



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

Αναπαράσταση Σχήματος, Ταξινόμηση και
Ανάκτηση Εικόνων

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

Κωνσταντίνος Α. Ραυτόπουλος

Αθήνα, Ιούλιος 2011



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

ΑΝΑΠΑΡΑΣΤΑΣΗ ΣΧΗΜΑΤΟΣ, ΤΑΞΙΝΟΜΗΣΗ ΚΑΙ ΑΝΑΚΤΗΣΗ ΕΙΚΟΝΩΝ

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

Κωνσταντίνος Α. Ραυτόπουλος

Συμβουλευτική Επιτροπή : Στέφανος Κόλλιας
Ανδρέας-Γεώργιος Σταφυλοπάτης
Νικόλαος Μήτρου

Εγκρίθηκε από την επταμελή εξεταστική επιτροπή την 25η Ιουλίου 2011.

.....
Σ. Κόλλιας
Καθηγητής ΕΜΠ

.....
Α.Γ. Σταφυλοπάτης
Καθηγητής ΕΜΠ

.....
Ν. Μήτρου
Καθηγητής ΕΜΠ

.....
Π. Μαραγκός
Καθηγητής ΕΜΠ

.....
Γ. Στάμου
Λέκτορας ΕΜΠ

.....
Π. Τσανάκας
Καθηγητής ΕΜΠ

.....
Μ. Ζερβάκης
Καθηγητής
Πολυτεχνείου Κρήτης

Αθήνα, Ιούλιος 2011

.....
Κωνσταντίνος Α. Ραυτόπουλος, Διδάκτωρ Ε.Μ.Π.

Copyright © Κωνσταντίνος Α. Ραυτόπουλος, 2011

Με επιφύλαξη παντός δικαιώματος - All rights reserved.

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς τον συγγραφέα. Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν τον συγγραφέα και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις επίσημες θέσεις του Εθνικού Μετσόβιου Πολυτεχνείου.

Περίληψη

Η Διατριβή επικεντρώνεται σε θέματα αναπαράστασης, κωδικοποίησης, ταξινόμησης και ανάκτησης της οπτικής πληροφορίας είτε αυτή είναι σε μορφή σχήματος είτε σε μορφή εικόνας με μία προσέγγιση που είναι συμβατή με βιολογικά αλλά και λειτουργικά δεδομένα σε σχέση με την ανθρώπινη αντίληψη. Στην περίπτωση της αναπαράστασης και κωδικοποίησης σχημάτων εισάγονται έννοιες βιολογικής αλλά και αντιληπτικής συμβατότητας. Στην περίπτωση των εικόνων εισάγονται έννοιες γνωσιακής συμβατότητας που χρησιμοποιούνται για τον σκοπό της καλύτερης ταξινόμησης και ανάκτησης εικόνων.

Λέξεις Κλειδιά: Οπτικές Οδοί, Σκελετός Καμπυλότητας, Αναπαράσταση Σχήματος, Ολοτοπικός Μετασχηματισμός, Ολοτοπικό Μέτρο, Ολοτοπική Επιφάνεια, Ολοτοπικές Εξι-σώσεις, Αναπαράσταση Εμβαδών Απόψεων, Μαρκοβιανή Σημασιολογική Ταξινόμηση.

Abstract

The dissertation is focused on issues of representation, classification, indexing and retrieval of visual information either in the form of shape or in the form of image in an approach that is compatible to biological but also functional data concerning human perception. In the case of shape representation and image classification, concepts of biological as well as perceptual compatibility to the human perceptual system are introduced. In the case of indexing and classification of images, concepts of adaptive knowledge representation through dynamic human interaction are introduced and implemented for better annotation based image retrieval.

Key-words: Visual Pathways, Curvature Skeleton, Shape Representation, Global-Local Transformation, Global-Local Surface, View Area Representation, Global-Local Measure, Global-Local Equations, Annotation based Image Retrieval, Markovian Semantic indexing.

Στην Οικογένειά μου
Στους Φίλους μου
Στους Δασκάλους μου

μάλιστα ἔμοιγε καταφαίνεται ὧδε ἔχειν: τά μὲν εἶδη
ταῦτα ὡσπερ παραδείγματα ἐστάναι ἐν τῇ φύσει,
τά δέ ἄλλα τούτοις εἰκέναι καί εἶναι ὁμοιώματα,
καί ἢ μέθεξις αὕτη τοῖς ἄλλοις γίνεσθαι τῶν εἰδῶν
οὐκ ἄλλη τις ἢ εἰκασθῆναι αὐτοῖς.

Πλάτων, *Παρμενίδης*, 370 π.Χ.

Η Διατριβή αυτή δεν θα είχε πραγματοποιηθεί χωρίς την παρέμβαση σε κρίσιμες στιγμές των συνάδελφων και φίλων μου Ν. Παπαδάκη, Δ. Σκούτα και Κ. Νταλιάνη, χωρίς την Ακαδημαϊκή πληρότητα του Καθηγητή Τ. Σελλή αλλά κυρίως χωρίς την επί πέντε ετών στήριξη και άψογη συνεργασία του Επιβλέποντος Καθηγητού μου, Στέφανου Κόλλια. Ένα ευχαριστώ σε αυτούς δεν αρκεί.

Ο υπερατλαντικός συνοδοιπόρος και φίλος Καθηγητής Δ. Σουρλάς είχε μία πολύπλευρη συνεισφορά όλον αυτό τον καιρό, μέσα από τις Μαθηματικές μας συζητήσεις, τις παρατηρήσεις του και ενίοτε την παρέμβασή του στα κείμενα. Τον ευχαριστώ. Ο φίλος μου Δημήτρης Παπαδημητρίου μοιράστηκε μαζί μου τις δύσκολες αλλά και τις ευχάριστες στιγμές όλον αυτόν τον καιρό. Τον ευχαριστώ.

Ευχαριστώ τα υπόλοιπα μέλη της Τριμελούς Συμβουλευτικής Επιτροπής μου, Καθηγητές Α.-Γ. Σταφυλοπάτη και Ν. Μήτρου. Οι συμβουλές τους, σε συνδυασμό με την ήρεμη και φιλική προσωπικότητά τους με βοήθησαν να συνεχίσω. Ευχαριστώ τα υπόλοιπα μέλη της επιταμελούς επιτροπής μου Καθηγητές Π. Μαραγκό, Μ. Ζερβάκη και Γ. Σιάμου για τις συμβουλές και το ενδιαφέρον τους.

Ευχαριστώ επίσης τα μέλη του Εργαστηρίου Ψηφιακής Επεξεργασίας Σημάτων και Εικόνων που διατηρούν ένα όμορφο και εποικοδομητικό περιβάλλον.

Τέλος, σε αντίστροφη σειρά σημαντικότητας, θέλω να αναφερθώ στους γονείς μου Αντώνη και Βασιλική και την αδερφή μου Κυριακή. Λέγεται, πώς στην ζωή, αυτοί που θυσιάζουν τα περισσότερα, λαμβάνουν την μικρότερη αναγνώριση. Αυτό δεν θα ισχύσει εδώ μιας και αναγνωρίζω πως η στήριξή τους ήταν η σημαντικότερη συνεισφορά σε αυτήν την πορεία και συνεπώς η Διατριβή αυτή τούς ανήκει εξ ίσου.

Κωνσταντίνος Α. Ραυτόπουλος
Αθήνα, Ιούλιος 2011

Περιεχόμενα

1	ΕΙΣΑΓΩΓΗ	17
2	ΑΝΑΔΡΟΜΗ ΣΕ ΘΕΩΡΙΕΣ ΚΑΙ ΤΕΧΝΙΚΕΣ ΣΧΕΤΙΚΑ ΜΕ ΤΗΝ ΟΠΤΙΚΗ ΠΛΗΡΟΦΟΡΙΑ	19
2.1	Αναδρομή σε τεχνικές αναπαράστασης και ανάκτησης σχήματος	19
2.2	Η ανθρώπινη αντίληψη για την οπτική μορφή	35
2.2.1	Κλασικές Θεωρίες για την Οπτική Αντίληψη	35
2.2.2	Σύγχρονες Θεωρίες Οπτικής Αντίληψης	38
3	ΒΙΟΛΟΓΙΚΑ ΣΥΜΒΑΤΕΣ ΤΕΧΝΙΚΕΣ ΓΙΑ ΑΝΑΠΑΡΑΣΤΑΣΗ ΣΧΗΜΑΤΟΣ	43
3.1	Η Μέτρηση Καμπυλότητας σαν διαδικασία των Φλοιωδών Νευρώνων	44
3.2	Το Προτεινόμενο Μοντέλο για Μέτρηση Επίπεδης Καμπυλότητας	45
3.3	Κωδικοποίηση της Επίπεδης Καμπυλότητας στις Τιμές Έντασης Φωτεινότητας	50
3.4	Η Διαδικασία Εξαγωγής του Σκελετού Καμπυλότητας	53
3.5	Πειραματικά Αποτελέσματα	56
4	Ο ΟΛΟΤΟΠΙΚΟΣ ΜΕΤΑΣΧΗΜΑΤΙΣΜΟΣ	59
4.1	Ορισμός	62
4.2	Ιδιότητες του Ολοτοπικού Μετασχηματισμού	64
4.2.1	Καινοτομία του GLT σε Σχέση με Πολυπλοκότητα και Ανθεκτικότητα σε Θόρυβο	65
4.3	Τοπικές δεσμεύσεις των συναρτήσεων άποψης	67
4.4	Η Αναπαράσταση Εμβαδών Απόψεων (VAR)	69
4.4.1	Ιδιότητες της VAR	71
4.5	Προσανατολισμός των σημείων της καμπύλης	73
4.6	Αντιστοίχιση καμπυλότητας	75
4.7	Σύγκριση της VAR με τις μεθόδους καμπυλότητας	77
4.8	Πειραματική επικύρωση	79
4.8.1	VAR έναντι LAII ως περιγραφείς σχήματος ανθεκτικοί στο θόρυβο	79
4.8.2	VAR έναντι LAII σε αντιστοίχιση θορυβωδών σχημάτων	81
4.8.3	LAII έναντι VAR σε απόδοση, εφαρμοσιμότητα και πολυπλοκότητα	86
4.8.4	GLS/VAR έναντι Shape Contexts	87
4.8.5	GLS/VAR έναντι SC ως προς την πολυπλοκότητα και την ταχύτητα εκτέλεσης	91
4.9	Κωδικοποίηση Ολικών Μορφομετρικών Χαρακτηριστικών στα Τοπικά Μέτρα της Ολοτοπικής Επιφάνειας	94
4.9.1	Ταξινόμηση Σχημάτων με το Ολοτοπικό Μέτρο	97
5	ΜΑΡΚΟΒΙΑΝΗ ΣΗΜΑΣΙΟΛΟΓΙΚΗ ΤΑΞΙΝΟΜΗΣΗ	100
5.1	Ταξινόμηση Χρηστών Πληροφοριακών Συστημάτων	101
5.1.1	Το Μαρκοβιανό Μοντέλο για τις Λέξεις Κλειδιά	102
5.1.2	Η Παροδική Συμπεριφορά Σύγκλισης	103
5.1.3	Η Διαδικασία Σύγκλισης σαν Βάση για Ομαδοποίηση των Καταστάσεων της Μαρκοβιανής Ανέλιξης	105
5.1.4	Στατιστικά Μέτρα για τις Αρχικές Συνθήκες	109
5.1.5	Η Απόσταση Ανάμεσα στις Λέξεις Κλειδιά	111

5.1.6	Η απόσταση ανάμεσα στους χρήστες	114
5.2	Αυτόματη Περιγραφή και Ταξινόμηση Εικόνων μέσω Διάδρασης με τους Χρήστες Πληροφοριακών Συστημάτων	115
5.2.1	Σχετικές Εργασίες	116
5.2.2	Συνεισφορά της μεθόδου MSI σε εργασίες ABIR	117
5.2.3	Η προτεινόμενη προσέγγιση	118
5.2.4	Σύγκριση με LSI	120
6	ΣΥΜΠΕΡΑΣΜΑΤΑ ΚΑΙ ΜΕΛΛΟΝΤΙΚΗ ΕΡΕΥΝΑ	129
7	ΑΝΑΦΟΡΕΣ	131
A	ΣΧΕΤΙΚΕΣ ΔΗΜΟΣΙΕΥΣΕΙΣ ΣΥΓΓΡΑΦΕΑ	151
A.1	Περιοδικά Ταξινομημένα κατά Science Citation Index (SCI)	151
A.2	Συνέδρια με Κριτές	151
A.3	Σε Στάδιο Κρίσης	151

Ευρετήριο Σχημάτων

1	Το οπτικό σύστημα των πρωτεύοντων	46
2	Μέτρηση της επίπεδης καμπυλότητας με διαδοχικές περιστροφές κατάλληλων αποδεκτικών πεδίων.	47
3	Δύο σχήματα και οι αντίστοιχες εικόνες μετά την κωδικοποίηση της επίπεδης καμπυλότητας στις τιμές της έντασης. Τα εικονοστοιχεία που βρίσκονται σε ακμές με υψηλή καμπυλότητα έχουν μεγαλύτερη ένταση (φωτεινότερα) στις προκύπτουσες εικόνες.	49
4	Τα στρώματα για τον υπολογισμό της καμπυλότητας.	52
5	Τα στρώματα για τον υπολογισμό του σκελετού καμπυλότητας.	54
6	Ο υπολογισμός του σκελετού καμπυλότητας από κυκλικά αποδεκτικά πεδία διαφόρων διαμέτρων.	55
7	Δείγμα εκπαίδευσης και δείγμα ελέγχου.	56
8	Στις περιττές στήλες (πρώτη και τρίτη) βλέπουμε τα σχήματα του επιπέδου L3 με την καμπυλότητα του περιγράμματος να έχει κωδικοποιηθεί στην φωτεινότητα Στις άρτιες στήλες βλέπουμε τα σχήματα του επιπέδου L4 όπως εμφανίζονται με σκελετό καμπυλότητας.	58
9	Σε αυτό το παράδειγμα απεικονίζεται το ανεπιθύμητο αποτέλεσμα της εξομάλυνσης. Ενώ τα σχήματα A και B ήταν αρχικά διαφορετικά (περιπτώσεις "α" και "β"), με την εξομάλυνση των θορυβωδών εκδοχών τους (γ) και (δ), τα δύο σχήματα γίνονται πανομοιότυπα (περιπτώσεις "ε" και "στ"), εξαλείφοντας τον κρίσιμο σχηματισμό (σταυρός) που χαρακτήριζε το σχήμα B.	61
10	Μία κλειστή επίπεδη καμπύλη και η αντίστοιχη επιφάνεια GLT. Οι τιμές απόστασης επιπέδου κανονικοποιούνται στη διάμετρο της καμπύλης. Η απόσταση στο περίγραμμα μοιράζεται σε 100 σημεία για πιο εύκολη επιθεώρηση της αντιστοίχισης με τα πραγματικά σημεία του περιγράμματος.	63
11	Τοπικές δεσμεύσεις των συναρτήσεων άποψης. Δύο μέρη του ίδιου περιγράμματος, ο εγγύτατος κύκλος, το σύστημα αναφοράς του Frenet και η συνάρτηση άποψης $v_x(s_*) = \ \mathbf{r} \ $	68
12	Σιλουέτες από την συλλογή του KIMIA και οι αντίστοιχες αναπαραστάσεις εμβαδών απόψεων.	70
13	Καμπυλότητα και Αναπαράσταση Εμβαδών Απόψεων (VAR) για δύο σχήματα και τις εξομαλυμένες εκδοχές τους. Συγκρίνοντας τις θορυβώδεις με τις εξομαλυμένες εκδοχές αυτών των σχημάτων, παρατηρούμε σε σύγκριση με την καμπυλότητα, την αυξημένη αντοχή της VAR στις υψηλόσυχνες αναταραχές περιγράμματος	75

- 14 Η καμπυλότητα, η VAR και η πρώτη και η δεύτερη παράγωγος της VAR για δύο τυπικά εξομαλυμένα σχήματα παρουσιάζονται σε μία κάθετη διάταξη για εύκολη σύγκριση. Προσέξτε την αντιστοιχία ανάμεσα στη δεύτερη παράγωγο της VAR (συνεχής γραμμή στα γραφήματα της τελευταίας σειράς) και την καμπυλότητα (γραφήματα της δεύτερης σειράς) ως προς τη θέση των τοπικών ακρότατων. Τα αντίστοιχα σημεία επισημαίνονται με '*' και 'ο'. Αυτά που επισημαίνονται με το 'ο' απεικονίζουν την περίπτωση στην οποία τα τοπικά ακρότατα της $\ddot{\phi}_\alpha$ (η $\ddot{\phi}_\beta$) που αντιστοιχούν επίσης στα τοπικά ακρότατα της κ_α (η κ_β) δεν είναι απαραίτητα τοπικά ακρότατα της ϕ_α (η ϕ_β). Να σημειωθεί επίσης ο τρόπος που η θέση και ο προσανατολισμός (υπό την έννοια της ενότητας 4.5) των σημείων του περιγράμματος ποσοτικοποιούνται στις τιμές VAR (περιγράμματα τρίτης σειράς) και την πρώτη παράγωγο της VAR (διακεκομμένη γραμμή στα περιγράμματα της τελευταίας σειράς) αντίστοιχα. 78
- 15 Οι θορυβώδεις εκδοχές ενός δεδομένου σχήματος για σταδιακά αυξανόμενες τιμές διασποράς μιας διαταραχής Gauss και οι αντίστοιχοι LAII και VAR περιγραφείς. Ο αριθμός πάνω από κάθε γράφημα χρησιμοποιείται για τη μέτρηση του αποτελέσματος του θορύβου. Πρόκειται για το άθροισμα των κατά σημείο απόλυτων διαφορών ανάμεσα στο γράφημα που αντιστοιχεί στο θορυβώδες σχήμα και το γράφημα που αντιστοιχεί στο αρχικό σχήμα. Κάθε κατα σημείο διαφορά θεωρείται ως κλάσμα της αρχικής τιμής. Όσο πιο υψηλός είναι αυτός ο αριθμός, τόσο πιο παραμορφωμένο είναι το αντίστοιχο θορυβώδες γράφημα σε σύγκριση με το αντίστοιχο αρχικό γράφημα. Οι αριθμοί δικαιολογούν την οπτική εντύπωση ότι η VAR είναι πιο ανθεκτική στο θόρυβο από τη LAII. 80
- 16 Παρουσιάζονται οι έξι εγγύτερες αντιστοιχίσεις σε θορυβώδης εκδοχές του συνόλου των 24 σχημάτων του KIMIA για δύο εφαρμογές δύο διαφορετικών μεθόδων σύγκρισης. Η πραγματική απόσταση του σχήματος από τη θορυβώδη εκδοχή του είναι ο αριθμός ακριβώς πάνω από κάθε σχήμα. Στην πρώτη σειρά είναι οι θορυβώδεις εικόνες (εικόνες ερώτησης) που προκύπτουν με διαταραχή Gauss των περιγραμμάτων τους με διασπορά $\sigma = 0.5$. Σε καθεμία από τις άλλες σειρές υπάρχει, η σχετική με τον αντίστοιχο δείκτη διαβάθμισης (πρώτη στήλη) πιο κοντινή αντιστοιχία με καθένα από τα θορυβώδη σχήματα για δύο διαφορετικές μεθόδους, VAR και LAII και δύο διαφορετικές εφαρμογές δυναμικού προγραμματισμού. Η πρώτη εφαρμογή χρησιμοποιεί την VAR για την επιλογή της αρχικής αντιστοιχίας και για τις δύο μεθόδους (VAR/VAR και LAII/VAR) ενώ η δεύτερη εφαρμογή εξετάζει εξαντλητικά την αντιστοιχία σε κάθε 5ο σημείο του περιγράμματος και για τις δύο μεθόδους (VAR/5P και LAII/5P). Η VAR παρουσιάζει καλύτερα αποτελέσματα από την LAII αφού και οι πρώτες και οι δεύτερες αντιστοιχίσεις είναι σωστές για τη VAR (τα θορυβώδη σχήματα αντιστοιχούν σε σχήματα της σωστής κατηγορίας), ενώ τρεις από τις δεύτερες αντιστοιχίσεις δεν είναι σωστές για το LAII. Τα αποτελέσματα για τις άλλες αντιστοιχίσεις είναι επίσης καλύτερα για την VAR σε σύγκριση με αυτές της LAII. Συγκρίνοντας τις εφαρμογές και για τις δύο μεθόδους, βλέπουμε ότι επιλέγοντας την αρχική αντιστοιχία μέσω της VAR βελτιώνεται η απόδοση αλλά και η ακρίβεια αναγνώρισης αφού η VAR επιλέγει σημεία εκκίνησης που έχουν διαισθητική σύνδεση με ολόκληρο το σχήμα. 82

- 17 Πίνακες απόστασης για τις τέσσερις καλύτερες αντιστοιχίσεις για σταδιακά αυξανόμενο θόρυβο περιγράμματος. Συγκρίνονται οι VAR/VAR (VAR στο Σχήμα) και LAII/VAR (LAII στο Σχήμα). Οι βαθμολογίες αντιστοίχισης για την LAII/VAR υποβαθμίζονται με ταχύτερο ρυθμό σε σύγκριση με τις βαθμολογίες της VAR/VAR, γεγονός που αποδεικνύει ότι η VAR/VAR έχει αυξημένη αντίσταση στον θόρυβο. 85
- 18 Ένα ζήτημα σχετικά με την επιλογή της ακτίνας του κυκλικού δίσκου στην εφαρμογή της μεθόδου LAII παρουσιάζεται εδώ. Ο κυκλικός δίσκος με κέντρο στο σημείο A τέμνει το περίγραμμα σε πολλαπλές περιοχές, συνεπώς, η περιοχή του που βρίσκεται μέσα στο περίγραμμα δεν μπορεί να αναπαραστήσει σωστά την καμπυλότητα στο σημείο A . Αυξάνεται λοιπόν η πολυπλοκότητα σε περιπτώσεις όπου πρέπει να υπολογιστεί μία έγκυρη ακτίνα κυκλικού δίσκου για ολόκληρο το σχήμα ή όπου πρέπει να πραγματοποιηθεί μία δυναμική τροποποίηση της ακτίνας σύμφωνα με τις μεταβολές του σχήματος. 87
- 19 Ένα σχήμα με μία συρρίκνωση και δύο διατάσεις. 96
- 20 Σχήματα από την βάση του Kimia και τα τρία πιο κοντινά τους σύμφωνα με το ολοτοπικό μέτρο 98
- 21 Γεωμετρική αναπαράσταση της r -ισοεπιφάνειας της προτεινόμενης MSI απόστασης στις τρεις διαστάσεις. Το τρισδιάστατο ελλειψοειδές καταδεικνύει ως ισοαπέχοντα τα αντιδιαμετρικά ζεύγη σημείων στους άξονές του, συνεπώς οι αποστάσεις στην διεύθυνση του ιδιοδιανύσματος v_1 βαραίνουν περισσότερο και αυτές στου v_3 λιγότερο. 113
- 22 Η συνολική Μαρκοβιανή διαδικασία του πειράματος της Ενότητας 5.2.4. Μόνο το μέρος του δικτύου που διαμορφώνει συστάδες γύρω από τις έννοιες *Greek*, *Hawaiian* και *Islands* παρουσιάζεται γραφικά. Οι πιθανότητες μετάβασης που οδηγούν έξω από αυτό το μέρος του δικτύου έχουν τεθεί μηδέν και το βάρος τους διανεμήθηκε ομοιόμορφα στις αντίστοιχες συνδέσεις εντός του εμφανούς δικτύου. Οι πιθανότητες μετάβασης μικρότερες από 0.05 παρελήφθησαν. 122
- 23 Ο πίνακας απόστασης μεταξύ 30 εικόνων δύο κατηγοριών. Οι εικόνες $G1 - G15$ είναι οι πρώτες 15 εικόνες του πίνακα 4 και ανήκουν στην κατηγορία *Greek* ενώ οι εικόνες $H33 - H47$ αντιπροσωπεύονται από τις αντίστοιχες 15 σειρές του ίδιου πίνακα και ανήκουν στην κατηγορία *Hawai*. Η προτεινόμενη απόσταση MSI (άνω αριθμός) και η απόσταση LSI (κάτω αριθμός) συγκρίνονται. Οι εννέα καλύτερες αντιστοιχίες είναι μαρκαρισμένες με τους αντίστοιχους εκθέτες. Η MSI επιτυγχάνει καλύτερα αποτελέσματα δεδομένου ότι όλες οι πρώτες εννέα αντιστοιχίες είναι σωστές (οι εικόνες ταιριάζουν με τις εικόνες στην ίδια κατηγορία), ενώ πολλές από τις πρώτες εννέα αντιστοιχίες δεν είναι σωστές για την LSI . Το διάγραμμα ακρίβειας / ανάκλησης για αυτό το πείραμα, για όλες τις 64 εικόνες του πίνακα 4, παρουσιάζεται στο Σχ. 24. 124

- 24 Σύγκριση ακρίβειας / ανάκλησης μεταξύ LSI και του προτεινόμενου συστήματος MSI για διάφορες παραμέτρους των δύο αλγορίθμων. Το πείραμα εκτελείται στις 64 εικόνες του πίνακα 4. Οι πρώτες 32 εικόνες ανήκουν στην κατηγορία Greek και οι επόμενες 32 στην κατηγορία Hawaiian. Οι γραφικές παραστάσεις διαμαντιών αντιπροσωπεύουν τα αποτελέσματα της προτεινόμενης απόστασης (που υπονοείται από την γενική Μαρκοβιανή διαδικασία του πίνακα 3) για τις περιπτώσεις $n = 1, n = 5, n = 10, n = 12, n = 14, n = 15$. Οι γραφικές παραστάσεις κύκλων αντιπροσωπεύουν τα αποτελέσματα όταν χρησιμοποιείται η LSI απόσταση με τον πίνακα συχνότητας εικόνας/λέξης κλειδιού του πίνακα 4 για τις παραμέτρους $k = 10$ και $k = 21$. Περισσότερα στο κείμενο.

Ευρετήριο Πινάκων

- 1 Βαθμολογίες Bullseye, πολυπλοκότητα και χρόνοι εκτέλεσης για διάφορες μεθόδους που χρησιμοποιούν τη βάση MPEG7 CE-Shape-1 Part B. 90
- 2 Παρουσιάζεται ο αριθμός των σωστών αντιστοιχίσεων από όλες τις πιθανές 1400 αντιστοιχίσεις σύμφωνα με το πείραμα της ενότητας 4.8.4. Στην πρώτη στήλη υπάρχει η σειρά της αντιστοιχίας (1η πιο κοντινή, 2η πιο κοντινή κ.λ.π.), στη δεύτερη και τρίτη στήλη υπάρχει ο αριθμός των σωστών αντιστοιχιών για κάθε διαβάθμιση για τις μεθόδους GLS/VAR και GLS/5P αντίστοιχα. Σημειώνουμε ότι η επιλεκτική – με βάση τις διαισθητικές ιδιότητες του σχήματος – μείωση του χώρου αναζήτησης για αντιστοιχίσεις των σημείων εκκίνησης μέσω της VAR δεν μειώνει μόνον την πολυπλοκότητα και βελτιώνει τον χρόνο εκτέλεσης (Πίνακας 1) αλλά οδηγεί επίσης και σε καλύτερες βαθμολογίες αντιστοιχίσης. Το αποτέλεσμα αυτό συμφωνεί με το πείραμα του Σχ. 16. 92
- 3 Ο πίνακας μετάβασης πρώτης τάξης της Μαρκοβιανής αλυσίδας του Σχ. 22. 121
- 4 64 εικόνες τακτοποιημένες όσον αφορά την απόστασή τους από την ερώτηση *Greek Islands*. Κάθε σειρά αντιπροσωπεύει μια εικόνα. Η πρώτη στήλη κρατά τον αριθμό ετικετών εικόνας, οι δεύτερη και τρίτη στήλες κρατούν τον αριθμό ταξινόμησης όσον αφορά τις αποστάσεις *MSI* και *LSI* από την ερώτηση *Greek Islands* αντίστοιχα για αυτήν την εικόνα. Το υπόλοιπο των καταχωρήσεων αντιστοιχεί στις πιθανότητες ισορροπίας της αντίστοιχης Μαρκοβιανής αλυσίδας που αντιπροσωπεύει κάθε εικόνα. Εμφανίζονται μόνο οι μη μηδενικές καταχωρήσεις ομαδοποιημένες κατά κατηγορία για καλύτερη επισκόπηση. Η ταξινόμηση για την απόσταση *MSI* υπονοείται από τον *AMC* του Σχ. (22). 30 από αυτές τις εικόνες παρουσιάζονται στον πίνακα απόστασης του Σχ. (23) για τις σειρές 1-15 και 33-47 αυτού του πίνακα. . . 128

Απόδοση Ακρονυμίων - Όρων

HVS(ΑΟΣ)	: Human Visual System - Ανθρώπινο Οπτικό Σύστημα
CAR	: Circular Autoregressive Model - Κυκλικό Αυτοπαλίνδρομο Μοντέλο
AR	: AutoRegressive - Αυτοπαλίνδρομο
AHF	: Arc Height Function - Συνάρτηση Ύψους Τόξου
MAT	: Medial Axis Transform - Μετασχηματισμός Μεσαίου Άξονα
GCF	: Geometric Correlation Function - Συνάρτηση Γεωμετρικής Συσχέτισης
GLT	: Global-Local Transformation - Ολοτοπικός Μετασχηματισμός
GLS	: Global-Local Surface - Ολοτοπική Επιφάνεια
VAR	: View Area Representation - Αναπαράσταση Εμβαδών Απόψεων
LAII	: Local Area Integral Invariant
RF	: Receptive Field - Αποδεκτικό Πεδίο - Δεκτικό Πεδίο - Πεδίο Υποδοχής
SC	: Shape Contexts - Μέθοδος Περιβάλλοντος Σχήματος
TPS	: Thin Plate Splines
MSI	: Markovian Semantic Indexing - Μαρκοβιανή Σημασιολογική Ταξινόμηση
LSI	: Latent Semantic Indexing - Λανθάνουσα Σημασιολογική Ταξινόμηση
pLSI	: probabilistic Latent Semantic Indexing - Πιθανοθεωρητική Λανθάνουσα Σημασιολογική Ταξινόμηση
LDA	: Latent Dirichlet Allocation
ABIR	: Annotation Based Image Retrieval - Ανάκτηση Εικόνων από Κειμενικές Περιγραφές
AMC	: Aggregate Markovian Chain - Συνολική Μαρκοβιανή Αλυσίδα - Μαρκοβιανή Αλυσίδα Γενικής Συμπεριφοράς
LGN(ΠΓΠ)	: Lateral Geniculate Nucleus-Πλάγιος Γονατώδης Πυρήνας

1. ΕΙΣΑΓΩΓΗ

Η Διατριβή επικεντρώνεται σε θέματα αναπαράστασης, κωδικοποίησης, ταξινόμησης και ανάκτησης της οπτικής πληροφορίας είτε αυτή είναι σε μορφή σχήματος είτε σε μορφή εικόνας με μία προσέγγιση που είναι συμβατή με βιολογικά αλλά και λειτουργικά δεδομένα σε σχέση με την ανθρώπινη αντίληψη. Στην περίπτωση της αναπαράστασης και κωδικοποίησης σχημάτων εισάγονται έννοιες βιολογικής αλλά και αντιληπτικής συμβατότητας. Στην περίπτωση των εικόνων εισάγονται έννοιες γνωσιακής συμβατότητας που χρησιμοποιούνται για τον σκοπό της καλύτερης ταξινόμησης και ανάκτησης εικόνων.

Η απόδοση σημασιολογίας σε μία ψηφιακή εικόνα είναι ένα κεντρικό πρόβλημα της τεχνητής νοημοσύνης με ιδιαίτερο ενδιαφέρον στα συστήματα τεχνητής όρασης. Η διαδικασία της κατανόησης και της ταξινόμησης του περιεχομένου μιας ψηφιακής εικόνας ακολουθεί πολλά στάδια τα οποία στην παρούσα εργασία αναλύονται. Παρουσιάζονται συνεισφορές με έμφαση σε βιολογική και αντιληπτική με τον άνθρωπο συμβατότητα. Τα στάδια ανάλυσης και επεξεργασίας μέχρι την κατανόηση μιας ψηφιακής εικόνας ταξινομούνται ως εξής:

- Εξαγωγή περιγραφών από τα αριθμητικά δεδομένα των εικονοστοιχείων της εικόνας.
- Διαχωρισμός και αναπαράσταση του σχήματος η του περιεχομένου της εικόνας.
- Σύγκριση και ταξινόμηση των εικόνων.
- Σημασιολογική κατηγοριοποίηση των εικόνων.

Κάθε ένα από τα παραπάνω στάδια προϋποθέτει το προηγούμενο και σε μεγάλο βαθμό οι δυνατότητές του καθορίζονται από τα δεδομένα που έχουν διαμορφωθεί στα προηγούμενα στάδια. Το πρώτο στάδιο αναφέρεται σε μεθόδους πάνω στα αριθμητικά δεδομένα για την εξαγωγή περιοχών με στόχο την κατανόηση του περιεχομένου της εικόνας. Τυπικές τεχνικές είναι η ανίχνευση ακμών, και εν συνεχεία συγκεκριμένων γεωμετρικών σχημάτων όπως ευθειών, κύκλων, καμπύλων η και πιο σύνθετων προτύπων με διάφορα φίλτρα και ταίριασμα προτύπων.

Το δεύτερο στάδιο χρησιμοποιεί τις εξαγόμενες από το στάδιο 1 περιοχές για να αναπαραστήσει το περιεχόμενο της εικόνας μέσω κάποιου περιγραφέα. Υπάρχουν πληθώρα τέτοιων περιγραφέων που παρουσιάζονται αναλυτικά και μετασχηματισμοί που εφαρμόζονται απ ευθείας στις ψηφιακές εικόνες. Οι περιγραφείς στην συνέχεια χρησιμοποιούνται στο στάδιο 3 για το ταίριασμα και ταξινόμηση των εικόνων. Η κατασκευή μιας απόστασης ανάμεσα στις εικόνες που θα βασίζεται στους αντίστοιχους περιγραφείς είναι κεντρική εργασία αυτού του σταδίου.

Τέλος επειδή καμία ταξινόμηση εικόνων δεν μπορεί να είναι χρήσιμη εάν δεν λάβει υπόψη την εφαρμογή για την οποία προορίζεται, στο στάδιο 4 εμπλουτίζεται η απόσταση

ανάμεσα σε εικόνες με σημασιολογικές συσχετίσεις γνωσιακού χαρακτήρα. Θεωρείται επίσης πολύ σημαντικό στην κατασκευή των τεχνητών μεθόδων και μοντέλων να τηρείται συμβατότητα είτε με τα διαισθητικά γνωρίσματα της ανθρώπινης αντίληψης είτε με τις λειτουργίες του Ανθρώπινου Οπτικού Συστήματος (ΑΟΣ) που αποτελεί το πιο εξελιγμένο σύστημα αντίληψης που γνωρίζουμε μέχρι σήμερα. Στην συνέχεια γίνεται ανάλυση των σταδίων (Κεφάλαια 2 και 3) καθώς και της συμβατότητάς τους με το ΑΟΣ και παρουσιάζονται οι ακόλουθες συνεισφορές:

- Για τα στάδια 1 και 2 παρουσιάζεται στο κεφάλαιο 3 μία μέθοδος ανίχνευσης, μέτρησης καμπυλότητας και εξαγωγής του *σκελετού καμπυλότητας* που είναι συμβατή με την λειτουργία του ΑΟΣ. Για το Σταδιο 2 ειδικότερα παρουσιάζεται μια μέθοδος κωδικοποίησης του σκελετού (μεταχηματισμός μεσαίου άξονα) ενός κλειστού σχήματος. Ο σκελετός εμπεριέχει και πληροφορίες καμπυλότητας περιγράμματος. Η μέθοδος είναι επίσης συμβατή με το ΑΟΣ.
- Για το στάδιο 3 παρουσιάζεται στο κεφάλαιο 4 ένας νέος μετασχηματισμός, ο *ολοτοπικός*, που εφαρμόζεται στο περίγραμμα, που έχει εξαχθεί μέσω διαδικασιών εξαγωγής περιγράμματος στα δύο προηγούμενα στάδια, για να παράγει περιγραφείς σχήματος. Αποδεικνύεται ότι ο μετασχηματισμός αυτός προσφέρει μεγάλες δυνατότητες στην αναγνώριση σχήματος και είναι συμβατός με την ανθρώπινη αντίληψη.
- Για το στάδιο 4 παρουσιάζεται στο κεφάλαιο 5 μια μέθοδος απόδοσης σημασιολογικών διασυνδέσεων στις λέξεις κλειδιά που αποδίδουν στις εικόνες οι χρήστες ενός πληροφοριακού συστήματος. Κατασκευάζεται μια απόσταση ανάμεσα στις λέξεις κλειδιά που βασίζεται σε ένα στοχαστικό μοντέλο συμπεριφοράς των χρηστών του πληροφοριακού συστήματος. Με αυτόν τον τρόπο κατασκευάζεται μια σημασιολογική συσχέτιση στο χώρο των λέξεων κλειδιών. Η ίδια τεχνική χρησιμοποιείται στην συνέχεια για να ορισθεί μια ομαδοποίηση Μαρκοβιανού χαρακτήρα στον χώρο των λέξεων κλειδιών. Μέσω της ομαδοποίησης αυτής ορίζεται μία απόσταση ανάμεσα στις εικόνες αυτήν την φορά. Οι κειμενικές περιγραφές που αποδόθηκαν στις εικόνες αυτόματα απο το σύστημα, μέσω της αλληλεπίδρασής του με τους χρήστες που διάλεξαν αυτές τις εικόνες, χρησιμοποιούνται στην συνέχεια σε ένα Μαρκοβιανό μοντέλο ανάκτησης που δείχνουμε ότι έχει μεγαλύτερες δυνατότητες από τις μεθόδους *LSI*.

