμα 5-7:** Εξέλιξη της καμπύλης (curve evolution)

Στην συνέχεια προκύπτει η CSS εικόνα η οποία έχει την ακόλουθη μορφή:



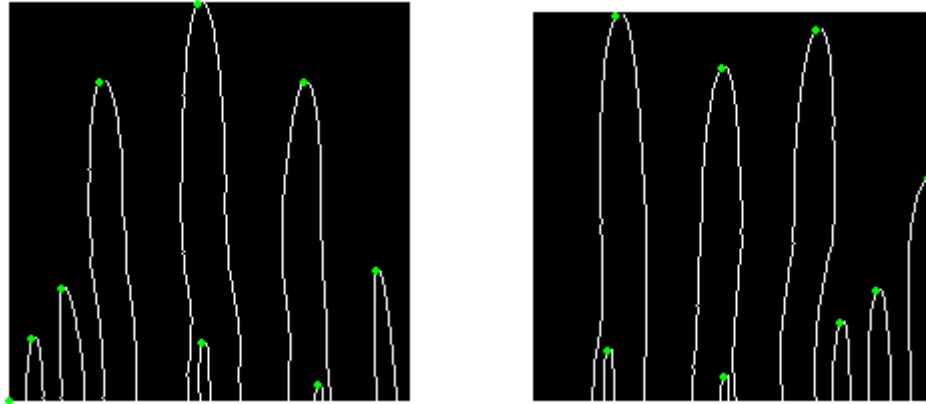
**Σχήμα 5-8:** CSS απεικόνιση αρχικής και μετασχηματισμένης καμπύλης

Με μαύρο χρώμα αναπαριστούμε την CSS εικόνα της αρχικής μας καμπύλης, ενώ με μοβ χρώμα την CSS εικόνα της μετασχηματισμένης καμπύλης. Όπως ήδη αναφέραμε, τα μέγιστα που μας ενδιαφέρουν είναι αυτά με τιμή μεγαλύτερη από ένα *κατώφλι*  $\tau$ , η τιμή του οποίου έχει ορισθεί στο 0.2 της μέγιστης τιμής του  $s$ , όποτε προκύπτει η παρακάτω εικόνα:



**Σχήμα 5-9:** CSS απεικόνιση αρχικής και μετασχηματισμένης καμπύλης μετά την εφαρμογή κατωφλίου  $\tau = 0.2$

Τέλος, παραθέτουμε την αναπαράσταση των δύο εικόνων όπως αυτές έχουν αποτυπωθεί σε έναν δυαδικό αραιό πίνακα, έχοντας σημειώσει τις τιμές των μεγίστων που βρήκαμε με την μέθοδο που προαναφέραμε:



**Σχήμα 5-10:** CSS απεικόνιση αρχικής και μετασχηματισμένης καμπύλης μετά την εύρεση των μεγίστων

### 5.2.2 Πειράματα με affine μετασχηματισμούς της ίδιας αρχικής καμπύλης της ίδιας κατηγορίας σχήματος

Επιλέγοντας από την κατηγορία των αεροπλάνων την εικόνα του αεροπλάνου A\_F15 σαν αρχική καμπύλη, υπολογίζουμε το *κόστος ταιριάσματος* (matching\_cost) που προκύπτει σε σχέση με 10 affine μετασχηματισμένες καμπύλες της αρχικής (οι affine παράμετροι εφαρμόζονται τυχαία). Στη συνέχεια συνοψίζουμε τα αποτελέσματα στον ακόλουθο πίνακα:



Καμπύλες	OR	A_F15#1	A_F15#2	A_F15#3	A_F15#4	A_F15#5	A_F15#6	A_F15#7	A_F15#8	A_F15#9	A_F15#10
OR	0,00	38,32	26,63	26,13	38,23	15,95	37,55	39,37	9,08	37,89	40,63
A_F15#1	38,32	0,00	78,01	59,41	28,21	29,60	6,53	27,90	35,97	23,31	9,06
A_F15#2	26,63	78,01	0,00	52,25	134,89	65,00	68,85	133,52	57,21	94,76	73,38
A_F15#3	26,13	59,41	52,25	0,00	35,70	34,06	60,08	41,39	42,44	12,41	22,55
A_F15#4	38,23	28,21	134,89	35,70	0,00	35,79	16,46	3,60	35,31	19,31	11,23
A_F15#5	15,95	29,60	65,00	34,06	35,79	0,00	29,03	66,57	8,47	28,13	32,06
A_F15#6	37,55	6,53	68,85	60,08	16,46	29,03	0,00	30,30	35,08	18,98	16,94
A_F15#7	39,37	27,90	133,52	41,39	3,60	66,57	30,30	0,00	38,49	27,42	12,36
A_F15#8	9,08	35,97	57,21	42,44	35,31	8,47	35,08	38,49	0,00	35,89	38,86
A_F15#9	37,89	23,31	94,76	12,41	19,31	28,13	18,98	27,42	35,89	0,00	16,65
A_F15#10	40,63	9,06	73,38	22,55	11,23	32,06	16,94	12,36	38,86	16,65	0,00

Πίνακας 1: Αποτελέσματα ταιριάσματος της καμπύλης A\_F15 με 10 μετασχηματισμένες κατά affine καμπύλες της.

Τα αποτελέσματα του παραπάνω πειράματος μας δίνουν το κόστος ταιριάσματος μεταξύ αρχικών και (affine) μετασχηματισμένων καμπυλών, καθώς και το κόστος μεταξύ διαφορετικών μετασχηματισμένων καμπυλών.

Η μέση τιμή των τιμών κόστους προκύπτει:

average	34,60
---------	-------

Από τον πίνακα 1 βλέπουμε ότι η τιμή ταιριάσματος μεταξύ των ίδιων καμπυλών (στοιχεία κύριας διαγωνίου) είναι μηδενική (οι καμπύλες ταιριάζουν απόλυτα).

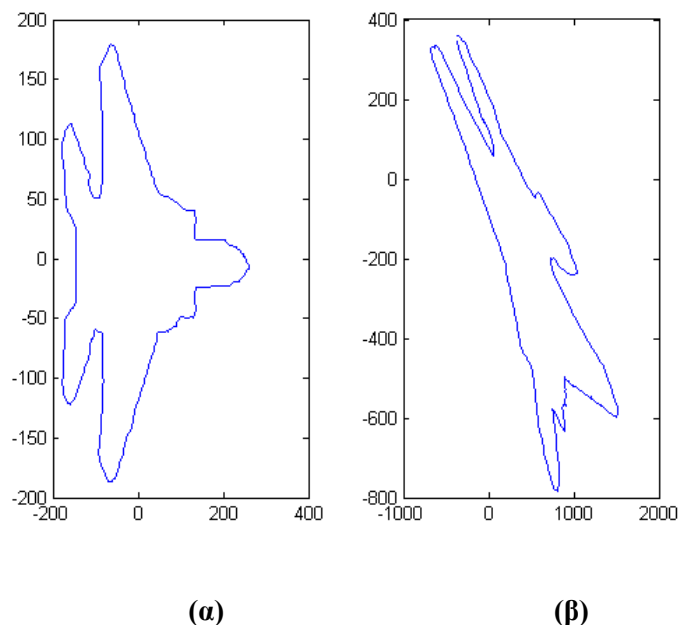
Οι τιμές ταιριάσματος μεταξύ διαφορετικών affine καμπυλών είναι κατά κύριο λόγο λίγο μεγαλύτερες από αυτές που προέκυψαν για την αρχική καμπύλη και τους affine μετασχηματισμούς της, κάτι που δείχνει την αποτελεσματικότητα της μεθόδου όσον αφορά την σταθερότητα της σε affine μετασχηματισμούς της αρχικής καμπύλης.

Πιο συγκεκριμένα παραθέτουμε δύο τέτοιες ενδεικτικές περιπτώσεις ταιριάσματος:

- **Περίπτωση 1η:**  $matching\_cost=9,08$

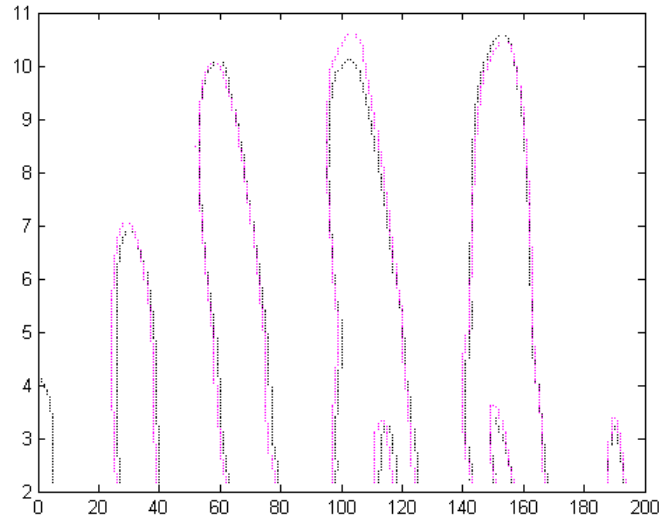
Στην περίπτωση αυτή εξετάζουμε τα αποτελέσματα της μεθόδου για την αρχική καμπύλη του αεροπλάνου A\_F15 και μιας μετασχηματισμένης κατά affine, της A\_F15#8.