2. ΑΝΑΔΡΟΜΗ ΣΕ ΘΕΩΡΙΕΣ ΚΑΙ ΤΕΧΝΙΚΕΣ ΣΧΕΤΙΚΑ ΜΕ ΤΗΝ ΟΠΤΙΚΗ ΠΛΗΡΟΦΟΡΙΑ

Δεδομένου ότι τις τελευταίες δεκαετίες έχει προκύψει ένα τεράστιος όγκος εργασιών σχετικά με την ανάλυση σχημάτων και η πορεία της εξέλιξης έχει επηρεαστεί από επιτεύγματα σε άλλους σχετικούς τομείς έρευνας καθώς και από εφαρμογές ανάλυσης εικόνας, θα γίνει εδώ μια προσπάθεια να καλυφθεί μία ποικιλία των πιο χαρακτηριστικών μεθόδων για την περιγραφή σχημάτων αλλά και θεωριών για την οπτική αντίληψη. Επίσης θα παρουσιαστεί ένα σύνολο αναφορών στην βιβλιογραφία που ο αναγνώστης μπορεί να χρησιμοποιήσει για περαιτέρω έρευνα.

2.1. Αναδρομή σε τεχνικές αναπαράστασης και ανάκτησης σχήματος

Ένα τυπικό σύστημα επεξεργασίας και ανάλυσης εικόνας δέχεται μια ψηφιακή εικόνα που περιέχει τα αντικείμενα ενδιαφέροντος. Προκειμένου να γίνει κατανοητό το περιεχόμενο της εικόνας πρέπει να αναγνωριστούν τα αντικείμενα που βρίσκονται σε αυτήν. Υπάρχουν πολλές εφαρμογές απεικόνισης όπου η ανάλυση εικόνας μπορεί να αφορά σε ανάλυση σχήματος, (πχ όργανα, κύτταρα, μέρη μηχανών, χαρακτήρες). Οι μέθοδοι αναπαράστασης σχήματος μπορεί να οδηγούν σε μια μη αριθμητική αντιπροσώπευση του αρχικού σχήματος έτσι ώστε τα σημαντικά χαρακτηριστικά του διατηρούνται. Η περιγραφή σχήματος αναφέρεται στις μεθόδους που οδηγούν σε έναν αριθμητικό περιγραφέα του σχήματος και είναι ένα βήμα επόμενο για να διαμορφώσει την τελική αναπαράσταση. Οι μέθοδοι περιγραφής σχήματος παράγουν ένα διάνυσμα χαρακτηριστικών για την εικόνα. Ο στόχος της περιγραφής είναι να χαρακτηριστεί μονοσήμαντα το σχήμα χρησιμοποιώντας το διάνυσμα περιγραφέα. Μία πολύ σημαντική ιδιότητα ενός περιγραφέα σχήματος είναι η ανεξαρτησία στην μεταφορά, κλίμακα και την περιστροφή, διότι αυτοί οι τρεις μετασχηματισμοί, εξ ορισμού, δεν αλλάζουν το σχήμα του αντικειμένου. Η είσοδος στους αλγόριθμους ανάλυσης σχήματος είναι σχήματα, όπως ψηφιακές εικόνες. Το ταίριασμα σχημάτων αναφέρεται στις μεθόδους που χρησιμοποιούνται για την αναγνώριση αντικειμένου όταν ένα σύνολο γνωστών προτύπων αντικειμένων συγκρίνεται με ένα άγνωστο αντικείμενο που ανιχνεύεται σε μια εικόνα. Τυπικά, μία μέθοδος αναπαράστασης σχήματος χρησιμοποιείται για να καθορίσει το διάνυσμα περιγραφέα σχήματος για κάθε πρότυπη μορφή, όπως και για την άγνωστη μορφή στη σκηνή. Η άγνωστη μορφή αντιστοιχείται σε μια από τις πρότυπες μορφές με τη σύγκριση των διανυσμάτων περιγραφέα σχήματος χρησιμοποιώντας μία συνάρτηση απόστασης.

Το πρόβλημα της ανάλυσης σχήματος έχει για δεκαετίες αποτελέσει θέμα έρευνας [1, 2, 3] και πολλά βιβλία έχουν πραγματευτεί το θέμα αυτό. [4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16] Αναλυτικότερες τεχνικές μπορούν να βρεθούν στο [17]. Οι μέθοδοι ανάλυσης σχήματος μπορούν να ταξινομηθούν σύμφωνα με διαφορετικά κριτήρια. Ο Pavlidis [1] έχει

προτείνει τις ακόλουθες ταξινομήσεις. Η πρώτη ταξινόμηση είναι βασισμένη στη χρήση των σημείων περιγράμματος σχήματος σε αντιδιαστολή με το εσωτερικό του αντικειμένου. Οι δύο προκύπτουσες κατηγορίες αλγορίθμων είναι γνωστές ως περιμετρικές η μέθοδοι περιγράμματος (επίσης αποκαλούμενες εξωτερικές) και ολικές (ή εσωτερικές), αντίστοιχα. Οι περιμετρικές μέθοδοι είναι αλγόριθμοι που αναλύουν το περίγραμμα του σχήματος [18, 19, 20, 21, 22, 23, 24] καθώς και διάφοροι μετατροπείς του περιγράμματος με χρήση του μετασχηματισμού Fourier [25, 26, 27, 28, 29]. Στις ολικές μεθόδους ανήκουν, η διάμεση μετατροπή άξονα (επίσης αποκαλούμενου συμμετρικού άξονα) (MAT) που έχει προταθεί από τον Blum και που περιγράφεται στα [30, 31, 32, 18, 33, 34, 21, 22], οι μετασχηματισμοί που βασίζονται στις ροπές [35, 36, 37, 38], και οι μέθοδοι αποσύνθεσης σχήματος σε άλλες πρωτογενείς μορφές [39, 40, 41, 42].

Μια άλλη ταξινόμηση των αλγορίθμων ανάλυσης σχήματος μπορεί να γίνει με βάση το εάν το αποτέλεσμα της ανάλυσης είναι αριθμητικό ή όχι. Παραδείγματος χάριν, ο MAT παράγει μια νέα εικόνα που περιέχει έναν συμμετρικό άξονα. Άλλες τεχνικές παράγουν ως αποτέλεσμα βαθμωτά μεγέθη ή διανύσματα. Πιο πρόσφατα παραδείγματα περιλαμβάνουν διάφορες βασισμένες στον μετασχηματισμό Fourier μεθόδους [25, 26, 27, 28, 29] ή διαδικασίες βασισμένες σε ροπές [35, 36, 37, 38].

Μια τρίτη ταξινόμηση των μεθόδων ανάλυσης σχήματος μπορεί να γίνει με βάση την συντήρηση πληροφοριών. Οι μέθοδοι που επιτρέπουν την ακριβή (ή τουλάχιστον αρκετά ακριβή) αναδημιουργία σχήματος από τον περιγραφέα του καλούνται μέθοδοι που συντηρούν την πληροφορία, σε αντιδιαστολή με τις μεθόδους που είναι μόνο ικανές για μερική αναδημιουργία οι οποίες καλούνται μέθοδοι μη-συντηρητικές της πληροφορίας. Πολλοί περιγραφείς σχήματος υποφέρουν από το πρόβλημα μη-συντήρησης της πληροφορίας.

Διάφορες αναπαραστάσεις σχήματος που βασίζονται στο περίγραμμα έχουν παρουσιαστεί στην βιβλιογραφία. Οι Zahn και Roskies [25] και οι Bennet και McDonald [43] χρησιμοποίησαν μια συνάρτηση της εφαπτομένης γωνίας ως προς το μήκος τόξου. Η γωνία της εφαπτομένης σε κάποιο σημείο μετράται σε σχέση με την γωνία της εφαπτομένης στο αρχικό σημείο. Η συνάρτηση καλείται επίσης συνάρτηση ακρότατου και έχει χρησιμοποιηθεί από τους Arkin et al [44] για τη σύγκριση πολυγωνικών σχημάτων.

Ο Granlund [26], οι Richards και Hemami [27] και οι Persoon και Fu [28] επεξέτειναν την μέθοδο χρησιμοποιώντας μια πιο σύνθετη συνάρτηση. Μία άλλη προσέγγιση για την παραγωγή μονοδιάστατης αναπαράστασης του περιγράμματος είναι η χρήση του κέντρου βάρους (centroid) του σχήματος. Στην πρώτη παραλλαγή της ιδέας, οι τιμές της μονοδιάστατης συνάρτησης ισούνται με τις αποστάσεις ανάμεσα στο κέντρο βάρους του σχήματος και τα σημεία του περιγράμματος. Τα σημεία περιγράμματος επιλέγονται με σκοπό οι κεντρικές γωνίες να είναι ίσες μεταξύ τους. Μία άλλη παραλλαγή της ιδέας είναι η χρήση των αποστάσεων ανάμεσα σε διαδοχικά σημεία περιγράμματος ως τιμών της μονοδιάστατης συνάρτησης. Στην τρίτη παραλλαγή, τα σημεία περιγράμματος ενός

σχήματος παρουσιάζονται κατά προσέγγιση με ένα πολύγωνο κατά τρόπο ώστε όλες οι πλευρές να έχουν το ίδιο μήκος. Οι τιμές της μονοδιάστατης συνάρτησης είναι ίσες με την γωνία ανάμεσα στις πλευρές του πολυγώνου και μία γραμμή αναφοράς. Οι Chang, Hwang Buehrer [45] δημιούργησαν μία συνάρτηση για την απόσταση του κέντρου βάρους από τα χαρακτηριστικά σημεία. Στην μέθοδο που πρότειναν, τα χαρακτηριστικά σημεία είναι τα σημεία μεγάλης καμπυλότητας. Υπάρχουν δύο προσεγγίσεις για τον εντοπισμό τέτοιων σημείων. Η πρώτη βασίζεται στον υπολογισμό της καμπυλότητας άμεσα από την καμπύλη του περιγράμματος. Η δεύτερη προσέγγιση βασίζεται στην πραγματοποίηση μίας πολυγωνικής προσέγγισης του σχήματος και στην χρήση των κορυφών του πολυγώνου ως χαρακτηριστικών σημείων όπου βασίζεται στο γεγονός ότι οι γωνίες είναι σημεία μεγάλης καμπυλότητας. Οι υπολογισθείσες αποστάσεις αποθηκεύονται σε μία συνδεδεμένη λίστα. Οι αποστάσεις κατανέμονται για να επιτευχθεί ανεξαρτησία ως προς περιστροφή. Οι αποστάσεις χωρίζονται επίσης με βάση την ελάχιστη απόσταση για να επιτευχθεί ανεξαρτησία ως προς κλίμακα. Η ανεξαρτησία ως προς μεταφορά επιτυγχάνεται αυτόματα μέσω της χρήσης του κέντρου βάρους.

Οι Wang et al. [46] χρησιμοποίησαν μία αλληλουχία ροπών ευθύγραμμων τμημάτων ως μία μονοδιάστατη συνάρτηση. Τα ευθύγραμμα τμήματα προκύπτουν από τον διαχωρισμό της ακτίνας από το κέντρο της μάζας στο σημείο περιγράμματος. Τα τμήματα χωρίζονται σε μέρη εντός του σχήματος και μέρη εκτός του σχήματος. Η μέθοδος των Zahn και Roskies [25] χρησιμοποιεί την αναπαράσταση του περιγράμματος του σχήματος της εφαιπόμενης γωνίας προς το μήκος τόξου. Το περίγραμμα εξετάζεται διαδοχικά με φορά προς τα δεξιά παράγοντας αρνητικές γωνίες σε σχέση με το αρχικό σημείο. Εφαρμόζεται τότε ο μετασχηματισμός Fourier στην συνάρτηση του περιγράμματος και οι συντελεστές που προκύπτουν χρησιμοποιούνται για την περιγραφή του σχήματος. Χάρη στην κανονικοποίηση του μήκους του τόξου, ο περιγραφέας του σχήματος παραμένει αμετάβλητος παρά τις αλλαγές στην κλίμακα. Ο περιγραφέας του σχήματος παραμένει αμετάβλητος στις μεταφορές γιατί η συνάρτηση της εφαιπόμενης γωνίας είναι αμετάβλητη ως προς την θέση του σχήματος. Η περιστροφή του αντικειμένου (δηλ. παραλλαγή του σημείου εκκίνησης) προκαλεί αλλαγή φάσης στον προκύπτοντα μετασχηματισμό Fourier, επομένως η εξέταση μόνο του μεγέθους των συντελεστών Fourier θα εξασφαλίσει την ανεξαρτησία της μεθόδου και ως προς τις περιστροφές. Το βασικό πλεονέκτημα αυτής της μεθόδου είναι ότι είναι εύκολη η εφαρμογή της και ότι βασίζεται σε μία καλά αναπτυγμένη θεωρία της ανάλυσης Fourier. Τα μειονέκτημα είναι ότι ο μετασχηματισμός Fourier δεν προσφέρει τοπικές πληροφορίες για το σχήμα. Μετά τον μετασχηματισμό Fourier, οι τοπικές πληροφορίες του σχήματος κατανέμονται σε όλους τους συντελεστές και δεν τοποθετούνται στο πεδίο συχνότητας. Οι αναπαραστάσεις της εφαιπόμενης της γωνίας προς το μήκος του τόξου υποφέρουν από πολύ μεγάλη ευαισθησία στον θόρυβο καθώς είναι δύσκολος ο καθορισμός της εφαιπόμενης της γωνίας σε θορυβώδες περίγραμμα.

Οι Young, Walker και Bowie [19] εισήγαγαν μία ενδιαφέρουσα έννοια, αυτή της ενέργειας κάμψης. Σύμφωνα με αυτή την προσέγγιση, ένα σχήμα μπορεί να αναπαρασταθεί με την βοήθεια της ενέργειας κάμψης. Η ενέργεια κάμψης υπολογίζεται αποτελεσματικά χρησιμοποιώντας τους συντελεστές Fourier και την σχέση του Parseval. Επιπλέον, οι συγγραφείς απέδειξαν ότι το σχήμα με την ελάχιστη μέση ενέργεια κάμψης είναι ο κύκλος.

Οι Kashyap και Chellappa [20] πρότειναν την χρήση του κυκλικού αυτοπαλίνδρομου (CAR) μοντέλου για την αναπαράσταση του κέντρου βάρους στην συνάρτηση απόστασης των σημείων του περιγράμματος. Το μοντέλο CAR χαρακτηρίζεται από ένα σύνολο αγνώστων παραμέτρων σε μία ανεξάρτητη ακολουθία. Από την στιγμή που το περίγραμμα είναι κλειστό, η μονοδιάστατη συνάρτηση περιγράμματος είναι περιοδική συνάρτηση. Το μοντέλο CAR που χρησιμοποίησε και ο Huang [47], είναι μία στοχαστική διαδικασία, που οι Kashyap και Chellappa ερεύνησαν περαιτέρω για τρόπους κωδικοποίησης και ανακατασκευής και απέδειξαν ότι οι στοχαστικές μέθοδοι μπορούν να χρησιμοποιηθούν για την περιγραφή κλειστών περιγραμμάτων.

Οι Dubois και Glanz [48] χρησιμοποίησαν το ίδιο αυτοπαλίνδρομο μοντέλο με το [20] αλλά ερεύνησαν τρεις επιπλέον μεθόδους για τη βελτίωση της απόδοσης στην κατηγοριοποίηση σχημάτων. Η πρώτη βελτίωση ήταν ο υπολογισμός του περιγραφέα με τέτοιο τρόπο ώστε να δοθεί έμφαση στα στοιχεία που ανήκαν στην ίδια τάξη εκπαίδευσης. Αποδείχθηκε ότι το βέλτιστο βάρος του χαρακτηριστικού είναι αντιστρόφως ανάλογο με την τυπική απόκλιση του χαρακτηριστικού από το σύνολο της τάξης εκπαίδευσης [49]. Η δεύτερη βελτίωση επιτυγχάνεται με την περιστροφή του συστήματος συντεταγμένων και την κλιμάκωση των περιστρεφόμενων αξόνων [49] με τρόπο ώστε να συγκεντρώνονται τα μέλη μίας κατηγορίας που βρίσκονται πιο κοντά στο νέο σύστημα συντεταγμένων. Η τρίτη τροποποίηση που περιλαμβάνει την χρήση υπερεπιπέδων για τον διαχωρισμό του χώρου, παράγει τις βέλτιστες παραμέτρους υπερεπιπέδου με την χρήση μίας διαδικασίας Ελάχιστου Μέσου Τετραγωνικού Σφάλματος [50] και τα πειραματικά αποτελέσματα απέδειξαν ότι το κανονικοποιημένο μοντέλο AR δημιουργεί εύχρηστους περιγραφείς σχήματος. Μία τροποποίηση της μεθόδου των Dubois και Glanz περιελάμβανε την στατιστική γνώση σχετικά με τον θόρυβο που προκύπτει από την διαδικασία απεικόνισης και τον δειγματισμό του περιγράμματος [51].

Οι Das, Paulik και Loh [52] πρότειναν μία διμεταβλητή τεχνική για το αυτοπαλίνδρομο μοντέλο του περιγράμματος και τα πειράματά τους είχαν ακόμη καλύτερα αποτελέσματα. Το γραμμικό μοντέλο AR έχει συγκριθεί με το μη γραμμικό μοντέλο της δευτεροβάθμιας εξίσωσης Volterra. Σύμφωνα με τους Kartikeyan και Sarkar, τα γραμμικά μοντέλα AR δεν επαρκούν για την αναγνώριση πιο σύνθετων (μη κυρτών) σχημάτων. Σε αυτές τις περιπτώσεις ένα γραμμικό μοντέλο μεγαλύτερης τάξης για την αύξηση της διάστασης του ανύσματος του περιγραφέα σχήματος είναι απαραίτητο. Τα μη γραμμικά μοντέλα

επίσης μπορούν να παράσχουν την μεγαλύτερη ακρίβεια που είναι απαραίτητη για την περιγραφή πιο σύνθετων σχημάτων. Τα πειράματα αυτά επέδειξαν ότι τα δευτεροβάθμια μοντέλα του Volterra πετυχαίνουν καλύτερη κατάταξη από το γραμμικό μοντέλο AR [23]. Το μειονέκτημα των μοντέλων περιγράμματος AR είναι ότι στην περίπτωση πιο σύνθετων ορίων, δεν επαρκεί ένας μικρός αριθμός παραμέτρων AR. Για τον λόγο αυτό, οι He και Kundu [53] συνδύασαν την χρήση του μοντέλου AR με το κρυφό μοντέλο του Markov όπου το περίγραμμα του σχήματος τμηματοποιήθηκε και κάθε τμήμα περιγράφηκε με βάση το μοντέλο AR. Τα προκύπτοντα ανύσματα αναλύθηκαν χρησιμοποιώντας το κρυφό μοντέλο του Markov.

Οι Lin, Dou και Wang [54] χρησιμοποίησαν την Συνάρτηση Ύψους Τόξου (AHF) για να χαρακτηρίσουν το περίγραμμα του σχήματος. Ο Freeman [55] έχει προτείνει μία μέθοδο κωδικοποίησης σχημάτων που λέγεται κωδικοποίηση αλυσίδας. Μπορείτε να βρείτε μία πιο λεπτομερή σύνοψη των μεθόδων και αλγορίθμων του Freeman στο [56]. Στο [57], ο Freeman χρησιμοποίησε μία περιγραφή του περιγράμματος με βάση τον κώδικα αλυσίδας για να εξάγει κρίσιμα συμπεράσματα τα οποία χρησιμοποίησε για να παραγάγει την περιγραφή ενός σχήματος που είναι αμετάβλητη ως προς τη μεταφορά, την περιστροφή και την κλίμακα. Παρουσίασε επίσης μία άλλη προσέγγιση του κώδικα αλυσίδας βασισμένη στο κεντροειδές προφίλ του σχήματος. Το κεντροειδές προφίλ είναι η συνάρτηση της απόστασης του κέντρου βάρους από τα σημεία περιγράμματος.

Οι Parui και Majumder [58] πραγματοποίησαν μία ανάλυση συμμετρίας χρησιμοποιώντας έναν τροποποιημένο κώδικα αλυσίδας. Τα όρια του σχήματος παρουσιάζονται ιεραρχικά ούτως ώστε στο υψηλότερο επίπεδο να χρησιμοποιηθεί ένας μικρότερος αριθμός κορυφών πολυγώνου, ενώ στο χαμηλότερο επίπεδο να χρησιμοποιηθεί η καλύτερη προσομοίωση πολυγώνου. Η αναζήτηση της θέσης του συμμετρικού άξονα πραγματοποιείται ξεκινώντας από το υψηλότερο επίπεδο και αλλάζοντας με χαμηλότερα επίπεδα καθώς η θέση του άξονα συμμετρίας καθορίζεται με μεγαλύτερη ακρίβεια.

Μια άλλη κατηγορία μεθόδων για την περιγραφή σχήματος είναι οι συντακτικές μέθοδοι όπου δομικές περιγραφές μπορούν να παρουσιαστούν ως γραφήματα που εκφράζονται σε όρους τυπικής γλώσσας [59, 60, 61, 62, 63]. Οι συντακτικές μέθοδοι έχουν πολλά πλεονεκτήματα. Εκτός του ότι επιχειρούν να μιμηθούν την δομική και ιεραρχική φύση του Ανθρώπινου Οπτικού Συστήματος (HVS), η θεωρία της τυπικής γλώσσας, στην οποία βασίζονται, είναι ένα πεδίο που έχει αναπτυχθεί αρκετά. Το βασικό μειονέκτημα αυτής της προσέγγισης είναι ότι πρέπει να κρυπτογραφηθεί αρχικά το σχήμα (ή το περίγραμμα του σχήματος) για να εισαχθούν δεδομένα στον συντακτικό αναλυτή. Συνήθως πρέπει να πραγματοποιηθεί τμηματοποίηση της εικόνας σε κάποιο βαθμό για να εξαχθούν διάφοροι τύποι τμημάτων ευθείας, καμπύλης και γωνιών [64]. Πέρα από τις ντετερμινιστικές (μη στοχαστικές) γραμματικές, ερευνήθηκαν και οι στοχαστικές γραμματικές [59]. Ο Fu [65] παρουσίασε μία τέτοια μέθοδο.

Άλλη κατηγορία μεθόδων χρησιμοποιεί προσομοίωση του περιγράμματος με μαθηματικές συναρτήσεις. Τα δύο πιο δημοφιλή σχέδια προσομοίωσης είναι η πολυγωνική και spline προσομοίωση. Οι προσομοιώσεις πολυγώνου χρησιμοποιούνται για την προσομοίωση του περιγράμματος χρησιμοποιώντας την πολυγωνική γραμμή. Αυτές οι μέθοδοι βασίζονται στην χρήση του ελάχιστου σφάλματος, της ελάχιστης περιμέτρου πολυγώνου, της μέγιστης εσωτερικής επιφάνειας πολυγώνου ή της ελάχιστης εξωτερικής επιφάνειας πολυγώνου ως κριτήρια προσομοίωσης. Τα πιο συχνά χρησιμοποιούμενα μέτρα σφάλματος είναι το μέγιστο σφάλμα (συγκέντρωση διάφορων μεθόδων *minimax*) και του εσωτερικού τετραγωνικού σφάλματος. Μία από τις πιο δημοφιλείς μεθόδους αυτής της κατηγορίας είναι ο αλγόριθμος Διαίρεσης και Ένωσης (*split-and-merge*) [66]. Με βάση αυτή την προσέγγιση, μία καμπύλη χωρίζεται σε τμήματα μέχρι να προκύψει ένα αποδεκτό σφάλμα. Συγχρόνως ενώνονται διαιρεμένα τμήματα πολυγώνου, εάν το προκύπτον τμήμα προσομοιάζει την καμπύλη εντός των ορίων κάποιου μέγιστου σφάλματος. Ο Παυλίδης [67] χρησιμοποίησε μερικά παράγωγα της συνάρτησης του εσωτερικού τετραγωνικού σφάλματος για να κατευθύνει τη Νευτώνεια μέθοδο στην έρευνα των καλύτερων σημείων τμηματοποίησης (*breakpoints*).

Οι Wu και Leou [68] πρότειναν ένα διαφορετικό κριτήριο βελτιστοποίησης για την παραγωγή πολυγωνικών προσομοιώσεων. Στο έργο τους χρησιμοποίησαν τις πολυγωνικές προσομοιώσεις εσωτερικής μέγιστης επιφάνειας, εξωτερικής ελάχιστης επιφάνειας ή απόκλισης ελάχιστης επιφάνειας. Οι Bengtsson και Eklundh [69] παρουσίασαν μία ιεραρχική μέθοδο στην οποία το περίγραμμα του σχήματος παρουσιάζεται μέσω της πολυγωνικής προσέγγισης. Ο αλγόριθμος διαίρεσης και ένωσης χρησιμοποιήθηκε για την δημιουργία πολυγωνικής προσομοίωσης. Η προσέγγιση χώρου-κλίμακας *scale-space* [70] χρησιμοποιήθηκε για τον εντοπισμό της θέσης των σημείων καμπής του περιγράμματος. Τα σταθερά χαρακτηριστικά του σχήματος είναι αυτά που παραμένουν αμετάβλητα ανά κλίμακα. Οι εφαπτόμενες στο περίγραμμα εκτιμούνται χρησιμοποιώντας μία προσέγγιση πολυωνύμου για να δημιουργηθεί μία αναπαράσταση της καμπύλης σε πολλαπλές κλίμακες.

Οι *splines* ήταν ήδη πολύ δημοφιλής για την παρεμβολή συναρτήσεων και την προσέγγιση καμπυλών. Οι Ikebe και Miyamoto [71] συνέγραψαν μία σύνοψη των εφαρμογών των *spline* για τον σχεδιασμό, την αναπαράσταση και την αποκατάσταση σχημάτων. Η μαθηματική αντιμετώπιση των *splines* απαντάται σε πολλά συγγράμματα [72, 73, 74], ενώ η προοπτική τους στα γραφικά των υπολογιστών απαντάται στα [75, 76, 77, 78]. Οι *splines* συγκεντρώνουν την ευεργετική ιδιότητα να ελαχιστοποιούν την καμπύλη. Με άλλα λόγια, προσεγγίζουν μία δεδομένη συνάρτηση με μία καμπύλη που έχει την ελάχιστη μέση καμπυλότητα. Αυτό τα καθιστά ιδανικούς υποψηφίους για την "φυσική" αναπαράσταση των καμπυλών. Το μειονέκτημα των *splines* σε προβλήματα παρεμβολών είναι ότι η τροποποίηση της τιμής της τοπικής συνάρτησης αλλάζει ολόκληρη την αναπαράσταση

των splines. Τα B-splines κατασκευάζονται ούτως ώστε η αλλαγή της τιμής της τοπικής συνάρτησης να επεκταθεί στα υπόλοιπα διαστήματα. Τα B-splines μπορούν να χρησιμοποιηθούν στην παρεμβολή επίπεδων καμπύλων που προκύπτουν από παραμετρικές εξισώσεις $(x(t), y(t))$. Σε αυτή την περίπτωση, κάθε παραμετρική εξίσωση επιδέχεται μιας ανεξάρτητης παρεμβολής. Οι Cohen et al. [79] πρότειναν μία τεχνική για την αναπαράσταση καμπυλών και την αντιστοίχιση αυτών χρησιμοποιώντας τα B-splines.

Οι επιφάνειες αναπαριστώνται συχνά μέσω μία οικογένειας καμπυλών. Η πιο απλή λύση είναι απλά μία γραμμική παρεμβολή σημείου προς σημείο ανάμεσα στις καμπύλες. Η τεχνική αυτή αποκαλείται *lofting* και χρησιμοποιείται ευρέως στην ναυσιπλοΐα και στη βιομηχανία αεροσκαφών [71]. Οι Chung et al. [80] ανέπτυξαν μία μέθοδο για την πολυγωνική προσομοίωση ενός σχήματος μέσω του νευρωνικού δικτύου Hopfield. Η προσομοίωση διαμορφώνεται ως η ελαχιστοποίηση της συνάρτησης της ενέργειας του δικτύου η οποία ορίζεται ως η παράκαμψη τόξου-προς-χορδή ανάμεσα στην καμπύλη και το πολύγωνο.

Άλλη σημαντική ομάδα αναπαραστάσεων περιλαμβάνει μεθόδους οι οποίες βασίζονται στην αναπαράσταση χώρου-κλίμακας (scale-space). Ο Witkin [70] πρότεινε μία προσέγγιση φίλτρου χώρου-κλίμακας που προσφέρει μία χρήσιμη αναπαράσταση για την αναπαράσταση σημαντικών χαρακτηριστικών σχήματος. Η αναπαράσταση δημιουργήθηκε εντοπίζοντας την θέση των σημείων καμπής σε σήματα που πέρασαν από βαθυπερατά φίλτρα Gauss διαφορετικού εύρους. Τα σημεία καμπής που παρέμειναν στην αναπαράσταση αναμένεται να είναι 'σημαντικά' χαρακτηριστικά σχήματος. Οι Babaud et al. [81] απέδειξαν την μοναδικότητα του φίλτρου kernel χώρου- κλίμακας καθώς το φίλτρο (kernel) του Gauss έχει την επιθυμητή ιδιότητα να σώζει σημεία καμπής όταν αυξάνεται το εύρος του φίλτρου. Με άλλα λόγια, η περίμετρος σε εικόνες χώρου-κλίμακας δεν μπορεί να εξαφανιστεί όταν αυξηθεί το εύρος του φίλτρου. Το φίλτρο του Gauss είναι το μόνο φίλτρο με αυτή την ιδιότητα.

Οι Asada και Brady [82] πρότειναν μία νέα προσέγγιση βασισμένη στις ιδέες των Marr et al. για να εισάγουν μία αναπαράσταση που λέγεται αρχικό σχεδιάγραμμα καμπυλότητας (curvature primal sketch). Πρόκειται για μία προσέγγιση χώρου-κλίμακας για την αναπαράσταση της καμπυλότητας και μοιάζει με την προσέγγιση τύπου ακμών και σκελετών του Marr για τον εντοπισμό των ακμών. Το περίγραμμα περνάει από τα φίλτρα με τις συναρτήσεις που περιέγραψε ο Gauss για την αύξηση του εύρους για να προκύψει μία αναπαράσταση του περιγράμματος βασισμένη σε πολλές κλίμακες. Έτσι, η καμπυλότητα υπολογίζεται σε διαφορετικές κλίμακες για να προκύψει το αρχικό σχεδιάγραμμα καμπυλότητας.

Οι Mokhtarian και Mackworth [83, 84] εφήρμοσαν την προσέγγιση χώρου-κλίμακας στην περιγραφή στατικών σχημάτων χρησιμοποιώντας το περίγραμμα. Στην συνέχεια υπολογίστηκε και αμβλύθηκε η καμπυλότητα στην περίμετρο με φίλτρα Gauss διαφορετικού

εύρους. Η εικόνα χώρου-κλίμακας της συνάρτησης καμπυλότητας χρησιμοποιήθηκε ως ένας ιεραρχικός περιγραφέας σχήματος που είναι αμετάβλητος ως προς τη μεταφορά, την κλίμακα και την περιστροφή. Η έννοια του φιλτραρίσματος σε πολλές κλίμακες είναι παρούσα και στην μαθηματική μορφολογία. Οι Chen και Yan [85] χρησιμοποίησαν ένα μεταβλητό στοιχείο δόμησης μεγέθους για να πραγματοποιήσουν διάφορες μορφολογικές λειτουργίες. Οι Dill et al. [86] μελέτησαν τον ρόλο της μετακίνησης λευκοκυττάρων. Δημιούργησαν πολλές αμβλυμμένες εκδόσεις των ορίων του λευκοκυττάρου για να εξάγουν σκελετούς. Το όριο αναπαραστάθηκε με την χρήση του 1 κώδικα αλυσίδας. Χρησιμοποιήθηκε ένα φίλτρο Gauss διαφορετικού εύρους για να δημιουργηθούν αμβλυμμένες εκδοχές του περιγράμματος. Ο σκελετός υπολογίστηκε με βάση μία τεχνική που βασίζεται στον αλγόριθμο του Arcella [87]. Αυτός ο αλγόριθμος βελτιώνει την ευαισθησία στον θόρυβο και την ευαισθησία σε κυρτά σημεία του περιγράμματος.

Άλλες μέθοδοι αποσυνθέτουν το όριο του σχήματος σε τμήματα καμπύλης. Οι H. Liu και M. Srinath [88] ανέπτυξαν μία διαδικασία για κατάταξη σχημάτων χρησιμοποιώντας την αντιστοίχιση περιγράμματος. Το στοιχείο που εισάγεται στην διαδικασία ήταν το σχήμα ενός αντικειμένου. Χρησιμοποιήθηκε ο ανιχνευτής ακμών του Sobel για τον υπολογισμό της διαβάθμισης της κατεύθυνσης και της εφαπτόμενης γωνίας στο περίγραμμα. Η συνάρτηση της εφαπτόμενης γωνίας εξελίσσεται μαζί με το παράγωγο της συνάρτησης του Gauss για τον εντοπισμό της συνάρτησης της αμβλυμμένης καμπυλότητας. Το περίγραμμα χωρίστηκε σε σημεία μεγάλης καμπυλότητας. Η αντιστοίχιση καμπυλών πραγματοποιήθηκε σε δύο βήματα. Στο πρώτο βήμα, συγκρίθηκαν μεμονωμένα τμήματα για δύο σχήματα. Στο δεύτερο βήμα συγκρίθηκαν ομάδες τμημάτων και μία ομάδα θεωρούνταν ακατάλληλη εάν αντιστοιχούσαν λιγότερα από τρία διαδοχικά τμήματα. Το τελικό βήμα χρησιμοποιήθηκε για να μετρηθεί η απόσταση ανάμεσα σε δύο σχήματα. Αυτό συνέβη μέσω του μετασχηματισμού της απόστασης 3/4 του Chamfer [89, 90] που μοιάζει πολύ με τον μετασχηματισμό της Ευκλείδειας απόστασης, αλλά είναι λιγότερο εντατικός από υπολογιστικής άποψης. Ο μετασχηματισμός της απόστασης 3/4 του Chamfer υπολογίστηκε για το πρώτο όριο, το δεύτερο όριο επιβλήθηκε στον μετασχηματισμό της απόστασης του πρώτου ορίου και υπολογίστηκε η μέση τιμή απόστασης. Τα αποτελέσματα των πειραμάτων κατέδειξαν την σκοπιμότητα της μεθόδου για την αντιστοίχιση σχημάτων.

Οι μέθοδοι των ροπών υπολογίζουν ένα βαθμωτό αποτέλεσμα με βάση το καθολικό σχήμα. Οι μέθοδοι που βασίζονται στην ροπή είναι μεταξύ των πιο δημοφιλών της ομάδας των μεθόδων ολικού (ή εσωτερικού) βαθμωτού μετασχηματισμού. Η χρήση της ροπής για την περιγραφή των σχημάτων ξεκίνησε από τον Hu [35], ο οποίος απέδειξε ότι η περιγραφή σχημάτων με βάση την ροπή βοηθά στην διατήρηση πληροφοριών. Μεταγενέστερες μελέτες στον τομέα έγιναν από τους Prokop και Reeves [38] και τον Weiss [91] που έχει μια πιο γενική προσέγγιση. Η ροπή μηδενικής τάξης m_{00} είναι ίση με την επιφάνεια του σχήματος εάν υποθέσουμε ότι η $f(x, y)$ είναι μία συνάρτηση περιγράμματος με τιμή

ένα εντός του σχήματος και μηδέν εκτός του σχήματος. Η ροπή πρώτης τάξης μπορεί να χρησιμοποιηθεί για τον υπολογισμό των συντεταγμένων του κέντρου της μάζας με τη μορφή $xc = \frac{m_{10}}{m_{00}}$ και $yc = \frac{m_{01}}{m_{00}}$. Η ροπή δεύτερης τάξης αποκαλείται ροπή αδράνειας και μπορεί να χρησιμοποιηθεί για τον καθορισμό των αρχικών αξόνων του σχήματος. Οι αρχικοί άξονες είναι οι άξονες σε σχέση με τους οποίους υπάρχει μέγιστη και ελάχιστη ροπή δεύτερης τάξης. Η ροπή των προβολών είναι στην ουσία μονοδιάστατη ροπή συναρτήσεων προβολής. Η ροπή που ορίζεται με αυτόν τον τρόπο δεν είναι ιδανική για την περιγραφή του σχήματος εφόσον δεν είναι αμετάβλητη ως προς τη μεταφορά, την περιστροφή και την κλίμακα. Για να ξεπεραστεί αυτή η δυσκολία, ο Hu [35] πρότεινε επτά αμετάβλητες ροπές (επίσης αποκαλούνται και αμετάβλητοι παράγοντες ροπής (moment invariants)). Οι ροπές αυτές δεν εξαρτώνται από την θέση, την κατεύθυνση ή την κλίμακα του σχήματος. Είναι πιθανό να γίνει μια γενίκευση του μετασχηματισμού της ροπής σε άλλους πυρήνες (kernels) μετασχηματισμού αντικαθιστώντας έναν συμβατικό πυρήνα μετασχηματισμού με μία πιο γενική μορφή. Είναι ιδιαίτερα ελκυστική η χρήση ορθογώνιων πολυωνύμων ως πυρήνες μετασχηματισμού της ροπής [36]. Σε αυτή την περίπτωση, οι ροπές παράγουν ελάχιστες περιττές πληροφορίες [38]. Αυτό είναι σημαντικό για την καλύτερη δυνατή χρήση των διαθέσιμων πληροφοριών σε ένα δεδομένο αριθμό ροπών. Μερικά από το ορθογώνια πολυωνυμικά συστήματα που έχουν ερευνηθεί [36] είναι τα πολυώνυμα Legendre, Zernike και τα ψευδο-Zernike πολυώνυμα. Το πλεονέκτημα των μεθόδων ροπής είναι ότι είναι μαθηματικά ακριβείς. Το μειονέκτημα είναι ότι είναι δύσκολος ο συσχετισμός ροπών υψηλής τάξης με τα χαρακτηριστικά του σχήματος. Όπως συμβαίνει και με τις περισσότερες μεθόδους βαθμωτού ολικού μετασχηματισμού, οι τοπικές πληροφορίες και τα χαρακτηριστικά του σχήματος δεν είναι δυνατόν να ανιχνευτούν. Μία εναλλακτική προσέγγιση μετασχηματισμού του σχήματος είναι ο μετασχηματισμός Fourier. Το μειονέκτημα αυτού είναι ότι είναι αδύνατο να "εντοπίσει" τα τοπικά χαρακτηριστικά του σχήματος και είναι απαιτητικός. Εντοπίζεται και σε αυτή την προσέγγιση το πρόβλημα της αμεταβλητότητας ως προς τη μεταφορά, την περιστροφή και την κλίμακα.

Πολύ σημαντική ομάδα περιγραφών σχήματος είναι αυτή που βασίζεται στην μαθηματική μορφολογία η οποία έχει εξελιχθεί σε ένα χρήσιμο εργαλείο για διάφορες εργασίες επεξεργασίας εικόνας [92, 93, 94, 95, 96]. Είναι αρχικά απαραίτητο να δοθούν κάποιοι βασικοί ορισμοί που συνδέονται με τη μαθηματική μορφολογία. Οι μορφολογικές λειτουργίες ορίζονται με τους όρους της θεωρίας των συνόλων. Οι λειτουργίες μορφολογικής συστολής (erosion) και διαστολής (dilation) είναι βασικές και μεταξύ των πιο χρήσιμων λειτουργιών για την επεξεργασία εικόνας. Οι περισσότερες άλλες λειτουργίες προέρχονται από αυτές τις δύο. Το μορφολογικό άνοιγμα (opening) και κλείσιμο (closing) είναι μεταξύ των πιο ισχυρών λειτουργιών στην μαθηματική μορφολογία. Έχουν ερευνηθεί πολλές προσεγγίσεις της περιγραφής σχημάτων που βασίζονται στην μαθηματική μορφολογία. Σε αυτή την παράγραφο, παρουσιάζονται μέθοδοι περιγραφής σχημάτων που βασίζονται

στην επιφάνεια των μορφολογικά επεξεργασμένων εικόνων. Ο Μαραγκός [97] πρότεινε την έννοια του φάσματος προτύπου (pattern spectrum). Το φάσμα προτύπου του συνόλου Q προκύπτει από το άνοιγμα Q με ένα δίσκο διαφόρων μεγεθών. Μετρήθηκαν οι επιφάνειες των συνόλων που προέκυψαν. Το φάσμα προτύπου ορίζεται ως η παράγωγος της συνάρτησης επιφάνειας ως προς την ακτίνα του δομικού στοιχείου του δίσκου r . Εκτός από την συνεχή περίπτωση ο Μαραγκός έδωσε και ορισμό για το φάσμα προτύπου σε διακριτές εικόνες και σε μορφολογία κλίμακας των τόνων του γκρι.

Η προσέγγιση του φάσματος προτύπου συνδέεται με την έννοια της διακριτομετρίας την οποία μελέτησε αρχικά ο Matheron [92] και πιο πρόσφατα ο Dougherty [98, 99]. Η διακριτομετρίας είναι αποτελέσματα της λειτουργίας διακριτοποίησης που εφαρμόζεται στις ψηφιακές εικόνες. Κατά αυτήν την λειτουργία η δομή της εικόνας φιλτράρεται ανάλογα με το μέγεθος των στοιχείων. Το αποτέλεσμα είναι μία ακολουθία που προκύπτει από το άνοιγμα του σχήματος μέσω της ακολουθίας των δομικών στοιχείων. Η αύξουσα σειρά των περιοχών διαδοχικών ανοιγμάτων ονομάζεται κατανομή μεγέθους. Να σημειωθεί ότι η αρνητική παράγωγος (ή η διαφορά στην περίπτωση των διακριτών εικόνων) των διαδοχικών επιφανειών είναι το φάσμα προτύπων του Μαραγκού.

Οι Shih και Pu [100] πρότειναν ένα άλλον μετασχηματισμό "φάσματος" που λέγεται G-spectrum και αποτελεί επέκταση του έργου των Γκουτσιά και Schonfeld. Οι συγγραφείς απέδειξαν ότι η αναπαράσταση που προτείνουν αυτοί είναι λιγότερο πλεονασματική από τις κατανομές μεγέθους της διακριτομετρίας ή το φάσμα προτύπων.

Μία άλλη ομάδα μεθόδων προέκυψε από την έννοια της μορφολογικής συνδιακύμανσης [93]. Οι Loui et al. χρησιμοποίησαν την συνάρτηση γεωμετρικού συσχετισμού (GCF) για την αναπαράσταση δισδιάστατων σχημάτων. Η GCF βασίζεται στην μορφολογική συνδιακύμανση [93]. Οι συγγραφείς χρησιμοποίησαν την GCF για την περιγραφή και την αντιστοίχιση σχημάτων. Η μέθοδος αυτή παραμένει αμετάβλητη ως προς την περιστροφή και την μεταφορά. Εάν χρειαστεί η αμεταβλητότητα ως προς την κλίμακα, μπορεί να προστεθεί ένα στάδιο επανακλιμάκωσης πριν την επεξεργασία. Η αμεταβλητότητα ως προς την περιστροφή επετεύχθη μέσω της χρήσης συναρτήσεων ελάχιστου ή μέγιστου συσχετισμού. Τα αποτελέσματα των πειραμάτων έδειξαν ότι αυτή η μέθοδος είναι χρήσιμη για την αντιστοίχιση σχημάτων.

Ο Μαραγκός [101, 102] συνέδεσε το κριτήριο μέσου απόλυτου σφάλματος για την αντιστοίχιση σήματος με την μορφολογική συνάρτηση συσχετισμού. Η μορφολογική συνάρτηση διασυσχετισμού μπορεί να συνδεθεί με την αντιστοίχιση ελάχιστου απόλυτου σφάλματος. Οι Shapiro et al. [103] χρησιμοποίησαν την υπολειμματική προσέγγιση για την αντιστοίχιση σχημάτων. Ο αλγόριθμος χρησιμοποίησε μία ανάλυση και ένα μόνο σχήμα δομικού στοιχείου (δίσκο). Για την αναπαράσταση της εικόνας χρησιμοποιήθηκαν τα υπολειμματικά κέντρα βάρους της εικόνας, η επιφάνεια και ο λόγος του μικρότερου προς τον μεγαλύτερο άξονα της ελλειπτικής καμπύλης που εφαρμόζει καλύτερα από όλες.

Η πιο δημοφιλής και μελετημένη μέθοδος ολικού περιγραφέα είναι ο μετασχηματισμός μεσαίου άξονα (medial axis transform, MAT) που εισήγαγε αρχικά ο Blum [30, 32]. Η ιδέα αυτής της προσέγγισης είναι η αναπαράσταση του σχήματος χρησιμοποιώντας ένα γράφημα και ελπίζοντας ότι θα διατηρηθούν στο γράφημα τα σημαντικά χαρακτηριστικά του σχήματος. Από άποψη ορολογίας, ο μεσαίος άξονας χρησιμοποιήθηκε αρχικά από τον Blum. Ο Calabi [104] χρησιμοποίησε ένα ζευγάρι σκελετών που αποτελείται από τον σκελετό και την συνάρτηση "quench". Οι όροι μετασχηματισμός φωτιάς λιβαδιού (prairie fire transform), μετασχηματισμός συμμετρικού άξονα (symmetric axis transform) και μετασχηματισμός σκελετού (skeleton transform) έχουν χρησιμοποιηθεί στην βιβλιογραφία και αναφέρονται στην ίδια προσέγγιση. Μπορείτε να βρείτε επιπλέον υλικό στο [105, 31, 18, 34, 21, 22]. Αυτή η προσέγγιση ενθαρρύνεται από την μελέτη της νευροφυσιολογίας και της οπτικής ψυχολογίας. Συγκεκριμένα, ο Blum υπέθεσε ότι η επεξεργασία της μορφής της εικόνας στον αμφιβληστροειδή του ματιού αποτελεί μία αλυσιδωτή αντίδραση υπό την παρακάτω έννοια. Όταν η εικόνα ενός αντικειμένου παίρνει μορφή στον αμφιβληστροειδή του ματιού, πυροδοτείται ένας αριθμός διεγερμένων νευρώνων, υποβαθμίζοντας τα επίπεδα ερεθισμού των γειτονικών νευρώνων και προκαλώντας την πυροδότησή τους μετά από ένα μικρό χρονικό διάστημα (καταστολή). Η διαδικασία αυτή επαναλαμβάνεται μέχρι να "χτιστεί" ολόκληρη η επιφάνεια του αντικειμένου με τους διεγερμένους νευρώνες. Οι νευρώνες που βρίσκονται σε καταστολή δεν μπορούν να πυροδοτήσουν ξανά για ένα μικρό διάστημα λόγω των υποκείμενων νευροφυσιολογικών διαδικασιών [106]. Επομένως το μέτωπο κυμάτων των πυροδοτημένων κυττάρων δεν μπορεί να κινηθεί πίσω προς τις περιοχές του αμφιβληστροειδούς που περιέχουν τις νευρώνες που βρίσκονται σε καταστολή. Ο μηχανισμός αυτός είναι παρόμοιος με την επέκταση της φωτιάς σε ένα λιβάδι. Στην ουσία, η πρώτη προσέγγιση που χρησιμοποίησε ο Blum ήταν μία προσωρινή συνάρτηση που έδειχνε το μήκος του τόξου του μετώπου κύματος προς τον χρόνο. Η προσέγγιση αυτή δεν αποδείχτηκε πολύ χρήσιμη για την περιγραφή σχημάτων. Η δεύτερη σύλληψη του Blum, η έννοια του μεσαίου άξονα, αποδείχτηκε πιο χρήσιμη για την περιγραφή σχημάτων. Ο σκοπός του μετασχηματισμού του μεσαίου άξονα, είναι η εξαγωγή ενός σκελετού, ενός σχήματος που μοιάζει με ράβδο, από το αντικείμενο. Ελπίζουμε ότι αυτό το σχήμα θα μπορέσει να χρησιμοποιηθεί για την αναπαράσταση και την περιγραφή του σχήματος του αντικειμένου. Η μορφή του σκελετού μπορεί να εξηγηθεί από το παρακάτω παράδειγμα. Έστω ότι το εσωτερικό του αντικειμένου αποτελείται από ένα εύφλεκτο στεγνό γρασίδι ενώ το εξωτερικό του αντικειμένου αποτελείται από μη εύφλεκτο υγρό γρασίδι. Εάν τεθεί φωτιά ταυτόχρονα σε όλα τα σημεία περιγράμματος του σχήματος θα επεκταθεί προς το κέντρο του σχήματος. Ωστόσο, σε μερικές θέσεις το μέτωπο των κυμάτων από τη μία κατεύθυνση θα δημιουργήσει τομή με το μέτωπο του πύρινου κύματος που προέρχεται από άλλη κατεύθυνση και επομένως θα σβήσει την φωτιά. Τα σημεία της σύγκρουσης του μετώπου αποκαλούνται σημεία κατάσβεσης

(quench points). Ο σκελετός της εικόνας ορίζεται ως το σύνολο των σημείων κατάσβεσης. Είναι πιθανή η αναδόμηση του σχήματος χρησιμοποιώντας την συνάρτηση του σκελετού του και των σημείων κατάσβεσης. Η συνάρτηση των σημείων κατάσβεσης $q(x)$ σε κάποιο σημείο x του σκελετού S έχει μία τιμή ίση με την ακτίνα του κύκλου επαφής σε αυτό το σημείο. Για την αναδόμηση του αντικειμένου πρέπει να τοποθετηθεί ένα δίσκος ακτίνας $q(x)$ στην θέση x , για κάθε $x \in S$. Η ένωση όλων των δίσκων ισούται με το αρχικό σχήμα. Η διάδοση του μετώπου κύματος μπορεί αν υπολογιστεί χρησιμοποιώντας τις μεθόδους μετασχηματισμού της απόστασης [107, 108] ή τη λειτουργία σμίκρυνσης [31] ή επεξεργασία των τελεστών min-max [109, 34, 110] ή τελεστών συρρίκνωσης τύπου Boolean [111, 112, 113, 114, 115, 116, 117, 118]. Το μειονέκτημα του μετασχηματισμού μεσαίου άξονα είναι ότι είναι πολύ ευαίσθητο σε θορύβους στο όριο του σχήματος. Ακόμη και μικρές αλλαγές στο σχήμα μπορούν να προκαλέσουν σημαντικές αλλαγές στην τοπολογία του γραφήματος MAT. Μία άλλη δυσκολία βρίσκεται στην πρακτική εφαρμογή στα διακριτά διαστήματα [119]. Η αλλαγή από συνεχή σε διακριτό διάστημα προκαλεί μερικές δυσκολίες. Για παράδειγμα, το γράφημα MAT ενός συνδεδεμένου αντικειμένου μπορεί να μην συνδέεται.

Για να επιλυθεί το πρόβλημα του θορύβου, οι Blum και Nagel [33] εισήγαγαν έναν γενικευμένο μετασχηματισμό μεσαίου άξονα (generalized medial axis transform), βασισμένο στο γεγονός ότι ο κύκλος επαφής ορίζεται ως ένας κύκλος που εφάπτεται στα όρια του σχήματος χωρίς να δημιουργεί τομή με αυτόν. Ο συμμετρικός άξονας r ενός σχήματος ορίζεται ως η ένωση όλων των σημείων που έχουν έναν κύκλο επαφής ακτίνας μεγαλύτερης από το r και τουλάχιστον δύο από αυτά τα σημεία αγγίζουν τα όρια. Η απαίτηση ότι η ακτίνα είναι μεγαλύτερη από το r αποτρέπει μικρές θορυβώδεις αιχμές να παράγουν τμήματα σκελετού. Η απαίτηση των δύο σημείων επαφής περιορίζει την επιλογή σε μόνο σημεία σκελετού. Ο γενικευμένος σκελετός έχει αποδειχτεί ότι έχει καλύτερη ευρωστία παρουσία θορύβου [33]. Η έννοια της συμμετρίας και του συμμετρικού άξονα έχει αναπτυχθεί περαιτέρω στα έργα των Brooks [120] και Brady. Το πρόβλημα του καθορισμού του άξονα συμμετρίας συνδέεται με ένα αντίστροφο πρόβλημα - την παραγωγή σχημάτων από τους συμμετρικούς άξονές τους. Τα αντικείμενα που παράγονται κατ' αυτό τον τρόπο ονομάζονται κορδέλες (ribbons). Η σύνθεση ενός αντικειμένου που μοιάζει με κορδέλα γίνεται μέσω του άξονα και του γεννήτορα. Ο γεννήτορας είναι ένα γεωμετρικό σχήμα που κινείται πάνω στον άξονα και αλλάζει μέγεθος καθώς κινείται. Κατά την προσέγγιση του Blum, χρησιμοποιήθηκε ένας δίσκος ως γεννήτορας και τα σχήματα που κατασκευάστηκαν λέγονται κορδέλες του Blum (Blum ribbons). Οι L-ribbons παράγονται χρησιμοποιώντας το τμήμα ευθείας. Οι L-ribbons με μία σταθερή γωνία ανάμεσα στο τμήμα ευθείας και τον άξονα λέγονται κορδέλες του Brooks (Brooks ribbons) [120]. Οι L-ribbons με μία σταθερή γωνία ανάμεσα στο τμήμα ευθείας και την πλευρά της κορδέλας λέγονται κορδέλες του Brady (Brady ribbons) [121]. Οι Rosenfeld [122] και Ponce [123] συνέκριναν διάφορες

διαδικασίες παραγωγής κορδελών.

Ο Brady [121] ανέπτυξε την έννοια της ομαλής τοπικής συμμετρίας (smooth local symmetry). Σύμφωνα με αυτή, δύο σημεία διαμορφώνουν μία τοπική συμμετρία, εάν οι γωνίες ανάμεσα στα ομαλά σημεία της καμπύλης σε δύο δεδομένα σημεία και στην γραμμή που συνδέει τα σημεία είναι τα ίδια. Μία ιεραρχική προσέγγιση (πολλαπλής ανάλυσης) στην ανάλυση σκελετού παρουσιάστηκε από τους Pizer, Oliver και Bloomberg [124]. Στην προσέγγισή τους, η ιεραρχία κατά κλίμακα χρησιμοποιήθηκε για την κατασκευή μίας σειράς σκελετών. Αυτή η προσέγγιση πολλαπλής κλίμακας έχει αποδειχτεί ότι ξεπερνά την εξάρτηση της αναπαράστασης από τον χρόνο [124]. Οι Rom και Medioni [125] πρότειναν μία ιεραρχική αναπαράσταση του σχήματος χρησιμοποιώντας την αξονική περιγραφή του σχήματος. Το πρωτότυπο σχήμα έσπασε σε τμήματα σε αρνητική ελάχιστη καμπυλότητα των ορίων του σχήματος [126]. Τα τμήματα που προέκυψαν αναπαραστάθηκαν χρησιμοποιώντας κορδέλες ομαλής τοπικής συμμετρίας. Η παράλληλη συμμετρία χρησιμοποιήθηκε για την σύλληψη της συνολικής σχέσης ανάμεσα σε τμήματα.

Οι Leymarie και Levine [127] ανέπτυξαν μία νέα μέθοδο για την εξαγωγή συμμετρικού άξονα ο οποίος δεν υποφέρει από προβλήματα δημιουργίας διακριτών συνόλων από τα οποία υποφέρουν πολλοί άλλοι αλγόριθμοι. Αυτή η μέθοδος βασίστηκε στην χρήση φιδιών (snakes) για την ενεργητική αναπαράσταση του περιγράμματος, τα σημεία υψηλής καμπυλότητας και τον μετασχηματισμό των συμμετρικών αξόνων. Το αποτέλεσμα ήταν μία δυναμική αναπαράσταση του σκελετού πολλαπλής κλίμακας. Οι άξονες συμμετρίας εξήχθησαν αρχικά από ψηφιακές εικόνες, όμως οι Gauch και Pizer [128] πρότειναν μία μέθοδο εξαγωγής των αξόνων έντασης (intensity axis) της συμμετρίας από εικόνες τόνων του γκρι. Η μέθοδος αυτή εφαρμόστηκε σε τμηματοποίηση εικόνας βασισμένη στο σχήμα όπου ήταν πιθανός ο προσδιορισμός τμημάτων του αντικειμένου που αντιστοιχεί σε διαφορετικά στοιχεία του άξονα έντασης της συμμετρίας. Οι Han και Fan [129] πρότειναν μία μέθοδο εξαγωγής σκελετών χρησιμοποιώντας μία οριακή αναπαράσταση του σχήματος. Στο πρώτο βήμα, εξήχθη το περίγραμμα του σχήματος χρησιμοποιώντας έναν αλγόριθμο δημιουργίας ανυσμάτων του περιγράμματος. Στο δεύτερο βήμα, πραγματοποιήθηκε η δημιουργία ζευγών των περιγραμμάτων χρησιμοποιώντας έναν αλγόριθμο αντιστοίχισης περιγράμματος. Στο τελικό βήμα, εντοπίστηκε ο σκελετός. Με τη μέθοδο αυτή αποφεύγονται ορισμένα από τα κλασικά προβλήματα στους συμβατικούς αλγόριθμους σκελετοποίησης. Οι Shih και Pu [130] ανέπτυξαν έναν αλγόριθμο σκελετοποίησης εντοπίζοντας τις μέγιστες τιμές του μετασχηματισμού της Ευκλείδειας απόστασης. Η μέθοδος αυτή έχει πολλά πλεονεκτήματα όπως διατήρηση της σύνδεσης, εύρος μόνον ενός εικονοστοιχείου και ο σκελετός που προκύπτει είναι παρόμοιος με τους περισσότερους άλλους συμμετρικούς άξονες. Οι Ogniewicz και Kuebler [131] πρότειναν την σκελετοποίηση ενός στατικού σχήματος που βασίζεται στο Ευκλείδειο μετρικό σύστημα και διατηρείται η σύνδεση. Η μέθοδος αυτή χρησιμοποιεί το διάγραμμα Voronoi για την αναπαράσταση

του σχήματος.

Οι Μαραγκός και Schafer [132] χρησιμοποίησαν τη μαθηματική μορφολογία για να εξάγουν υποσύνολα σκελετού για την αποτελεσματική κωδικοποίηση των ψηφιακών εικόνων. Η μέθοδος αυτή βασίστηκε στην εργασία των Lantuejoul και Serra, οι οποίοι απέδειξαν [93] ότι ο σκελετός θα μπορούσε να αναπαρασταθεί ως μία ένωση των στοιχείων και ότι το αρχικό σχήμα θα μπορούσε να αναδομηθεί χρησιμοποιώντας αυτά τα στοιχεία. Τα στοιχεία του σκελετού κωδικοποιήθηκαν χρησιμοποιώντας μεθόδους αποκλεισμού και κωδικοποίησης τρέχοντος μήκους [47]. Η κωδικοποίηση ήταν αποτελεσματική γιατί τα υποσύνολα του σκελετού ήταν ψηφιακές εικόνες.

Στις τεχνικές αποσύνθεσης σχήματος, ένα σχήμα αναπαριστάται ως συνδυασμός των συστατικών σχημάτων. Η ιδέα είναι η αναπαράσταση σύνθετων σχημάτων υπό όρους απλών συστατικών στοιχείων. Ο Παυλίδης δήλωσε το πρόβλημα της ανάλυσης καθολικών σχημάτων (αποσύνθεση) ως εξής: Μεταξύ των σημείων περιγράμματος βρίσκονται σύνολα σημείων που συνδέονται στενά. Αυτά τα σύνολα μπορεί να χρησιμοποιηθούν για να ανατεθούν ετικέτες σε αντίστοιχα τμήματα του αντικειμένου [2]. Σε αυτή την προσέγγιση, η αποσύνθεση του σχήματος βασίζεται στις ιδιότητες των σημείων περιγράμματος. Αρκετοί συγγραφείς έχουν χρησιμοποιήσει αυτή την προσέγγιση για την αποσύνθεση σχημάτων. Τα κριτήρια αποσύνθεσης μπορούν να οριστούν με τους όρους του μετασχηματισμού μεσαίου άξονα και μπορεί να απαιτούν κυρτά συστατικά στοιχεία ή ορατότητα των σημείων του περιγράμματος. Στην προσέγγιση του μετασχηματισμού του μεσαίου άξονα, δύο σημεία περιγράμματος χαρακτηρίζονται συνδεδεμένα εάν και τα δύο βρίσκονται στον κύκλο που περιλαμβάνεται στο σχήμα του αντικειμένου. Το κριτήριο της αποσύνθεσης μπορεί να διαμορφωθεί ούτως ώστε να απαιτεί το τμήμα της ευθείας ανάμεσα σε δύο σημεία στο περίγραμμα να περιλαμβάνονται στο σχήμα που περιγράφεται. Αυτό το είδος κριτηρίων οδηγεί στην αποσύνθεση σε κυρτά συστατικά στοιχεία. Η ομαδοποίηση των σημείων του περιγράμματος [133, 134] είναι μία στοχαστική μέθοδος η οποία απαιτεί να γίνεται η αποσύνθεση ούτως ώστε κάθε σημείο στο περίγραμμα να είναι ορατό από τα περισσότερα από τα άλλα σημεία. Η αποσύνθεση στη βάση των *k*-nearest neighbors (*k* πιο κοντινών γειτόνων) [135] αντιστοιχεί στην προσέγγιση πλευρών του πολυγώνου αντί για τα σημεία περιγράμματος. Οι σημασιολογικές απόψεις σχετικά με τα σχήματα που περιγράφονται λαμβάνονται υπόψη στη μέθοδο για την αποσύνθεση σχήματος με βάση την πολυσυγγραμμικότητα [41]. Η μέθοδος αυτή βασίστηκε στις ευρετικές μεθόδους που βασίζονται στις ανθρώπινες έννοιες της πολυσυγγραμμικότητας. Οι Vanderheydt et al. [136] χρησιμοποίησαν την θεωρία του ασαφούς υποσυνόλου [137, 138] για να κατευθύνουν την αποσύνθεση βάσει των κυρτών και κοίλων σημείων περιγράμματος και της πολυγωνικής προσομοίωσης του περιγράμματος. Η μέθοδος εφαρμόστηκε στο πρόβλημα της ανάλυσης του σχήματος των χρωμοσωμάτων. Οι Sanniti και Thiel [139] ανέπτυξαν μία μέθοδο για την αποσύνθεση του σχήματος μέσω της αποσύνθεσης το ζυγισμένου σκελετού του. Η

μέθοδος αυτή είναι η καλύτερη δυνατή για επιμηκυσμένα αντικείμενα που έχουν τη μορφή κορδέλας. Οι Pitas et al. [140, 141, 142, 143, 144] ανέπτυξαν μία εναλλακτική μορφολογική μέθοδο για την αποσύνθεση σχημάτων. Αυτή η τεχνική αποσύνθεσης διασπά το σχήμα σε κορδέλες του Blum. Πρόκειται για ένα πλεονέκτημα εφόσον μία από τις επιθυμητές ιδιότητες της αποσύνθεσης είναι τα στοιχεία που προκύπτουν να είναι ενστικτωδώς απλά. Οι Schonfeld και Γκουτσιάς [145, 146, 147] εισήγαγαν ένα άλλο σχέδιο για την μορφολογική αποσύνθεση του σχήματος η οποία χρησιμοποιεί μία ακολουθία δομικών στοιχείων και μία ακολουθία μετασχηματισμών συνόλου. Η προκύπτουσα αποσύνθεση ονομάζεται γενικευμένη μορφολογική τμηματοποίηση. Προκύπτει σαν μία ειδική περίπτωση αποσύνθεσης χρησιμοποιώντας τον μετασχηματισμό του συνόλου ταυτότητας και ονομάζεται γενικευμένος μειωμένος μορφολογικός σκελετός. Οι συγγραφείς έχουν αποδείξει ότι η τελευταία αποσύνθεση είναι η “καλύτερη δυνατή” στην κατηγορία των αντιστρέψιμων μορφολογικών αποσυνθέσεων. Είναι η καλύτερη δυνατή υπό την έννοια της ελαχιστοποίησης της πιθανότητας σφάλματος για την αναδόμηση παρουσία θορύβου.

Οι T. B. Sebastian, P. N. Klein και B. B. Kimia [148] διορθώνοντας τα σοκ γραφήματα [149] με τρόπο ώστε να ενσωματώνουν στον σκελετό του σχήματος στοιχεία από το περίγραμμα επέδειξαν την καλύτερη από όλες τις προγενέστερες, αναγνώριση σχήματος ως προς τον αριθμό των διαδοχικά συνεχών επιτυχών αναγνωρίσεων σε συγκεκριμένο σύνολο σιλουετών, γνωστού και ως βάση σιλουετών του Kimia.

Μία σκελετική αναπαράσταση με βάση σταθερές ιδιότητες του σχήματος που προκύπτουν από την υπερβολική κανονικοποίηση είναι μία άλλη πρόσφατη προσέγγιση που αποδίδει επιτυχή αποτελέσματα σε διάφορες βάσεις δεδομένων με επίπεδα σχήματα [150].

Οι Mohammad Reza Daliri και Vincent Torre προτείνουν την μέθοδο (Kernel-edit distance) για ταίριασμα περιγραμμάτων [151] ενώ οι Adrian Ion et al [152] εισάγουν τον μετασχηματισμό εκκεντρότητας (eccentricity transform).

Η inner-distance [153], είναι μία άλλη μέθοδος που στοχεύει στο να υποκαταστήσει την Ευκλείδεια απόσταση ως προς τη χρήση της με ορισμένες μεθόδους αναπαράστασης σχήματος, βελτιώνοντας τα αποτελέσματα αναγνώρισης σε πολλά σύνολα δεδομένων δίδοντας ιδιαίτερη έμφαση στην αμεταβλητότητα άρθρωσης. Από την άλλη μεριά, η summation invariant [154], είναι μία μέθοδος κατασκευής ορισμένων αμετάβλητων στοιχείων επιλύοντας τις εξισώσεις κανονικοποίησης όσον αφορά ορισμένες ομάδες μετασχηματισμού. Πρόσφατα παρουσιάστηκαν στατιστικά μοντέλα σχημάτων για μετρήσεις απόστασης σχήματος [155, 156, 157] όπου μία γενίκευση των μοντέλων Markov για τοποθέτηση σχήματος σε έντονα παρεμποδισμένες εικόνες είναι μία ακόμα νέα προσέγγιση [158]. Η *θεωρία μεγέθους* και οι *συναρτήσεις μεγέθους* αναπαριστούν μία προσέγγιση υψηλού επιπέδου για την ολική σύγκριση ανάμεσα σε σχήματα που είναι ανθεκτικά στο θόρυβο και περιέχει επίσης τοπικές πληροφορίες [159, 160, 161]. Μεταξύ άλλων πρόσφατων μεθόδων είναι η αντιστοίχιση με δέντρα σχήματος shape trees [162], η ομοιότητα διαδρομής path similarity [163], η μετα-

τροπή γραφήματος *graph transduction* [164] ή η αντιστοίχιση γραφήματος με επίπεδα *layered graph matching* [165], ενώ απόπειρες ανθεκτικής καμπυλότητας *robust curvature* [166], εστιάζουν σε τοπικές ιδιότητες αλλά υπάρχει έλλειψη ενσωμάτωσης ολικών ιδιοτήτων.

Η συνιστώμενη προσέγγιση είναι εγγενώς χαμηλού επιπέδου και συνδέεται κυρίως με τις απόπειρες ανάπτυξης περιγραφών περιγράμματος σχήματος που είναι ανθεκτικοί στον θόρυβο. Έχουν κυρίως τη μορφή συναρτήσεων καμπύλων που βασίζονται είτε σε ακτινικά ιστογράμματα [167, 168] είτε σε ολοκληρωτικούς τελεστές [169, 170, 171]. Στο [171], εισάγετε ένα πλαίσιο εργασίας για ολοκληρωτικά αναλλοιώτα στοιχεία και ορίζεται ένας εντοπισμένος περιγραφέας *Local Area Integral Invariant (LAI)*. Στη συνιστώμενη προσέγγιση, αποδεικνύουμε ότι η ενσωμάτωση ολικών και τοπικών ιδιοτήτων σχήματος οδηγεί σε βελτιωμένη δυνατότητα αναγνώρισης όσον αφορά τον περιγραφέα *Local Area Integral Invariant*. Χρησιμοποιείται επίσης ο συνδυασμός *GLT* και *VAR* για να καθορίσει μία νέα μέθοδο η οποία σε σύγκριση με την μέθοδο *Shape Contexts* (Περιβάλλον του σχήματος) [167] αποκαλύπτει πλεονεκτήματα ως προς την ταχύτητα και την απόδοση. Περαιτέρω, οι διαισθητικές ιδιότητες των προτεινόμενων περιγραφέων αποδεικνύονται σημαντικές για την μείωση της πολυπλοκότητας αλλά και την βελτίωση της ικανότητας ανάκτησης μίας ορισμένης κατηγορίας μεθόδων που βρίσκονται στην τρέχουσα στάθμη της τεχνικής.

2.2. Η ανθρώπινη αντίληψη για την οπτική μορφή

Το Ανθρώπινο Οπτικό Σύστημα (ΑΟΣ) είναι ένα από τα πιο εξελιγμένα και ευέλικτα στη φύση. Η ικανότητά του να καταλαβαίνει την οργάνωση του περιβάλλοντος χώρου είναι αζεπέραστη σε σχέση με τα τεχνητά συστήματα. Υπάρχει μεγάλη κινητικότητα στην έρευνα που επικεντρώνεται στη μελέτη του ΑΟΣ. Ένας στόχος των προσπαθειών αυτών είναι η ανακάλυψη ενός μοντέλου για ανάπτυξη τεχνητών συστημάτων οπτικής αντίληψης και γνωστικής. Παλαιότερες έρευνες στο πεδίο των τεχνητών συστημάτων για ανάλυση σχήματος ήταν απομονωμένες από το πεδίο της ανθρώπινης οπτικής αντίληψης, σε αντίθεση με πιο πρόσφατες έρευνες που βασίζονται περισσότερο στα αποτελέσματα της μελέτης του ΑΟΣ και της αντίληψης. Από το ευρύ πεδίο της γνωστικής επιστήμης, οι περιοχές της οπτικής γνωστικής και αντίληψης έχουν ιδιαίτερο ενδιαφέρον για τη μελέτη της περιγραφής σχήματος. Εάν η δομή του ανθρώπινου συστήματος ανάλυσης σχήματος ήταν γνωστή, θα ήταν δυνατή η ανάπτυξη ανάλογων τεχνητών συστημάτων. Για το λόγο αυτό η μελέτη των μεθόδων ανάλυσης σχήματος συχνά χρησιμοποιεί και κινητοποιείται από τα αποτελέσματα της έρευνας στην περιοχή της ανθρώπινης οπτικής αντίληψης. Μια εξαντλητική επισκόπηση της έρευνας για την ανθρώπινη οπτική αντίληψη είναι πέρα από τους σκοπούς του παρόντος. Κάποια εισαγωγικά και κάποια πιο προχωρημένα βιβλία και άρθρα σχετικά με την οπτική αντίληψη και γνωστική περιλαμβάνονται στη βιβλιογραφία [4, 172, 173, 174, 175, 176, 177]. Σε αυτή την ενότητα, παρουσιάζεται μια σύντομη γενική επισκόπηση της έρευνας για την οπτική αντίληψη που σχετίζεται με την περιγραφή σχήματος.

2.2.1. Κλασικές Θεωρίες για την Οπτική Αντίληψη

Αρκετές σχολές ψυχολογίας έχουν προσπαθήσει να κατανοήσουν και να περιγράψουν, σε γενικές γραμμές, τους μηχανισμούς της συμπεριφοράς και τη συγκεκριμένη διάσταση της οπτικής αντίληψης. Ο Hake [178] πρότεινε αρκετές προσεγγίσεις για την αναπαράσταση των φυσικών μορφών. Ο Zusne [173] παρουσίασε μια επισκόπηση των σύγχρονων θεωριών της οπτικής μορφής. Μια εις βάθος συζήτηση των θεωριών της ψυχολογίας για την οπτική αντίληψη είναι πέρα από τους σκοπούς της Διατριβής, επομένως περιοριζόμαστε σε μια σύντομη συζήτηση επιλεγμένων θεμάτων. Η σχολή Gestalt της ψυχολογίας [179] έχει διαδραματίσει έναν επαναστατικό ρόλο με την καινοτόμο προσέγγιση στην οπτική μορφή. Μια λεπτομερής παρουσίαση του αντικειμένου της θεωρίας Gestalt μπορεί να βρεθεί στα [180, 181, 182, 183]. Η θεωρία Gestalt είναι μια μη-υπολογιστική θεωρία της οπτικής μορφής, το οποίο αποτελεί μειονέκτημα στις πρακτικές εφαρμογές. Παρ' όλα αυτά, σύμφωνα με τον Zusne είναι ακόμη η μόνη θεωρία που πραγματεύεται τη μορφή με κατανοητό τρόπο ([173], σελ. 108). Έχουν γραφτεί πολλά βιβλία πάνω στους νόμους Gestalt που παρουσιάζουν διάφορες λίστες με βασικές αρχές. Αυτές οι λίστες κυμαίνονται

σε αριθμό από έξι μέχρι περισσότερες από εκατό. Εδώ, παραθέτουμε μια λίστα νόμων για τις οπτικές μορφές όπως προτάθηκαν από τον Zusne [173].