Οι καμπύλες είναι της μορφής:



**Σχήμα 5-11:** (α) αρχική καμπύλη A\_F15, (β) μετασχηματισμένη καμπύλη A\_F15#8

Ακολουθεί η CSS αναπαράσταση των δύο καμπυλών:



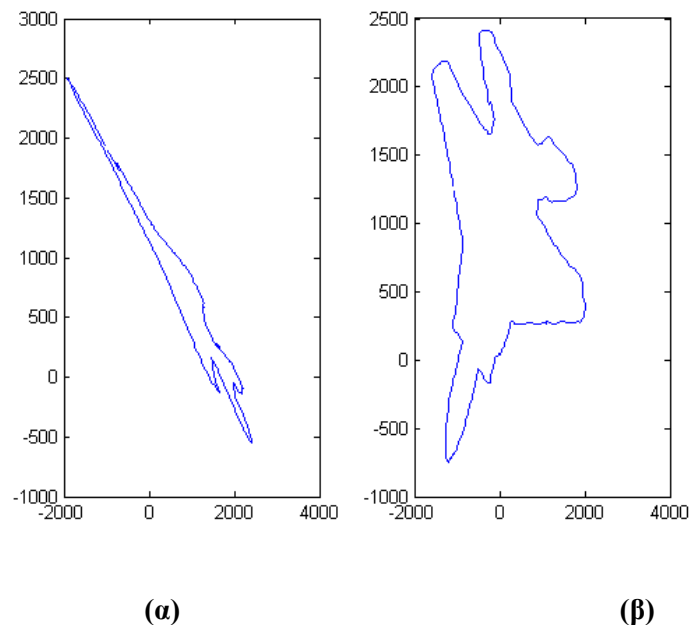
**Σχήμα 5-12:** CSS απεικόνιση των δύο καμπυλών

Παρατηρούμε ότι αυτές συμπίπτουν σχεδόν η μια πάνω στην άλλη, σε μικρούς αλλά και μεγαλύτερους λοβούς. Από το σχήμα 5-11 παρατηρούμε ότι η μετασχηματισμένη καμπύλη έχει υποστεί περιστροφή, αλλαγή στην κλίμακα και μετάθεση σε σχέση με την αρχική καμπύλη. Συμπεραίνουμε επομένως πως η μέθοδος της CSS έχει ικανοποιητική απόδοση σε περιπτώσεις παρουσίας affine μετασχηματισμών.

- **Περίπτωση 2η:**  $matching\_cost=134,89$

Στην περίπτωση αυτή εξετάζουμε τα αποτελέσματα της μεθόδου για δύο affine μετασχηματισμένες καμπύλες της ίδιας αρχικής.

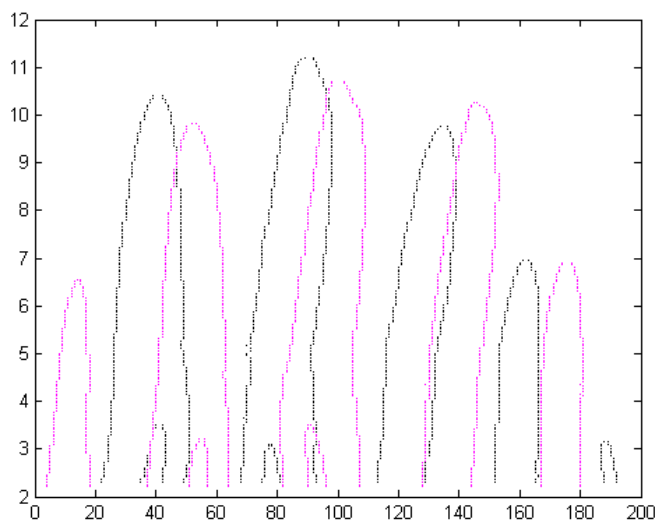
Οι καμπύλες είναι της μορφής:



**Σχήμα 5-13:** (α) affine καμπύλη A\_F15#2, (β) affine καμπύλη A\_F15#4

Από το σχήμα 2 βλέπουμε ότι οι μετασχηματισμένες καμπύλες διαφέρουν αρκετά όσον αφορά τις affine παραμέτρους τους. Πιο συγκεκριμένα, διαφέρουν κυρίως ως προς την περιστροφή αλλά και ως προς τον προσανατολισμό τους.

Ακολουθεί η CSS αναπαράσταση των δύο καμπυλών:



**Σχήμα 5-14:** CSS απεικόνιση των δύο καμπυλών

Από την CSS απεικόνιση παρατηρούμε την απόκλιση των δύο CSS εικόνων, όχι σε μεγάλο βαθμό καθώς προέρχονται από την ίδια αρχική καμπύλη. Η απόκλιση αυτή οφείλεται στην διαφορά των δύο μετασχηματισμένων καμπυλών, κυρίως όσον αφορά τον προσανατολισμό, γεγονός που απαιτεί την εφαρμογή κατάλληλης κυκλικής ολίσθησης.

Είναι φανερό επομένως, ότι οι εικόνες μπορούν να ταιριάξουν με την εφαρμογή κατάλληλης κυκλικής ολίσθησης της μιας εικόνας πάνω στην άλλη. Η κυκλική ολίσθηση αποτελεί σημαντική παράμετρο, στις οποίας την επίδραση θα αναφερθούμε στην ενότητα των παρατηρήσεων (5.3).

### 5.2.3 Πειράματα με affine μετασχηματισμούς διαφορετικής αρχικής καμπύλης της ίδιας κατηγορίας σχήματος

Επιλέγοντας την κατηγορία των αεροπλάνων (A\_F15, Jas7, A2, Eagle09, Tomcat1) δημιουργούμε αντίστοιχους πίνακες με αυτόν της προηγούμενης ενότητας για διαφορετικές αρχικές καμπύλες της κατηγορίας αυτής και υπολογίζουμε το κόστος ταιριάσματος.

Τα αποτελέσματα συνοψίζονται στον ακόλουθο πίνακα:

Καμπύλες	Jas7	A2	Tomcat1	Eagle09	A_F15
Jas7	44,31	55,69	55,85	56,41	58,23
A2	55,69	40,62	56,37	56,42	63,53
Tomcat1	55,85	56,37	45,57	56,14	65,86
Eagle09	56,41	56,42	56,14	46,37	59,53
A_F15	58,23	63,53	65,86	59,53	34,60

**Πίνακας 2:** Αποτελέσματα ταιριάσματος διαφορετικών καμπυλών της ίδιας κατηγορίας σχήματος

Τα στοιχεία του πίνακα – εκτός της κύριας διαγωνίου του – προέκυψαν με ταίριασμα μεταξύ affine μετασχηματισμένων καμπυλών διαφορετικών αρχικών καμπυλών.

Από τα αποτελέσματα που προέκυψαν, παρατηρούμε ότι όσον αφορά το ταίριασμα των μετασχηματισμένων καμπυλών διαφορετικής αρχικής καμπύλης, οι τιμές του πίνακα είναι παραπλήσιες, όπως άλλωστε περιμέναμε, και λίγο μεγαλύτερες από εκείνες που προέκυψαν από το ανάλογο ταίριασμα για την ίδια αρχική καμπύλη. Γεγονός όμως είναι ότι η τιμή του ταιριάσματος παραμένει σταθερή σε μια τιμή, όπως προκύπτει και από το μέσο όρο του Πίνακα 2:

<b>average</b>	55,18
----------------	-------

Μπορούμε επομένως να πούμε ότι η μέθοδος παρουσιάζει ανάλογη συμπεριφορά για καμπύλες της ίδιας κατηγορίας σχήματος, και είναι σε θέση να καθορίσει αν μια affine μετασχηματισμένη καμπύλη ανήκει σε μια κατηγορία σχήματος ή όχι σε ικανοποιητικό βαθμό.

#### 5.2.4 Πειράματα με affine μετασχηματισμούς διαφορετικής αρχικής καμπύλης διαφορετικής κατηγορίας σχήματος

Στο πείραμα αυτό εξετάζουμε την αποτελεσματικότητα της μεθόδου για καμπύλες διαφορετικών κατηγοριών σχήματος. Επιλέγουμε τις κατηγορίες αεροπλάνα, αυτοκίνητα, ποτήρια και 3 διαφορετικές αρχικές καμπύλες αυτών. Στην συνέχεια δημιουργούμε μια σειρά affine μετασχηματισμών τους και εφαρμόζουμε την CSS μέθοδο ταιριάσματος.

Τα αποτελέσματα συνοψίζονται στον ακόλουθο πίνακα:

Κατηγορίες	Αεροπλάνα	Αυτοκίνητα	Ποτήρια
Αεροπλάνα	55,18	75,26	82,33
Αυτοκίνητα	75,26	52,11	73,44
Ποτήρια	82,33	73,44	59,67

**Πίνακας 3:** Αποτελέσματα ταιριάσματος διαφορετικών καμπυλών διαφορετικής κατηγορίας σχήματος

Η μέση τιμή του κόστους ταιριάσματος του παραπάνω πίνακα προκύπτει:

<b>average</b>	69,89
----------------	-------

Οι τιμές του κόστους που προέκυψαν ήταν οι αναμενόμενες, καθώς πλέον έχουμε να κάνουμε με καμπύλες διαφορετικών σχημάτων, με αποτέλεσμα να αυξάνεται και η τιμή του κόστους ταιριάσματος.

### 5.3 Συμπεράσματα

Από τα παραπάνω τρία διαφορετικά πειράματα διαπιστώσαμε την αποδοτικότητα της CSS μεθόδου. Πιο συγκεκριμένα, προέκυψαν τα εξής:

1. Η μέθοδος CSS φαίνεται να αντιμετωπίζει με αποτελεσματικό τρόπο την περίπτωση των τυχαίων affine μετασχηματισμών. Πιο συγκεκριμένα η παραμετροποίηση που εφαρμόζει με βάση το affine μήκος, επιτυγχάνει την διατήρηση της αρχικής μορφής της καμπύλης, ανεξαρτήτως των όποιων affine μετασχηματισμών μπορεί να έχει αυτή υποστεί.
2. Το κόστος ταιριάσματος ξεκινά από μια τιμή για το ταίριασμα affine μετασχηματισμένων καμπυλών που προέρχονται από την ίδια αρχική καμπύλη και αυξάνει σε περιπτώσεις διαφορετικών καμπυλών της ίδιας κατηγορίας. Τέλος, στην περίπτωση διαφορετικών κατηγοριών διαφοροποιείται αρκετά. Συνεπώς, η CSS μέθοδος μπορεί να χρησιμοποιηθεί με επιτυχία στην περίπτωση ανάκτησης ομοίων εικόνων από μια βάση δεδομένων.
3. Η μέθοδος εφαρμόζεται γρήγορα σε μεγάλο αριθμό καμπυλών (περιγραμμάτων) αποθηκευμένων σε μια βάση δεδομένων, δεν παρουσιάζει πολυπλοκότητα στους υπολογισμούς και έτσι ελαχιστοποιεί την πιθανότητα σφάλματος.
4. Όπως φαίνεται και από τις CSS απεικονίσεις, η CSS μέθοδος καταφέρνει να διατηρεί τις τοπικές πληροφορίες της υπό εξέταση εικόνας.

Κάθε κοίλο και κυρτό σημείο του σχήματος έχει το αντίστοιχο περίγραμμα στην CSS απεικόνιση, όπως και κάθε σημείο στον οριζόντιο άξονα της CSS απεικόνισης έχει το αντίστοιχο στο περίγραμμα της εικόνας.