- Η οπτική μορφή είναι η σημαντικότερη ιδιότητα μιας διαμόρφωσης.
- Η οπτική μορφή είναι είτε δυναμική ή το αποτέλεσμα υποκείμενων δυναμικών διαδικασιών.
- Όλες οι οπτικές μορφές έχουν τουλάχιστον δύο χαρακτηριστικά, ένα διαμορφωμένο τμήμα που καλείται σχήμα και ένα φόντο που καλείται βάση.
- Οι οπτικές μορφές μπορεί να έχουν ένα ή περισσότερα κέντρα βαρύτητας γύρω από τα οποία οργανώνεται η μορφή.
- Οι οπτικές μορφές είναι μεταθετές (όσον αφορά τη μεταφορά, το μέγεθος, τον προσανατολισμό και το χρώμα) χωρίς να χάνουν την ταυτότητά τους.
- Οι οπτικές μορφές αντιστέκονται στην αλλαγή. Τείνουν να διατηρούν τη δομή τους έναντι εξωτερικών δυνάμεων.
- Οι οπτικές μορφές είναι πάντα τόσο καλές (κανονικές, συμμετρικές, απλές, ομοιόμορφες, με την ελάχιστη καταπόνηση) όσο το επιτρέπουν οι συνθήκες (πρότυπο ερεθισμών).
- Οι μορφές έχουν τη δυνατότητα να ενώνονται για την παραγωγή νέων μορφών.
- Μια αλλαγή σε ένα τμήμα της μορφής επηρεάζει και τα υπόλοιπα τμήματά της (νόμος της αντιστάθμισης).
- Οι οπτικές μορφές τείνουν να εμφανίζονται και να εξαφανίζονται σαν σύνολα.
- Οι οπτικές μορφές αφήνουν ένα υπόλειμμα με το οποίο μπορούμε να τις θυμόμαστε ευκολότερα (νόμος της αναπαραγωγής).
- Ο χώρος είναι ανισοτροπικός, έχει διαφορετικές ιδιότητες σε διαφορετικές κατευθύνσεις.

Μια διαφορετική προσέγγιση της θεωρίας της οπτικής μορφής βρίσκεται στο έργο του Hebb. Ο Hebb παρουσίασε μια νευρο-ψυχολογική θεωρία για την συμπεριφορά στο βιβλίο του 'Η Οργάνωση της Συμπεριφοράς' [184]. Στη θεωρία του, ο Hebb έδωσε έμφαση στο ρόλο των νευρωνικών δομών στο μηχανισμό της οπτικής αντίληψης. Η εργασία του επηρέασε αρκετούς ερευνητές στο πεδίο των τεχνητών νευρωνικών δικτύων. Σε αντίθεση με τη σχολή Gestalt, ο Hebb επιχειρηματολογεί ότι η μορφή δεν γίνεται αντιληπτή ως σύνολο αλλά αποτελείται από τμήματα. Για την επιτυχή αναγνώριση, πρέπει να γίνει γνωστή η οργάνωση

και η αμοιβαία χωρική συσχέτιση των τμημάτων. Η απόκτηση γνώσης για αυτήν την πλευρά της αντίληψης είναι το βασικό στοιχείο στην θεωρία του Hebb.

Η κίνηση των ματιών είναι ο κύριος μηχανισμός που ενσωματώνει απλούστερα στοιχεία της αντίληψης. Τα απλούστερα στοιχεία είναι γωνίες και γραμμές. Ο Hebb εισήγαγε επίσης την έννοια των συγκροτημάτων κελιών, που είναι ομαδοποιημένοι νευρώνες με τέτοιο τρόπο ώστε επανειλημμένες ενεργοποιήσεις να μειώνουν την συναπτική αντίσταση, και έτσι οι νευρώνες της ομάδας να ερεθίζουν ο ένας τον άλλον. Η θεωρία του Hebb ήταν κυρίως ποιοτικού και όχι υπολογιστικού χαρακτήρα, και έτσι παρουσιάζει ένα μειονέκτημα στις πρακτικές εφαρμογές. Ο Gibson [185] ανέπτυξε μια διαφορετική πλήρη θεωρία για την οπτική αντίληψη. Η πρώτη βασική αρχή της θεωρίας του είναι ότι ο χώρος δεν είναι μια γεωμετρική ή αφηρημένη οντότητα, αλλά μια πραγματική οπτική οντότητα που χαρακτηρίζεται από τις μορφές που βρίσκονται σε αυτόν. Η θεωρία του Gibson κινείται γύρω από την αντίληψη πραγματικών τρισδιάστατων αντικειμένων, και όχι των δισδιάστατων προβολών τους. Η δεύτερη βασική αρχή είναι ότι πίσω από οποιαδήποτε απλή ή πολύπλοκη οπτική αντίληψη βρίσκεται ένα πραγματικό ερέθισμα. Αυτό το ερέθισμα είναι στη μορφή μιας διαβάθμισης και αποτελεί ιδιότητα της επιφάνειας. Παραδείγματα φυσικών διαβαθμίσεων είναι η αλλαγή του μεγέθους των στοιχείων της υψής (διάσταση βάρους), ο βαθμός σύγκλισης των παράλληλων ακμών (προοπτική), η απόχρωση και η καθαρότητα των χρωμάτων, και η σκίαση. Ο Gibson τονίζει ότι η σχολή Gestalt ασχολείται με τη μελέτη των δισδιάστατων προβολών του τρισδιάστατου κόσμου και ότι ο δυναμισμός της δεν είναι τίποτα άλλο από την δισημία στην ερμηνεία των προβαλλομένων εικόνων. Στην κατηγοριοποίησή του υπάρχουν δέκα διαφορετικά είδη μορφών.

- Συμπαγής μορφή. (Βλέποντας ένα αντικείμενο βλέπουμε μια συμπαγή μορφή.)
- Επιφανειακή μορφή. (Μορφές κεκλιμένες και με ακμές.)
- Περιγραμματική μορφή. (Απεικόνιση των ακμών μιας συμπαγούς μορφής.)
- Μορφή Εικόνας. (Αναπαραστάσεις που σχεδιάζονται, φωτογραφίες, πίνακες, κλπ.)
- Μορφή σχεδίου. (Απεικόνιση των ακμών μια επιφάνειας που προβάλλεται σε μια άλλη επίπεδη επιφάνεια.)
- Προοπτική μορφή. (Μια προοπτική απεικόνιση μιας μορφής.)
- Μορφή χωρίς νόημα. (Απεικονίσεις που δεν αναπαριστούν ένα πραγματικό αντικείμενο.)
- Επίπεδη γεωμετρική μορφή. (Μια αφηρημένη γεωμετρική μορφή που δεν επιχειρεί να κάνει ορατή μια συμπαγή μορφή ή δεν παράγεται από μια συμπαγή μορφή.)

- Συμπαγής γεωμετρική μορφή. (Ένα αφηρημένο τμήμα ενός τρισδιάστατου χώρου που δεσμεύεται από φανταστικές επιφάνειες.)
- Προβαλλόμενη μορφή. (Μια επίπεδη γεωμετρική μορφή που είναι η προβολή μιας μορφής.)

Αυτές οι μορφές ομαδοποιούνται σε τρεις κατηγορίες όπως ακολούθως.

1. Πραγματικά αντικείμενα: συμπαγείς και επιφανειακές μορφές.
2. Αναπαραστάσεις πραγματικών μορφών: περιγραμματικές, εικόνας, σχεδίου, προοπτικές και χωρίς νόημα μορφές.
3. Αφηρημένα (μη-πραγματικά) αντικείμενα: Επίπεδες γεωμετρικές μορφές, συμπαγείς γεωμετρικές μορφές και προβαλλόμενες μορφές.

Η πρώτη κατηγορία είναι η "πραγματική" και αποτελείται από αντικείμενα του πραγματικού κόσμου. Η δεύτερη κατηγορία αποτελείται από αναπαραστάσεις πραγματικών αντικειμένων. Η τρίτη κατηγορία αποτελείται από αφηρημένες μορφές που μπορούν να αναπαρασταθούν χρησιμοποιώντας σύμβολα αλλά δεν αντιστοιχούν σε πραγματικά αντικείμενα (γιατί δεν έχουν αντίστοιχο ερέθισμα στον πραγματικό κόσμο). Μια σειρά μαθημάτων πάνω στην οπτική γνωστική με μια έμφαση στην αναγνώριση σχήματος έχει γραφτεί από τον Pinker [186].

2.2.2. Σύγχρονες Θεωρίες Οπτικής Αντίληψης

Ο Marr et al [187, 188, 189, 190, 191, 192] συνεισέφεραν σημαντικά στη μελέτη του ανθρώπινου συστήματος οπτικής αντίληψης. Ακολουθώντας το παράδειγμα του Marr [193], το επίκεντρο της έρευνας μετατοπίζεται από τις εφαρμογές σε θέματα σχετικά με τα δομικά στοιχεία του ανθρώπινου οπτικού συστήματος. Αυτό φαίνεται στην αποκαλούμενη *σχήμα από x* έρευνα, η οποία αποτελεί ένα σημαντικό κομμάτι της συνολικής έρευνας στην όραση υπολογιστή [3]. Οι μελέτες που ασχολούνται με τις τεχνικές σχήμα από x συμπεριλαμβάνουν τις ακόλουθες: σχήμα από σκίαση [194], σχήμα από περίγραμμα [195, 196], σχήμα από υφή [185], σχήμα από στερεοσκοπία [197], και σχήμα από γεωμετρία φράκταλ [198]. Στην βιβλιογραφική αναφορά [188] ο Marr ανέπτυξε ένα παράδειγμα πρωταρχικού σκαριφήματος για την πρώτη επεξεργασία οπτικής πληροφορίας. Στη μέθοδό του, χρησιμοποιείται μια ομάδα από μάσκες για την μέτρηση των ασυνεχειών στις πρώτες και δεύτερες παραγώγους της αρχικής εικόνας. Κατόπιν, κάποιες διαδικασίες επεξεργάζονται αυτή την πληροφορία για να δημιουργήσουν ένα πρωταρχικό σκαρίφημα του σκηνικού. Το πρωταρχικό σκαρίφημα περιέχει τις τοποθεσίες των ακμών στην εικόνα και χρησιμοποιείται από ακόλουθα στάδια της διαδικασίας ανάλυσης σχήματος. Οι Marr και

Hildreth [191] ανέπτυξαν την ιδέα του πρωταρχικού σκαριφήματος και πρότειναν ένα νέο φίλτρο ανίχνευσης ακμών που βασίζεται στους μηδενισμούς της Λαπλασιανής της δισδιάστατης Γκαουσιανής συνάρτησης κατανομής. Σε αυτή την προσέγγιση, οι μηδενισμοί της Λαπλασιανής υποδεικνύουν τα σημεία καμπής ώστε να εντοπιστούν οι τοποθεσίες των ακμών. Ο Koenderink [199] και οι Koenderink van Doorn [200] έχουν μελετήσει τις ψυχολογικές πλευρές της οπτικής αντίληψης και έχουν προτείνει αρκετά ενδιαφέροντα παραδείγματα. Οι συμβατικές προσεγγίσεις του σχήματος είναι συχνά στατικές με την έννοια ότι αντιμετωπίζουν όλες τις λεπτομέρειες του σχήματος με τον ίδιο τρόπο όπως τα συνολικά χαρακτηριστικά του σχήματος [200]. Έχει αναπτυχθεί ένα δυναμικό μοντέλο σχήματος στο οποίο η οπτική αντίληψη εκτελείται σε διάφορα επίπεδα ανάλυσης. Παρόμοιες αντιλήψεις ιεραρχίας και σχετικότητας υπάρχουν στην οπτική ψυχολογία αλλά απουσιάζουν από τις συμβατικές γεωμετρικές θεωρίες του σχήματος. Στη βιβλιογραφική αναφορά [200] υπάρχει η επιχειρηματολογία ότι υπάρχουν εγχειρίδια για τη θεωρία της τέχνης (όπως είναι το [201]), στα οποία όμως δεν έχει δοθεί η δέουσα προσοχή, και τα οποία περιέχουν πρακτικές γνώσεις συσσωρευμένες ανά τους αιώνες. Στην τέχνη όπως και στην αντίληψη, ένα σχήμα θεωρείται ως μια ιεραρχική δομή. Στη βιβλιογραφική αναφορά [200] αναπτύσσεται μια διαδικασία μορφογένεσης βασισμένη σε πολλαπλά επίπεδα ανάλυσης. Οποιοδήποτε σχήμα μπορεί να ενσωματωθεί σε μια "μορφογενετική ακολουθία" ανάλογα με τη λύση της μερικής διαφορικής εξίσωσης που περιγράφει την εξέλιξη του σχήματος μέσω πολλαπλών αναλύσεων. Πολλοί συγγραφείς συμφωνούν στη σημασία των σημείων υψηλής καμπυλότητας για την οπτική αντίληψη. Οι Attneave και Arnoult [202, 203] εκτέλεσαν ψυχολογικά πειράματα ώστε να εξερευνήσουν τη σημασία των γωνιών για την αντίληψη. Στο διάσημο πείραμα με τη γάτα του Attneave χρησιμοποιήθηκε ένα σχέδιο μιας γάτας για τον εντοπισμό σημείων με υψηλή καμπυλότητα τα οποία κατόπιν συνδέθηκαν για να δημιουργηθεί ένα απλοποιημένο σχέδιο της γάτας. Μετά από μια σύντομη παρουσίαση η γάτα αναγνωριζόταν με αξιοπιστία στο σχέδιο. Έχει προταθεί ότι τέτοια σημεία έχουν υψηλό πληροφοριακό περιεχόμενο. Το έργο του Attneave έχει προκαλέσει περαιτέρω έρευνα πάνω στο θέμα της κατάτμησης καμπυλών [204, 205, 206, 207, 208, 209, 210]. Για την προσέγγιση καμπυλών από ευθείες γραμμές, τα σημεία υψηλής καμπυλότητας είναι τα καλύτερα για το σπάσιμο των γραμμών, ώστε η προκύπτουσα εικόνα να περιέχει τη μέγιστη ποσότητα πληροφορίας που είναι απαραίτητη για την επιτυχή αναγνώριση σχήματος. Για την περιγραφή σχήματος, ως σημεία υψηλής καμπυλότητας χρησιμοποιούνται οι γωνίες και το σχήμα μπορεί να προσεγγισθεί από ένα πολύγωνο. Έχει αναπτυχθεί μια ποικιλία αλγορίθμων για πολυγωνική προσέγγιση σχήματος [67, 1, 211, 212, 213, 214, 68]. Ο Davis [215] συνδύασε τη χρήση σημείων υψηλής καμπυλότητας και προσεγγίσεων με τμήματα γραμμών για πολυγωνική προσέγγιση σχήματος. Οι Stokely και Shang [216] εξέτασαν μεθόδους μέτρησης της καμπυλότητας τρισδιάστατων επιφανειών, οι οποίες εξελίσσονται σε πολλές εφαρμογές (π.χ. ιατρικές τομογραφικές εικόνες). Οι Hoffman

και Richards [126, 217] επιχειρηματολογούν ότι το οπτικό σύστημα αποσυνθέτει αντικείμενα σε σημεία υψηλής αρνητικής καμπυλότητας. Αυτό συνάδει με την αρχή της εγκαρσιότητας [218] που βρίσκεται στη φύση. Η αρχή της εγκαρσιότητας υποστηρίζει ότι όταν δύο τυχαία κυρτά αντικείμενα αλληλοδιεισδύονται, τότε το σημείο επαφής είναι ένα σημείο περιγράμματος με κοίλη ασυνέχεια που ανήκει στα εφαπτόμενα επίπεδά τους. Ο Leyton [219] απέδειξε το θεώρημα Συμμετρίας-Καμπυλότητας, το οποίο υποστηρίζει ότι οποιοδήποτε τμήμα καμπύλης με ένα ακρότατο καμπυλότητας έχει ένα μοναδικό άξονα συμμετρίας που καταλήγει στο ακρότατο αυτό. Αυτό το αποτέλεσμα είναι πολύ σημαντικό γιατί εδραιώνει τη σύνδεση ανάμεσα σε δύο σημαντικές έννοιες της οπτικής αντίληψης. Στη βιβλιογραφική αναφορά [220], ο Leyton ανέπτυξε μια θεωρία που υποστηρίζει ότι όλα τα σχήματα είναι βασικά κύκλοι, οι οποίοι έχουν αλλάξει μορφή ως αποτέλεσμα διαφόρων παραμορφώσεων, όπως είναι η πίεση, η έλξη, η επιμήκυνση κλπ., επιβαλλόμενων από εξωτερικές δυνάμεις. Εμφανίστηκαν δύο προβλήματα: το πρώτο ήταν το συμπέρασμα του ιστορικού του σχήματος από ένα απλό σχήμα, και το δεύτερο ήταν το συμπέρασμα της εξέλιξης από το ένα σχήμα στο άλλο. Η έννοια της συμμετρίας-καμπυλότητας χρησιμοποιήθηκε για να εξηγηθεί η διαδικασία παραμόρφωσης του αντικειμένου. Οι άξονες συμμετρίας δείχνουν τις πιθανές κατευθύνσεις παραμόρφωσης. Στη βιβλιογραφική αναφορά [219], ο Leyton πρότεινε μια θεωρία δομών ελέγχου, οι οποίες, υποστηρίζει, ότι καθορίζουν τη λειτουργία του ανθρώπινου συστήματος αντίληψης. Είναι ένα ιεραρχικό σύστημα όπου σε κάθε επίπεδο ελέγχου, όλα τα επίπεδα κάτω από οποιοδήποτε δοθέν επίπεδο συμπεριλαμβάνονται στην επεξεργασία της πληροφορίας. Ο Pentland [221, 222] εξέτασε μεθόδους για την αναπαράσταση φυσικών μορφών μέσω της γεωμετρίας φράκταλ [223, 224, 225]. Υποστήριξε ότι οι συναρτήσεις φράκταλ είναι κατάλληλες για την αναπαράσταση φυσικών μορφών επειδή πολλές φυσικές διεργασίες παράγουν σχήματα επιφανειών φράκταλ. Αυτό συμβαίνει επειδή οι φυσικές μορφές επαναλαμβάνονται όποτε αυτό είναι δυνατό και τα αντικείμενα, εξαιρουμένων των ζώων, έχουν περιορισμένο αριθμό δυνατών μορφών [226, 227]. Τα περισσότερα μοντέλα για την αναπαράσταση του σχήματος είχαν αναπτυχθεί για πρακτικές εφαρμογές και όχι αναγκαστικά για τη μελέτη της αντίληψης. Οι αναπαραστάσεις φράκταλ παράγουν αντικείμενα που αντιστοιχούν πολύ καλύτερα στο ανθρώπινο μοντέλο της οπτικής αντίληψης και γνωστικής. Ο Lowe [228] πρότεινε ένα σύστημα όρασης υπολογιστή που μπορεί να αναγνωρίζει τρισδιάστατα αντικείμενα από άγνωστες οπτικές γωνίες και απλές δισδιάστατες εικόνες. Η διαδικασία είναι μη-τυπική και χρησιμοποιεί τρεις μηχανισμούς της αντίληψης για ομαδοποίηση ώστε να προσδιορίζει τρισδιάστατη επίγνωση του αντικειμένου σε αντίθεση με μια συμβατική προσέγγιση από το ειδικό στο γενικό. Το μειονέκτημα των προσεγγίσεων από το ειδικό στο γενικό είναι ότι χρειάζονται μεγάλο όγκο πληροφοριών για να αναγνωρίσουν ένα αντικείμενο. Αντίθετα, το ανθρώπινο οπτικό σύστημα μπορεί να αναγνωρίσει αντικείμενα ακόμα και με διάσπαρτα δεδομένα ή μερικώς εγκλεισμένα αντικείμενα. Οι συνθήκες που

πρέπει να ικανοποιηθούν από τις διαδικασίες της αντίληψης για την ομαδοποίηση είναι οι ακόλουθες:

1. Η συνθήκη του αναλλοίωτου της οπτικής γωνίας. Αυτό σημαίνει ότι τα παρατηρούμενα πρωτογενή χαρακτηριστικά πρέπει να παραμένουν σταθερά σε διάφορες οπτικές γωνίες.
2. Η συνθήκης εντοπισμού. Πρέπει να είναι διαθέσιμη αρκετή πληροφορία ώστε να αποφευχθούν συμπτωματικές παρερμηνείες.

Οι διαδικασίες ομαδοποίησης που χρησιμοποίησε ο Lowe είναι οι ακόλουθες: Η ομαδοποίηση βασισμένη στην εγγύτητα των ακραίων σημείων των γραμμών χρησιμοποιήθηκε ως μια διαδικασία αναλλοίωτου οπτικής γωνίας. Η δεύτερη διαδικασία ήταν ομαδοποίηση βασισμένη στον παραλληλισμό, η οποία είναι επίσης ανεξάρτητη της οπτικής γωνίας. Η τρίτη διαδικασία ήταν ομαδοποίηση βασισμένη στη συγγραμικότητα. Η διαδικασία προεπεξεργασίας αποτελούταν από ανίχνευση ακμών, χρησιμοποιώντας τους μηδενισμούς του Marr στην εικόνα, σε συνέλιξη με μια Λαπλασιανή ενός Γκαουσιανού φίλτρου. Στο επόμενο βήμα εκτελούνταν μια κατάτμηση γραμμής. Οι διαδικασίες ομαδοποίησης εκτελούνταν στα δεδομένα από την κατάτμηση γραμμής ώστε να προσδιοριστούν οι πιθανές τοποθεσίες των αντικειμένων.

Τελευταία η ιδέα της χρήσης σημείων αναφοράς με υψηλή καμπυλότητα για την περιγραφή σχημάτων έχει εκτιμηθεί από πολλούς ερευνητές σε διαφορετικές περιοχές της νευροψυχολογίας [229, 230, 231]. Οι A.Dobbins, S.W.Zucker και M.S.Cynader παρουσίασαν αποδείξεις για το γεγονός ότι η ανίχνευση καμπυλότητας σχετίζεται με τερματικούς νευρώνες και επίσης παρουσίασαν ένα μαθηματικό μοντέλο που υποστηρίζει την απόδειξη [232]. Τα υπερπολύπλοκα τερματικά κύτταρα έχουν αποδεκτικά πεδία με ανασταλτικές ζώνες που σταματούν την απόκριση του κυττάρου σε ερεθίσματα που εισχωρούν στις τερματικές ζώνες [233]. Ο A.Dobbins και οι υπόλοιποι εξηγούν το φαινόμενο αυτό ως ένα συνδυασμό των αποδεκτικών πεδίων διαφορετικών μεγεθών επιλεγόμενου προσανατολισμού των απλών κυττάρων που τελικά κωδικοποιούν σε πρώτη προσέγγιση τόσο την καμπυλότητα όσο και τον προσανατολισμό. Το μοντέλο μας μετρά την καμπυλότητα συνδυάζοντας τις τοπικές αποκρίσεις των απλών κυττάρων (επίσης αποκρίσεις επιλογής προσανατολισμού και τερματικές αποκρίσεις) σε μια μεγαλύτερη τοπική περιοχή όπως θα εξηγηθεί παρακάτω όπου περιγράφεται η τεχνητή νευρωνική στρωματική αρχιτεκτονική.

Περαιτέρω έρευνα στο πεδίο της γνωστικής έχει αποκαλύψει ότι η εξαγωγή σημείων αναφοράς είναι μια κρίσιμη διαδικασία στην ανθρώπινη αντίληψη και η βάση για πιθανούς μηχανισμούς ταυτοποίησης και αναγνώρισης σχημάτων [234]. Στη βιομετρική της βιολογίας ο Bookstein όρισε σημεία αναφοράς σε διάφορα βιολογικά σχήματα για

ταξινόμηση των ειδών [235]. Η χρήση σημείων αναφοράς θεωρείται, από τις μεθόδους αναπαράστασης και κωδικοποίησης πληροφοριών του σχήματος, η πιο συμβατή με τη γνωστική. Ο Berreti εισήγαγε την αποσύνθεση των σχημάτων σε αρχέγονα βάσει της καμπυλότητας [236]. Ο Atneave και ο Pomerantz και οι υπόλοιποι παρατήρησαν ότι η καμπυλότητα μιας καμπύλης έχει εξέχοντα χαρακτηριστικά αντίληψης [237, 238] και ότι αποδεικνύεται χρήσιμη στην αναγνώριση σχημάτων [239, 240]. Οι Asada και Brandy έχουν αναπτύξει τον περιγραφέα "καμπυλότητας πρωταρχικού σκαριφήματος" [241], μια κλιμακωτή κατασκευή πολλών επιπέδων βασισμένη στην εξαγωγή αλλαγών στην καμπυλότητα. Από τα χαρακτηριστικά της καμπυλότητας κατασκευάζεται μια περιγραφή του περιγράμματος όσον αφορά τα κατασκευαστικά αρχέγονα (π.χ. άκρες, γωνίες κ.λπ.). Οι Mokhtarian και Mackworth [242, 243] έδειξαν ότι τα σημεία αλλαγής μονοτονίας της καμπυλότητας που εξάγονται χρησιμοποιώντας ένα Γκαουσιανό χώρο κλιμάκων μπορούν να χρησιμοποιηθούν στην αναγνώριση καμπύλων αντικειμένων. Οι Dudek και Tsotsos [244] παρουσίασαν μια τεχνική για αναπαράσταση σχημάτων και αναγνώριση αντικειμένων βασισμένη σε πληροφορία καμπυλότητας πολλών κλιμάκων. Μια παρόμοια τεχνική βασισμένη σε σημεία αναφοράς υψηλής καμπυλότητας παρουσιάζεται στη βιβλιογραφική αναφορά [245].

3. ΒΙΟΛΟΓΙΚΑ ΣΥΜΒΑΤΕΣ ΤΕΧΝΙΚΕΣ ΓΙΑ ΑΝΑΠΑΡΑΣΤΑΣΗ ΣΧΗΜΑΤΟΣ

Οι ειδικοί δεκτικοί τομείς προσανατολισμού των απλών νευρώνων που ανακαλύφθηκαν από τους κατόχους του βραβείου Νομπέλ για τον λόγο αυτόν, Hubel και Wiesel, εμπνέουν την περιγραφή μιας πολυστρωματικής νευρωνικής αρχιτεκτονικής για τον υπολογισμό της τοπικής καμπυλότητας σε κάθε σημείο μιας επίπεδης μορφής χωρίς εξαγωγή του ελλοχεύοντος περιγράμματος. Η αρχιτεκτονική που προτείνεται εδώ ομοιάζει με την οπτική νεύρωση των πρωτευόντων, δεδομένου ότι καταδεικνύεται ένας τρόπος κατά τον οποίο οι ειδικοί δεκτικοί τομείς προσανατολισμού των απλών νευρώνων μπορούν να εκτελέσουν τοπικούς υπολογισμούς καμπυλότητας της επίπεδης μορφής που προβάλλεται στον αμφιβληστροειδή χιτώνα του ματιού. Στην συνέχεια η μέθοδος αυτή χρησιμοποιείται για την κωδικοποίηση της επίπεδης καμπυλότητας στις τιμές έντασης φωτεινότητας μιας ασπρόμαυρης εικόνας.

Η διάμεση μετατροπή άξονα (MAT) (ή μετατροπή σκελετού) είναι μια από τις πιο μελετημένες τεχνικές αντιπροσώπευσης μορφής με εγκατεστημένα πλεονεκτήματα σε σχέση με αναγνώριση μορφής. Ιδιαίτερα, η ενσωμάτωση των τοπικών πληροφοριών του περιγράμματος της μορφής στο σκελετό, έχει αποδειχθεί ότι βελτιώνει την ικανότητα αναγνώρισης μορφής στα επίπεδα της τρέχουσας στάθμης της τεχνικής. Στην διατριβή παρουσιάζονται οπτικές δίοδοι φλοιωδών νευρώνων για την εξαγωγή ενός, ανάλογου του σκελετού, περιγραφέα μορφής, που όμως περιέχει επίσης τοπικές πληροφορίες καμπυλότητας περιγράμματος. Αναφερόμαστε σε αυτήν την δομή με τον όρο *σκελετός καμπυλότητας*. Η προτεινόμενη αρχιτεκτονική εμπνέεται από τα βιολογικά ευρήματα σχετικά με τους φλοιώδεις νευρώνες του οπτικού φλοιού και των ειδικής χρήσης αποδεκτικών πεδίων τους (RFs). Τα σημεία της υψηλής καμπυλότητας προσδιορίζονται αρχικά και συνδυάζονται στη συνέχεια με τη βοήθεια μιας οπτικής διάβασης που επιτυγχάνει μία ανάλογη του σκελετού της μορφής κωδικοποίηση αλλά και ενσωματώνει στο σκελετό τις τοπικές πληροφορίες καμπυλότητας του περιγράμματος Παρουσιάζουμε πειραματικά αποτελέσματα που δείχνουν ότι μια τέτοια κωδικοποίηση μπορεί να βελτιώσει την ικανότητα αναγνώρισης των ταξινομητών νευρικών δικτύων.

3.1. Η Μέτρηση Καμπυλότητας σαν διαδικασία των Φλοιωδών Νευρώνων

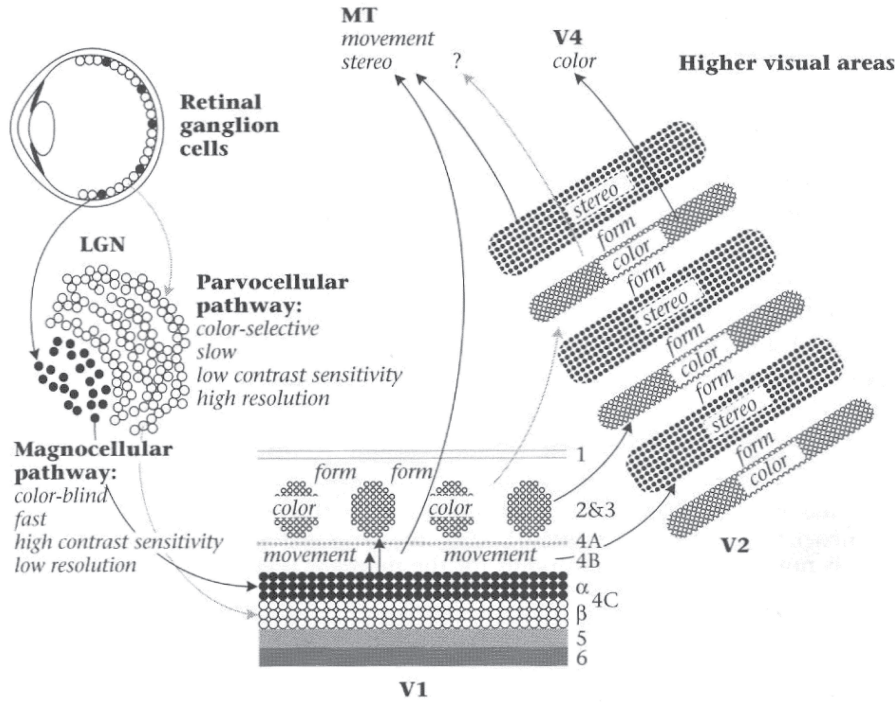
Η γνώση γύρω από τους φλοιώδεις νευρώνες και τα αποδεκτικά τους πεδία οφείλεται στους νομπελίστες Hubel και Wiesel. Η 25ετής τους συνεργασία επέφερε μια χωρίς προηγούμενο πρόοδο στην κατανόηση των αποκρίσεων των φλοιωδών νευρώνων. Στις εργασίες τους προσδιορίζουν τις διαφορές ανάμεσα στα αποδεκτικά πεδία περιοχής VI και στα αποδεκτικά πεδία Πλάγιου Γονατώδη Πυρήνα (ΠΓΠ) χρησιμοποιώντας ερεθίσματα σχετικά με την όραση. Το Σχ. 1 δείχνει τις περιοχές του οπτικού συστήματος των πρωτευόντων. Οι ποιοτικές μέθοδοι που χρησιμοποίησαν για τη μελέτη του φλοιού είναι ακόμη κυρίαρχες στην πειραματική φυσιολογία [246, 247, 248]. Οι Hubel και Wiesel

κατέγραψαν τη δραστηριότητα των φλοιωδών νευρώνων κατά την προβολή διατεταγμένων ερεθισμάτων, κυρίως τμήματα γραμμών και σημεία, σε μια οθόνη. Η εικόνα της οθόνης έφτανε στον αμφιβληστροειδή του πειραματόζωου μέσω του κερατοειδούς χιτώνα και του φακού. Καθώς το μικρο-ηλεκτρόδιο εισχωρούσε στον οπτικό φλοιό, προβάλλονταν τμήματα γραμμών με ρυθμιζόμενο πλάτος και μήκος. Αφού άλλαζαν τη θέση του ερεθίσματος στην οθόνη, αναζητώντας το αποδεκτικό πεδίο του νευρώνα, κατόπιν μέτρησαν την απόκριση του νευρώνα στις γραμμές, στις παύλες και στα σημεία ξεχωριστά. Μια από τις κύριες ανακαλύψεις τους ξεκίνησε με μια εισχώρηση που έφτασε στο φλοιό με γωνία περίπου 45° και μήκος 2.5 mm. Καθώς το ηλεκτρόδιο προχωρούσε οι προσανατολισμοί των διαδοχικά καταγεγραμμένων κυττάρων μεταβάλλονταν με μικρά βήματα, περίπου 10° για κάθε 50 mm καταγράφοντας με αυτόν τον τρόπο 53 διαδοχικούς προσανατολισμούς χωρίς μεγάλη διαφορά μεταξύ τους [246]. Η συγκεκριμένη λοιπόν συγκεκριμένου προσανατολισμού διάταξη των κυττάρων εντοπίζει ή είναι δεδομένη για την κατεύθυνση ενός μικρού τμήματος γραμμής, κατά πόσο όμως η αντίληψη μιας γραμμής ή μιάς καμπύλης εξαρτάται από αυτά και το πώς συγκεντρώνεται η πληροφορία από τέτοιες ομάδες κυττάρων σε ακόλουθα στάδια του συστήματος, ώστε να σχηματιστεί αυτό που αποκαλούμε αντίληψη γραμμών ή καμπυλών, είναι ακόμα άγνωστο.

Η βασική συνεισφορά αυτού του κεφαλαίου είναι η περιγραφή ενός βιολογικά εμπνευσμένου, σε συμφωνία με την από τους Hubel και Wiesel, περιγραφή των περιστρεφόμενων αποδεκτικών πεδίων επιλεγόμενου προσανατολισμού των φλοιωδών νευρώνων, νευρωνικού συστήματος για τον υπολογισμό της επίπεδης καμπυλότητας ενός σχήματος [23], το οποίο παρουσιάζεται στον αμφιβληστροειδή χιτώνα του ματιού. Ειδικότερα, κατασκευάζεται μια νευρωνική αρχιτεκτονική τύπου αποδεκτικού πεδίου για την κωδικοποίηση της επίπεδης καμπυλότητας στις τιμές έντασης φωτεινότητας των ασπρόμαυρων εικόνων χρησιμοποιώντας τα πλεονεκτήματα της πολυστρωματικής αρχιτεκτονικής νευρωνικών δικτύων για την ανίχνευση σημείων υψηλής καμπυλότητας στο περίγραμμα και προτείνοντας ένα ιεραρχικό μοντέλο που συμπεριλαμβάνει οπτική απόκτηση και μέτρηση καμπυλότητας σε διακεκριμένα στρώματα τηρώντας την συμβατότητα με το ΑΟΣ.

3.2. Το Προτεινόμενο Μοντέλο για Μέτρηση Επίπεδης Καμπυλότητας

Τα απλά κύτταρα έχουν προσανατολισμένα αποδεκτικά πεδία και επομένως αντιδρούν στα ερεθίσματα ανάλογα με τον προσανατολισμό. Αυτή η ιδιότητα των αποδεκτικών πεδίων καλείται επιλεκτικότητα προσανατολισμού. Ο προσανατολισμός του ερεθίσματος που προκαλεί την ισχυρότερη απόκριση καλείται ο κύριος προσανατολισμός του κυττάρου. Η επιλεκτικότητα προσανατολισμού των φλοιωδών νευρώνων είναι μια κρίσιμη ιδιότητα των αποδεκτικών πεδίων. Οι νευρώνες ΠΓΠ και αυτοί του αμφιβληστροειδή έχουν αποδεκτικά πεδία με κυκλική συμμετρία και αντιδρούν σχεδόν εξίσου καλά σε όλους τους προ-



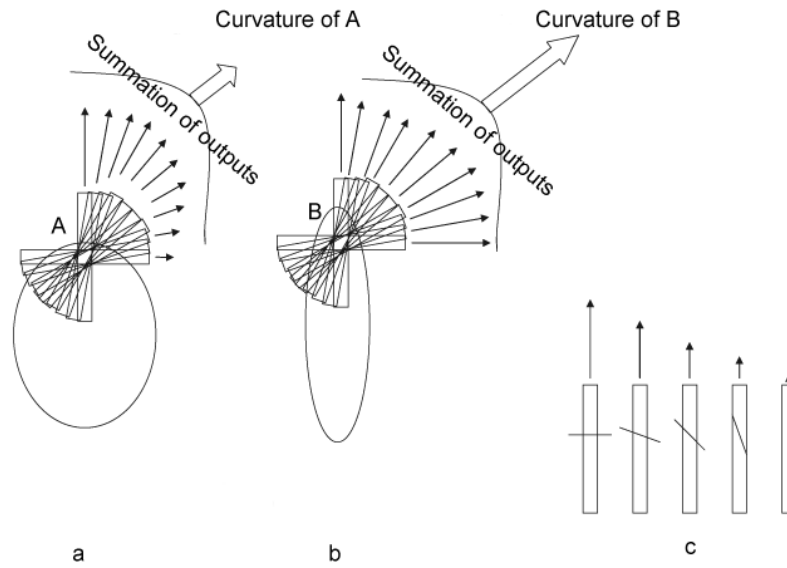
Σχήμα 1: Το οπτικό σύστημα των πρωτεύοντων

σανατολισμούς των ερεθισμάτων. Οι νευρώνες επιλογής προσανατολισμού βρίσκονται στα στρώματα 2 και 3 του οπτικού φλοιού, και σπανιότερα στις κύριες εισόδους του στρώματος 4C. Εδώ θα δειχθεί ότι συνεχείς διαδοχικοί προσανατολισμοί ενός φίλτρου επιλογής προσανατολισμού, όπως αυτοί που ανακάλυψαν οι Hubel και Wiesel στον οπτικό φλοιό των πρωτεύοντων, μπορεί να αποτελέσει μηχανισμό μέτρησης της επίπεδης καμπυλότητας. Στην επόμενη ενότητα θα περιγραφεί ο μηχανισμός και το μοντέλο διασύνδεσης των νευρώνων που με τα παραπάνω δεδομένα κωδικοποιεί την καμπυλότητα του επίπεδου σχήματος στις τιμές έντασης φωτεινότητας της εικόνας.