Παρόλα αυτά, σε ορισμένες περιπτώσεις παρατηρήσαμε πως η CSS μέθοδος μπορεί να παρουσιάσει μειονεκτήματα, που μειώνουν την αποδοτικότητα της. Πιο συγκεκριμένα, οι διαφορετικές τιμές κόστους που προκύπτουν κατά την διαδικασία του ταιριάσματος οφείλονται πιθανώς στους εξής παράγοντες:

- **Στην επιλογή της παραμέτρου κυκλικής ολίσθησης shift.**

Όπως προαναφέραμε στον αλγόριθμο ταιριάσματος, ο μόνος τρόπος να βρούμε την καλύτερη ολίσθηση είναι να εκτιμήσουμε όλα τα κόστη ταιριάσματος για κάθε πιθανή ολίσθηση, και να επιλέξουμε το μικρότερο εξ αυτών. Επειδή κάτι τέτοιο όμως απαιτούσε πολυπλοκότητα στους υπολογισμούς και κατά συνέπεια θα αύξανε η πιθανότητα σφάλματος, ακολουθήσαμε την εξής λογική:

Αρχικά, υπολογίσαμε όλες τις πιθανές Ευκλείδειες αποστάσεις μεταξύ των δύο ομάδων μεγίστων των υπό εξέταση καμπυλών. Στη συνέχεια, επιλέξαμε την μικρότερη Ευκλείδεια απόσταση και βάση αυτής πραγματοποιήσαμε την ολίσθηση. Πιο συγκεκριμένα, οι συντεταγμένες των ζευγών μεγίστων που της αντιστοιχούν, χρησιμοποιήθηκαν για τον υπολογισμό της παραμέτρου ολίσθησης.

Η μέθοδος αυτή φάνηκε να αποδίδει, υπήρχαν όμως και περιπτώσεις όπου η επιλεγμένη ολίσθηση δεν ανταποκρινόταν στην καλύτερη δυνατή, με αποτέλεσμα τα υψηλά κόστη ταιριάσματος.

- **Στις διαφορετικές τιμές των affine παραμέτρων.**

Ο παράγοντας αυτός υπεισέρχεται στο ταιρίασμα μεταξύ affine μετασχηματισμένων καμπυλών, ίδιας αλλά και διαφορετικής καμπύλης (στοιχεία Πίνακα 2 εκτός της κυρίας διαγωνίου), όπου – όπως φαίνεται και από την ανάλυση που έγινε στην προηγούμενη ενότητα (5.2.2) – οι τιμές των affine παραμέτρων επηρεάζουν και τις τιμές ταιριάσματος. Άλλωστε, οι τιμές ταιριάσματος μεταξύ affine μετασχηματισμένων καμπυλών, υπολογίζονται κυρίως με σκοπό να πιστοποιήσουν αν οι καμπύλες ανήκουν στην ίδια κατηγορία σχήματος ή όχι.

- **Στη εφαρμογή του αλγορίθμου ταιριάσματος.**

Ο αλγόριθμος ταιριάσματος ενδέχεται να μην αποδίδει το αναμενόμενο σε περιπτώσεις ταιριάσματος μετασχηματισμένων καμπυλών. Αυτό οφείλεται στο γεγονός ότι πλέον δεν ταιριάζουμε μια μετασχηματισμένη καμπύλη με την αρχική της, αλλά με μια άλλη μετασχηματισμένη της ίδιας ή και διαφορετικής κατηγορίας σχήματος. Έτσι το κόστος διαφοροποιείται ανάλογα με το ποια από τις υπό εξέταση καμπύλες ορίζουμε ως «πρότυπη» (αρχική). Και αυτό γιατί ο αλγόριθμος ταιριάσματος αναφέρεται σε σύγκριση αρχικών (πρότυπων) και μετασχηματισμένων εκδοχών της αρχικής καμπύλης. Η διαφοροποίηση στο κόστος εξαρτάται από τον αριθμό των μεγίστων που παραμένουν αταίριαστα, και προστίθενται στην τελική τιμή του κόστους ταιριάσματος.

Τα παραπάνω μειονεκτήματα της CSS μεθόδου προσπαθήσαμε να αντιμετωπίσουμε με την χρήση διαφόρων εναλλακτικών μεθόδων, μια προσπάθεια που παρουσιάζεται αναλυτικά στο επόμενο κεφάλαιο.

## 6. Μελέτη της CSS μεθόδου σε σχέση με άλλες μεθόδους

### 6.1 Εισαγωγικά

Όπως διαπιστώσαμε από την παραπάνω μελέτη, η CSS μέθοδος φαίνεται να είναι αποδοτική στο ταίριασμα εικόνων με ανάλυση σχήματος. Η αποδοτικότητα της έγκειται στο κατά πόσο αυτή παραμένει σταθερή σε affine μετασχηματισμούς, αλλά και στην διαδικασία ταίριασματος που ακολουθεί. Μπορούμε επομένως, μελετώντας τις δύο αυτές παραμέτρους, να βελτιστοποιήσουμε την απόδοση της μεθόδου. Μια τέτοια προσπάθεια γίνεται σε αυτήν την ενότητα, όπου ασχολούμαστε με το κατά πόσο μπορούμε να δημιουργήσουμε μια αναπαράσταση καμπυλών, αναλλοίωτες σε affine μετασχηματισμούς. Πιο συγκεκριμένα, εξετάσαμε την μέθοδο κανονικοποίησης περιγραμμάτων που προτάθηκε το 2001 από τους Γ. Αβρίθη, Γ. Ξηρουχάκη και Σ. Κόλλια [65], και σκοπό είχε την επίλυση του προβλήματος των τυχαίων affine μετασχηματισμών στην ανάκτηση και αναζήτηση εικόνων με βάση την ανάλυση σχήματος.

### 6.2 Παρουσίαση εναλλακτικής μεθόδου κανονικοποίησης

Η εναλλακτική μέθοδος κανονικοποίησης έχει σκοπό να παρέχει περιγραφή της καμπύλης ενός αντικειμένου που να είναι αναλλοίωτη σε affine μετασχηματισμούς καθώς και μετασχηματισμούς παραμέτρου (κυκλική ολίσθηση, φορά διαγραφής). Επιπλέον η μέθοδος δεν παρουσιάζει απώλεια πληροφορίας, αφού το περίγραμμα του αντικειμένου διατηρείται ακέραιο.

Η διαδικασία της μεθόδου κανονικοποίησης αποτελείται από μια σειρά βημάτων, που διαδοχικά προσπαθούν να φέρουν την affine μετασχηματισμένη καμπύλη στην αρχική της μορφή. Συνοπτικά, στο πρώτο στάδιο, με χρήση ροπών μέχρι δεύτερου βαθμού πραγματοποιείται ορθοκανονικοποίηση, δηλαδή κανονικοποίηση ως προς τους μετασχηματισμούς μετατόπισης αλλαγής κλίμακας (scaling) και κλίσης (skew). Έπειτα, με χρήση της φάσης κατάλληλων συντελεστών του μετασχηματισμού Fourier της καμπύλης κανονικοποιείται το σημείο αναφοράς της καμπύλης, το οποίο ενδέχεται να διαφέρει λόγω μετασχηματισμού κυκλικής ολίσθησης. Τέλος, κανονικοποιείται η περιστροφή και ο κατοπτρισμός και πάλι με χρήση πληροφορίας του μετασχηματισμού Fourier.

Πιο συγκεκριμένα, το πρώτο στάδιο της κανονικοποίησης είναι μια διαδικασία *ορθοκανονικοποίησης* (orthogonalization), η οποία κανονικοποιεί την καμπύλη ως προς τους μετασχηματισμούς *μετατόπισης* (translation), *αλλαγής κλίμακας* (scaling) και *κλίσης* (skew). Με αυτόν τον τρόπο, κάθε affine μετασχηματισμός ανάγεται ουσιαστικά σε ορθογώνιο (orthogonal) μετασχηματισμό, δηλαδή σε μετασχηματισμό που περιλαμβάνει μόνο περιστροφή (rotation) ή /και κατοπτρισμό (mirroring).

Σε αντίθεση με την ορθοκανονικοποίηση, η διαδικασία κανονικοποίησης περιστροφής και κατοπτρισμού που πραγματοποιείται στο δεύτερο στάδιο εμπλέκει ποσότητες που εξαρτώνται από το επιλεγμένο σημείο αναφοράς της κλειστής



καμπύλης. Ειδικότερα, η εκτίμηση περιστροφής εξαρτάται από το σημείο αναφοράς ενώ η εκτίμηση κατοπτρισμού εξαρτάται από την περιστροφή. Για αυτό το λόγο, πραγματοποιείται *κανονικοποίηση του σημείου αναφοράς* που βασίζεται στον διακριτό μετασχηματισμό Fourier.

Τέλος, για να ολοκληρωθεί η κανονικοποίησης περιστροφής και κατοπτρισμού, απαιτούνται δύο ακόμη βήματα. Κατά πρώτον, εφαρμόζεται μια επιπρόσθετη κυκλική ολίσθηση ενός αριθμού δειγμάτων με σκοπό να εξαλειφθεί κάθε παράγοντας αβεβαιότητας που τυχόν έχει προκύψει. Κατά δεύτερον, κανονικοποιείται η φορά διαγραφής της καμπύλης ώστε να είναι δεξιόστροφη. Η διαδικασία αυτή πραγματοποιείται οπωσδήποτε μετά την κανονικοποίηση κατοπτρισμού, γιατί ο κατοπτρισμός αντιστρέφει και την φορά διαγραφής.