Έστω X μία επίπεδη καμπύλη σε παραμετρική μορφή ως προς το βαθμωτό μέγεθος t , της μορφής $\vec{c}(t) = (x(t), y(t))$, όπου $x(t)$ και $y(t)$ οι συντεταγμένες συναρτήσεις. Εάν το κέντρο των αξόνων περιστραφεί κατά θ ένα σημείο της καμπύλης θα γίνει $\vec{c}_\theta(t) = (x_\theta(t), y_\theta(t))$ όπου $x_\theta(t) = x(t)\cos(\theta) + y(t)\sin(\theta)$, $y_\theta(t) = -x(t)\sin(\theta) + y(t)\cos(\theta)$ και θ η γωνία περιστροφής.

Για κάθε ζεύγος (t, θ) όπου θ είναι η γωνία της εφαπτόμενης της καμπύλης στο σημείο $\vec{c}(x(t), y(t))$ είναι:

$$\begin{aligned} \tan(\theta) &= \frac{\dot{y}(t)}{\dot{x}(t)} \iff \dot{y}(t)\cos(\theta) = \dot{x}(t)\sin(\theta) \\ &\iff \dot{y}(t)\cos(\theta) - \dot{x}(t)\sin(\theta) = 0 \iff \dot{y}_\theta(t) = 0 \end{aligned} \tag{1}$$



Σχήμα 2: Μέτρηση της επίπεδης καμπυλότητας με διαδοχικές περιστροφές κατάλληλων αποδεκτικών πεδίων.

Επομένως η παράγωγος της δεύτερης συνάρτησης συντεταγμένης σε κάποιο σημείο t μηδενίζεται όταν ο άξονας περιστρέφεται κατά μια γωνία ίση με τη γωνία της εφαπτομένης στην καμπύλη στο ίδιο σημείο. Η καμπυλότητα ορίζεται ως ο ρυθμός μεταβολής ανά μονάδα απειροστού μήκους τόξου (ds) της γωνίας της εφαπτομένης στην καμπύλη στο συγκεκριμένο σημείο. Για την ευθεία γραμμή δεν υπάρχει μεταβολή και για τον κύκλο υπάρχει συνεχής μεταβολή. Γενικά η καμπυλότητα $\kappa(t)$ της X στο σημείο $\vec{c}(t)$ είναι:

$$\kappa(t) = \frac{\dot{x}(t)\ddot{y}(t) - \dot{y}(t)\ddot{x}(t)}{\|\dot{\vec{c}}(t)\|^3} \quad (2)$$

που προκύπτει από την Εξ. 1 λύνοντας ως προς $d\theta/ds$, αλλά αν οι άξονες περιστραφούν γύρω από την αρχή κατά μια γωνία θ ίση με τη γωνία της εφαπτομένης στην καμπύλη σε εκείνο το σημείο τότε η Εξ. 2 της καμπυλότητας σε αυτό το σημείο, αφού $\dot{y}_\theta(t) = 0$, γίνεται:

$$\kappa(t) = \frac{\dot{y}_\theta(t)}{\dot{x}_\theta(t)^2} \quad (3)$$

Συνεπώς ο αναλυτικός υπολογισμός της καμπυλότητας στο σημείο $\vec{c}_\theta(x_\theta(t), y_\theta(t))$ έχει απλοποιηθεί λόγω της περιστροφής των αξόνων κατά γωνία θ . Επιπλέον από την Εξ. 3 φαίνεται ότι όταν οι άξονες περιστρέφονται κατά αυτή τη γωνία, η καμπυλότητα σε αυτό το σημείο μπορεί να προκύψει ως μια προσέγγιση της παραγώγου της πρώτης συνάρτησης

συντεταγμένης $\dot{x}_\theta(t)$ και της δεύτερης παραγώγου της δεύτερης συνάρτησης συντεταγμένης $\ddot{y}_\theta(t)$. Τώρα θα περιγραφεί ο τρόπος κατά τον οποίο η παράγωγος $\dot{x}_\theta(t)$ μπορεί να προσεγγιστεί από ένα κατάλληλο περιστρεφόμενο αποδεκτικό πεδίο και η δεύτερη παράγωγος $\ddot{y}_\theta(t)$ μέσω μιας μεθόδου συνένωσης σε μια ευρύτερη περιοχή γύρω από το σημείο \vec{c}_θ . Επίσης θα περιγραφεί ο τρόπος κατά τον οποίο αυτά μπορούν να συνδυαστούν σε μια τεχνική για τον υπολογισμό την καμπυλότητας μιας επίπεδης καμπύλης με μια σειρά απλών πράξεων που συνάδουν με την παραπάνω ανάλυση και είναι ταυτόχρονα συμβατές με την υπάρχουσα γνώση γύρω από τη λειτουργικότητα των φλοιωδών κυττάρων.

Τα φλοιώδη κύτταρα στα στρώματα 2 και 3 του οπτικού φλοιού έχουν αποδεκτικά πεδία επιλεγόμενου προσανατολισμού, ο οποίος αλλάζει συνεχόμενα σε διαδοχικά επίπεδα. Με διαδοχικές περιστροφές ενός φίλτρου επιλεγόμενου προσανατολισμού μπορεί πράγματι να μετρηθεί η καμπυλότητα σε κάθε σημείο του περιγράμματος. Η ιδέα είναι να χρησιμοποιηθεί η επιλεκτικότητα προσανατολισμού ώστε να εντοπισθεί η κατεύθυνση της εφαπτομένης σε ένα σημείο του περιγράμματος και ταυτόχρονα να μετρηθεί η καμπυλότητα στο σημείο αυτό μέσω του αριθμού των ενεργοποιήσεων των διαδοχικών στρωμάτων στα οποία το περιστρεφόμενο πεδίο παραμένει κοντά στην κατεύθυνση της εφαπτομένης. Έτσι χρησιμοποιείται η περιστροφή, που συμβαίνει στα κύτταρα του οπτικού φλοιού, και η επιλεκτικότητα του προσανατολισμού ώστε να μετρηθεί η τοπική καμπυλότητα με τρόπο συμβατό με την λειτουργικότητα των κυττάρων στα στρώματα 2 και 3 του οπτικού φλοιού.

Μια περιστροφή του αποδεκτικού πεδίου επιλεγόμενου προσανατολισμού αντιστοιχεί σε ένα βήμα της κατεύθυνσης της εφαπτομένης και του μήκους τόξου. Εάν επομένως περιστραφεί το αποδεκτικό πεδίο επιλεγόμενου προσανατολισμού και οι αντίστοιχοι νευρώνες συνεχίσουν να ενεργοποιούνται αυτό σημαίνει ότι προσεγγίζεται το περίγραμμα από την κατεύθυνση της εφαπτομένης παρά την μετατόπιση του ελεγχόμενου σημείου κατά μια μονάδα μήκους τόξου. Όσο το περίγραμμα συνεχίζει να προσεγγίζεται περιστρέφοντας την κατεύθυνση της εφαπτομένης και προχωρώντας στο τόξο κατά μια μονάδα μήκους, τόσο μεγαλύτερη θα είναι η καμπυλότητα στο συγκεκριμένο σημείο και μπορεί να μετρηθεί μέσω του αριθμού των ενεργοποιήσεων των στρωμάτων, όπου κάθε στρώμα αντιστοιχεί σε μια περιστροφή του αποδεκτικού πεδίου κατά ένα μικρό βήμα, για παράδειγμα 10 μοίρες. Η μέθοδος αυτή υπολογισμού της καμπυλότητας μπορεί να εφαρμοστεί από μια στρωματική αρχιτεκτονική σχεδιασμένη με ανάλογο τρόπο ώστε να μοιάζει με τη διάταξη των φύλλων του φλοιού. Τώρα θα περιγραφεί αναλυτικά η προτεινόμενη διαδικασία για τον υπολογισμό της επίπεδης καμπυλότητας.

Έστω $\Theta = \{\theta_1, \theta_2 \dots \theta_N\}$ ένα σύνολο γωνιών καθεμία από τις οποίες σχετίζονται με κάποιο στρώμα L_θ στο οποίο πραγματοποιείται μια ανίχνευση ακμών της αρχικής εικόνας I υπό τη συγκεκριμένη γωνία. Έστω $p_i, i \in I$ ένα εικονοστοιχείο (νευρώνας) στο όριο του σχήματος της παρουσιαζόμενης εικόνας. Έστω ότι η δυνατότητα του p_i να βρίσκεται σε ακμή γωνίας θ συμβολίζεται με $\epsilon_\theta(i)$. Σε κάθε επίπεδο L_θ , το $\epsilon_\theta(i)$ εκτιμάται από την πράξη

του στοιχειώδους νευρώνα :

$$\epsilon_{\theta}(i) = \sum_{k \in A_i} w_{\theta,k} p_k \quad (4)$$

όπου $\{p_k, k \in A_i\}$ είναι το σύνολο εικονοστοιχείων (νευρώνων) σε μια περιοχή του p_i και $w_{\theta,k}$ είναι το βάρος που σχετίζεται με το εικονοστοιχείο p_i στο επίπεδο $w_{\theta,k}$. Εδώ πρέπει να τονιστεί η διαφοροποίηση του συμβολισμού ανάμεσα στα κ και k , όπου το πρώτο συμβολίζει την καμπυλότητα ενώ το δεύτερο είναι δείκτης. Σε κάθε επίπεδο L_{θ} οι ποσότητες $\epsilon_{\theta}(i)$ εκτιμώνται από αυτή την πράξη στοιχειώδους νευρώνα για όλα τα εικονοστοιχεία p_i . Σύμφωνα με τη προτεινόμενη μέθοδο, η καμπυλότητα κ_i της επίπεδης καμπύλης στο σημείο p_i είναι ανάλογη με τον αριθμό των διαφορετικών κατευθύνσεων των ανιχνευομένων ακμών στην περιοχή του p_i . Για να μετρηθεί ο αριθμός των διαφορετικών κατευθύνσεων που ανιχνεύονται για τις ακμές στην περιοχή του p_i πρέπει να αθροιστούν οι ποσότητες $\epsilon_{\theta}(k)$ για όλες τις κατευθύνσεις θ και όλα τα γειτονικά εικονοστοιχεία p_k :

$$\kappa_i = \sum_{k \in A_i, \theta \in \Theta} \epsilon_{\theta}(k) = \sum_{\theta \in \Theta} \sum_{k \in A_i} \sum_{r \in A_k} w_{\theta,r} p_r \quad (5)$$

όπου A_i και A_k συμβολίζουν τα σύνολα αρίθμησης για τα γειτονικά εικονοστοιχεία των p_i και p_k αντίστοιχα. Εφόσον το p_k βρίσκεται στην περιοχή του p_i και το p_r στη περιοχή του p_k η προτεινόμενη μέτρηση καμπυλότητας στο p_i εκτελείται μέσω τοπικών υπολογισμών σε μια ευρύτερη έκταση που ορίζεται από τις δύο περιοχές γύρω από το σημείο p_i .



Σχήμα 3: Δύο σχήματα και οι αντίστοιχες εικόνες μετά την κωδικοποίηση της επίπεδης καμπυλότητας στις τιμές της έντασης. Τα εικονοστοιχεία που βρίσκονται σε ακμές με υψηλή καμπυλότητα έχουν μεγαλύτερη ένταση (φωτεινότερα) στις προκύπτουσες εικόνες.

Μια γραφική απεικόνιση δείχνει το Σχ. 4. Οι πράξεις των στοιχειωδών νευρώνων στα

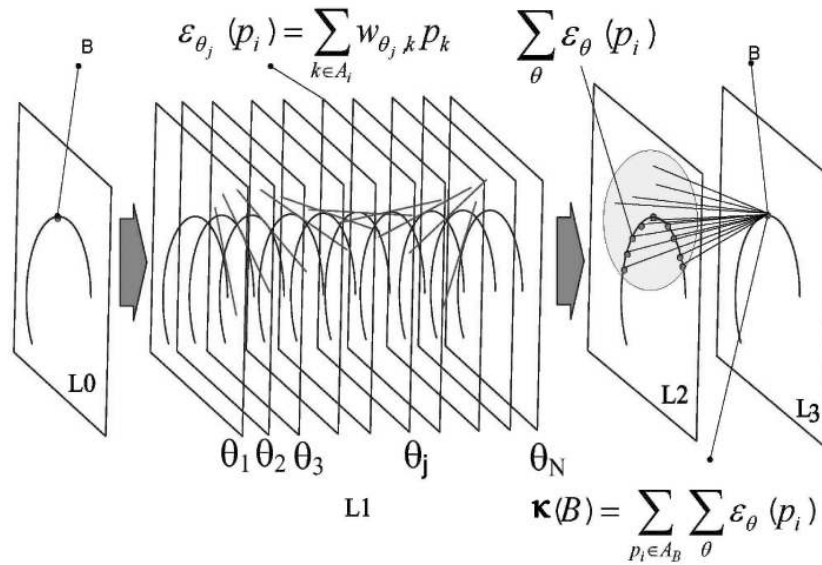
διάφορα επίπεδα L_θ είναι ανεξάρτητες και επομένως εκτελούνται παράλληλα. Έτσι η συνολική καθυστέρηση του υπολογισμού της καμπυλότητας με την προτεινόμενη μέθοδο είναι ισοδύναμη με μια πράξη στοιχειώδους νευρώνα ή μια πράξη φιλτραρίσματος σε λογισμικό επεξεργασίας εικόνας. Παρατηρήστε στην Εξ.5 ότι μόνο το εσωτερικό άθροισμα συμπεριλαμβάνει ένα σταθμισμένο άθροισμα και επομένως πράξεις πολλαπλασιασμού. Τα άλλα δύο αθροίσματα μπορούν να εκτελεστούν μέσω παράλληλων απλών αθροίσεων και επομένως δεν συνεισφέρουν στην πολυπλοκότητα της συνολικής πράξης.

Στο Σχ. 2 φαίνεται ένα αποδεκτικό πεδίο επιλεγόμενου προσανατολισμού και οι διαδοχικές περιστροφές του σε δύο περιπτώσεις διαφορετικών καμπυλοτήτων. Στην περίπτωση υψηλής καμπυλότητας το πεδίο προσεγγίζει την κατεύθυνση της εφαπτομένης στο περίγραμμα καλύτερα για αρκετές διαδοχικές περιστροφές. Η καμπυλότητα είναι ανάλογη με το άθροισμα των εξόδων των περιστρεφόμενων αποδεκτικών πεδίων. Στο (c) φαίνεται το αποδεκτικό πεδίο επιλεγόμενου προσανατολισμού ως μια ορθογώνια περιοχή και οι αντίστοιχες έξοδοι ως βέλη για διάφορους προσανατολισμούς από 0 έως 90 μοίρες. Το μήκος του βέλους είναι ανάλογο με την ένταση της εξόδου για τον δεδομένο προσανατολισμό. Στο (a) και (b) απεικονίζεται ο υπολογισμός της καμπυλότητας για δύο σημεία A και B. Και στις δύο περιπτώσεις το αποδεκτικό πεδίο επιλεγόμενου προσανατολισμού περιστρέφεται από 0 έως 90 μοίρες με βήμα 10 μοιρών και η έξοδος για κάθε βήμα περιστροφής είναι ανάλογη με τον βαθμό προσέγγισης της καμπύλης στην επιτρεπόμενη, από το αποδεκτικό πεδίο για τη συγκεκριμένη περιστροφή, κατεύθυνση. Οι έξοδοι στην περίπτωση (b) έχουν μεγαλύτερη ένταση από τις εξόδους στην περίπτωση (a) για τις περισσότερες περιστροφές, αφού η αλλαγή στον προσανατολισμό του αποδεκτικού πεδίου προσεγγίζει την αλλαγή κατεύθυνσης της εφαπτομένης της καμπύλης στην περιοχή του B καλύτερα. Συγκεκριμένα έχει δειχθεί ότι περιστρέφοντας ένα διδιάστατο αποδεκτικό πεδίο επιλεγόμενου προσανατολισμού επιτυγχάνεται η μέτρηση της καμπυλότητας σε κάθε ένα από τα σημεία μιας επίπεδης καμπύλης χωρίς να χρειάζεται η εξαγωγή του περιγράμματος. Στην επόμενη ενότητα παρουσιάζεται μια πολυστρωματική νευρωνική αρχιτεκτονική αποδεκτικών πεδίων για τον υπολογισμό της καμπυλότητας, η οποία κωδικοποιεί την καμπυλότητα του περιγράμματος στις τιμές έντασης φωτεινότητας των ασπρόμαυρων εικόνων.

3.3. Κωδικοποίηση της Επίπεδης Καμπυλότητας στις Τιμές Έντασης Φωτεινότητας

Προτείνεται μια στρωματική αρχιτεκτονική ως μοντέλο του οπτικού συστήματος για την κωδικοποίηση των μετρήσεων της τοπικής καμπυλότητας στις τιμές έντασης φωτεινότητας των ασπρόμαυρων εικόνων, βασισμένη σε διαδοχικές περιστροφές ενός αποδεκτικού πεδίου επιλεγόμενου προσανατολισμού, όπως παρουσιάστηκε στην προηγούμενη ενότητα. Το κατώτερο στρώμα της προτεινόμενης αρχιτεκτονικής είναι το L0 και αποτελείται από μο-

νάδες αντίληψης που καλούνται νευρώνες-L0 και έχουν το ρόλο των φωτοευαίσθητων κυττάρων του αμφιβληστροειδούς χιτώνα. Αυτοί οι νευρώνες δέχονται ως είσοδο την ένταση της εικόνας και ενεργοποιούνται με μεγάλες αλλαγές στην ένταση. Χρησιμοποιούνται αυτοί οι νευρώνες-L0 για την ανίχνευση του περιγράμματος του σχήματος του παρουσιάζεται. Οι νευρώνες-L0 τροφοδοτούν τις εισόδους των νευρώνων-L1 στο στρώμα L1. Το στρώμα-L1 αποτελείται από ένα σύνολο υποστρώματων. Κάθε νευρώνας-L0 στέλνει την έξοδό του στον αντίστοιχο νευρώνα-L1 σε κάθε υπόστρωμα-L1. Μπορεί να φανταστεί κανείς όλα τα στρώματα διαβαθμισμένα και ευθυγραμμισμένα και τις συνδέσεις να υπάρχουν μόνο μεταξύ νευρώνων που αντιστοιχούν στις ίδιες συντεταγμένες στα αντίστοιχα στρώματά τους. Μια εικόνα που περιέχει ένα δεδομένο τυχαίο σχήμα λαμβάνεται μέσω της ευαισθησίας των νευρώνων-L0 στις διαφορές έντασης φωτεινότητας της εικόνας. Όταν ένας L0 νευρώνας λαμβάνει ένα κατάλληλο ερέθισμα, στέλνει μια ενεργή έξοδο σε όλους τους αντίστοιχους διαβαθμισμένους νευρώνες στα υποστρώματα του στρώματος L1. Με αυτό τον τρόπο όταν παρουσιάζεται μια εικόνα στο στρώμα L0, όλα τα L1 υποστρώματα δέχονται μια ενεργή είσοδο από τους L0 νευρώνες που τυχαίνει να βρίσκονται στο όριο του περιγράμματος του σχήματος. Στην προτεινόμενη αρχιτεκτονική τα L1-υποστρώματα ενεργούν ως κατευθυνόμενοι ανιχνευτές ακμών. Σε αυτό το στρώμα, ανιχνεύονται οι ακμές του περιγράμματος για δεδομένο προσανατολισμό. Κάθε L1-υπόστρωμα αποτελείται από νευρώνες με το απαραίτητο αποδεκτικό πεδίο για ανίχνευση ακμών σε δεδομένο προσανατολισμό. Η προτίμηση προσανατολισμού του αποδεκτικού πεδίου των νευρώνων μεταβάλλεται ομαλά από κάθε L1-υπόστρωμα στο επόμενο και καλύπτει μια πλήρη σειρά γωνιών περιστροφής τουλάχιστον 360 μοιρών. Τα L1-υποστρώματα είναι ανάλογα με τις υπερ-στήλες του ανθρώπινου οπτικού φλοιού [246]. Οι L1-νευρώνες ενεργοποιούνται όποτε υπάρχει μια έντονη ακμή στην συγκεκριμένη κατεύθυνση του αποδεκτικού τους πεδίου. Το απαραίτητο για την ενεργοποίηση αποδεκτικό πεδίο υπό αυτές τις συνθήκες μπορεί να κατασκευαστεί συνδυάζοντας μη-προσανατολισμένα αποδεκτικά πεδία με κυκλική συμμετρία [249]. Επίσης τα τερματικά αποδεκτικά πεδία που αποτελούνται από αποδεκτικά πεδία επιλεγόμενου προσανατολισμού [232] μπορούν να χρησιμοποιηθούν για το σκοπό αυτό ώστε να εντοπιστούν μεγάλες ακμές με χαμηλή τοπική καμπυλότητα. Το υπολογιστικό ισοδύναμο είναι ένα δισδιάστατο φίλτρο με προσαρμοσμένα βάρη ώστε να εντοπίζει ακμές σε διάφορες κατευθύνσεις. Το ίδιο σημείο στο όριο του παρουσιαζόμενου αντικειμένου θα προκαλέσει ενεργοποιήσεις μόνο σε κάποια L1-υποστρώματα, ανάλογα με την περιστροφή των αποδεκτικών πεδίων των L1-νευρώνων του στρώματος, αλλά όσο μεγαλύτερη η καμπυλότητα στο σημείο αυτό τόσα περισσότερα L1-υποστρώματα θα ενεργοποιούνται στην περιοχή του σημείου γιατί θα ανιχνεύονται ισχυρές ακμές σε διάφορες κατευθύνσεις. Το L2-στρώμα είναι ο συλλέκτης όλων των ενεργοποιήσεων που προκλήθηκαν στα υποστρώματα του στρώματος L1 και το στρώμα L3 είναι το στρώμα εξόδου. Κάθε νευρώνας (εικονοστοιχείο) στο στρώμα εξόδου L3 έχει ένα αποδεκτικό πεδίο γύρω από τον αντίστοιχο νευρώνα του



Σχήμα 4: Τα στρώματα για τον υπολογισμό της καμπυλότητας.

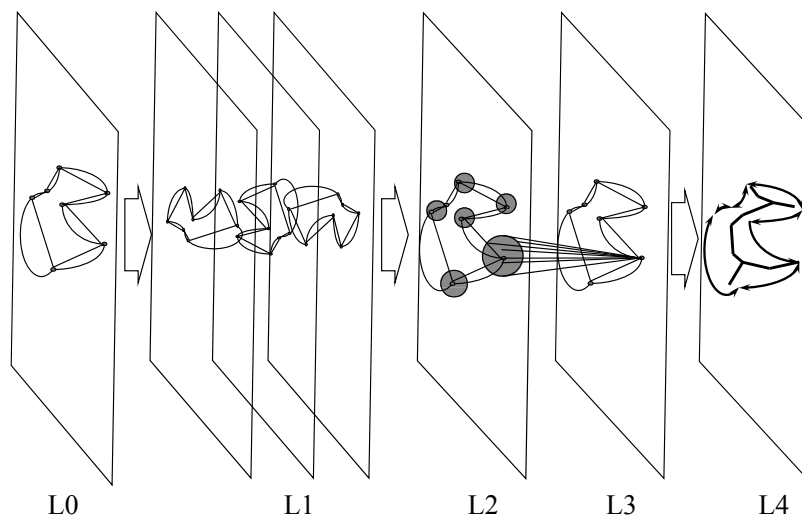
στρώματος L2. Γύρω από ένα σημείο μεγάλης καμπυλότητας θα ανιχνεύονται πολλές ακμές για πολλά διαδοχικά περιστροφικά βήματα και επομένως θα προκαλείται μια σειρά L1-ενεργοποιήσεων σε διαδοχικά L1-υποστρώματα που θα καταλήγουν στην περιοχή του αντίστοιχου L2-νευρώνα. Αυτό με τη σειρά του σημαίνει ότι ο αντίστοιχος L3 (έξοδος) νευρώνας (εικονοστοιχείο), ο οποίος έχει αποδεκτικό πεδίο που περιβάλλει τον συγκεκριμένο L2-νευρώνα, θα αποκτήσει μια τιμή ανάλογη με το άθροισμα των τιμών των νευρώνων στην περιοχή του αντίστοιχου L2-νευρώνα, το οποίο άθροισμα είναι ανάλογο με το μέγεθος της καμπυλότητας στο σημείο αυτό. Από τη στιγμή που για κάθε σημείο περιγράμματος, τουλάχιστον για μια κατεύθυνση, θα ανιχνεύεται μια ακμή, τα στρώματα L2 και L3 θα λαμβάνουν ένα πλήρες αντίγραφο του περιγράμματος, αλλά με διαφορετικά μεγέθη εισόδου για κάθε σημείο ανάλογα με την αντίστοιχη καμπυλότητα. Επομένως, οι διαφορετικές τοπικές καμπυλότητες του περιγράμματος έχουν μετασχηματιστεί σε ισοδύναμες διαφορές στην διέγερση των L3-νευρώνων, χωρίς να χαθεί η χωρική αντιστοιχία των σχετικών L0-νευρώνων. Ένα σημείο περιγράμματος του παρουσιαζόμενου αντικειμένου με υψηλή καμπυλότητα θα προκαλέσει ενεργοποιήσεις σε αρκετά διαδοχικά L1-υποστρώματα που θα καταλήγουν στην περιοχή του σημείου και ο σχετικός L3-νευρώνας (ο οποίος έχει αποδεκτικό πεδίο σε αυτή την περιοχή) θα διεγείρεται περισσότερο σε σύγκριση με τις περιοχές που αποτελούνται από σημεία μικρότερης καμπυλότητας. Η στρωματική αρχιτεκτονική για ανίχνευση καμπυλότητας παρουσιάζεται γραφικά στο Σχ. 4. Η προτεινόμενη αρχιτεκτονική κωδικοποιεί την καμπυλότητα, κάθε σημείου του περιγράμματος του παρου-

σιαζόμενου αντικειμένου, στις τιμές έντασης φωτεινότητας της εικόνας που το περιέχει. Στο Σχ. 3 απεικονίζονται δύο σχήματα και οι εικόνες που προκύπτουν μετά την κωδικοποίηση της τοπικής καμπυλότητας στις τιμές έντασης φωτεινότητας αυτών. Σε αυτό το παράδειγμα χρησιμοποιήθηκαν μόνο 4 διαφορετικές κατευθύνσεις για την ανίχνευση των ακμών. Είναι ξεκάθαρο ότι τα σημεία ακμών με υψηλή καμπυλότητα εμφανίζονται πιο άσπρα (πιο φωτεινά) από τα σημεία ακμών που βρίσκονται σε ευθεία τμήματα του περιγράμματος από τη στιγμή που η μέτρηση της τοπικής καμπυλότητας έχει κωδικοποιηθεί στην τιμή έντασης φωτεινότητας του σημείου. Ο απαιτούμενος χρόνος για την κωδικοποίηση της επίπεδης καμπυλότητας ισούται με τον απαιτούμενο χρόνο για ένα πέρασμα της εικόνας με ένα φίλτρο, αφού η ανίχνευση των ακμών για όλες τις διαφορετικές κατευθύνσεις γίνεται με παράλληλες πράξεις και οι απλές αθροίσεις (σε αντίθεση με τα σταθμισμένα αθροίσματα) θεωρούνται στιγμιαίες. Δεδομένου ότι σχεδόν σε όλες τις εργασίες επεξεργασίας εικόνας εκτελείται τουλάχιστον μια πράξη φιλτραρίσματος για εξομάλυνση, η προτεινόμενη μέθοδος λειτουργεί πρακτικά χωρίς κόστος.

3.4. Η Διαδικασία Εξαγωγής του Σκελετού Καμπυλότητας

Έχει ήδη περιγραφεί στα προηγούμενα ότι οι συνεχείς διαδοχικοί προσανατολισμοί ενός δεκτικού πεδίου επιλεκτικού προσανατολισμού, όπως αυτά που ανακαλύφθηκαν από τους Hubel και Wiesel στον οπτικό φλοιό των πρωτεύοντων, μπορεί να είναι ένας μηχανισμός υπολογισμού καμπυλότητας. Τα κύτταρα ορισμένου εύρους, αφ' ετέρου, χαρακτηρίζονται από την παρουσία ανασταλτικών άκρων της ζώνης (RF) αποδεκτικού πεδίου τους, που «σταματάει» την απόκριση του κυττάρου στα ερεθίσματα που είναι αρκετά εκτεταμένα ώστε να παρεισφρύσουν στις άκρες της ζώνης. Σε αυτό το τμήμα θα περιγράψουμε το μηχανισμό και το πρότυπο συνδεσιμότητας νευρώνων που υπό τις παραπάνω προϋποθέσεις κωδικοποιεί το σκελετό της επίπεδης μορφής στις τιμές έντασης μαζί με τις μετρήσεις καμπυλότητας που είναι ήδη διαθέσιμες από το προηγούμενο επίπεδο.

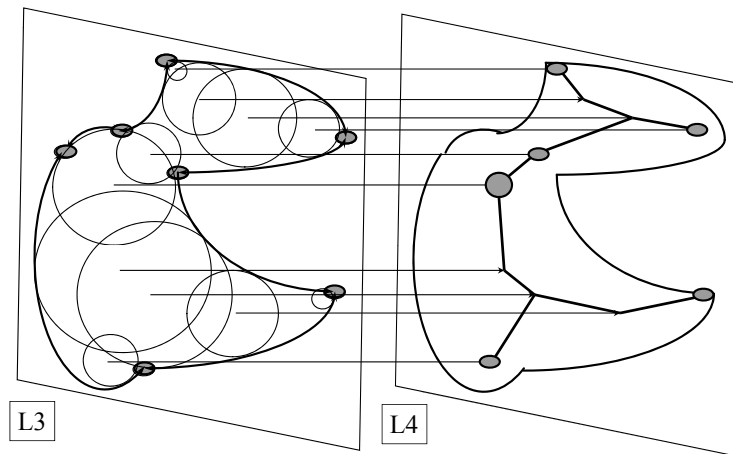
Οι φλοιώδης νευρώνες στα στρώματα 2 και 3 του οπτικού φλοιού έχουν επιλεκτικά προσανατολισμού αποδεκτικά πεδιά και αυτός ο προσανατολισμός αλλάζει κατεύθυνση συνεχώς σε διαδοχικά στρώματα. Με διαδοχικές περιστροφές ενός φίλτρου επιλεκτικού προσανατολισμού μπορούμε πράγματι να μετρήσουμε την καμπυλότητα σε κάθε σημείο μίας καμπύλης, όπως έχει ήδη δειχθεί στα προηγούμενα όπου η επιλεκτικότητα προσανατολισμού χρησιμοποιείται για να εντοπιστεί η κατεύθυνση που είναι εφαπτομένη στην καμπύλη σε ένα συγκεκριμένο σημείο. Η καμπυλότητα σε αυτό το σημείο μετράται μέσω της συσσώρευσης των διεγέρσεων των διαδοχικών στρωμάτων στα οποία ο μεταβαλλόμενος προσανατολισμός παραμένει κοντά στην κατεύθυνση της εφαπτομένης. Αυτή η μέθοδος υπολογισμού καμπυλότητας εφαρμόστηκε στα προηγούμενα από μια στρωματική αρχιτεκτονική που ομοιάζει με την διάταξη των φλοιωδών νευρώνων. Επεκτείνουμε τώρα αυτήν



Σχήμα 5: Τα στρώματα για τον υπολογισμό του σκελετού καμπυλότητας.

την αρχιτεκτονική με την προσθήκη ενός άλλου στρώματος με σκοπό να εξαγάγουμε το σκελετό της προβαλλόμενης μορφής και να ενσωματώσουμε τα προηγούμενα υπολογισμένα σημεία της υψηλής καμπυλότητας στην αντιπροσώπευση σκελετών. Η πλήρης στρωματική αρχιτεκτονική παρουσιάζεται στο Σχ. 5 ενώ ο μηχανισμός για την εξαγωγή σκελετών παρουσιάζεται στο Σχ. 6.

Ο διάμεσος άξονας μιας μορφής είναι το σύνολο όλων των εσωτερικών σημείων που έχουν περισσότερα από ένα εγγύτερα σημεία στο όριο του αντικειμένου. Αρχικά αναφερόμενος ως τοπολογικός σκελετός, εισήχθη από τον Blum ως εργαλείο για τη βιολογική αναγνώριση μορφής. Στις δύο διαστάσεις, ο διάμεσος άξονας μιας επίπεδης καμπύλης S είναι ο γεωμετρικός τόπος των κέντρων των κύκλων που εφάπτονται της S σε δύο ή περισσότερα σημεία και που κείνται στο εσωτερικό της S . Είναι επομένως φυσικό να υποθέσουμε τα κυκλικά - συμμετρικά αποδεκτικά πεδία των απλών νευρώνων αυτού του επιπέδου ως πρότυπο των εγγεγραμμένων κύκλων υποθέτοντας ένα off-center, on-surround τύπο αποδεκτικού πεδίου. Αυτό θα εξασφαλίσει διέγερση του κατάλληλου νευρώνα, που βρίσκεται στο κέντρο του κύκλου, μόνο εάν το περίγραμμα εμπέσει στην περιβάλλουσα ενισχυτική ζώνη και όχι στην κεντρική ανασταλτική ζώνη του κυκλικού αποδεκτικού πεδίου. Έτσι μόνο τα κυκλικά αποδεκτικά πεδία που είναι εγγεγραμμένα στο περίγραμμα θα προκαλέσουν διέγερση και όχι αυτά που κόβουν το περίγραμμα σε ένα ή περισσότερα σημεία. Αυτό γιατί στην τελευταία περίπτωση το περίγραμμα θα εισαχθεί στην ανασταλτική ζώνη του αποδεκτικού πεδίου. Επιπλέον, η διέγερση θα είναι ανάλογη με την υποκίνηση στην ενισχυτική περιβάλλουσα ζώνη (περίμετρος του δίσκου) με τέτοιο

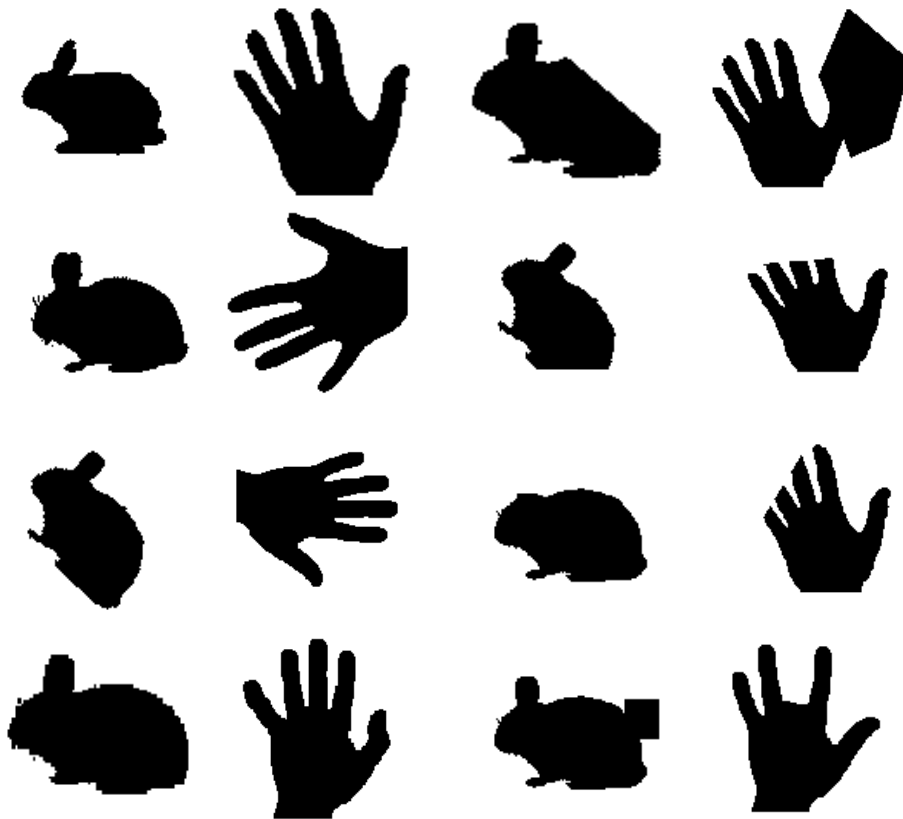


Σχήμα 6: Ο υπολογισμός του σκελετού καμπυλότητας από κυκλικά αποδεκτικά πεδία διαφόρων διαμέτρων.