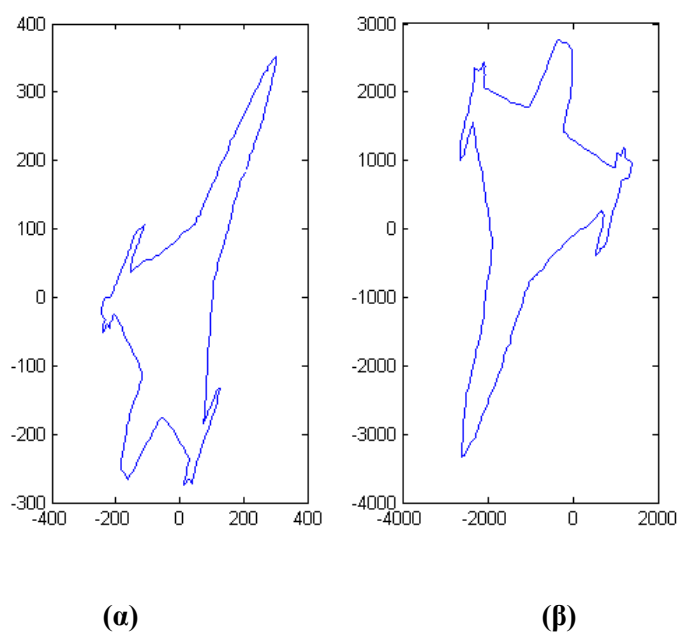
Προκύπτει έτσι, στο τέλος της μεθόδου κανονικοποίησης, μια περιγραφική σχήματος αναλλοίωτη σε affine μετασχηματισμούς (μετατόπιση, αλλαγή κλίμακας, κλίση, περιστροφή, κατοπτρισμός) καθώς και στο αρχικό σημείο αναφοράς και την φορά διαγραφής.

### **6.3 Προσαρμογή της εναλλακτικής μεθόδου κανονικοποίησης στην CSS μέθοδο – πειράματα, αποτελέσματα και παρατηρήσεις**

Για να εκτελέσουμε το πείραμα που ακολουθεί, προσαρμόσαμε την παραπάνω μέθοδο κανονικοποίησης στην μέθοδο της CSS. Πιο συγκεκριμένα, αντί να εφαρμόσουμε την κανονικοποίηση της καμπύλης με βάση το affine μήκος, χρησιμοποιήσαμε την εναλλακτική μέθοδο κανονικοποίησης.

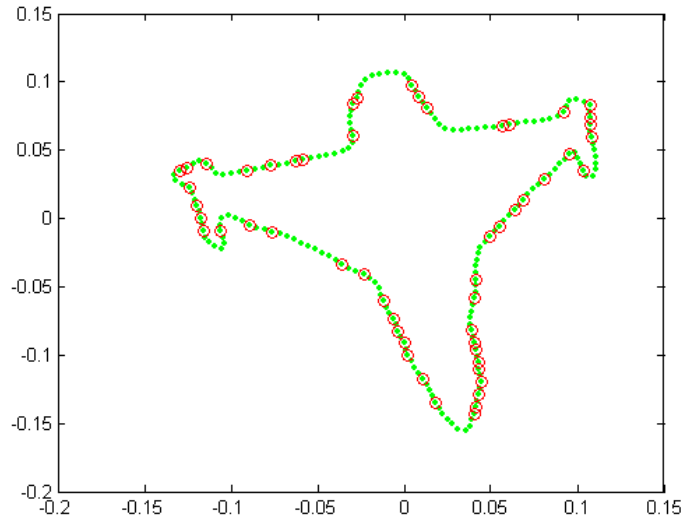
Επιλέγοντας μια αρχική καμπύλη από την κατηγορία των αεροπλάνων δημιουργήσαμε μια affine μετασχηματισμένη εκδοχή της. Τα αποτελέσματα του πειράματος παρουσιάζονται στα ακόλουθα σχήματα.

Αρχικά, βλέπουμε την αρχική και την μετασχηματισμένη καμπύλη:

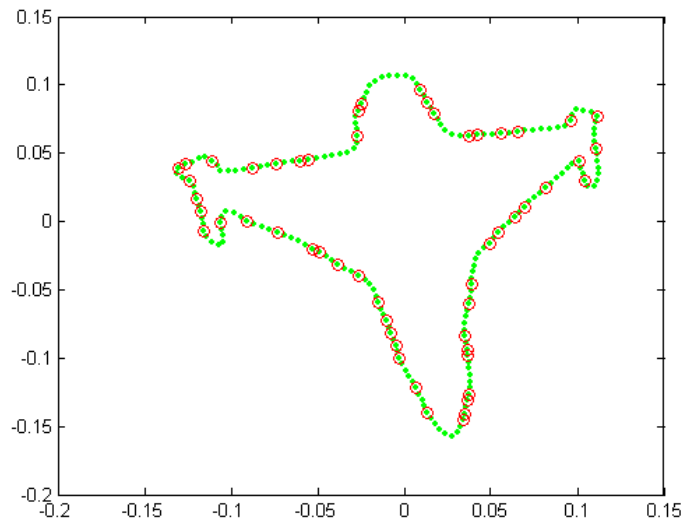


**Σχήμα 6-1:** (α) αρχική καμπύλη jas7, (β) μετασχηματισμένη καμπύλη jas7#1

Ακολουθούν τα πρώτα στάδια της εξέλιξης των δύο καμπυλών. Επειδή αυτά έπονται της μεθόδου της κανονικοποίησης, παρατηρούμε ότι έχουν την ίδια μορφή, εφόσον η μετασχηματισμένη καμπύλη έχει κανονικοποιηθεί και έχει έρθει στην μορφή της αρχικής. Επιπλέον φαίνεται η κανονικοποίηση ως προς το σημείο αναφοράς, πράγμα που επιτρέπει την σύγκριση των δύο εικόνων με μεγαλύτερη ακρίβεια.

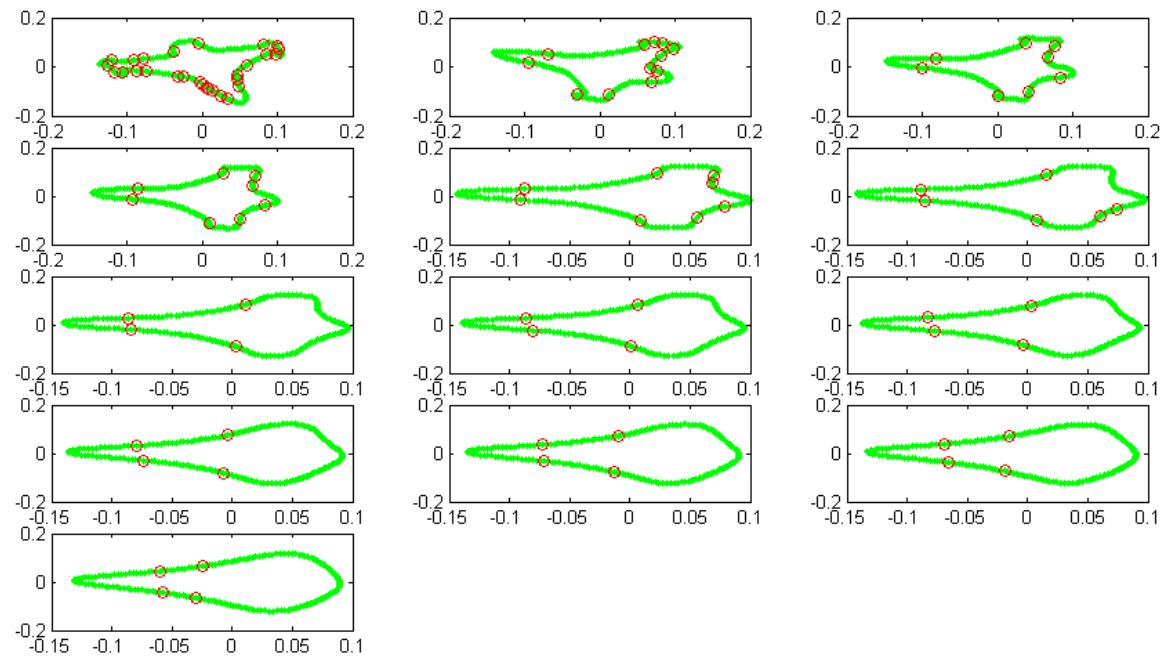


**Σχήμα 6-2:** Πρώτο στάδιο της εξέλιξης της αρχικής καμπύλης



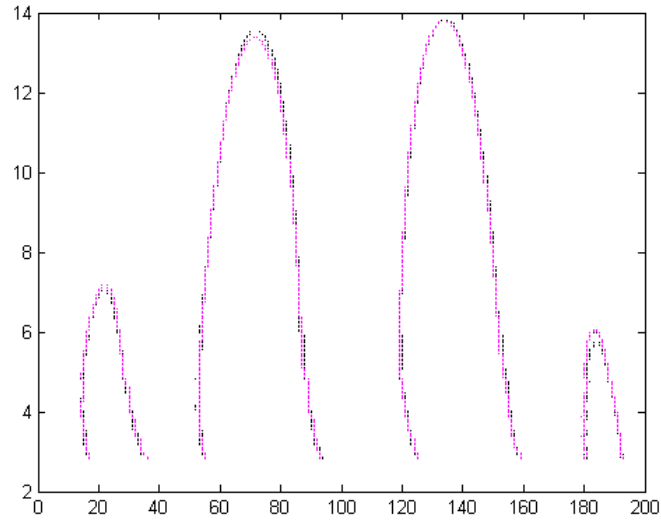
**Σχήμα 6-3:** Πρώτο στάδιο της εξέλιξης της μετασχηματισμένης καμπύλης

Σαν αποτέλεσμα, η εξέλιξη της καμπύλης που ακολουθεί σχεδόν ταυτίζεται και για τις δύο εικόνες και φαίνεται στο παρακάτω σχήμα:



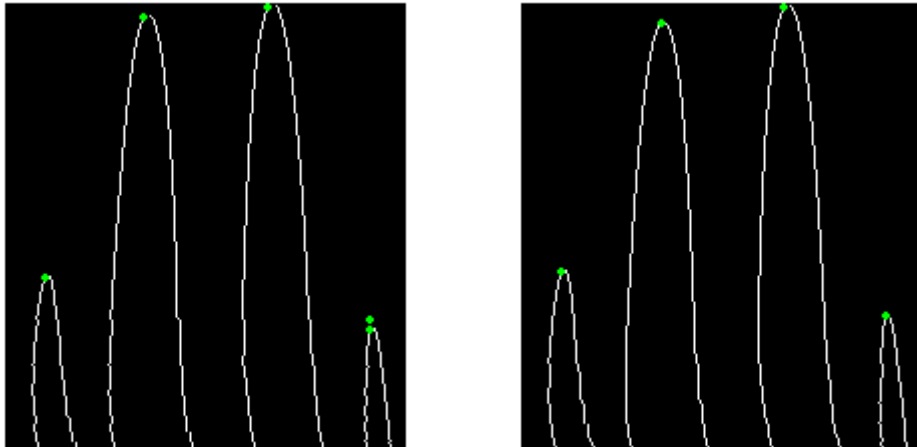
Σχήμα 6-4: Εξέλιξη της καμπύλης

Ακολουθεί η CSS αναπαράσταση των δύο εικόνων, όπου παρατηρούμε ότι πάλι σχεδόν ταυτίζονται με μικρές αποκλίσεις. Η ταύτιση επιτυγχάνεται ακόμα και στους μικρότερους λοβούς.