τρόπο ώστε εάν ο δίσκος είναι εφαπτομένος στο περίγραμμα σε περισσότερα του ενός σημεία, η διέγερση του αντίστοιχου νευρώνα στο κέντρο του δίσκου, θα είναι ισχυρότερη. Δεδομένου ότι οι δίσκοι που είναι εφαπτόμενοι στο περίγραμμα σε περισσότερα του ενός σημεία είναι κεντροθετημένοι στο σκελετό της μορφής, οι αντίστοιχοι νευρώνες (που έχουν αποδεκτικά πεδία σε αυτούς τους δίσκους) στο επόμενο στρώμα, θα λάβουν ισχυρότερη διέγερση.

Σε αυτό το σημείο ανακαλούμε ότι η εικόνα περιγράμματος που εμφανίζεται στο στρώμα $L3$ (Σχ. 6) λαμβάνει ισχυρότερες εισαγωγές στα σημεία υψηλής καμπυλότητας. Αυτά τα σημεία απεικονίζονται στο $L3$ με τα μεγάλα σημεία στο περίγραμμα. Μια συνέπεια αυτού είναι ότι ένας κύκλος που είναι εφαπτόμενος σε σημεία υψηλής καμπυλότητας θα λάβει μια ισχυρότερη διέγερση με τέτοιο τρόπο ώστε ότι ο αντίστοιχος νευρώνας $L4$, που έχει αποδεκτικό πεδίο σε αυτόν τον κύκλο, θα λάβει ισχυρότερη διέγερση επίσης. Αλλά αυτός ο νευρώνας βρίσκεται στο κέντρο του κυκλικού αποδεκτικού πεδίου και επομένως το ισχυρό σήμα λόγω της υψηλής καμπυλότητας που ανιχνεύεται στο στρώμα $L3$ θα εμφανιστεί στο σκελετό της μορφής στο στρώμα $L4$. Αυτά τα σημεία απεικονίζονται στο $L4$ με τα μεγάλα σημεία στον σκελετό. Με αυτόν τον τρόπο, στο στρώμα $L4$, κωδικοποιούμε όχι μόνο την γενική τοπολογία μορφής (σκελετός) αλλά και τις τοπικές πληροφορίες του περιγράμματος (καμπυλότητα) του στρώματος $L3$ στον ίδιο τον σκελετό. Ο προτεινόμενος μηχανισμός *σκελετού καμπυλότητας* μιας μορφής είναι σύμφωνος με τα νευροανατομικά ευρήματα του οπτικού φλοιού, όπως εξηγήθηκε πριν σχετικά με τους τερματικούς, επιλεκτικού προσανατολισμού νευρώνες, αλλά και τα κυκλικά off-center, on-surround αποδεκ-

τικά πεδία διαφόρων μεγεθών. Συγχρόνως το προτεινόμενο πρότυπο αφαίρεσης σχήματος είναι σύμφωνο με τις επιτυχέστερες προσπάθειες στην αντιπροσώπευση και την αναγνώριση μορφής όπως εξηγήθηκε στο σχετικό τμήμα της διατριβής.



Σχήμα 7: Δείγμα εκπαίδευσης και δείγμα ελέγχου.

3.5. Πειραματικά Αποτελέσματα

Για να αξιολογήσουμε την προτεινόμενη αντιπροσώπευση εξετάσαμε την δυνατότητα ενός νευρονικού δικτύου που εκπαιδεύθηκε από ένα αποδοτικό σχήμα οπισθοδιάδοσης [250] να γενικεύσει πέρα από ένα σύνολο πεντακοσίων μορφών όταν υπάρχει επικάλυψη και παραμόρφωση, όπως αυτές που παρουσιάζονται στο Σχ. 7. Οι επικαλύψεις και οι παραμορφώσεις αλλάζουν σημαντικά το σήμα και επομένως οι παραδοσιακοί ταξινομητές δεν αποδίδουν ικανοποιητικά σε αυτές τις συνθήκες.

Για το πείραμα χρησιμοποιήσαμε ασπρόμαυρες εικόνες, χεριών και κουνελιών από τη βάση σχημάτων του KIMIA και ορίσαμε αυτές τις δύο κατηγορίες. 500 τέτοιες εικόνες των δύο κατηγοριών, όπως αυτές του Σχ. 7, κωδικοποιήθηκαν αρχικά με την προτεινόμενη αντιπροσώπευση σκελετού καμπυλότητας και χρησιμοποιήθηκαν για την εκπαίδευση ενός νευρονικού δικτύου δύο ενδιάμεσων στρωμάτων με αριθμό νευρώνων στο πρώτο στρώμα μεταξύ 2 και 8 και στο δεύτερο στρώμα από 1 έως 3. Το δίκτυο ήταν σε θέση να μάθει τη

σωστή διαδική ταξινόμηση για όλες τις παρουσιαζόμενες 500 εικόνες. Εξετάσαμε στην συνέχεια την δυνατότητα γενίκευσης αυτού του δικτύου με την παρουσίαση 50 νέων εικόνων όπως αυτές στο Σχ. 7 οι οποίες είναι εκδόσεις των αρχικών 500 εικόνων μετά από μερική επικάλυψη, παραμόρφωση και απώλειας μερών. Το σύνολο δοκιμής κωδικοποιήθηκε επίσης με τη βοήθεια της αντιπροσώπευσης σκελετού καμπυλότητας πριν παρουσιαστεί στον νευρωνικό ταξινομητή. Η σωστή αναγνώριση επιτεύχθηκε σε 91.6 % αυτών των περιπτώσεων που καθιστά τη μέθοδο εφαρμόσιμη σε πραγματικά σενάρια. Στο Σχ. 8 οι εικόνες των σκελετών καμπυλότητας που χρησίμευσαν ως εισαγωγές στον ταξινομητή παρουσιάζονται για διάφορες εικόνες δοκιμής και εκπαίδευσης. Οι στήλες 1 και 3 παρουσιάζουν τις εικόνες που παραλαμβάνονται από το στρώμα 3 (με την κωδικοποιημένη καμπυλότητα στις τιμές γκριζας κλίμακας), οι στήλες 2 και 4 παρουσιάζουν εικόνες σκελετών καμπυλότητας που παράγονται στο στρώμα 4 και χρησιμοποιούνται ως εισαγωγή στον ταξινομητή.

Ο ίδιος νευρωνικός ταξινομητής εκπαιδεύτηκε επίσης χρησιμοποιώντας τις ίδιες 500 εικόνες αλλά χωρίς κωδικοποίηση σκελετού καμπυλότητας στις τιμές του γκρι. Το δίκτυο ήταν σε θέση πάλι να μάθει όλες τις 500 εικόνες. Στην συνέχεια όμως όταν εξετάσαμε την ικανότητα γενίκευσης με την παρουσίαση του ίδιου συνόλου δοκιμής των 50 εικόνων αλλά χωρίς κωδικοποίηση σκελετού καμπυλότητας, η αναγνώριση σε αυτό το πείραμα ήταν σωστή σε 61.5 % αυτών των περιπτώσεων που είναι χαμηλό έναντι του 91.6 % που επιτυγχάνεται με την κωδικοποίηση του σκελετού καμπυλότητας.

Κατά συνέπεια μπορούμε να καταλήξουμε στο συμπέρασμα ότι η κωδικοποίηση του περιεχομένου της εικόνας μέσω ολικών (σκελετός) και τοπικών περιγραφέων (καμπυλότητα) στην ενέργεια του ίδιου του σήματος (κλίμακες έντασης του γκρι) βελτιώνει την δυνατότητα αναγνώρισης των ταξινομητών νευρωνικών δικτύων σε συνθήκες επικάλυψης και παραμόρφωσης.

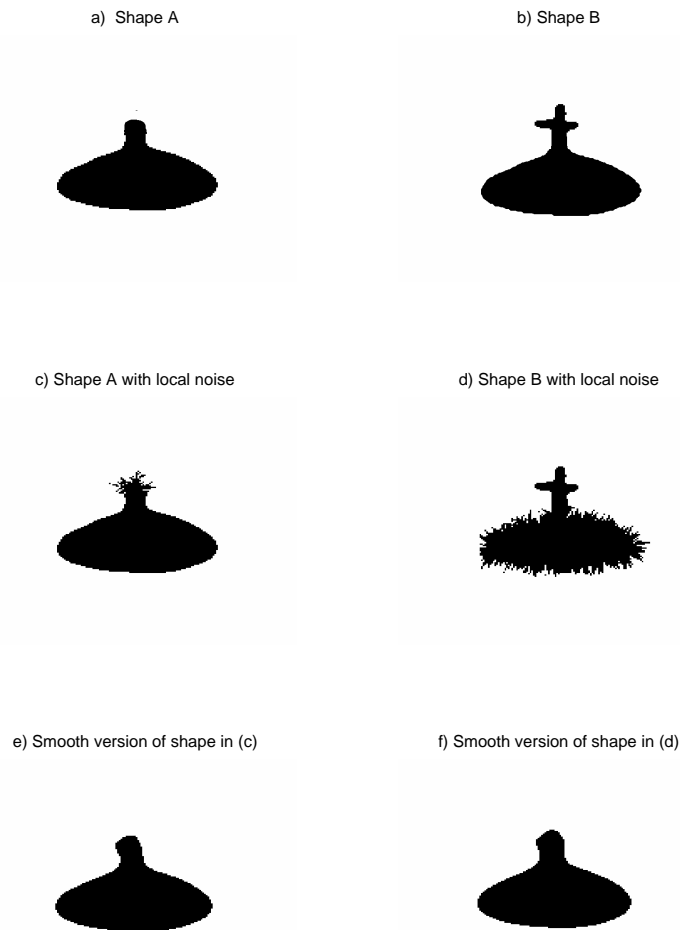


Σχήμα 8: Στις περιττές στήλες (πρώτη και τρίτη) βλέπουμε τα σχήματα του επιπέδου L3 με την καμπυλότητα του περιγράμματος να έχει κωδικοποιηθεί στην φωτεινότητα. Στις άρτιες στήλες βλέπουμε τα σχήματα του επιπέδου L4 όπως εμφανίζονται με σκελετό καμπυλότητας.

4. Ο ΟΛΟΤΟΠΙΚΟΣ ΜΕΤΑΣΧΗΜΑΤΙΣΜΟΣ

Η εξομάλυνση επάνω στο περίγραμμα ενός θορυβώδους σχήματος είναι μία μέθοδος που χρησιμοποιείται ευρέως για την εξαγωγή τοπικών ιδιοτήτων του σχήματος που είναι απαραίτητες για την επακόλουθη διάκριση και αντιστοίχιση. Ωστόσο, η χρήση φίλτρων εξομάλυνσης ενδέχεται να οδηγήσει σε παραμορφώσεις στο σχήμα που επηρεάζουν τόσο το σχήμα όσο και τη θέση των περιγραμμάτων του αντικειμένου. Αυτό είναι ένα γνωστό πρόβλημα σε πολλά πεδία της μηχανικής όρασης, π.χ. στην εμφάνιση τρισδιάστατων όψεων σχήματος από δεδομένα εικόνας ή την εκτίμηση του προσανατολισμού της επιφάνειας σε τεχνικές σχήματος από υφή. Στην ιατρική απεικόνιση, τα σχετικά μετρικά στοιχεία του αντικειμένου και οι σχηματισμοί υψηλής συχνότητας πρέπει να παραμένουν άθικτοι. Ο προσδιορισμός τοπικών περιγραφέων του σχήματος, όπως η καμπυλότητα, σε ένα τέτοιο συγκεκριμένο είναι δύσκολο έργο εφόσον το περίγραμμα είναι γενικά θορυβώδες και δεν είναι επιθυμητός ο υπολογισμός της μέσης τιμής των σημείων επάνω στο περίγραμμα, που τροποποιεί τα μετρικά στοιχεία του σχήματος. Στην αντιστοίχιση σχημάτων, η εξομάλυνση του περιγράμματος του σχήματος μπορεί να κάνει διαφορετικά σχήματα να γίνουν πανομοιότυπα, όπως στο Σχ. 9. Σε αυτή την περίπτωση, βλέπουμε ότι η απομάκρυνση του θορύβου μέσω της εξομάλυνσης του σχήματος A και του σχήματος B οδηγεί σε απώλεια κρίσιμων πληροφοριών σχήματος (τα αποτελέσματα στο (γ) και το (δ) είναι πανομοιότυπα).

Ο μετασχηματισμός GLT, μία νέα τεχνική αναπαράστασης που παρουσιάζεται σε αυτή τη μελέτη, αντιμετωπίζει αυτό το ζήτημα ορίζοντας ένα εγγενές πλαίσιο εργασίας, ανθεκτικό στο θόρυβο, για την ταυτόχρονη κωδικοποίηση τοπικών και ολικών ιδιοτήτων σχήματος. Καθιστά δυνατή τη συγκράτηση των υβριδικών ιδιοτήτων του σχήματος, ακόμα και μετά από υψηλής συχνότητας αναταραχές στο περίγραμμα, αποτρέποντας την απώλεια σημαντικών πληροφοριών σχήματος λόγω των λειτουργιών εξομάλυνσης. Ενώ οι ολικοί περιγραφείς σχήματος είναι σχετικά ανθεκτικοί στον θόρυβο, έχουν περιορισμένη χρησιμότητα όσον αφορά την αλληλοεπικάλυψη και/ή την μη ύπαρξη τμημάτων. Από την άλλη μεριά, οι μέθοδοι που δίνουν έμφαση στις τοπικές ιδιότητες είναι πιο ανθεκτικές στην αλληλοεπικάλυψη αλλά όχι στον θόρυβο και την παραμόρφωση. Ο GLT είναι ένας ολικός περιγραφέας υπό την έννοια ότι ο υπολογισμός του πραγματοποιείται με βάση ολόκληρο το σχήμα, ωστόσο, οι τοπικές ιδιότητες του σχήματος όπως η καμπυλότητα ή άλλα νέα υβριδικά χαρακτηριστικά, όπως ο προσανατολισμός που ορίζεται παρακάτω, καταγράφονται στις παραγώγους του GLT. Αυτός ο υβριδικός χαρακτήρας του GLT είναι συνεπώς μία επιθυμητή ιδιότητα. Μία άλλη επιθυμητή ιδιότητα του GLT είναι η συμβατότητα ανάμεσα στην αναπαράσταση του σχήματος που προσφέρει και την ανθρώπινη ενόραση του σχήματος. Αυτή η ιδιότητα είναι σημαντική γιατί μία διαισθητική αναπαράσταση του σχήματος, επιτρέπει την έξυπνη ανάλυση, την διάδραση και την ερμηνεία από ειδικούς. Η αναπαράσταση Fourier, για παράδειγμα, προσφέρει έναν εξαιρετικό ολικό περιγραφέα σχήματος, αλλά το μετασχηματισμένο σχήμα περιγράφεται μέσω συντελεστών συχνότητας που



Σχήμα 9: Σε αυτό το παράδειγμα απεικονίζεται το ανεπιθύμητο αποτέλεσμα της εξομάλυνσης. Ενώ τα σχήματα A και B ήταν αρχικά διαφορετικά (περιπτώσεις "α" και "β"), με την εξομάλυνση των θορυβωδών εκδοχών τους (γ) και (δ), τα δύο σχήματα γίνονται πανομοιότυπα (περιπτώσεις "ε" και "στ"), εξαλείφοντας τον κρίσιμο σχηματισμό (σταυρός) που χαρακτήριζε το σχήμα B.

δεν διαθέτουν καμία διαισθητική ερμηνεία. Από την άλλη μεριά, οι αναπαραστάσεις που δίνουν έμφαση στις τοπικές ιδιότητες του σχήματος, όπως η εφαπτομένη, η καμπυλότητα κλπ, παρέχουν μία πιο *σημαντική* περιγραφή των τοπικών σημείων αλλά δεν καταφέρνουν να αναπαραστήσουν το σχήμα ως σύνολο. Η μέτρηση της καμπυλότητας μίας καμπύλης σε ένα δεδομένο σημείο, για παράδειγμα, δεν προσφέρει καμία πληροφορία σχετικά με τη σχετική τοποθέτηση αυτού του τμήματος της καμπύλης σε σχέση με την υπόλοιπη καμπύλη.

4.1. Ορισμός

Ο GLT για επίπεδες καμπύλες μπορεί να περιγραφεί ως εμβαπτισμός μετρικών, ανεξάρτητος του συστήματος συντεταγμένων. Από ένα σημείο εκκίνησης στην καμπύλη, δημιουργούμε πρώτα μία *συνάρτηση άποψης*, ως την απόσταση στο επίπεδο έναντι της απόστασης σε μήκος τόξου, ανάμεσα στο σημείο εκκίνησης και ένα άλλο σημείο που διατρέχει το περίγραμμα της καμπύλης. Στην συνέχεια επιτρέπουμε και στο σημείο εκκίνησης να διατρέξει ολόκληρη την καμπύλη και με τον συνδυασμό των αντίστοιχων συναρτήσεων απόψεων, παράγεται μία επιφάνεια, η *ολοτοπική επιφάνεια* (GLS στο εξής), που είναι εμβαπτισμένη στον μοναδιαίο κύβο. Ένα παράδειγμα παρουσιάζεται στο Σχ. 10. Μπορεί να γίνει ανακατασκευή της καμπύλης, ανεξάρτητα από την Ευκλείδεια ομάδα μετασχηματισμών (περιστροφή, κλιμάκωση, μετατόπιση, κατοπτρισμός), μέσω μόνον τριών από τις συναρτήσεις άποψης. Το σύνολο όλων των συναρτήσεων άποψης, που παράγουν την GLS, περιέχει όλες τις πιθανές κανονικές αναπαραμετρήσεις της καμπύλης με έναν άκρως εκφραστικό και πλεονάζοντα τρόπο. Οι τοπικές παραμορφώσεις της καμπύλης μεταφράζονται σε μία ολική παραμόρφωση στον χώρο μετασχηματισμού, συνεπώς, η καμπυλότητα της αρχικής καμπύλης μπορεί να καταγραφεί με τη μορφή συνάρτησης ολοκληρώματος που διαθέτει συγκεκριμένη γεωμετρική ερμηνεία στον χώρο μετασχηματισμού. Ας ορίσουμε τώρα αυστηρά αυτές τις έννοιες:

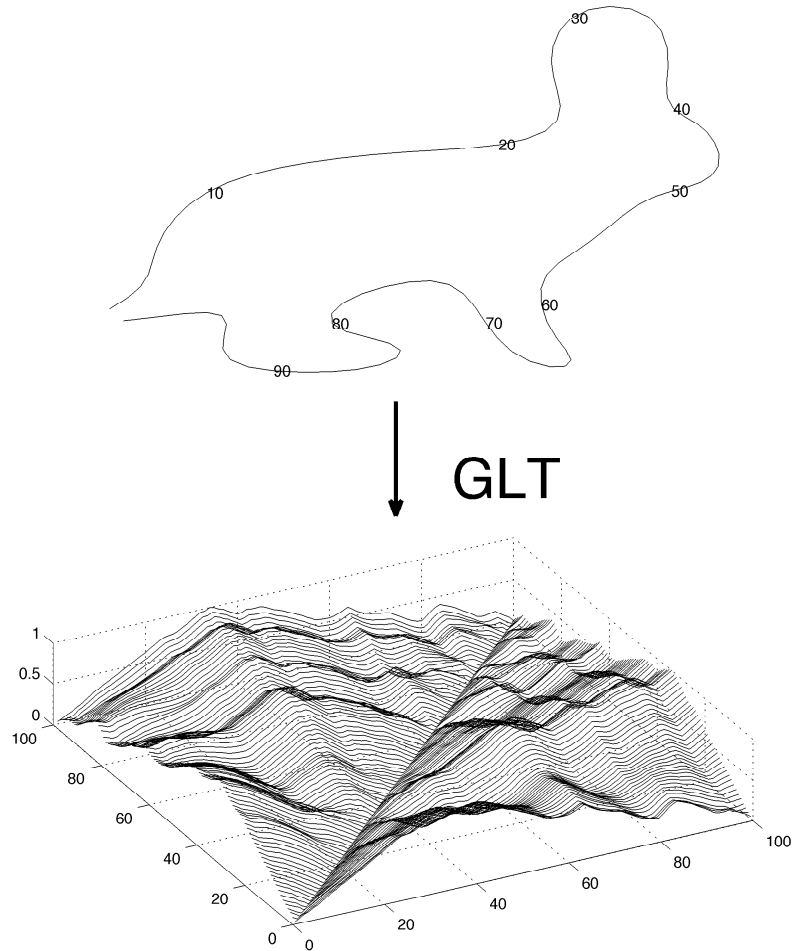
Ας είναι $[0, \lambda] \subset \mathbb{R}$ και $\alpha : [0, \lambda] \rightarrow \mathbb{R}^2$ μία συνεχής και 1-1¹, *επίπεδη καμπύλη μήκους* λ στον \mathbb{R}^2 , παραμετρημένη ως προς το μήκος τόξου s . Στο εξής θεωρούμε τις εξής επιπλέον *συνθήκες κανονικότητας* για την α :

1. $\alpha(0) = \alpha(\lambda)$, έτσι ώστε α είναι μία *κλειστή* καμπύλη στον \mathbb{R}^2 και συνεπώς ισοδύναμη με μία αμφιμονοσήμαντη απεικόνιση από τον μοναδιαίο κύκλο S^1 στο πραγματικό επίπεδο \mathbb{R}^2 .
2. $\alpha(s) \in C^2(S^1, \mathbb{R}^2)$, η κλειστή καμπύλη α είναι τουλάχιστον δύο φορές παραγωγίσιμη² σε κάθε σημείο του μοναδιαίου κύκλου εξασφαλίζοντας ότι η καμπύλη είναι *αρκετά* ομαλή σε όλα της τα σημεία.

¹Ισοδυναμεί με το ότι δεν τέμνει τον εαυτό της.

²Η αιτία πίσω από τον συγκεκριμένο βαθμό παραγωγισιμότητας θα φανεί στην Ενότητα (4.3).

THE GLOBAL-LOCAL TRANSFORMATION



Σχήμα 10: Μία κλειστή επίπεδη καμπύλη και η αντίστοιχη επιφάνεια GLT. Οι τιμές απόστασης επιπέδου κανονικοποιούνται στη διάμετρο της καμπύλης. Η απόσταση στο περίγραμμα μοιράζεται σε 100 σημεία για πιο εύκολη επιθεώρηση της αντιστοίχισης με τα πραγματικά σημεία του περιγράμματος.

Εάν τώρα θεωρήσουμε τον \mathbb{R}^2 εφοδιασμένο με την συνήθη μετρική που προκύπτει από την συνήθη νόρμα, $\|\cdot\|$ και το γράφημα της α σαν $\alpha([0, \lambda])$ (την εικόνα του διαστήματος $[0, \lambda]$ δια μέσου της διανυσματικής συνάρτησης α), τότε η διάμετρος της καμπύλης στον \mathbb{R}^2 είναι:

$$\Delta = \sup_{\mathbf{x}, \mathbf{y} \in \alpha([0, \lambda])} \|\mathbf{x} - \mathbf{y}\| \quad (6)$$

Καθώς η α είναι συνεχής και 1-1, είναι εμφύτευση³ και μεταφέρει την ολική διάταξη του $[0, \lambda] \subset \mathbb{R}$ στο $\alpha([0, \lambda]) \subset \mathbb{R}^2$, και συνεπώς το γράφημα μπορεί να θεωρηθεί εφοδιασμένο και με την συνήθη μονοδιάστατη μετρική που προκύπτει από την απόλυτη τιμή $|\cdot|$ στο ισομορφικό του $[0, \lambda]$ και αυτό είναι το μήκος τόξου της καμπύλης.

Ορισμός 1 *Ας είναι $\mathbf{x}_0 = \alpha(s_0)$ ένα σημείο της α . Την συνάρτηση:*

$$v_{s_0} : [0, \lambda] \rightarrow [0, \Delta], s \mapsto v_{s_0}(s) := \|\mathbf{x}_0 - \alpha(s)\| \quad (7)$$

καλούμε *άποψη της α από το σημείο $\mathbf{x}_0 \equiv \alpha(s_0)$.*

Η συνάρτηση *άποψης* v_{s_0} είναι καλώς ορισμένη για κάποιο σημείο \mathbf{x}_0 , είναι επίσης συνεχής σαν νόρμα στον \mathbb{R}^2 .

Ορισμός 2 Έστω $\alpha : [0, \lambda] \rightarrow \mathbb{R}^2$ μία συνεχής, 1-1 κλειστή επίπεδη καμπύλη μήκους λ ως άνω. Για κάθε $s \in [0, \lambda]$ θεωρούμε την συνάρτηση άποψης $v_s(\xi)$, $s, \xi \in [0, \lambda]$. Η αντίστοιχη οικογένεια συναρτήσεων $v_s(\xi)$, $s \in [0, \lambda]$ παράγει (καθώς το s διατρέχει το πεδίο ορισμού της α) μία επιφάνεια εμβαπτισμένη στον μοναδιαίο κύβο. Ο *Ολοτοπικός Μετασχηματισμός* της καμπύλης α ορίζεται η συνάρτηση $S_\alpha(s, \xi) := v_s(\xi)$, $s, \xi \in [0, \lambda]$. Το γράφημα της S_α είναι η *Ολοτοπική Επιφάνεια (GLS)* της κλειστής καμπύλης α .

4.2. Ιδιότητες του Ολοτοπικού Μετασχηματισμού

Ο παραπάνω ορισμός εισάγει μία κατασκευή επιφάνειας από μία παραμετρική οικογένεια βαθμωτών συναρτήσεων. Αυτή η προσέγγιση έχει πλεονεκτήματα όσον αφορά τη δυνατότητα αναδρομικής επέκτασής της σε περισσότερες διαστάσεις.

Η GLS είναι συμμετρική, έτσι: $S_\alpha(s, \xi) = S_\alpha(\xi, s) = v_\xi(s) = v_s(\xi)$. Ένα σημείο μπορεί να θεωρηθεί εμβαπτισμένο είτε στην GLS, $\{s, \xi, S_\alpha(s, \xi)\}$, είτε σε οποιαδήποτε από τις συναρτήσεις άποψης, $\{s, v_\xi(s)\}$ ή $\{\xi, v_s(\xi)\}$. Και τα δύο χρησιμοποιούνται στα επόμενα. Οι μερικές παράγωγοι της GLS θα θεωρηθούν είτε ως μερικές παράγωγοι είτε ως απλές

³Σαν εμφύτευση, είναι επίσης αντιστρέψιμη ως $\alpha^{-1} : \alpha([0, \lambda]) \rightarrow [0, \lambda] \subset \mathbb{R}$.

παράγωγοι των αντίστοιχων συναρτήσεων άποψης, για παράδειγμα :

$$\begin{aligned}\frac{\partial S_\alpha}{\partial s}(s, \xi) &= \frac{d(v_\xi)}{ds}(s) \\ \frac{\partial S_\alpha}{\partial \xi}(s, \xi) &= \frac{d(v_s)}{d\xi}(\xi)\end{aligned}\tag{8}$$

Εκτός από την παροχή του πλαισίου εργασίας για ενδογενή αναπαράσταση σχήματος, ο GLT έχει την ιδιότητα μετασχηματισμού τοπικών σε ολικών περιγραφών και αντίστροφα, ανάμεσα στον \mathbb{R}^2 και την GLS. Αυτό μπορεί να αποδειχθεί αναφέροντας δύο ιδιότητες :

- Μία τοπική παραμόρφωση του περιγράμματος αντιστοιχεί σε μία ολική παραμόρφωση της GLS. Στην Ενότητα 4.4, εισάγουμε την Αναπαράσταση Εμβαδών Απόψεων VAR, η οποία είναι ένας ολικός περιγραφέας που ορίζεται ως ένα ολοκλήρωμα στον χώρο μετασχηματισμού και κωδικοποιεί την τοπική παραμόρφωση της καμπύλης σε ολική παραμόρφωση του χώρου μετασχηματισμού.
- Τα τοπικά ακρότατα της GLS αντιστοιχούν σε μορφομετρικούς⁴ σχηματισμούς στο περίγραμμα του σχήματος. Περισσότερες πληροφορίες για τον καθορισμό και τη μέτρηση αυτών των σχηματισμών καθώς και κάποια αρχικά αποτελέσματα σε αντιστοιχισι σχημάτων βάσει μορφομετρικών ιδιοτήτων θα δούμε στο κεφάλαιο 4.9.

Στην συνέχεια εστιάζουμε στην αντοχή στο θόρυβο, παρουσιάζοντας σημεία σύνδεσης ανάμεσα στις συναρτήσεις άποψης και την επίπεδη καμπυλότητα.

4.2.1. Καινοτομία του GLT σε Σχέση με Πολυπλοκότητα και Ανθεκτικότητα σε Θόρυβο

Παρόλο που υπάρχει ένας τεράστιος αριθμός μεθόδων για την αναπαράσταση σχημάτων, μόνον μερικές εξετάζουν το ζήτημα της αναπαράστασης τοπικών ιδιοτήτων περιγραμμάτων, όπως η καμπυλότητα, σε θορυβώδες περιβάλλον χωρίς τη καταφύγουν στην εξομάλυνση και ακόμα λιγότερες εξετάζουν την ενσωμάτωση τοπικών και ολικών ιδιοτήτων στο ίδιο θορυβώδες περιβάλλον χωρίς να μετατραπεί το σχήμα μέσω εξομάλυνσης. Όλες οι μέθοδοι που χρησιμοποιούν την εφαιπόμενη κατεύθυνση σε κάθε σημείο του περιγράμματος για τον καθορισμό τοπικών συντεταγμένων ή σχετικών γωνιών όπως η αναλλοίωτη από την περιστροφή εκδοχή του Shape Contexts (Περιβάλλοντος σχήματος) [167] ή της Inner Distance (Εσωτερικής απόστασης) [153] λειτουργούν σε μία ομαλή εκδοχή του αρχικού σχήματος, συνεπώς υπάρχει μία σιωπηρή υπόθεση ότι όλα τα σχήματα είναι θορυβώδη. Το πρόβλημα με αυτή την υπόθεση είναι ότι και τα ομαλά σχήματα φιλτράρονται / τροποποιούνται εφόσον δεν υπάρχει κάποιος εμφανής τρόπος για να γνωρίζει η μέθοδος εάν πρόκειται για

⁴Χρησιμοποιούμε τον όρο *μορφομετρικούς* για να διαφοροποιηθούμε από τον όρο *μορφολογικούς* ο οποίος παραδοσιακά χρησιμοποιείται για μεθόδους που βασίζονται σε πράξεις συνόλων. Με τον όρο αυτό θα αναφερόμαστε σε μεθόδους που χρησιμοποιούν διαφορική γεωμετρία σε μετρικούς χώρους.

Θορυβώδες σχήμα ή όχι, ή την ποσότητα του θορύβου κι αν αυτό είναι αποδεκτό ή μη. Από την άλλη, στα πλαίσια του συνιστώμενου GLT, η υπόθεση είναι ότι όλα τα σχήματα είναι ομαλά. Αυτή η υπόθεση δεν είναι καταστροφική διότι ακόμα κι αν το σχήμα δεν είναι ομαλό, οι τοπικές και ολικές ιδιότητες εξακολουθούν να εξεικονίζονται καταλλήλως. Έχοντας αυτό υπόψη, η συμβολή μας μπορεί να παρουσιαστεί περιληπτικά με τα εξής στοιχεία:

- Παρουσιάζεται ένα πλαίσιο εργασίας (GLT) το οποίο λόγω των έμφυτων ιδιοτήτων του αποτελεί τη βάση για την ανάπτυξη *μη επεμβατικών* περιγραφών σχήματος, που δεν χρειάζεται να αλλοιώνουν το σχήμα με εξομάλυνση ή άλλα φίλτρα. Σε αυτή την εργασία παρουσιάζεται και μελετάται ένας περιγραφέας τέτοιου είδους, η VAR. Γίνεται μαθηματική διερεύνηση και σύνδεση με τις διαισθητικές ιδιότητες του σχήματος. Ορίζεται επίσης ένας άλλος περιγραφέας μέσω του Ολοτοπικού μέτρου. Η σύνδεσή του με διαισθητικές μορφομετρικές ιδιότητες του σχήματος παρουσιάζονται και αναφέρονται αποτελέσματα σε ανάκτηση σχημάτων.
- Χρησιμοποιώντας το πλαίσιο εργασίας του GLT, αποδεικνύονται οι μαθηματικές ιδιότητες του περιγραφέα VAR. Αποδεικνύεται ότι η VAR συνδέεται με σημαντικές ιδιότητες σχήματος με έναν αρκετά διαισθητικό τρόπο. Η μεγάλη ποικιλία χρήσεων του συνολικού πλαισίου εργασίας και η VAR παρουσιάζεται περαιτέρω μέσω μίας πειραματικής έρευνας όπου αποδεικνύεται ότι υπερτερεί μίας σχετικής, στην τρέχουσα στάθμη της τεχνικής, μεθόδου [171].
- Η VAR μπορεί να συμβάλλει στη μείωση της πολυπλοκότητας εκμεταλλευόμενη πληροφορίες του περιβάλλοντος σχήματος. Αποδεικνύεται ότι οι ιδιότητες της αναπαράστασης VAR μπορούν να συμβάλλουν στον περιορισμό σε σχετικές με το σχήμα αντιστοιχίσεις, αντί για τη χρήση εξαντλητικών ή "τυφλών" ελέγχων. Ως αποτέλεσμα, επιτυγχάνεται σημαντικό όφελος στον χρόνο υπολογισμού χωρίς να χάνεται δυνατότητα ανάκτησης.

Μία άλλη συμβολή είναι σε μια νέα μέθοδο αντιστοίχισης η οποία ορίζεται ως ένας συνδυασμός του GLT και της VAR:

- Το πλαίσιο εργασίας του GLT σε συνδυασμό με τον περιγραφέα VAR οδηγεί σε μία νέα μέθοδο αντιστοίχισης για την ταυτοποίηση σχημάτων. Σε σύγκριση με ευρέως χρησιμοποιούμενες σχετικές εναλλακτικές μεθόδους, επιτυγχάνει σημαντικά μικρότερο χρόνο εκτέλεσης για παρόμοιες βαθμολογίες ανάκτησης.

Περαιτέρω, μπορεί να καθοριστούν ανεξάρτητα ενδιαφέρουσες συνέπειες της μαθηματικής επεξεργασίας που τυποποιεί τις έννοιες που εισάγονται σε αυτή την εργασία.

- Η μαθηματική προσέγγιση της καμπυλότητας μέσω μίας διαφορικής οντότητας που βασίζεται στην απόσταση (βλ. Συναρτήσεις άποψης) λειτουργεί ως γέφυρα μεταξύ των τοπικών (διαφορικές) και των ολικών (ολοκληρωτικές) ιδιοτήτων σχήματος.
- Η περιγραφή της επιφάνειας GLT (2Δ αντικείμενο) ως μία συλλογή βαθμωτών συναρτήσεων (Μοναδιάστατα αντικείμενα) και του μετασχηματισμού ως μία κατασκευή αυτής της επιφάνειας από τις συναρτήσεις αυτές παρέχει δυνατότητα αναδρομικής επέκτασης σε υπερ-αντικείμενα περισσότερων διαστάσεων.
- Οι *Ολοτοπικές Εξισώσεις* και ιδιαίτερα η εξίσωση (31), στην οποία αποδεικνύεται μία κλειστή μορφή της σχέσης ανάμεσα στις ολικές και τις τοπικές περιγραφές σχήματος, μπορεί να εκτιμηθεί ανεξάρτητα ως θεωρητικό αποτέλεσμα.

4.3. Τοπικές δεσμεύσεις των συναρτήσεων άποψης

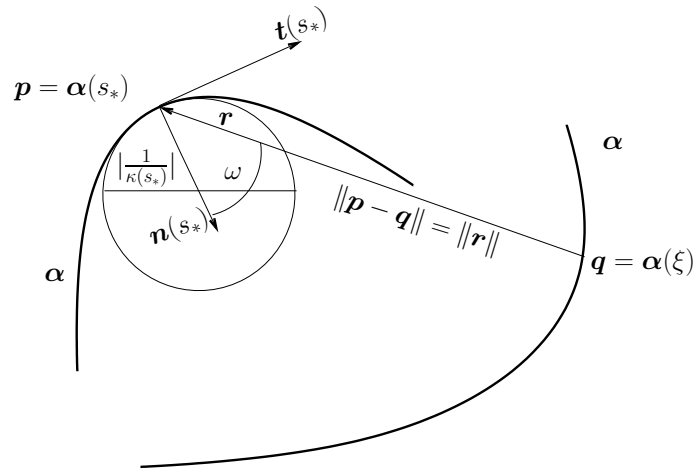
Οι συναρτήσεις άποψης απεικονίζουν το μήκος τόξου στο μήκος χορδής της καμπύλης και συνεπώς είναι 1-Lipschitz, εφόσον ο λόγος χορδής προς τόξο είναι πάντα μικρότερος από 1. Οι συνθήκες Lipschitz είναι σημαντικές διότι επιτρέπουν πολυωνυμική παρεμβολή με αποδεκτό τρόπο. Μέσω της αλγεβρικής ανάλυσης πολυωνύμων, αυτό θα μπορούσε να οδηγήσει σε επέκταση της παρουσιαζόμενης θεωρίας σε άλλα Μαθηματικά πεδία. Για το λόγο αυτό, κάθε φορά που στο εξής συναντάται μία τέτοια συνθήκη θα αναφέρεται ρητώς. Σε αυτή την υπο-ενότητα, θα δειχθεί μία σχέση ανάμεσα στις συναρτήσεις άποψης και την επίπεδη καμπυλότητα του σχήματος.

Έστω $\dot{\alpha}(s_*) = \mathbf{t}$ (η τελεία δηλώνει παράγωγο ως προς το μήκος τόξου s) και \mathbf{n} το μοναδιαίο εφαπτόμενο και το πρώτο κάθετο *εσωστρεφές*⁵ επίσης μοναδιαίο διάνυσμα στο s_* αντιστοίχως με

$$\frac{d\mathbf{t}}{ds} = -\kappa(s_*)\mathbf{n}, \quad \frac{d\mathbf{n}}{ds} = \kappa(s_*)\mathbf{t} \quad (9)$$

οι *τύποι του Frenet* για την καμπύλη α στο $\mathbf{p} = \alpha(s_*)$, με $\kappa(s_*)$ η καμπυλότητα στο \mathbf{p} , όπως φαίνεται στο σχήμα 11. Για την καμπυλότητα $\kappa(s_*)$, έχουμε υιοθετήσει τη συνθήκη πρόσημου σύμφωνα με την οποία, το $\kappa(s)$ είναι θετικό όταν το $\mathbf{n}(s)$ δείχνει *μακριά* από το κέντρο καμπυλότητας. Χρησιμοποιούμε τον συμβολισμό s_* για την τιμή της παραμέτρου όπου εξετάζεται το σύστημα Frenet. Επίσης έστω $\xi \in [0, \lambda], \xi \neq s_*$ μία τυχαία τιμή της μεταβλητής τόξου και το αντίστοιχο σημείο $\mathbf{q} = \alpha(\xi)$. Καλούμε $\mathbf{r}(s_*, \xi)$ το διάνυσμα $\alpha(s_*) - \alpha(\xi)$ και $\omega(s_*, \xi)$ την γωνία από το $\mathbf{n}(s_*)$ στο $-\mathbf{r}(s_*, \xi)$, όπως φαίνεται στο Σχ. 11, η οποία και μετράται αριστερόστροφα. Η γωνία ω είναι μια βαθμωτή συνάρτηση δύο μεταβλητών, s και ξ και το \mathbf{r} είναι μία διανυσματική συνάρτηση των ίδιων δύο μεταβλητών. Και στις δύο περιπτώσεις οι μεταβλητές θα παραλείπονται στο συμβολισμό στο εξής.

⁵Με κατεύθυνση προς το εσωτερικό της κλειστής περιοχής που περικλείεται από την α .



Σχήμα 11: Τοπικές δεσμεύσεις των συναρτήσεων άποψης. Δύο μέρη του ίδιου περιγράμματος, ο εγγύτατος κύκλος, το σύστημα αναφοράς του Frenet και η συνάρτηση άποψης $v_\xi(s_*) = \|\mathbf{r}\|$.

Πρόταση 1 Έστω $\alpha(s) \in C^2([0, \lambda], \mathbb{R}^2)$ μία κλειστή, επίπεδη καμπύλη μήκους λ , παραμετρημένη στο μήκος τόξου, τουλάχιστον δύο φορές διαφορίσιμη σαν συνάρτηση ορισμένη στον μοναδιαίο κύκλο, όπως συζητήθηκε στις υποθέσεις κανονικότητας στην Ενότητα 4.1. Ας είναι $\mathbf{p} = \alpha(s_*)$ ένα σημείο της καμπύλης και $\mathbf{q} = \alpha(\xi)$, $\xi \neq s_*$ ένα άλλο σημείο της καμπύλης με $v_\xi(s)$ την αντίστοιχη συνάρτηση άποψης, Τότε:

$$\frac{d^2 v_\xi}{ds^2}(s_*) = \kappa(s_*) \cos(\omega) \Big|_{s=s_*} + \frac{\cos^2(\omega)}{\|\mathbf{r}\|} \Big|_{s=s_*} \quad (10)$$

Απόδειξη 1 Για διευκρινιστικούς λόγους, η κουκίδα και η διπλή κουκίδα θα υποδηλώνουν την διανυσματική πρώτη και δεύτερη παράγωγο ως προς s , ενώ οι αντίστοιχες βαθμωτές παράγωγοι θα υποδεικνύονται από τα $\frac{d}{ds}$ και $\frac{d^2}{ds^2}$.

Για τη συνάρτηση άποψης $v_\xi(s_*) = \|\alpha(s_*) - \alpha(\xi)\| \equiv \|\mathbf{r}\|$ ισχύει:

$$\frac{d}{ds} v_\xi(s_*) = \frac{d}{ds} \|\mathbf{r}\| = \frac{1}{\|\mathbf{r}\|} \mathbf{r} \cdot \dot{\mathbf{r}} \quad (11)$$

όπου \cdot είναι το εσωτερικό γινόμενο. Θέτοντας $\mathbf{r} = \|\mathbf{r}\|(-\sin(\omega), \cos(\omega))$, $\dot{\mathbf{r}} = (1, 0)$ προκύπτει:

$$\frac{d}{ds} v_\xi(s_*) = -\sin(\omega) \Big|_{s=s_*} \quad (12)$$

που είναι και μία απόδειξη ότι οι συναρτήσεις άποψης είναι 1-Lipschitz. Είναι επίσης:

$$\begin{aligned} \frac{d^2}{ds^2}v_\xi(s_*) &= \frac{d^2}{ds^2}\|\mathbf{r}\| = \left(\frac{1}{\|\mathbf{r}\|}\dot{\mathbf{r}}\right) \cdot \dot{\mathbf{r}} + \frac{\mathbf{r}}{\|\mathbf{r}\|} \cdot \ddot{\mathbf{r}} \\ &= \left[\mathbf{r} \frac{d}{ds}\left(\frac{1}{\|\mathbf{r}\|}\right) + \frac{1}{\|\mathbf{r}\|}\dot{\mathbf{r}}\right] \cdot \dot{\mathbf{r}} + \frac{\mathbf{r}}{\|\mathbf{r}\|} \cdot \ddot{\mathbf{r}} \\ &= -\frac{(\mathbf{r} \cdot \dot{\mathbf{r}})^2}{\|\mathbf{r}\|^3} + \frac{\|\dot{\mathbf{r}}\|^2}{\|\mathbf{r}\|} + \frac{\mathbf{r} \cdot \ddot{\mathbf{r}}}{\|\mathbf{r}\|} \end{aligned} \quad (13)$$

και θέτοντας $\mathbf{r} = \|\mathbf{r}\|(-\sin(\omega), \cos(\omega))$, $\dot{\mathbf{r}} = (1, 0)$, $\ddot{\mathbf{r}} = -\kappa(s_*)(0, -1) = (0, \kappa(s_*))$ προκύπτει:

$$\frac{d^2}{ds^2}v_\xi(s_*) = \kappa(s_*) \cos(\omega) \Big|_{s=s_*} + \frac{\cos^2(\omega)}{\|\mathbf{r}\|} \Big|_{s=s_*} \quad (14)$$

και η απόδειξη είναι πλήρης.

Ελέγχουμε την ορθότητα της παραπάνω πρότασης παρατηρώντας ότι εάν $\omega = 0$ και $\|\mathbf{r}\| = -\frac{1}{\kappa(s_*)}$, που σημαίνει ότι το \mathbf{q} είναι το κέντρο του εγγύτατου κύκλου στο \mathbf{p} , τότε $\frac{d^2v_\xi}{ds^2}(s_*) = 0$, (εφόσον σε αυτήν την περίπτωση $\kappa(s_*) = -\frac{1}{\|\mathbf{r}\|}$) το οποίο είναι αναμενόμενο εφόσον στο s_* η ακτίνα του εγγύτατου κύκλου είναι σταθερή ως προς ds . Το ίδιο ισχύει και στην περίπτωση όπου $\omega = \pi$, $\|\mathbf{r}\| = \frac{1}{\kappa(s_*)}$, $\kappa(s_*) = \frac{1}{\|\mathbf{r}\|}$.

Αυτή η πρόταση συνδέει την επίπεδη καμπυλότητα της α στο \mathbf{p} με τη δεύτερη παράγωγο (κυρτότητα) κάποιας συνάρτησης άποψης στο $s_* = \alpha^{-1}(\mathbf{p})$ και παρατηρούμε ότι, δεδομένου του \mathbf{p} , το αποτέλεσμα ισχύει για οποιαδήποτε συνάρτηση άποψης, εφόσον, για κάθε s_* , η επιλογή του ξ είναι τυχαία. Αυτή η τελευταία παρατήρηση θα επιτρέψει στην Ενότητα 4.5, μία γενίκευση που θα παρέχει ανεξαρτησία από τις συναρτήσεις άποψης.

4.4. Η Αναπαράσταση Εμβαδών Απόψεων (VAR)

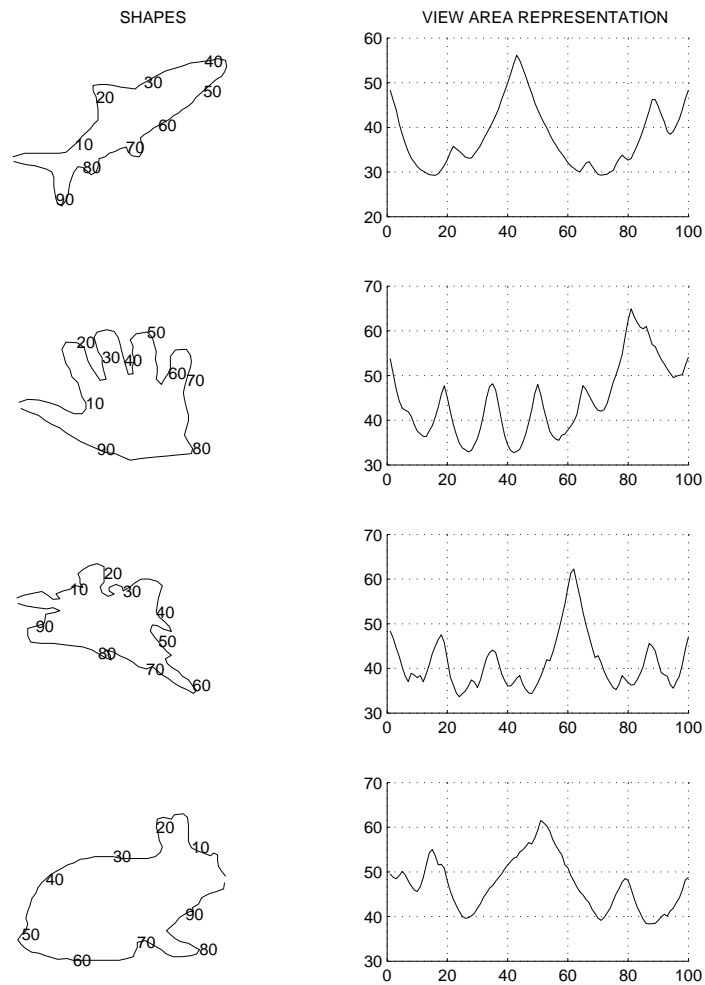
Ως συνήθως, ας είναι α μια κλειστή καμπύλη και $\mathbf{x} = \alpha(s)$ ένα σημείο της καμπύλης σε απόσταση δεξιόστροφου μήκους τόξου s από την αρχή της καμπύλης. Εδώ η Ολοτοπική Επιφάνεια GLS θεωρείται ως μία συλλογή των συναρτήσεων απόψεων που την παράγουν, εφοδιασμένη με τη νόρμα $\|\cdot\|_v$ ως:

$$\|\cdot\|_v : S_\alpha \rightarrow \mathbb{R}, v_s \mapsto \|v_s\|_v := \int_0^\lambda v_s(\xi) d\xi \quad (15)$$

Πρόκειται για την νόρμα μέτρησης των εμβαδών που περικλείουν οι συναρτήσεις απόψεων.

Ορισμός 3 Έστω ϕ_α μία συνάρτηση ορισμένη στο $[0, \lambda]$ με τιμές στο \mathbb{R} ως εξής:

$$\phi_\alpha : [0, \lambda] \rightarrow \mathbb{R} : s \mapsto \phi_\alpha(s) := \|v_s\|_v \quad (16)$$



Σχήμα 12: Σιλουέτες από την συλλογή του KIMIA και οι αντίστοιχες αναπαραστάσεις εμβαδών απόψεων.

Καλούμε την ϕ_α Αναπαράσταση Εμβαδών Απόψεων η σύντομα VAR της καμπύλης α .

Η συνάρτηση ϕ_α είναι παραμετρημένη με το μήκος τόξου και απεικονίζει ένα σημείο $\alpha(s)$ (μέσω της παραμέτρου μήκους τόξου s), στο εμβαδόν $\phi_\alpha(s)$ που περικλείεται από τη συνάρτηση άποψης v_s . Παρακάτω, αποδεικνύουμε ότι παρότι η VAR είναι ένας ολικός περιγραφέας, κωδικοποιεί την *σχετική θέση*, την *σχετική στρέψη* και την *επίπεδη καμπυλότητα* με τρόπο ώστε είναι ευαίσθητη στις διακυμάνσεις σχήματος και την ίδια στιγμή ανθεκτική στον θόρυβο.

4.4.1. Ιδιότητες της VAR

Η VAR μπορεί να ερμηνευτεί ως προτυποποίηση μίας έννοιας *ολικής απόστασης* ανάμεσα σε ένα σημείο της καμπύλης $\alpha(s)$ και την υπόλοιπη καμπύλη. Αυτό γίνεται φανερό αν σκεφτεί κανείς το διακριτό ανάλογο, ⁶ στην οποία περίπτωση η τιμή της VAR στο s είναι το άθροισμα των αποστάσεων του σημείου $\alpha(s)$ από όλα τα υπόλοιπα διακριτά σημεία της καμπύλης. Συνεπώς, η VAR αυξάνεται όσο μετακινούμαστε προς *απομακρυσμένα* σημεία της καμπύλης.

Αυτή η παρατήρηση οδηγεί στο συμπέρασμα ότι για κάθε σημείο της καμπύλης, η VAR κωδικοποιεί την θέση του ως προς την υπόλοιπη καμπύλη. Παρακάτω, θα δείξουμε ότι επίσης προτυποποιεί μία έννοια σχετικής στρέψης της καμπύλης, καθώς και την καμπυλότητα. Επιπλέον, θα δείξουμε ότι η VAR υπακούει στους παρακάτω κανόνες οι οποίοι είναι κρίσιμοι για την καταλληλότητά της VAR ως περιγραφέα σχήματος:

1. Μικρές αναταραχές στο σχήμα παράγουν μικρές αλλαγές στον περιγραφέα VAR.
2. Τα σχήματα διαφορετικών κατηγοριών έχουν διαφορετικούς περιγραφείς VAR.

Ο πρώτος κανόνας υποδεικνύει αντοχή στον θόρυβο και εξασφαλίζει ότι παρόμοια σχήματα θα έχουν παρόμοιες VAR (Πρόταση 2). Ο δεύτερος κανόνας εξασφαλίζει ότι η VAR είναι αρκετά επιλεκτική για σχήματα που είναι επαρκώς διαφορετικά ώστε να ανήκουν σε διαφορετικές κατηγορίες. Αυτό παρουσιάζεται στις ενότητες 4.5 και 4.6 παρέχοντας μία βασική σχέση ανάμεσα στην πρώτη και τη δεύτερη παράγωγο της VAR και την σχετική στρέψη της καμπύλης και την τοπική καμπυλότητα αντίστοιχα. Αυτή η σχέση εξασφαλίζει *εσωτερική* αναπαράσταση σχήματος και αντιληπτή προφάνεια μέσω της VAR. Περαιτέρω, η αμεταβλητότητα της VAR ως προς την Ευκλείδεια ομάδα μετασχηματισμών είναι η *εσωτερική* κληρονομιά της VAR από την GLT. Για τον πρώτο κανόνα, σε αυτό το σημείο δείχνουμε ότι η VAR είναι ανθεκτική ως προς τις θορυβώδεις εκδοχές της αρχικής καμπύλης.

⁶Θεωρήστε το γράφημα της καμπύλης σαν ένα πεπερασμένο σύνολο σημείων αντί για ένα συνεχές σύνολο άπειρων σημείων

Πρόταση 2 Έστω $\alpha \in C^2([0, \lambda], \mathbb{R}^2)$ μία κλειστή καμπύλη μήκους $\lambda \neq 0$ η οποία είναι επίσης C^2 σαν συνάρτηση του S^1 . Εάν $\alpha' \in C^2([0, \mu], \mathbb{R}^2)$ μία θορυβώδης εκδοχή της α μήκους $\mu \neq 0$ τότε για κάθε σημείο, $\mathbf{x} = \alpha(s), s \in [0, \lambda]$ της α υπάρχει ένα αντίστοιχο σημείο $\mathbf{x}' = \alpha'(s'), s' \in [0, \mu]$ της α' , τέτοιο ώστε:

$$\left| \frac{\mu}{\lambda} \phi_\alpha(s) - \phi_{\alpha'}(s') \right| \leq \mu \|\mathbf{x} - \mathbf{x}'\| + \frac{\mu}{\lambda} \int_0^\lambda \|\alpha(\xi) - \alpha'(\xi')\| d\xi \quad (17)$$

όπου ξ, ξ' παράμετροι μήκους τόξου των α και α' αντιστοίχως.

Απόδειξη 2 Αρχικά ορίζουμε μία αμφιμονοσήμαντη αντιστοιχία c ανάμεσα στα μήκη τόξων των δύο περιγραμμάτων, τέτοια ώστε: $s' = c(s) = \frac{\mu}{\lambda}s, s \in [0, \lambda], s' \in [0, \mu]$ και $\xi' = c(\xi) = \frac{\mu}{\lambda}\xi, \xi \in [0, \lambda], \xi' \in [0, \mu]$. Βεβαιώνουμε ότι αυτή η αντιστοιχία είναι κατ'ελάχιστον ορισμένη και αμφιμονοσήμαντη, την αντιστροφή της συμβολίζουμε με $c^{-1} : [0, \mu] \rightarrow [0, \lambda]$ και είναι τέτοια ώστε $s = c^{-1}(s') = \frac{\lambda}{\mu}s', s' \in [0, \mu], s \in [0, \lambda]$ και $\xi = c^{-1}(\xi') = \frac{\lambda}{\mu}\xi', \xi' \in [0, \mu], \xi \in [0, \lambda]$. Ας είναι τώρα $\mathbf{x} = \alpha(s)$ και $\mathbf{x}' = \alpha'(c(s)) = \alpha'(s')$ αντίστοιχα σημεία στις δύο καμπύλες. Από την τριγωνική ανισότητα της νόρμας $\|\cdot\|$ έχουμε:

$$\begin{aligned} \phi_\alpha(s) &= \int_0^\lambda v_s(\xi) d\xi = \int_0^\lambda \|\alpha(s) - \alpha(\xi)\| d\xi \leq \\ &\leq \int_0^\lambda \{\|\alpha(s) - \alpha'(s')\| + \|\alpha'(s') - \alpha'(\xi')\| + \\ &+ \|\alpha'(\xi') - \alpha(\xi)\|\} d\xi = \\ &= \lambda \|\mathbf{x} - \mathbf{x}'\| + \frac{\lambda}{\mu} \int_0^{c(\lambda)} \|\alpha'(c(s)) - \alpha'(c(\xi))\| d(c(\xi)) + \\ &+ \int_0^\lambda \|\alpha'(\xi') - \alpha(\xi)\| d\xi = \\ &= \lambda \|\mathbf{x} - \mathbf{x}'\| + \frac{\lambda}{\mu} \phi_{\alpha'}(s') + \int_0^\lambda \|\alpha'(\xi') - \alpha(\xi)\| d\xi \end{aligned} \quad (18)$$

όπου στον δεύτερο όρο του δεξιού μέλους της (18) αλληλόξαμε την μεταβλητή ως προς την οποία υπολογίζεται το ολοκλήρωμα από ξ σε $c(\xi) = \xi'$, και έτσι:

$$\phi_\alpha(s) - \frac{\lambda}{\mu} \phi_{\alpha'}(s') \leq \lambda \|\mathbf{x} - \mathbf{x}'\| + \int_0^\lambda \|\alpha'(\xi') - \alpha(\xi)\| d\xi \quad (19)$$

Ομοίως,

$$\begin{aligned}
\phi_{\alpha'}(s') &= \int_0^\mu v_{s'}(\xi') d\xi' = \int_0^\mu \|\alpha'(s') - \alpha'(\xi')\| d\xi' \leq \\
&\leq \frac{\mu}{\lambda} \int_0^{c^{-1}(\mu)} \|\alpha'(s') - \alpha(s)\| d(c^{-1}(\xi')) + \\
&+ \frac{\mu}{\lambda} \int_0^{c^{-1}(\mu)} \|\alpha(c^{-1}(s')) - \alpha(c^{-1}(\xi'))\| dc^{-1}(\xi') + \\
&+ \frac{\mu}{\lambda} \int_0^{c^{-1}(\mu)} \|\alpha(\xi) - \alpha'(\xi')\| d(c^{-1}(\xi')) = \\
&= \mu \|\mathbf{x} - \mathbf{x}'\| + \frac{\mu}{\lambda} \phi_\alpha(s) + \frac{\mu}{\lambda} \int_0^\lambda \|\alpha'(\xi') - \alpha(\xi)\| d\xi
\end{aligned} \tag{20}$$

όπου σε όλη τους τους όρους του δεξιού μέλους της (20) αλληλάξαμε την μεταβλητή ως προς την οποία υπολογίζονται τα ολοκληρώματα από ξ' σε $\xi = c^{-1}(\xi')$, και έτσι:

$$\phi_{\alpha'}(s') - \frac{\mu}{\lambda} \phi_\alpha(s) \leq \mu \|\mathbf{x} - \mathbf{x}'\| + \frac{\mu}{\lambda} \int_0^\lambda \|\alpha'(\xi') - \alpha(\xi)\| d\xi \tag{21}$$

Πολλαπλασιάζοντας και τα δύο μέλη της (19) με $\frac{\mu}{\lambda} > 0$ έχουμε:

$$\frac{\mu}{\lambda} \phi_\alpha(s) - \phi_{\alpha'}(s') \leq \mu \|\mathbf{x} - \mathbf{x}'\| + \frac{\mu}{\lambda} \int_0^\lambda \|\alpha'(\xi') - \alpha(\xi)\| d\xi \tag{22}$$

Από την (21) και την (22) τελικά έχουμε:

$$\left| \frac{\mu}{\lambda} \phi_\alpha(s) - \phi_{\alpha'}(s') \right| \leq \mu \|\mathbf{x} - \mathbf{x}'\| + \frac{\mu}{\lambda} \int_0^\lambda \|\alpha(\xi) - \alpha'(\xi')\| d\xi \tag{23}$$

και η απόδειξη ολοκληρώθηκε.

Το δεξιό μέλος της ανισότητας (17) περιλαμβάνει δύο όρους. Ο όρος $\mu \|\mathbf{x} - \mathbf{x}'\|$ είναι η τοπική παραμόρφωση λόγω θορύβου στο σημείο \mathbf{x} . Ο δεύτερος όρος $\frac{\mu}{\lambda} \int_0^\lambda \|\alpha(\xi) - \alpha'(\xi')\| d\xi$ είναι η ολική παραμόρφωση ολόκληρης της καμπύλης, λόγω του θορύβου και μπορεί να θεωρηθεί ως μια *ολική* απόσταση ανάμεσα στις δύο καμπύλες (με και χωρίς θόρυβο). Συνεπώς, η τοπική παραμόρφωση της VAR είναι άνω φραγμένη από τις τοπικές και ολικές παραμορφώσεις της καμπύλης, εν αντιθέσει με την καμπυλότητα, όπου μία τοπική παραμόρφωση της καμπύλης μπορεί να οδηγήσει σε τοπική παραμόρφωση χωρίς όριο στη συνάρτηση καμπυλότητας.

4.5. Προσανατολισμός των σημείων της καμπύλης

Για να γενικευτούν τα αποτελέσματα της Πρότασης (1), οι συναρτήσεις άποψης θεωρούνται πλέον εμβαπτισμένες στην επιφάνεια GLT (GLS). Εξ ορισμού, η $\phi_\alpha(s)$ είναι το εμβαδόν

κάτω από τη συνάρτηση άποψης v_s , συνεπώς το s υποδεικνύει τη θέση, στην GLS, της συνάρτησης άποψης. Για τις παρακάτω αποδείξεις, οι συναρτήσεις άποψης θα θεωρούνται εμβαπτισμένες στην GLS. Συνεπώς, χρειάζεται ένας βαθμός ελευθερίας για την επιλογή διαφορετικών συναρτήσεων άποψης από την GLS, έτσι θα χρησιμοποιείται ο συμβολισμός $S_\alpha(s, \xi)$ που σημαίνει $v_s(\xi)$, για να καταστεί σαφές ότι εργαζόμαστε έχοντας υπόψη την GLS. Εφόσον η GLS είναι συμμετρική, οι παράμετροι s και ξ είναι ισοδύναμες. Ως αποτέλεσμα ένα σημείο $S_\alpha(s, \xi)$, στην GLS, μπορεί να θεωρηθεί ότι ανήκει είτε στην επιφάνεια είτε στην συνάρτηση άποψης v_s είτε στην συνάρτηση άποψης v_ξ έτσι είναι: $v_s(\xi) = v_\xi(s) = S_\alpha(s, \xi) = S_\alpha(\xi, s)$. Στο εξής, η κουκίδα υποδεικνύει την παράγωγο ως προς s . Κάθε παράγωγος της $v_\xi(s)$ ή $v_s(\xi)$ και κάθε μερική παράγωγος της $S_\alpha(s, \xi)$ υπολογίζονται υπό την προϋπόθεση ότι $\xi \neq s$.

Έστω $\alpha \in C^2([0, \lambda], \mathbb{R}^2)$ κλειστή, επίπεδη καμπύλη μήκους λ , η οποία είναι τουλάχιστον δύο φορές παραγωγίσιμη σαν συνάρτηση που ορίζεται στον μοναδιαίο κύκλο όπως αναλύθηκε στις υποθέσεις κανονικότητας στην Ενότητα 4.1. Η καμπύλη $\alpha(s)$ είναι παραμετρημένη με το μήκος τόξου s , $\phi_\alpha(s)$ είναι η αναπαράσταση εμβαδών απόψεων και $s_* \in [0, \lambda]$. Τότε:

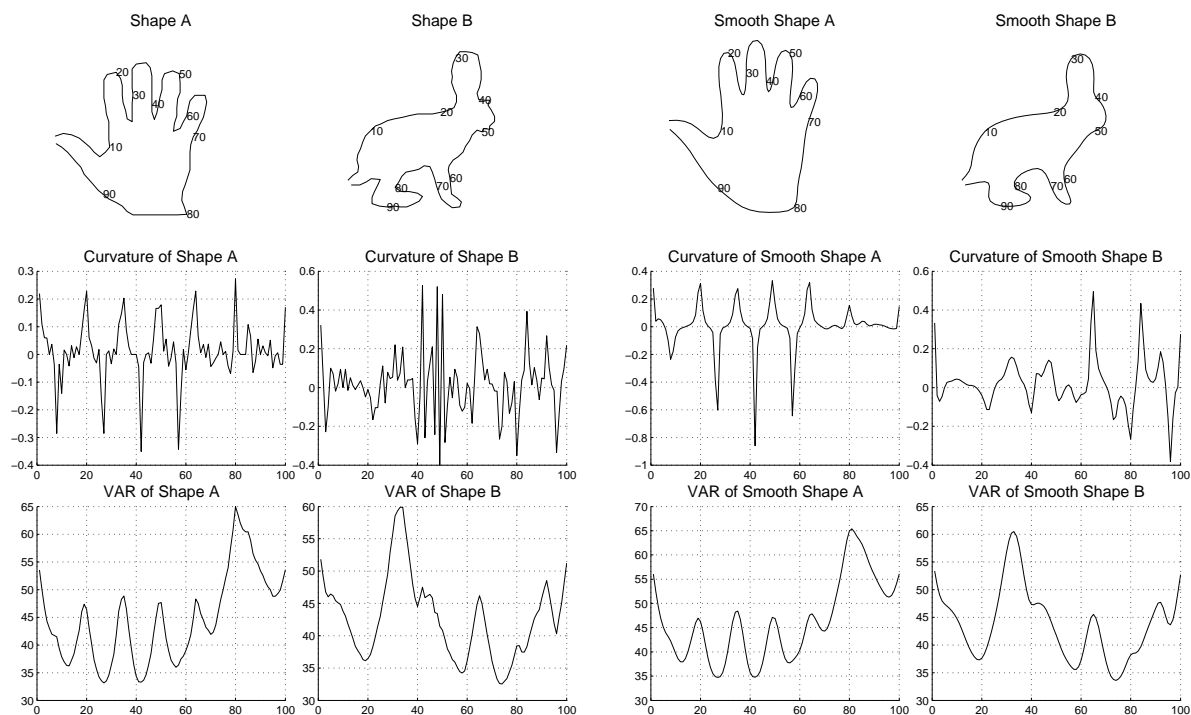
$$\begin{aligned} \dot{\phi}_\alpha(s_*) &= \frac{d}{ds} \int_0^\lambda v_s(\xi) d\xi \Big|_{s=s_*} = \int_0^\lambda \frac{\partial S_\alpha}{\partial s}(s, \xi) \Big|_{s=s_*} d\xi \\ &\stackrel{*}{=} \int_0^\lambda \frac{\partial S_\alpha}{\partial s}(\xi, s) \Big|_{s=s_*} d\xi = \int_0^\lambda \frac{dv_\xi}{ds}(s) d\xi \Big|_{s=s_*} = \int_0^\lambda \dot{v}_\xi(s_*) d\xi \end{aligned} \quad (24)$$

όπου έχουμε χρησιμοποιήσει τον κανόνα ολοκληρώματος Leibnitz. Το βήμα στις πράξεις που επισημαίνεται με * παραπάνω, ισχύει λόγω της συμμετρικής δομής της GLS, όπως έχει εξηγηθεί παραπάνω. Τα ολοκληρώματα έχουν νόημα αφού το σύνολο όπου το $v_\xi(s)$ (ή το $v_s(\xi)$) δεν είναι παραγωγίσιμο έχει μηδενικό μέτρο στο \mathbb{R} . Από την Πρόταση 1, εξάγουμε ότι το αποτέλεσμα του (12) είναι ανεξάρτητο από τη μεταβλητή ξ , για δεδομένη μεταβλητή s_* , όπου θεωρείται το τριεδρο Frenet. Το γεγονός αυτό επιτρέπει έγκυρη ολοκλήρωση ως προς ξ , συνεπώς:

$$\dot{\phi}_\alpha(s_*) = - \int_0^\lambda \sin(\omega) d\xi \Big|_{s=s_*} \quad (25)$$

που είναι και μία απόδειξη ότι η VAR είναι λ -Lipschitz.

Από το Σχ. 11, επαληθεύεται ότι η γωνία ω είναι μία συνάρτηση του ξ (δεδομένου s_*) και ότι το ολοκλήρωμα στο δεξιό μέλος του (25) ποσοτικοποιεί μία έννοια σχετικής θέσης της καμπύλης ως προς την κάθετο στο σημείο s_* . Για να γίνει αυτό φανερό, θεωρήστε ένα σημείο $\alpha(\xi)$ και παρατηρήστε ότι καθώς το ξ διατρέχει το πεδίο ορισμού της α , η γωνία ω μετρά την γωνιακή μετατόπιση αυτού του σημείου ως προς τη κάθετο στο s_* , συνεπώς το ολοκλήρωμα στο δεξιό μέλος της (25) μπορεί να θεωρηθεί ως μέτρηση της συνολικής



Σχήμα 13: Καμπυλότητα και Αναπαράσταση Εμβαδών Απόψεων (VAR) για δύο σχήματα και τις εξομαλυμένες εκδοχές τους. Συγκρίνοντας τις θορυβώδεις με τις εξομαλυμένες εκδοχές αυτών των σχημάτων, παρατηρούμε σε σύγκριση με την καμπυλότητα, την αυξημένη αντοχή της VAR στις υψηλόσυχνες αναταραχές περιγράμματος

γωνιακής μετατόπισης ολόκληρης της καμπύλης ως προς την κάθετο στο s_* . Η πρώτη παράγωγος της VAR στο s θα καλείται *προσανατολισμός ημιτόνου* ή *sine-orientation* του σημείου $\alpha(s)$, από το γεγονός ότι για τον υπολογισμό του, στην Εξ. (25), ολοκληρώνονται τα ημίτονα της γωνίας ω . Στην επόμενη Ενότητα θα δούμε τον ορισμό και τις συνθήκες με τις οποίες προκύπτει και ο *προσανατολισμός συνημιτόνου* (*cosine-orientation*).

4.6. Αντιστοίχιση καμπυλότητας

Σε αυτή την ενότητα, θα αποδείξουμε μία σχέση ανάμεσα στη δεύτερη παράγωγο της VAR και την καμπυλότητα του σχήματος.

Θεώρημα 1 Έστω $\alpha \in C^2([0, \lambda], \mathbb{R}^2)$ κλειστή, επίπεδη καμπύλη μήκους λ τουλάχιστον δύο φορές διαφορίσιμη ως συνάρτηση ορισμένη στον μοναδιαίο κύκλο, όπως έχει συζητηθεί στις συνθήκες κανονικότητας της Ενότητας 4.1. Η καμπύλη $\alpha(s)$ είναι παραμετρημένη με το μήκος τόξου s , $\phi_\alpha(s)$ η συνάρτηση εμβαδών απόψεων και $\kappa(s)$ η συνάρτηση καμπυλότητας. Εάν s_* είναι ένα σημείο τοπικού μεγίστου για την $\phi_\alpha(s)$, τότε $\kappa(s_*) \neq 0$ και

$$\frac{\ddot{\phi}_\alpha(s_*)}{\kappa(s_*)} = A(s_*) + \frac{B(s_*)}{\kappa(s_*)} \quad (26)$$

όπου $A(s_*) = \int_0^\lambda \cos(\omega) d\xi \Big|_{s=s_*}$ και $B(s_*) = \int_0^\lambda \frac{\cos^2(\omega)}{\|\mathbf{r}\|} d\xi \Big|_{s=s_*}$ ολικές ιδιότητες του σχήματος μετρημένες στο $\alpha(s_*)$.

Απόδειξη 3 Από (25) έχουμε:

$$\begin{aligned} \ddot{\phi}_\alpha(s_*) &= -\frac{d}{ds} \left(\int_0^\lambda \sin(\omega) d\xi \right) \Big|_{s=s_*} \\ &= -\int_0^\lambda \frac{\partial}{\partial s} \sin(\omega) d\xi \Big|_{s=s_*} \\ &= -\int_0^\lambda \cos(\omega) \dot{\omega} d\xi \Big|_{s=s_*} \end{aligned} \quad (27)$$

Από τον συνδυασμό των (12) και (14) έχουμε:

$$-\frac{\partial}{\partial s} \sin(\omega) = -\cos(\omega) \dot{\omega} = \cos(\omega) (\kappa(s) + \frac{\cos(\omega)}{\|\mathbf{r}\|}) \quad (28)$$

και αντικαθιστώντας στην (27):

$$\ddot{\phi}_\alpha(s_*) = \kappa(s_*) \left[\int_0^\lambda \cos(\omega) d\xi \Big|_{s=s_*} \right] + \int_0^\lambda \frac{\cos^2(\omega)}{\|\mathbf{r}\|} d\xi \Big|_{s=s_*} \quad (29)$$

Εάν s_* ένα σημείο τοπικού μεγίστου της $\phi_\alpha(s)$, τότε υπάρχει περιοχή U του s_* τέτοια ώστε $\phi_\alpha(s_*) \geq \phi_\alpha(s)$ για κάθε $s \in U$. Σε μία τέτοια περίπτωση $\ddot{\phi}_\alpha(s_*) \leq 0$ και αφού $\int_0^\lambda \frac{\cos^2(\omega)}{\|\mathbf{r}\|} d\xi > 0$ από (29) συνάγεται ότι:

$$\kappa(s_*) \left[\int_0^\lambda \cos(\omega) d\xi \Big|_{s=s_*} \right] < 0 \quad (30)$$

συνεπώς, $\kappa(s_*) \neq 0$, $\int_0^\lambda \cos(\omega) d\xi \Big|_{s=s_*} \neq 0$.

Από (29), θέτοντας $A(s) = \int_0^\lambda \cos(\omega) d\xi$ και $B(s) = \int_0^\lambda \frac{\cos^2(\omega)}{\|\mathbf{r}\|} d\xi$ και δεδομένου ότι $\kappa(s_*) \neq 0$, έχουμε τελικά:

$$\ddot{\phi}_\alpha(s_*) = \kappa(s_*) A(s_*) + B(s_*) \left(\text{i.e., } \frac{\ddot{\phi}_\alpha(s_*)}{\kappa(s_*)} = A(s_*) + \frac{B(s_*)}{\kappa(s_*)} \right) \quad (31)$$

και η απόδειξη είναι πλήρης.