**Σχήμα 6-4:** CSS απεικόνιση των δύο καμπυλών μετά την εφαρμογή κατωφλίου

Τέλος, ακολουθούν οι ίδιες απεικονίσεις μετά την εξαγωγή των μεγίστων:



**Σχήμα 6-6:** CSS απεικόνιση αρχικής και μετασχηματισμένης καμπύλης μετά την εύρεση των μεγίστων

Το αποτέλεσμα της εφαρμογής του αλγόριθμου ταιριάσματος προέκυψε:

<b>Matching_cost</b>	6,91
----------------------	------

Στο αντίστοιχο πείραμα, κατά την εφαρμογή της CSS μεθόδου με την κατά affine length κανονικοποίηση, το κόστος ταιριάσματος προέκυψε ως εξής:

<b>Matching_cost</b>	49,97
----------------------	-------

Παρατηρούμε πως η τιμή του κόστους που προκύπτει από την εφαρμογή της εναλλακτικής μεθόδου κανονικοποίησης είναι πολύ μικρή σε σχέση με την αντίστοιχη τιμή από την εφαρμογή της CSS μεθόδου με την affine length κανονικοποίηση. Για να έχουμε μια πιο ολοκληρωμένη εικόνα, εξετάζουμε την συμπεριφορά της νέας μεθόδου και σε άλλες περιπτώσεις μετασχηματισμών και παρουσιάζουμε τα αποτελέσματα στον ακόλουθο πίνακα:

<b>Καμπύλες Jas7</b>	<b>Αρχική</b>	<b>Jas7#1</b>	<b>Jas7#2</b>	<b>Jas7#3</b>	<b>Jas7#4</b>	<b>Jas7#5</b>
<b>Αρχική</b>	0,00	6,91	12,81	23,50	22,51	15,27
<b>Jas7#1</b>	6,91	0,00	6,10	16,91	15,92	8,31
<b>Jas7#2</b>	12,81	6,10	0,00	10,82	9,85	4,32
<b>Jas7#3</b>	23,50	16,91	10,82	0,00	8,14	8,85
<b>Jas7#4</b>	22,51	15,92	9,85	8,14	0,00	10,48
<b>Jas7#5</b>	15,27	8,31	4,32	8,85	10,48	0,00

**Πίνακας 4:** Αποτελέσματα ταιριάσματος της καμπύλης Jas7 με 5 affine μετασχηματισμένες καμπύλες (εφαρμογή της εναλλακτικής μεθόδου κανονικοποίησης)

Ο μέσος όρος του παραπάνω πίνακα προκύπτει:

<b>average</b>	10,04
----------------	-------

Στην συνέχεια ακολουθεί και ο αντίστοιχος πίνακας με την εφαρμογή της CSS μεθόδου με την affine length κανονικοποίηση:

Καμπύλες Jas7	Αρχική	Jas7#1	Jas7#2	Jas7#3	Jas7#4	Jas7#5
Αρχική	0,00	49,97	39,79	106,19	36,39	40,49
Jas7#1	49,97	0,00	37,43	86,93	71,14	20,41
Jas7#2	39,79	37,43	0,00	77,55	46,08	3,74
Jas7#3	106,19	86,93	77,55	0,00	65,72	71,56
Jas7#4	36,39	71,14	46,08	65,72	0,00	44,18
Jas7#5	40,49	20,41	3,74	71,56	44,18	0,00

**Πίνακας 5:** Αποτελέσματα ταιριάσματος της καμπύλης Jas7 με 5 affine μετασχηματισμένες καμπύλες (εφαρμογή της affine length μεθόδου κανονικοποίησης)

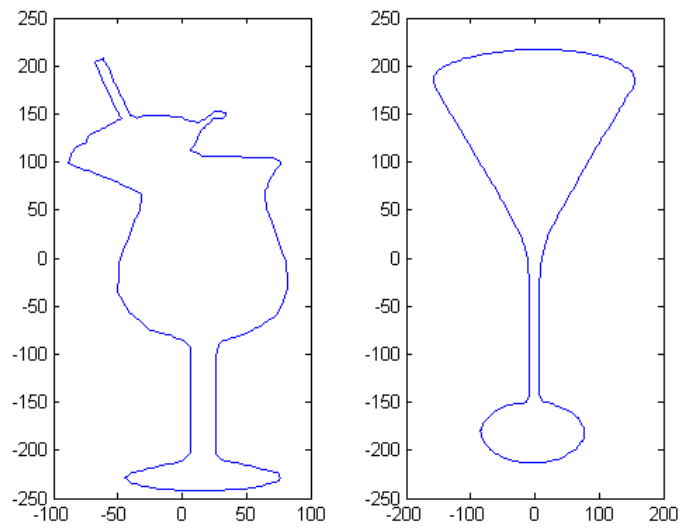
Ο αντίστοιχος μέσος όρος προκύπτει:

<b>average</b>	44,31
----------------	-------

Από τα αποτελέσματα των πειραμάτων είναι φανερό πως η εναλλακτική μέθοδος κανονικοποίησης φαίνεται να αυξάνει κατά πολύ την αποδοτικότητα της CSS μεθόδου. Από τον πίνακα 4 παρατηρούμε ότι οι τιμές κόστους ταιριάσματος είναι πολύ μικρότερες από τις αντίστοιχες του πίνακα 5, γεγονός που αποδεικνύει την αποτελεσματικότητα της εναλλακτικής μεθόδου κανονικοποίησης όσον αφορά τους affine μετασχηματισμούς.

Σημαντικό πλεονέκτημα της εναλλακτικής μεθόδου κανονικοποίησης είναι η επιπλέον κανονικοποίηση που αυτή εφαρμόζει όσον αφορά το σημείο αναφοράς της καμπύλης. Όπως ήδη αναφέραμε, το σημείο αναφοράς της καμπύλης, ενδέχεται να διαφέρει λόγω μετασχηματισμού κυκλικής ολίσθησης. Με την κανονικοποίηση του επομένως, δεν είναι απαραίτητος πλέον ο υπολογισμός της παραμέτρου κυκλικής ολίσθησης όπως απαιτούσε η affine length κανονικοποίηση. Σαν αποτέλεσμα, εκμηδενίζεται η πιθανότητα σφάλματος κατά την διαδικασία ταιριάσματος λόγω της παραμέτρου ολίσθησης, παράγοντας που επηρέαζε τα αποτελέσματα των μετρήσεων της CSS μεθόδου με affine length κανονικοποίηση (βλ. Ενότητα 5.3).

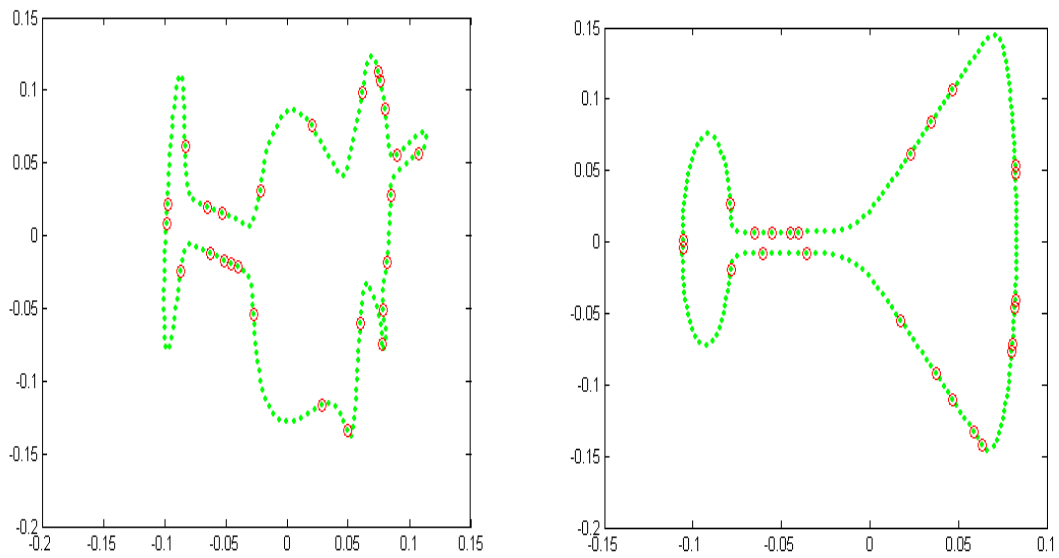
Στην συνέχεια εφαρμόζουμε την εναλλακτική μέθοδο κανονικοποίησης και στην περίπτωση διαφορετικών καμπυλών της ίδιας κατηγορίας και παρουσιάζουμε ενδεικτικά τα αποτελέσματα του πειράματος:



(α)

(β)

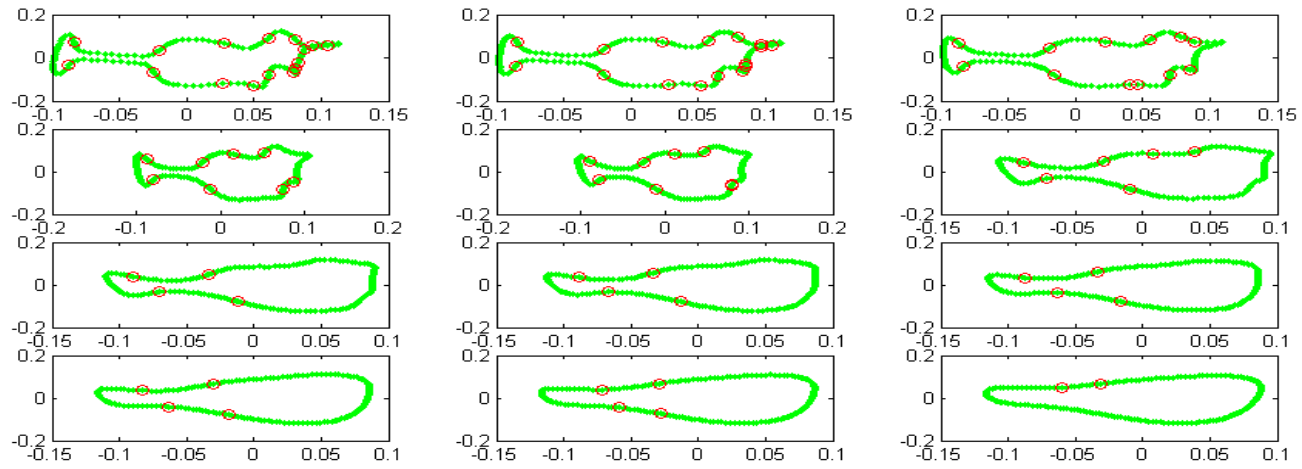
Σχήμα 6-7: (α) αρχική καμπύλη drink008, (β) αρχική καμπύλη martini



(α)

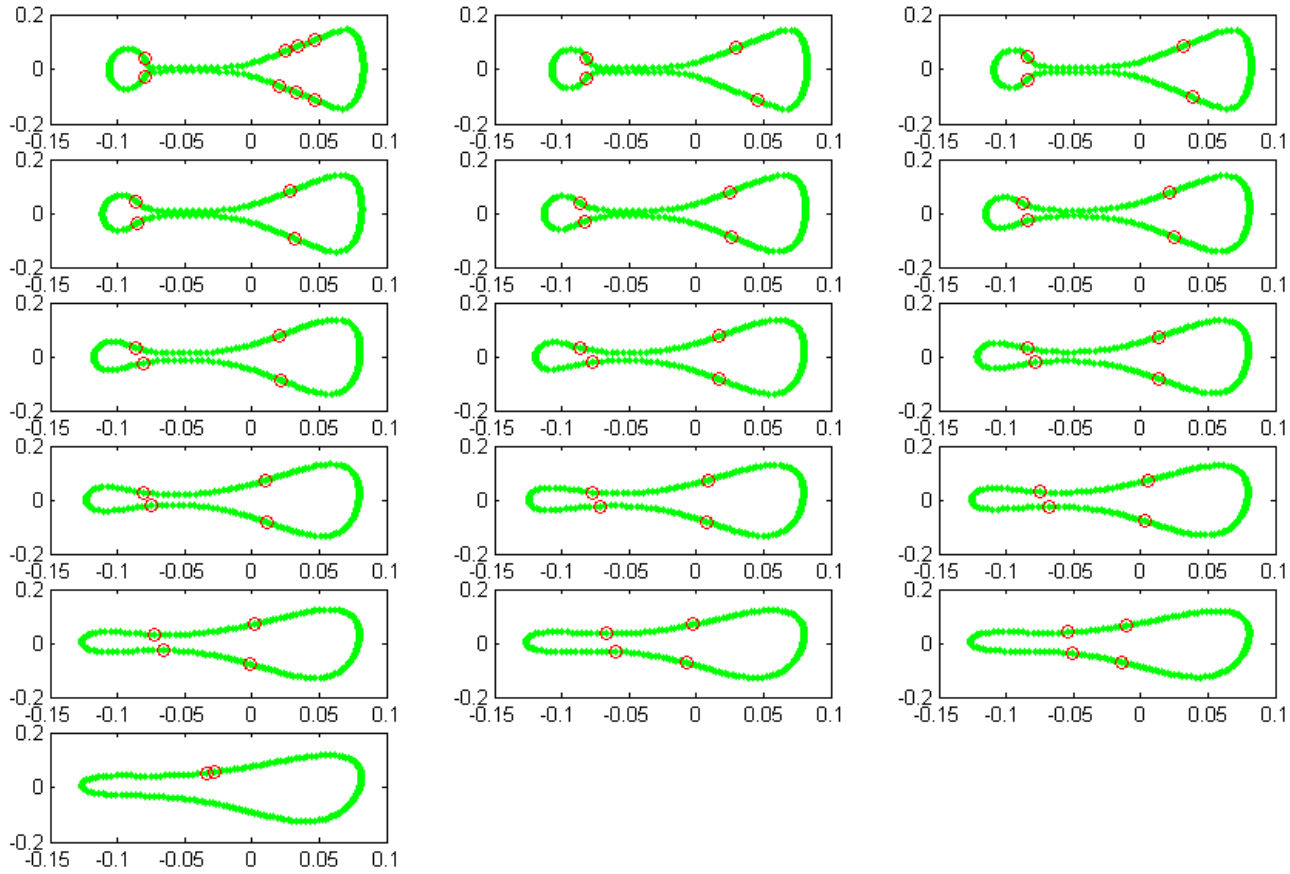
(β)

Σχήμα 6-7: πρώτο στάδιο της εξέλιξης των καμπυλών (α) drink008, (β) martini

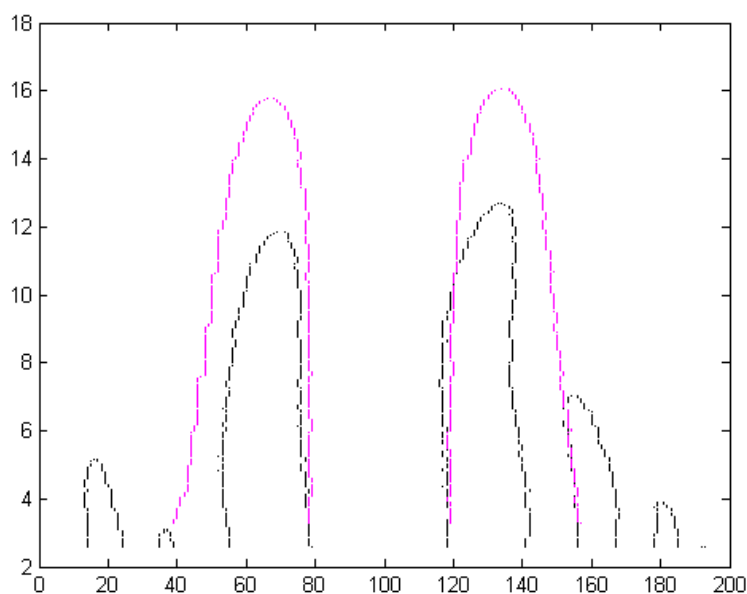


**Σχήμα 6-8:** Εξέλιξη της καμπύλης drink008





Σχήμα 6-9: Εξέλιξη της καμπύλης martini



**Σχήμα 6-10:** CSS απεικόνιση των δύο καμπυλών μετά την εφαρμογή κατωφλίου

Το αποτέλεσμα της εφαρμογής του αλγόριθμου ταιριάσματος προέκυψε:

<b>Matching_cost</b>	16,22
----------------------	-------

Στο αντίστοιχο πείραμα, κατά την εφαρμογή της CSS μεθόδου με την κατά affine length κανονικοποίηση, το κόστος ταιριάσματος προέκυψε ως εξής:

<b>Matching_cost</b>	38,15
----------------------	-------

Συγκρίνοντας την τιμή του κόστους που προκύπτει με το αντίστοιχο που προέκυψε από την εφαρμογή της CSS μεθόδου με την affine length κανονικοποίηση συμπεραίνουμε ότι η εναλλακτική μέθοδος κανονικοποίησης βελτιώνει αισθητά την απόδοση της CSS μεθόδου.

Πιο συγκεκριμένα, όπως φαίνεται και στο σχήμα 6-10, η κανονικοποίηση του σημείου αναφοράς των καμπυλών επιτυγχάνει να ελαχιστοποιήσει το σφάλμα λόγω κυκλικής ολίσθησης και κατά συνέπεια και την τιμή του κόστους ταιριάσματος. Οι δύο καμπύλες κανονικοποιούνται σε κοινό σημείο αναφοράς, με αποτέλεσμα στην διαδικασία του ταιριάσματος να μην απαιτείται ο υπολογισμός της παραμέτρου κυκλικής ολίσθησης.

Για να μπορέσουμε να εκτιμήσουμε καλύτερα την αποτελεσματικότητα της μεθόδου, υπολογίσαμε τα κόστη ταιριάσματος και σε περιπτώσεις μετασχηματισμένων καμπυλών της ίδιας κατηγορίας και συνοψίσαμε τα αποτελέσματα στον ακόλουθο πίνακα:

Καμπύλες	Drink008	Drink008#1	Drink008#2	Drink008#3
<b>martini</b>	16,22	23,29	19,85	17,15
<b>martini #1</b>	15,37	20,34	17,23	15,51
<b>martini #2</b>	17,90	23,98	20,75	18,11
<b>martini #3</b>	15,01	19,97	16,99	15,15

**Πίνακας 6:** Αποτελέσματα ταιριάσματος της καμπύλης martini και drink008 με 3 affine μετασχηματισμένες εκδοχές αυτών (εφαρμογή της εναλλακτικής μεθόδου κανονικοποίησης)

Ο μέσος όρος του παραπάνω πίνακα προκύπτει:

<b>average</b>	18,30
----------------	-------

Στην συνέχεια ακολουθεί και ο αντίστοιχος πίνακας με την εφαρμογή της CSS μεθόδου με την affine length κανονικοποίηση:

Καμπύλες	Drink008	Drink008#1	Drink008#2	Drink008#3
<b>martini</b>	38,15	29,70	44,69	35,84
<b>martini #1</b>	27,40	36,48	24,48	30,23
<b>martini #2</b>	23,87	35,00	35,40	32,11
<b>martini #3</b>	22,69	35,75	36,15	32,11

**Πίνακας 7:** Αποτελέσματα ταιριάσματος της καμπύλης martini και drink008 με 3 affine μετασχηματισμένες εκδοχές αυτών (εφαρμογή της affine length μεθόδου κανονικοποίησης)

Ο μέσος όρος του παραπάνω πίνακα προκύπτει:

<b>average</b>	32,50
----------------	-------

Από τις τιμές του κόστους που προέκυψαν παρατηρούμε πως το κόστος ταιριάσματος μειώθηκε κατά πολύ με την εφαρμογή της εναλλακτικής μεθόδου κανονικοποίησης. Μάλιστα, βλέπουμε πως η τιμή του μέσου όρου του κόστους ταιριάσματος μειώθηκε σχεδόν στο μισό. Επιπλέον, κάθε τιμή που προκύπτει από την εναλλακτική μέθοδο φαίνεται να είναι μικρότερη από την αντίστοιχη με την affine length κανονικοποίηση (πίνακας 6), πράγμα που δείχνει ουσιαστικά το μεγάλο ποσοστό επιτυχίας της εναλλακτικής μεθόδου κανονικοποίησης.

Τέλος, αναφέρουμε ένα παράδειγμα ταιριάσματος διαφορετικών καμπυλών, διαφορετικής κατηγορίας, έστω μιας καμπύλης της κατηγορίας των αεροπλάνων και μιας της κατηγορίας των ποτηριών. Τα αποτελέσματα προκύπτουν ως εξής

Μέθοδοι κανονικοποίησης	Matching_cost
Εναλλακτική μέθοδος	31,25
Με βάση το Affine length	59,49

**Πίνακας 8:** Κόστη ταιριάσματος καμπυλών jas7, drink008 για τις δύο μεθόδους κανονικοποίησης

Παρατηρούμε πως η εναλλακτική μέθοδος κανονικοποίησης επιτυγχάνει και σε αυτή την περίπτωση να μειώσει αρκετά το κόστος ταιριάσματος πάλι σχεδόν στο μισό της τιμής που προκύπτει με την affine length κανονικοποίηση.

## **6.4 Συμπεράσματα σύγκρισης των δύο μεθόδων – μελλοντικές επεκτάσεις**

Από τα παραπάνω πειράματα διαπιστώσαμε την βελτίωση της απόδοσης της CSS μεθόδου με την εφαρμογή της εναλλακτικής μεθόδου κανονικοποίησης. Πιο συγκεκριμένα η εναλλακτική μέθοδος κανονικοποίησης κατάφερε να μειώσει αισθητά το κόστος ταιριάσματος της CSS μεθόδου σε σχέση με τις αντίστοιχες τιμές κόστους που προέκυψαν για την CSS μέθοδο με affine length κανονικοποίηση.

Ο παράγοντας που έπαιξε καθοριστικό ρόλο στην βελτίωση της απόδοσης είναι η επιπρόσθετη κανονικοποίηση που έγινε όσον αφορά το σημείο αναφοράς της καμπύλης και τον προσανατολισμό της. Όπως ήδη αναφέραμε, η εναλλακτική μέθοδος κανονικοποίησης έχει σκοπό να παρέχει περιγραφή της καμπύλης ενός αντικειμένου που να είναι αναλλοίωτη σε affine μετασχηματισμούς (αλλαγή στην κλίμακα, μετάθεση και περιστροφή) αλλά επιπλέον και σε μετασχηματισμούς παραμέτρου (κυκλική ολίσθηση, προσανατολισμός). Αποτελεί επομένως μια πιο ολοκληρωμένη μέθοδο κανονικοποίησης, εφόσον εξετάζει όλους τους πιθανούς μετασχηματισμούς που μπορεί να υποστεί μια καμπύλη.

Από την άλλη πλευρά, η CSS μέθοδος με affine length κανονικοποίηση καταφέρνει να κανονικοποιήσει την καμπύλη όσον αφορά τους affine μετασχηματισμούς (αλλαγή στην κλίμακα, μετάθεση, περιστροφή) βασισμένη στην affine length παραμετροποίηση, αλλά μειονεκτεί στην κανονικοποίηση των μετασχηματισμών παραμέτρου. Αν και πριν τη διαδικασία του ταιριάσματος γίνεται μια απόπειρα τέτοιας κανονικοποίησης με την εύρεση της παραμέτρου κυκλικής ολίσθησης, αυτή φαίνεται να μην αποδίδει ικανοποιητικά. Έτσι, όπως ήδη αναφέραμε στην ενότητα 5.3 των παρατηρήσεων για την CSS μέθοδο με affine length κανονικοποίηση, η κυκλική ολίσθηση επιδρά στο ταίριασμα των CSS εικόνων και κατά συνέπεια και στην τιμή του κόστους που προκύπτει.

Όλα τα παραπάνω φαίνονται καλύτερα αν παρατηρήσουμε τα σχήματα που προέκυψαν κατά την εκτέλεση των πειραμάτων (σχήμα 6-2, 6-3, 6-4 για την ίδια καμπύλη και σχήμα 6-7, 6-10 για διαφορετικές καμπύλες).

Στα σχήματα 6-2, 6-3 και 6-7 φαίνεται πώς οι υπό εξέταση καμπύλες έχουν κανονικοποιηθεί και έχουν έρθει σε τέτοια μορφή, ώστε να έχουν κοινό σημείο αναφοράς. Με αυτόν τον τρόπο η σύγκριση τους έχει νόημα, πράγμα που φαίνεται και στα σχήματα των CSS εικόνων που προκύπτουν (σχήμα 6-4 και 6-10). Μάλιστα, στο σχήμα 6-4 όπου έχουμε τις CSS αναπαραστάσεις της αρχικής και μετασχηματισμένης της καμπύλης, βλέπουμε ότι αυτές σχεδόν ταυτίζονται – απόδειξη πως η κανονικοποίηση είχε επιτυχία και ως προς τυχόν μετασχηματισμούς κυκλικής ολίσθησης.

Καταλήγουμε επομένως στο συμπέρασμα πως ένας πολύ καλός τρόπος να βελτιώσουμε την απόδοση της CSS μεθόδου είναι η αντικατάσταση της affine length κανονικοποίησης με αυτή της εναλλακτικής μεθόδου κανονικοποίησης που προαναφέραμε. Η νέα CSS μέθοδος διατηρεί τα πλεονεκτήματα της αρχικής CSS μεθόδου με affine length κανονικοποίηση – παραμένει γρήγορη, χωρίς πολυπλοκότητα υπολογισμών, και επιπλέον υπολογίζει με μεγαλύτερη ακρίβεια το κόστος ταιριάσματος. Αυξάνεται επομένως η αξιοπιστία της μεθόδου, έτσι ώστε να μπορεί να χρησιμοποιηθεί αποτελεσματικά σε περιπτώσεις αναζήτησης αλλά και ανάκτησης ομοίων εικόνων από μια βάση δεδομένων.

Εξετάζοντας όλα τα παραπάνω, κανείς μπορεί να διαπιστώσει ότι η μέθοδος της curvature scale space επιδέχεται πιθανές βελτιώσεις αλλά και μελλοντικές επεκτάσεις. Στο κεφάλαιο αυτό ασχοληθήκαμε με την παράμετρο της κανονικοποίησης και προσπαθήσαμε να βελτιώσουμε την αποτελεσματικότητα της και κατά συνέπεια και την αποτελεσματικότητα της μεθόδου. Μια άλλη πιθανή βελτίωση θα ήταν ένας αποδοτικότερος και πιο ακριβής αλγόριθμος ταιριάσματος, βασισμένος πάνω σε αυτόν που ήδη υλοποιήσαμε. Ο αλγόριθμος αυτός θα μπορούσε να «διαχειρίζεται» καλύτερα την παράμετρο της κυκλικής ολίσθησης, με τέτοιο τρόπο ώστε να μην επηρεάζεται το κόστος ταιριάσματος που προκύπτει. Επιπλέον, ο αλγόριθμος θα μπορούσε να προσαρμοστεί κατάλληλα και σε περιπτώσεις ταιριάσματος διαφορετικών καμπυλών με σκοπό να προκύπτουν πιο ακριβή αποτελέσματα κατά την αναζήτηση καμπυλών από μια βάση δεδομένων.

Σημαντικό επίσης πλεονέκτημα της CSS μεθόδου αποτελεί το γεγονός ότι απαρτίζεται από επιμέρους στάδια, που το καθένα ξεχωριστά μπορεί να εξεταστεί και να βελτιωθεί. Εκτός της κανονικοποίησης και του αλγορίθμου ταιριάσματος επομένως, κάποιος μπορεί να μελετήσει την ομαλοποίηση της καμπύλης με την χρήση διαφορετικών συναρτήσεων (φίλτρων) εκτός της Γκαουσιανής, καθώς ίσως και την υλοποίηση ενός διαφορετικού τρόπου εξαγωγής μεγίστων.

Τέλος, αξίζει να αναφέρουμε ότι η CSS μέθοδος έχει την δυνατότητα, σε συνδυασμό με άλλες μεθόδους ανάκτησης εικόνων, να συνεισφέρει σημαντικά στην ανάλυση και επεξεργασία εικόνων έχοντας το πλεονέκτημα της εύκολης εφαρμογής, της ταχύτητας αλλά και – όπως προαναφέραμε – της πολύπλευρης δυνατότητας βελτίωσης της απόδοσης της.



## Βιβλιογραφία

1. F. Mokhtarian, S. Abbasi, “Shape similarity retrieval under affine transforms”, Centre for vision, speech and signal processing, Department of Electronic & Electrical Engineering, University of Surrey, Guildford UK, *Pattern recognition* 35, 31-41, (2002)
2. Y. Rui, Thomas S. Huang, S. Chang, “Image Retrieval: Past, present and future”, *Journal of Visual Communication and Image Representation*, 10:1-23, 1999.
3. S. K. Chang, A. Hsu, “Image information systems: where do we go from here?”, *IEEE Trans. On Knowledge and Data Engineer.*, vol. 4, no. 5, 1992.
4. G. Kowalski, M. T. Maybury, “Information storage and retrieval systems – theory and implementation”, 2<sup>nd</sup> ed. Kluwer Academic Publ., 2000.
5. J. Jeon, J. Lavrenko, R. Manmatha, “Automatic image annotation and retrieval using Cross – Media relevance models”, in *Proceedings of the 26th international ACM SIGIR conference on Research and development in information retrieval*, Toronto, Canada, 2003.
6. Pan, J., & Yang, H.J., & Duygulu, P., & Faloutsos, C., “Automatic Image Captioning”, In *Proceedings of the 2004 IEEE International Conference on Multimedia and Expo (ICME 2004)*, Taipei, Taiwan, 2004.
7. Flickner, M., Sawhney, H., Niblack, W., Ashley, J., Huang, Q., Dom, B., Gorkani, M., Hafner, J., Lee, D., Petkovic, D., Steele, D., and Yanker, P. (1997) “Query by Image and Video Content: The QBIC System.”, In Maybury (Ed.) *Intelligent Multimedia Information Retrieval*, p7-22.
8. P. D. Sherman, “Colour Vision in the Nineteenth Century: The Young – Helmholtz – Maxwell Theory”, Pp. xiv + 233. Adam Hilger, Bristol. 1981.
9. G. Pass, R. Zabih, “Comparing images using joint histograms”, *Multimedia systems* 7 (3) (1999) 234-240.
10. D. Androustos, K. N. Plataniotis, A.N. Venetsanopoulos, “A novel vector-based approach to colour image retrieval using a vector angular-based distance measure”, *Computer Vision Image Under.* 75 (1/2) (1999) 46-58.
11. Y. Deng, B. S. Manjunath, C. Kenney, M.S. Moore, H. Shin, “An efficient colour representation for image retrieval”, *IEEE Trans. Image Process.* 10 (1) (2001) 140-147.
12. E. Albuz, E. Kocalar, A. A. Khokhar, “Scalable colour image indexing and retrieval using vector wavelets”, *IEEE Trans. Knowl. Data Eng.* 13 (5) (2001) 851-861.
13. J. Berens, G. D. Finlayson, G. Gu, “Image indexing using compressed colour histogram”, *IEEE Proc. Vision Image Signal Process.* 147 (4) (2000) 349-353.
14. B. S. Manjunath, J. R. Ohm, V. V. Vasudevan, A. Yamada, “Colour and texture descriptors”, *IEEE Trans. Circuits Systems Video Technol.* 11 (6) (June 2001) 703-715.
15. Y. Gong, H. Chuan, M. Sakauchi, “An image database system with content capturing and fast image indexing abilities”, *Proc. IEEE International Conference on Multimedia Computing and Systems*, Boston, pp.121-130.
16. W. Hsu, T. Chua, H. Pung, “An intergrated colour-spatial approach to content-based image retrieval”, *Proc. Of Third ACM Inter. Mult. Conference*, San Francisco, 1995.

17. B. Julesz, "Textons, the Elements of Texture Perception, and their Interactions", *Nature*, vol.290, 12 March 1981.
18. B. Julesz, J. R. Bergen, "Textons, the Fundamental Elements in Preattentive Vision and Perception of Textures", *Bell System Tech. Journal*, vol.62, no.6, pp. 1619-1645, Jul-Aug 1983.
19. W. K. Pratt, O. D. Faugeras, A. Gagalowicz, "Applications of stochastic Texture Fields Models to Image Processing", *Proc IEEE*, vol.69, pp.542-551, 1981.
20. J. P. Havlicek, D. S. Harding, A. C. Bovik, "Multidimensional quasi-eigenfunction approximations and multicomponent AM -FM models", *IEEE Trans. Image Processing*, vol.9(2), pp. 227-242, 2000.
21. A. C. Bovik, N. Gopal, T. Emmoth, A. Restrepo, "Localized Measurement of Emergent Image Frequencies by Gabor Wavelets", *IEEE Trans. Info. Theory*, vol.38, pp.691-712, March 1992.
22. P. Maragos, "Computer Vision", NTUA, 2005.
23. J. K. Aggarwal, N. Nandhakumar, "On the computation of motion from sequences of images", *Proc. IEEE*, 76, pp. 917-935, May 1981.
24. T.S. Huang, R. Y. Tsai, "Image sequence analysis: motion estimation", in *Image sequence analysis*, T .S. Huang, Ed., Springer-Verlag, 1981.
25. R. Brockett, "Gramians, Generalized inverses, and the least-squares approximation of optical flow", *J. Visual Commun. Image Repres.*, 1, pp.3-11, Sep. 1990.
26. B. Horn, B. Schunk, "Determining optical flow", *Artificial Intelligence*, vol.17, pp.185-203, Aug. 1981.
27. J. Wu, R. Brockett, K. Wohn, "A contour-based recovery of image flow: Iterative method", *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 124-129, San Diego, June 1989.
28. J. R. Jain, A.K. Jain, "Displacement measurement and its applications in interframed coding", *IEEE Trans. Commun. Vol. COM-29*, pp. 1799-1808, Dec. 1981.
29. A. Rosenfeld, J. L. Pfaltz, "Distance functions on digital pictures", *Pattern Recognition*, 1, pp. 33-61, 1968.
30. G. Borgefors, "Distance transformations on digital images", *Comp. Vision, Graphics, Image Process.*, 34, pp. 344-371, 1986.
31. H. Blum, "A transformation for extracting new descriptors on shape", in *Models for the Perception of Speech and Visual Forms*, W. Wathen Dunn, Ed., MIT Press, Cambridge, MA, 1967.
32. A. Rosenfeld, "Axial representations of shapes", *Comput. Vision, Graphics, Image Process.*, 33, pp. 156-173, 1986.
33. P. Maragos, R. W. Schafer, "Morphological skeleton representation and coding of binary images", *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-34, pp. 1228-1244, Oct. 1986.
34. M. Bober, "MPEG-7 Visual shape descriptors", *IEEE Trans. Circuits&Systems Video Technology*, vol. 11, No 6, June 2001
35. M.-K. Hu, "Pattern recognition by moment invariants," *Proc. IRE*, vol. 49, p. 1428, Sept. 1961.
36. Cho-Huak Teh and Roland T. Chin, "On digital approximation of moment invariants," *Comput. Vision, Graphics, Image Processing*, vol 33,pp.318-326, 1986.



37. M. R. Teague, "Image analysis via the general theory of moments," *J. Opt. Soc. Amer.*, vol. 70, pp. 920-930, Aug. 1980.
38. Cho-Huak Teh and Roland T. Chin, "On Image Analysis by the Methods of Moments", *IEEE Trans. Pattern Anal. Machine Intell.*, vol 10, No. 4, July 1988
39. D. Zhang, G. Lu, "Shape-based image retrieval using generic Fourier descriptor", Gippsland School of Computing&Information Technology, Monash University, Vic. 3842, Australia, July 2002
40. D. Zhang, G. Lu, "A Comparative Study on Shape Retrieval Using Fourier descriptors with different Shape Signatures", Gippsland School of Computing&Information Technology, Monash University, Vic. 3842, Australia.
41. J. Flusser and T. Suk, "Pattern recognition by affine moment invariants", *Pattern Recognition* 26, 167-174 (1993).
42. J. Flusser and T. Suk, "Affine moment invariants: A new tool for character recognition", *Pattern Recognition* 15, 433-436 (1994).
43. C-C. Chen, "Improved moment invariants for shape discrimination", *Pattern Recognition* 26, 683-686 (1993).
44. Z. Huang, F.S. Cohen, "Affine invariant B-spline moments for curve matching", *Proc. IEEE Comput. Soc. Conf. Computer Vision and Pattern Recognition*, pp. 490-495, Seattle Washington, 1994.
45. D. Zhao, J. Chen, "Affine curve moment invariants for shape recognition", Department of Electrical & Computer Engineering, University of Michigan-Dearborn, Dearborn, MI 48128-1491, USA.
46. Zhang, D.S., Lu, G. J., 2001a, Shape retrieval using Fourier descriptors, In: *Proceedings of the International Conference on Multimedia and Distance Education (ICMADE'01)*, Fargo, ND, USA, pp. 1-9.
47. H. T.Chang, "Arbitrary affine transformation and their composition effects for 2-D fractal sets", Photonics and Information Laboratory, Department of Electrical Engineering, National Yunlin University of Science and Technology, Touliu Yunlin 64045, Taiwan, ROC.
48. B. S. Manjunath, P. Salembier, T. Sikora, "Introduction to MPEG-7: Multimedia Content Description Interface".
49. B. Guenter, R. Parent, "Computing the arc length of parametric curves", *IEEE Comp. Graph. Appl.* 5 (1990) 72-78.
50. S. Vincent, D. Forsey, "Fast and accurate parametric curve length computation", *J. Graph. Tools* 6 (4) (2002) 29-40.
51. F. Mokhtarian, S. Abbasi, "Affine Curvature Scale Space with Affine Length Parametrisation", Centre for Vision Speech and Signal Processing, Department of Electronic & Electrical Engineering, University of Surrey, Guildford, UK.
52. D. Cyganski, J. A. Orr (1985b), "3-D motion parameters from contours using a canonic differential", *Proc. IEEE Int. Conf. ASSP*, 917-920.
53. BE Kimia, K. Siddiqi, "Geometric heat equation and non linear diffusion of shapes and images", *Computer Vision and Image Understanding* 1996, 64(3):305-332.
54. D. Marr, S. Ullman, T. Poggio, "Bandpass channels, zero-crossings, and early visual information processing, *J. Opt. Soc. Am*, 69(1979) 914-916.
55. A. P. Witkin, "Scale-space filtering", *Proceedings of the international Joint Conference on Artificial Intelligence*, Palo Alto, CA, 1983, pp. 1019-1022.

56. A.C. Jalba, M. Wilkinson, J. Roedink, "Morphological hat – transform scale spaces and their use in pattern classification", Institute for Mathematics & Computing Science, University of Groningen, Groningen, The Netherlands, September 2003.
57. E. J. Breen, R. Jones, "Attribute openings, thinnings and granulometries", *Comput. Vision & Image Understanding* 64 (3) (1996) 377-389.
58. P. F. M. Nacken, "Chamfer metrics, the medial axis and mathematical morphology", *J. Math. Imaging Vision* 6 (1996) 235-248.
59. J. A. Bangham, P. D. Ling, R. Harvey, "Scale – space from nonlinear filters", *IEEE Trans. Pattern Anal. Mach. Intell.* 18 (1996) 520-528.
60. F. Leymarie, M. D. Levine, "Curvature morphology", Technical report TR-CIM-88-26, Computer Vision & Robotics Laboratory, McGill University, Montreal, Quebec, Canada, 1988.
61. A.C. Jalba, M. Wilkinson, J. Roedink, "Shape representation and recognition through morphological curvature scale spaces", *IEEE Trans. On Image Processing*.
62. Y. Avrithis, "Video Sequence Analysis for Content Description, Summarization and Content-Based Retrieval", Ph.D. Thesis, School of Electrical and Computer Engineering, National Technical University of Athens, 2001.
63. S. Adler, R. Krisnan, "Similarity and affine normalization of partially occluded planar curves using first and second derivatives", *Pattern Recognition*, vol. 31, no. 10, pp. 1551-1556, 1998.
64. F. Mokhtarian, S. Abbasi, J. Kittler, "Robust and efficient shape indexing through curvature scale space", Centre for Vision Speech and Signal Processing, Department of Electronic & Electrical Engineering, University of Surrey, Guildford, UK.
65. Y. Avrithis, Y. Xirouhakis and S. Kollias, "Affine-invariant curve normalization for object shape representation, classification, and retrieval," *Machine Vision and Applications* (2001) 13: 80–94.