Παρατηρούμε ότι το $A(s_*)$ δέχεται ερμηνεία ανάλογη με τον προσανατολισμό του σημείου $\alpha(s_*)$, όπως ορίζεται στην Ενότητα 4.5. Πράγματι, αντί της ολοκλήρωσης των

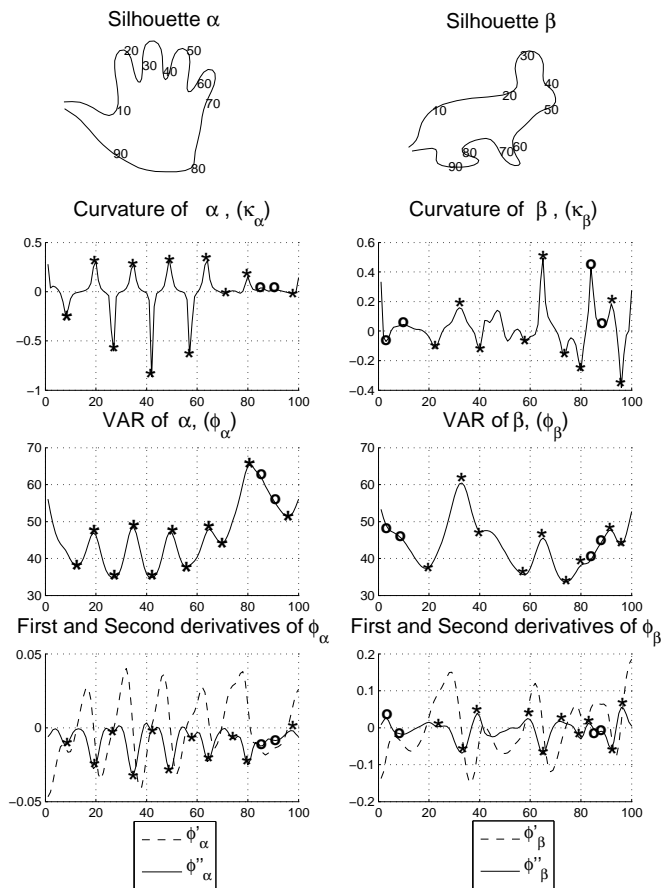
ημιτόνων των γωνιών που σχηματίζονται από την κάθετο στο $\alpha(s_*)$ και το ευθύγραμμο τμήμα που συνδέει τα σημεία $\alpha(s_*)$ και $\alpha(\xi)$ (καθώς το ξ διατρέχει το πεδίο ορισμού της α), ολοκληρώνονται τα συνημίτονα αυτών των γωνιών στον υπολογισμό του $A(s_*)$. Το $A(s)$ θα καλούμε *προσανατολισμό συνημιτόνου (cosine-orientation)* του σημείου $\alpha(s)$ ενώ το $B(s)$ θα καλούμε *προσανατολισμό κλίμακας (scale-orientation)* διότι εξαρτάται από την γωνία και την κλίμακα. Μπορεί να παρατηρήσει κανείς ότι τα $A(s_*)$ και $B(s_*)$ είναι ολικές μετρήσεις και δεν επηρεάζονται από την καμπυλότητα στο s_* . Οι εξισώσεις (31) και (25) θα ονομάζονται *Ολοτοπικές Εξισώσεις (Global-Local Equations)* διότι μέσω της VAR συνδέουν την καμπυλότητα με τους ολικούς περιγραφείς A και B οι οποίοι επιπλέον έχουν και διαισθητική ερμηνεία.

Στην εξίσωση (31) παρουσιάζεται μία σχέση ανάμεσα στη δεύτερη παράγωγο της VAR και την καμπυλότητα της καμπύλης α , σε σχέση με τις δύο ολικές ιδιότητες σχήματος $A(s)$ και $B(s)$. Στο Σχ. 13, βλέπουμε δύο σχήματα από τη βάση δεδομένων Kimia μαζί με τις ομαλές εκδοχές τους και την καμπυλότητα και τα γραφήματα VAR για το καθένα. Να σημειωθεί ότι τα τοπικά ακρότατα της VAR αντιστοιχούν στα σημεία της υψηλής καμπυλότητας στην καμπύλη, όπως επίσης και τον τρόπο που η θέση των σημείων έχει αποτέλεσμα στις τιμές της VAR. Στο σχήμα A, για παράδειγμα, η VAR αποκτά ολικό μέγιστο στο σημείο υπ' αριθμ. 80, γιατί αυτό το σημείο συνδυάζει μακρινή, σχετικά με τα υπόλοιπα σημεία της καμπύλης, θέση, υψηλή καμπυλότητα και σχεδόν *κάδειο* προσανατολισμό (υπό την έννοια της ενότητας 4.5). Το ίδιο ισχύει και για το σημείο 30 στο σχήμα B, το οποίο σε σύγκριση με το σημείο 65 έχει χαμηλότερη καμπυλότητα αλλά υψηλότερη τιμή VAR λόγω της απομακρυσμένης θέσης του στο περίγραμμα, παρόλα αυτά η διαφορά στην καμπυλότητα εξακολουθεί να καταγράφεται στη δεύτερη παράγωγο της VAR, όπως φαίνεται στο Σχ. 14. Να σημειωθεί επίσης η ανθεκτικότητα της VAR στον θόρυβο, εφόσον δεν υπάρχει σημαντική διαφορά ανάμεσα στις VAR των θορυβωδών και των ομαλών εκδοχών των σχημάτων σε αντίθεση με την καμπυλότητα όπου το αποτέλεσμα του θορύβου είναι σημαντικό. Επίσης στα Σχ. 12 και 15, φαίνεται η αντιστοιχία ανάμεσα στα τοπικά ακρότατα της VAR και τα σημεία της υψηλής καμπυλότητας διαφόρων σχημάτων, και μπορούν επίσης να επαληθευτούν οι παραπάνω παρατηρήσεις όσον αφορά την θέση.

4.7. Σύγκριση της VAR με τις μεθόδους καμπυλότητας

Τα παραπάνω αποτελέσματα που αφορούν την VAR αποδεικνύουν ότι η καμπυλότητα μίας επίπεδης καμπύλης, η οποία είναι τοπικός περιγραφέας, κωδικοποιείται στη δεύτερη παράγωγο ενός βαθμωτού ολικού περιγραφέα που είναι ανθεκτικός στο θόρυβο. Στο Σχ. 14, η αντιστοιχία των *κορυφών*⁷ της καμπύλης με τη δεύτερη παράγωγο της VAR παρουσιάζεται για δύο σχήματα. Οι κορυφές των περιγραμμάτων των A και B αντιστοιχούν

⁷Κορυφές είναι τα σημεία όπου η καμπυλότητα αποκτά τοπικό μέγιστο ή ελάχιστο



Σχήμα 14: Η καμπυλότητα, η VAR και η πρώτη και η δεύτερη παράγωγος της VAR για δύο τυπικά εξομαλυμένα σχήματα παρουσιάζονται σε μία κάθετη διάταξη για εύκολη σύγκριση. Προσέξτε την αντιστοιχία ανάμεσα στη δεύτερη παράγωγο της VAR (συνεχής γραμμή στα γραφήματα της τελευταίας σειράς) και την καμπυλότητα (γραφήματα της δεύτερης σειράς) ως προς τη θέση των τοπικών ακρότατων. Τα αντίστοιχα σημεία επισημαίνονται με '*' και 'ο'. Αυτά που επισημαίνονται με το 'ο' απεικονίζουν την περίπτωση στην οποία τα τοπικά ακρότατα της ϕ_α (η ϕ_β) που αντιστοιχούν επίσης στα τοπικά ακρότατα της κ_α (η κ_β) δεν είναι απαραίτητα τοπικά ακρότατα της ϕ_α (η ϕ_β). Να σημειωθεί επίσης ο τρόπος που η θέση και ο προσανατολισμός (υπό την έννοια της ενότητας 4.5) των σημείων του περιγράμματος ποσοτικοποιούνται στις τιμές VAR (περιγράμματα τρίτης σειράς) και την πρώτη παράγωγο της VAR (διακεκομμένη γραμμή στα περιγράμματα της τελευταίας σειράς) αντίστοιχα.

στα τοπικά ακρότατα της δεύτερης παραγώγου της VAR (γραφήματα τελευταίας σειράς). Τα περιγράμματα έχουν εξομαλυνθεί για να εξασφαλιστεί καλά ορισμένη καμπυλότητα για σύγκριση με τα γραφήματα της VAR. Τα αντίστοιχα σημεία επισημαίνονται κατάλληλα στο γράφημα. Στο Σχ. 13, η καμπυλότητα μαζί με την VAR παρουσιάζονται για δύο σχήματα. Παρά τα θορυβώδη σχήματα, οι κορυφές μπορούν να εκτιμηθούν από την VAR, η οποία σε σύγκριση με την καμπυλότητα, εμφανίζει υψηλή ανθεκτικότητα στον θόρυβο. Η αντιληπτή προφάνεια της αναπαράστασης VAR και η νέα ποσοτικοποίηση που παρέχει σε υβριδικές έννοιες σχήματος την καθιστούν κατάλληλη για ποιοτική σύνδεση με την ανθρώπινη αντίληψη ιδιοτήτων.

4.8. Πειραματική επικύρωση

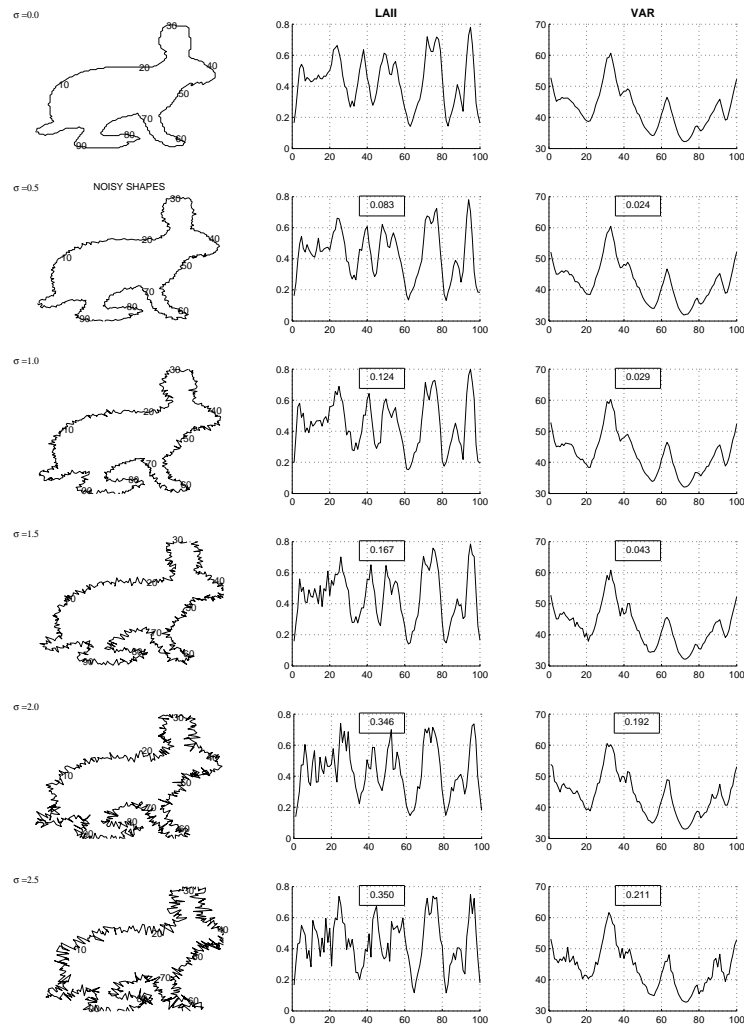
Σε αυτή την ενότητα, δείχνουμε ότι θεωρητικά αποτελέσματα έχουν αποτελέσματα και στην εφαρμογή. Στην συνέχεια, διεξάγουμε ένα σύνολο πειραμάτων όπου συγκρίνεται η VAR με τη Local Area Integral Invariant (LAI) [171]. Η LAI είναι μία σχετική με την VAR, όσον αφορά την αντοχή στον θόρυβο και την πολυπλοκότητα, μέθοδος της τρέχουσας στάθμης της τεχνικής. Η σύγκριση πραγματοποιήθηκε :

- Στην αντίσταση στο θόρυβο, Ενότητα 4.8.1.
- Στην δυνατότητα αναγνώρισης θορυβωδών σχημάτων χωρίς εξομάλυνση χρησιμοποιώντας το σύνολο δεδομένων KIMIA, Ενότητα 4.8.2.
- Στην πολυπλοκότητα, ταχύτητα εκτέλεσης, ευρύτητα εφαρμογής, Ενότητα 4.8.3.

Για την αξιολόγηση της VAR ως διαισθητικού περιγραφέα σχήματος και την αξιολόγηση της απόδοσης του προτεινόμενου πλαισίου, συνιστούμε μία νέα μέθοδο αντιστοίχισης που βασίζεται στη χρήση της επιφάνειας GLT (GLS) σε συνδυασμό με τη VAR (αυτή η νέα μέθοδος ονομάζεται GLS/ VAR). Στην Ενότητα 4.8.4, η νέα μέθοδος συγκρίνεται με την Shape Contexts (Περιβάλλον σχήματος) [167]. Οι συγκρίσεις πολυπλοκότητας εξετάζονται στην Ενότητα 4.8.5.

4.8.1. VAR έναντι LAI ως περιγραφείς σχήματος ανθεκτικοί στο θόρυβο

Η LAI είναι μία αναπαράσταση χαμηλής πολυπλοκότητας που γενικεύει την έννοια της καμπυλότητας προκειμένου να έχει νόημα και στα θορυβώδη τμήματα των καμπύλων. Χρησιμοποιείται ένα κύκλος μίας κάποιας ακτίνας, κεντραρισμένος σε κάθε σημείο και η καμπυλότητα υπολογίζεται ως ο λόγος του κυκλικού δίσκου που βρίσκεται στο εσωτερικό του κλειστού περιγράμματος. Στην περίπτωση μηδενικής καμπυλότητας, π. χ. μία θορυβώδης ευθεία γραμμή, ο μισός δίσκος θα βρίσκεται στο εσωτερικό του σχήματος, ενώ στην περίπτωση της άπειρης καμπυλότητας, αυτό το τμήμα θα τείνει προς το μηδέν ή



Σχήμα 15: Οι θορυβώδεις εκδοχές ενός δεδομένου σχήματος για σταδιακά αυξανόμενες τιμές διασποράς μιας διαταραχής Gauss και οι αντίστοιχοι LAI και VAR περιγραφείς. Ο αριθμός πάνω από κάθε γράφημα χρησιμοποιείται για τη μέτρηση του αποτελέσματος του θορύβου. Πρόκειται για το άθροισμα των κατά σημείο απόλυτων διαφορών ανάμεσα στο γράφημα που αντιστοιχεί στο θορυβώδες σχήμα και το γράφημα που αντιστοιχεί στο αρχικό σχήμα. Κάθε κατα σημείο διαφορά θεωρείται ως κλάσμα της αρχικής τιμής. Όσο πιο υψηλός είναι αυτός ο αριθμός, τόσο πιο παραμορφωμένο είναι το αντίστοιχο θορυβώδες γράφημα σε σύγκριση με το αντίστοιχο αρχικό γράφημα. Οι αριθμοί δικαιολογούν την οπτική εντύπωση ότι η VAR είναι πιο ανθεκτική στο θόρυβο από τη LAI.

το ένα ανάλογα με το πρόσημο της καμπυλότητας στό προς εξέταση σημείο. Συνεπώς η LAII χρησιμοποιεί μία *ευρύτερη* τοπικότητα για τον υπολογισμό της καμπυλότητας αλλά είναι τελικά ένας τοπικός περιγραφέας εφόσον υπολογίζεται μόνον η καμπυλότητα και όχι κάποια ολική ιδιότητα. Η VAR από την άλλη μεριά είναι ένας ολικός περιγραφέας που χρησιμοποιεί ολόκληρο το σχήμα και όπως έχουμε δει ήδη, καταγράφονται επίσης οι υβριδικές και οι τοπικές ιδιότητες σχήματος.

Στο Σχ. 15 ποσοτικοποιήθηκε μία σύγκριση ανάμεσα στη VAR και τη LAII όσον αφορά την αντίσταση στο θόρυβο. Οι θορυβώδεις εκδοχές ενός δεδομένου σχήματος με σταδιακά αυξανόμενες τιμές κανονικού θορύβου παρουσιάζονται μαζί με τις αναπαραστάσεις VAR και LAII για κάθε σχήμα. Για τη μέτρηση του αποτελέσματος του θορύβου στην ίδια την αναπαράσταση, χρησιμοποιείται η κατά σημείο απόσταση ανάμεσα στις θορυβώδεις και τις αρχικές αναπαραστάσεις ως κλάσμα της αρχικής τιμής (χωρίς θόρυβο). Εάν $r_\delta(i), i = 1, 2, \dots, n$ είναι η αναπαράσταση για μία συγκεκριμένη τιμή θορύβου δ και $r_0(i), i = 1, 2, \dots, n$ είναι η αντίστοιχη αναπαράσταση του αρχικού σχήματος (χωρίς τον επαγόμενο θόρυβο) τότε ο αριθμός πάνω από κάθε γράφημα υπολογίζεται από τον τύπο:

$$\sum_i \frac{|r_\delta(i) - r_0(i)|}{r_0(i)} \quad (32)$$

συνεπώς αυτός ο αριθμός παρουσιάζει τη συνολική εκτροπή της θορυβώδους αναπαράστασης ως κλάσμα της αρχικής αναπαράστασης. Τα αποτελέσματα στο Σχ. 15 αποκαλύπτουν ότι η VAR έχει αυξημένη αντοχή σε συνοριακό θόρυβο.

4.8.2. VAR έναντι LAII σε αντιστοίχιση θορυβωδών σχημάτων

Μία προσέγγιση δυναμικού προγραμματισμού, που χρησιμοποιούν πολλοί συγγραφείς για την αντιστοίχιση σχημάτων, [251], [171], [153], θα εισαχθεί τώρα για τη μέτρηση της απόστασης ανάμεσα σε σχήματα από τη βάση δεδομένων KIMIA που έχουν τροποποιηθεί με θόρυβο. Σε αυτό το σημείο διατυπώνουμε τη μεθοδολογία της αντιστοίχισης δύο καμπύλων στο πλαίσιο του δυναμικού προγραμματισμού. Έστω α_1 και α_2 δύο καμπύλες με μήκη κανονικοποιημένα σε 1 και c μία αντιστοίχιση ανάμεσα στα σημεία των δύο καμπυλών τέτοια ώστε $c(\alpha_1(s_1)) = \alpha_2(s_2)$ και C το σύνολο όλων αυτών των αντιστοιχίσεων. Μετρούμε την απόσταση ανάμεσα στις δύο καμπύλες ελαχιστοποιώντας το ενεργειακό συναρτησοειδές:

$$E(\alpha_1, \alpha_2, c) = \int_{s \in [0,1]} |\phi_{\alpha_1}(s) - \phi_{\alpha_2}(c(s))| ds \quad (33)$$

ως προς την αντιστοίχιση c . Συνεπώς, η απόσταση ανάμεσα στις δύο καμπύλες α_1 και

NOISY SHAPES						
RANK		HTS				
1	VAR/VAR	24	4.2 4.7 5.1 4.1 3.6 3.6	5.8 6.2 7.5 5.7 6.6 6.5	5.0 6.5 5.4 5.2 6.0 5.2	4.3 4.6 4.5 6.9 4.7 6.2
	VAR/5P	24	4.2 4.7 5.1 4.1 3.6 3.6	5.8 6.2 7.5 5.7 6.6 6.5	5.0 6.5 5.4 5.2 6.0 5.2	4.3 4.6 4.5 6.9 4.7 6.2
	LAII/VAR	24	1.8 2.4 1.7 2.0 2.1 1.9	2.3 2.6 2.3 2.4 2.2 1.9	2.4 3.5 2.9 2.1 2.9 2.6	2.1 2.3 1.8 2.1 1.9 1.7
	LAII/5P	24	1.8 2.4 1.7 2.0 2.1 1.9	2.3 2.6 2.3 2.4 2.2 1.9	2.4 3.5 2.9 2.1 2.9 2.6	2.1 2.3 1.8 2.0 1.9 1.7
2	VAR/VAR	24	5.1 7.6 9.0 5.8 6.4 8.2	5.9 6.5 7.5 11.0 6.7 12.3	10.8 11.3 9.5 12.5 7.0 6.4	7.9 6.3 9.6 7.5 5.4 6.7
	VAR/5P	24	8.6 7.6 9.8 5.8 6.4 8.1	5.9 6.5 7.6 11.8 6.7 11.8	10.7 11.3 9.5 12.5 7.0 6.4	7.8 6.3 5.6 7.4 5.3 6.7
	LAII/VAR	21	2.4 4.8 3.9 3.3 2.4 4.0	2.5 3.0 2.7 2.9 2.6 4.3	4.7 5.4 4.7 4.5 4.7 5.0	2.9 2.4 2.3 2.6 2.3 2.5
	LAII/5P	21	2.4 4.7 3.9 3.4 2.4 3.9	2.5 3.0 2.7 2.9 2.6 4.4	4.6 5.5 4.7 4.5 4.6 4.9	2.9 2.4 2.3 2.6 2.3 2.5
3	VAR/VAR	20	9.7 8.6 12.9 8.7 6.8 10.3	6.4 6.6 7.6 12.3 7.4 14.1	11.5 11.8 9.8 12.6 10.4 8.7	9.2 15.4 6.5 9.3 5.4 13.8
	VAR/5P	20	9.6 8.6 12.5 8.7 6.8 9.5	6.4 6.6 7.6 12.2 7.4 13.8	11.4 11.8 9.8 12.8 10.4 8.7	9.0 14.9 6.5 9.3 5.4 13.8
	LAII/VAR	17	3.8 4.9 4.1 3.9 3.8 4.2	2.9 3.1 3.0 3.1 2.6 4.4	5.1 5.4 5.5 4.5 4.7 5.1	3.0 4.8 2.7 2.7 2.8 4.2
	LAII/5P	17	3.8 4.8 4.1 3.6 3.8 4.1	2.9 3.1 3.0 3.0 2.6 4.4	5.1 5.5 5.5 4.5 4.7 5.1	3.0 4.7 2.7 2.6 2.6 4.2
4	VAR/VAR	18	10.2 11.1 17.7 9.5 10.0 11.9	9.4 8.9 8.8 13.2 8.8 14.2	12.2 12.1 13.5 14.9 11.0 12.3	9.7 15.8 7.8 9.3 6.4 15.5
	VAR/5P	17	10.0 11.1 17.6 9.5 10.1 11.7	9.4 8.9 8.8 13.2 8.8 14.1	12.2 12.1 12.6 14.8 11.0 12.3	9.7 15.4 7.8 9.3 6.4 15.5
	LAII/VAR	16	4.0 5.0 4.4 4.0 4.3 4.4	3.0 3.7 3.1 3.1 2.7 4.6	5.1 5.5 5.8 4.6 4.8 5.2	3.3 4.9 3.3 3.3 2.9 4.5
	LAII/5P	16	4.0 4.9 4.4 3.9 4.2 4.2	3.0 3.7 3.1 3.1 2.7 4.5	5.1 5.6 5.8 4.5 4.8 5.1	3.3 4.8 3.3 3.3 2.9 4.5
5	VAR/VAR	12	12.2 12.2 18.2 13.3 11.3 12.2	10.1 10.8 11.7 13.4 9.8 15.3	12.7 12.3 18.7 15.5 11.2 12.5	16.4 16.9 24.0 22.5 24.6 18.9
	VAR/5P	12	12.0 12.1 18.2 13.6 11.2 11.9	10.0 10.5 11.6 13.2 10.0 14.7	12.8 12.2 15.4 15.5 11.2 12.4	16.0 15.9 23.6 22.5 23.8 16.2
	LAII/VAR	8	4.5 5.1 4.6 4.6 4.5 4.4	3.3 3.9 3.8 3.9 3.4 4.8	5.1 5.8 5.8 4.6 4.8 5.2	4.1 5.0 4.5 4.2 4.8 4.8
	LAII/5P	8	4.5 5.1 4.6 4.5 4.5 4.2	3.3 3.9 3.8 3.9 3.4 4.6	5.1 5.8 5.8 4.6 4.8 5.1	3.6 4.9 4.1 4.1 4.4 4.8
6	VAR/VAR	12	13.1 12.6 18.9 14.2 14.0 12.8	10.4 11.3 12.1 13.5 10.2 16.6	12.9 12.4 16.9 15.6 11.3 12.6	19.8 16.7 24.5 22.7 25.7 16.9
	VAR/5P	11	13.0 12.6 18.9 13.7 13.9 12.0	10.1 11.3 12.0 13.4 10.2 16.5	13.0 12.4 15.7 15.5 11.3 12.6	19.8 16.5 24.1 22.7 24.8 16.9
	LAII/VAR	9	4.6 5.2 4.7 4.7 4.6 4.4	3.5 4.1 3.9 3.9 3.5 4.9	5.4 5.9 5.9 4.7 4.9 5.2	4.7 5.1 4.7 5.0 4.9 5.0
	LAII/5P	9	4.5 5.2 4.7 4.6 4.6 4.2	3.5 4.1 3.9 3.9 3.5 4.8	5.1 5.8 5.9 4.6 4.8 5.2	4.6 5.0 4.4 4.9 4.7 5.0

Σχήμα 16: Παρουσιάζονται οι έξι εγγύτερες αντιστοιχίσεις σε θορυβώδης εκδοχές του συνόλου των 24 σχημάτων του KIMIA για δύο εφαρμογές δύο διαφορετικών μεθόδων σύγκρισης. Η πραγματική απόσταση του σχήματος από τη θορυβώδη εκδοχή του είναι ο αριθμός ακριβώς πάνω από κάθε σχήμα. Στην πρώτη σειρά είναι οι θορυβώδεις εικόνες (εικόνες ερώτησης) που προκύπτουν με διαταραχή Gauss των περιγραμμάτων τους με διασπορά $\sigma = 0.5$. Σε καθεμία από τις άλλες σειρές υπάρχει, η σχετική με τον αντίστοιχο δείκτη διαβάθμισης (πρώτη στήλη) πιο κοντινή αντιστοιχίωση με καθένα από τα θορυβώδη σχήματα για δύο διαφορετικές μεθόδους, VAR και LAII και δύο διαφορετικές εφαρμογές δυναμικού προγραμματισμού. Η πρώτη εφαρμογή χρησιμοποιεί την VAR για την επιλογή της αρχικής αντιστοιχίας και για τις δύο μεθόδους (VAR/VAR και LAII/VAR) ενώ η δεύτερη εφαρμογή εξετάζει εξαντλητικά την αντιστοιχία σε κάθε 5ο σημείο του περιγράμματος και για τις δύο μεθόδους (VAR/5P και LAII/5P). Η VAR παρουσιάζει καλύτερα αποτελέσματα από την LAII αφού και οι πρώτες και οι δεύτερες αντιστοιχίσεις είναι σωστές για τη VAR (τα θορυβώδη σχήματα αντιστοιχούν σε σχήματα της σωστής κατηγορίας), ενώ τρεις από τις δεύτερες αντιστοιχίσεις δεν είναι σωστές για το LAII. Τα αποτελέσματα για τις άλλες αντιστοιχίσεις είναι επίσης καλύτερα για την VAR σε σύγκριση με αυτές της LAII. Συγκρίνοντας τις εφαρμογές και για τις δύο μεθόδους, βλέπουμε ότι επιλέγοντας την αρχική αντιστοιχία μέσω της VAR βελτιώνεται η απόδοση αλλά και η ακρίβεια αναγνώρισης αφού η VAR επιλέγει σημεία εκκίνησης που έχουν διαισθητική σύνδεση με ολόκληρο το σχήμα.

α_2 δίδεται από τον τύπο :

$$\min_{c \in C} E(\alpha_1, \alpha_2, c) \quad (34)$$

Εφόσον αυτή η μέθοδος χρησιμοποιείται ευρέως στην αντιστοίχιση σχημάτων, δεν θα μπούμε σε λεπτομέρειες. Η προσέγγιση μας δυναμικού προγραμματισμού είναι πανομοιότυπη με αυτή του [171], στη φάση αντιστοίχισης χρησιμοποιούμε πανομοιότυπες παραμέτρους και για τις δύο μεθόδους που συγκρίνονται. Υπό εξέλιξη έρευνα στην σύγκριση καμπύλων ελαχιστοποιώντας συναρτησοειδή αντιστοιχίσεων βρίσκεται στα [252, 253, 254, 255].

Παρουσιάζονται δύο πειράματα. Στο πρώτο πείραμα μετρούμε την επιτυχία και των δύο μεθόδων μετρώντας τον αριθμό των σωστών αντιστοιχίσεων από τις 6 πιο κοντινές αντιστοιχίσεις και για τα 24 θορυβώδη σχήματα. Χρησιμοποιούμε 24 σχήματα από τη βάση δεδομένων του KIMIA, με διαταραχή θορύβου $\sigma = 0.5$ στα περιγράμματα. Στη φάση κωδικοποίησης, η εφαρμογή της LAII ήταν ως εξής: Ξεκινώντας από μία δυαδική εικόνα του προς κωδικοποίηση σχήματος, πρώτα εξαγούμε το περίγραμμα. Κατόπιν, χρησιμοποιώντας έναν κυκλικό φίλτρο (με δομή δυαδικής εικόνας κύκλου ακτίνας 15, όπως προτείνεται στο [171]) εκτελούμε συνέλιξη του φίλτρου με την εικόνα του σχήματος μόνον στα σημεία του περιγράμματος. Οι τιμές της συνέλιξης σε καθένα από τα σημεία του περιγράμματος είναι και οι τιμές της LAII σε αυτά τα σημεία. Παρατηρήστε ότι εάν αρχικά δίδεται μόνον το περίγραμμα (ως λίστα σημείων) και όχι η δυαδική εικόνα του σχήματος, τότε η δυαδική εικόνα πρέπει να κατασκευαστεί πριν γίνει δυνατός ο υπολογισμός των τιμών της LAII. Αυτό παίζει κάποιο ρόλο στην πολυπλοκότητα της μεθόδου που θα συζητηθεί στην επόμενη ενότητα.

Διακριτοποιούμε τις καμπύλες με δειγματοληψία 100 σημείων τοποθετημένων σε ίση απόσταση σε καθεμία από αυτές. Η επιλογή της αρχικής αντιστοιχίας ανάμεσα σε ένα ζεύγος σημείων ανάμεσα στις δύο καμπύλες είναι ζήτημα κρίσιμης σπουδαιότητας για τον δυναμικό προγραμματισμό. Η εξαντλητική έρευνα για όλους τους πιθανούς συνδυασμούς για αυτή την αρχική αντιστοιχία αυξάνει την πολυπλοκότητα αντιστοίχισης από $O(n^2)$ σε $O(n^3)$ όπου n ο αριθμός των σημείων του περιγράμματος. Προτείνουμε μία εναλλακτική τεχνική, τη χρήση της VAR για τον περιορισμό της αναζήτησης των αρχικών αντιστοιχιών στα διαισθητικά σημεία. Για το λόγο αυτό, εξετάζουμε σε αυτό το σημείο δύο εφαρμογές του δυναμικού προγραμματισμού. Στην πρώτη εφαρμογή, χρησιμοποιούμε τα τοπικά ακρότατα της VAR ως τα μόνα ενδεχόμενα για μία αρχική αντιστοιχία ανάμεσα στις δύο καμπύλες και συνεπώς εξετάζουμε όλους τους συνδυασμούς αυτών των ακροτάτων σημείων στις δύο καμπύλες. Τα ακρότατα σημεία επιλέγονται λόγω της διαισθητικής ερμηνείας τους ως προς το αρχικό σχήμα. Όπως είδαμε, η πρώτη και η δεύτερη παράγωγος της VAR ποσοτικοποιούν άλλες διαισθητικές ερμηνείες οι οποίες μπορούν να χρησιμοποιηθούν για την επιλεκτική μείωση των πιθανών επιλογών σημείων που θα χρησιμοποιηθούν ως οι

αρχικές αντιστοιχίες στην φάση του δυναμικού προγραμματισμού.

Στη δεύτερη εφαρμογή, εξετάζουμε όλους τους συνδυασμούς για κάθε 5ο σημείο στις δύο καμπύλες ξεκινώντας από μία τυχαία θέση. Αυτή η εφαρμογή έχει χρησιμοποιηθεί κι από άλλους συγγραφείς (για κάθε 10ο ή 15ο σημείο σε περιγράμματα 100 σημείων) [153]. Οι VAR και LAII συγκρίνονται χρησιμοποιώντας τις δύο εφαρμογές δυναμικού προγραμματισμού με ίδιες παραμέτρους για τη φάση αντιστοίχισης και για τις δύο μεθόδους. VAR / VAR στο πείραμα σημαίνει χρήση της VAR ως περιγραφέα σχήματος και επίσης την χρήση της VAR για την επιλογή της αρχικής αντιστοιχίας για τον δυναμικό προγραμματισμό (πρώτη εφαρμογή). VAR /5P σημαίνει χρήση της VAR ως περιγραφέα σχήματος αλλά η αρχική αντιστοιχία για τον δυναμικό προγραμματισμό εξετάζεται για κάθε 5ο σημείο και στις δύο καμπύλες (δεύτερη εφαρμογή). Τα LAII/ VAR και LAII/5P υποδηλώνουν τη χρήση της LAII ως περιγραφέα σχήματος με τις ίδιες δύο όπως προηγουμένως εφαρμογές για την επιλογή των πιθανών σημείων εκκίνησης για δυναμικό προγραμματισμό.

Στον πίνακα του Σχ. 16, οι στήλες απαριθμούνται με θορυβώδης εκδοχές των 24 σχημάτων του KIMIA. Οι θορυβώδης αυτές εκδοχές προέκυψαν εφαρμόζοντας σε κάθε σημείο του περιγράμματος μία κάθετη στην εφαπτομένη του διαταραχή, το μέγεθος της οποίας προέκυψε από μία συνάρτηση Gauss με μέση τιμή 1 και διασπορά 0,5. Η βάση σχημάτων του KIMIA περιέχει 6 περιγράμματα για καθεμία από τις 4 κατηγορίες (ψάρι, χέρι, αεροπλάνο και λαγός). Για κάθε θορυβώδες σχήμα, οι 6 καλύτερες αντιστοιχίσεις (1 έως 6 είναι η διαβάθμιση της αντιστοίχισης) σε σχέση με τα αρχικά 24 σχήματα της βάσης δεδομένων παρουσιάζονται ως 6 τετράδες γραμμών. Κάθε τετράδα αναπαριστά το αποτέλεσμα και για τις δύο μεθόδους με τις δύο διαφορετικές εφαρμογές. Ο συνολικός αριθμός σωστών αντιστοιχίσεων για κάθε διαβάθμιση, μέθοδο και εφαρμογή τοποθετείται στην τρίτη στήλη. Η πραγματική απόσταση του σχήματος από την αντίστοιχη θορυβώδη εκδοχή του είναι ο μικρός αριθμός ακριβώς πάνω από κάθε σχήμα.

Βλέπουμε ότι η VAR παράγει καλύτερα αποτελέσματα από την LAII ανεξάρτητα από εφαρμογή αλλά επιβεβαιώνεται επίσης ότι η δεύτερη εφαρμογή δεν βελτιώνει την αναγνώριση καμίας μεθόδου. Η πρώτη εφαρμογή έχει τάξη πολυπλοκότητας $O(\kappa n^2) = O(n^2)$, αφού το κ εξαρτάται από τα διαισθητικά χαρακτηριστικά του σχήματος (καταγεγραμμένα στα τοπικά ακρότατα της VAR) και όχι απευθείας από τον αριθμό των σημείων. Η δεύτερη εφαρμογή από την άλλη πλευρά έχει πολυπλοκότητα $O(n/5 \times n^2) = O(n^3)$, συνεπώς η πρώτη εφαρμογή είναι πιο γρήγορη ειδικά σε περιπτώσεις απλών σχημάτων που έχουν μεγάλο αριθμό σημείων. Η δεύτερη εφαρμογή καλύπτει έναν μεγαλύτερο χώρο αναζήτησης για την αντιστοίχιση κάθε ζεύγους καμπύλων εφόσον πρέπει να εξεταστούν 20 πιθανά σημεία (υπό την υπόθεση ότι υπάρχουν 100 σημεία) για κάθε καμπύλη σε σύνολο $20 \times 20 = 400$ συνδυασμών, την στιγμή που ο μέσος αριθμός των τοπικών ακροτάτων για κάθε VAR είναι περίπου 5 σε σύνολο 25 συνδυασμών. Συγκρίνοντας τις προσεγγίσεις, επιβεβαιώνουμε ότι ο περιορισμός των επιλογών για τις αρχικές αντιστοιχίες μέσω της VAR οδηγεί σε

σ=1.5																									
RANK	HITS																								
1	VAR 24	6.5	7.8	9.2	8.5	7.1	6.6	8.7	8.0	8.8	8.6	10.6	10.2	6.8	10.8	4.8	10.5	8.6	9.0	16.3	6.5	7.9	9.2	7.3	5.6
	LAI 24	3.0	3.1	2.1	3.0	2.6	2.5	3.0	3.1	3.0	3.6	3.7	3.5	3.3	4.2	4.4	2.5	3.1	3.7	3.0	2.9	2.9	2.8	2.6	3.0
2	VAR 24	12.1	9.1	15.4	10.3	9.8	11.1	10.4	10.5	8.8	14.2	12.5	16.1	13.5	13.5	8.2	11.6	11.0	9.5	17.4	10.6	8.8	10.9	10.2	9.4
	LAI 21	3.1	5.2	3.9	4.0	2.9	3.6	3.5	3.3	3.2	4.0	3.8	5.0	4.7	5.4	4.9	3.5	4.1	4.2	3.5	3.8	3.1	3.5	3.0	3.3
3	VAR 18	12.1	11.2	18.7	12.6	10.2	12.7	11.0	10.8	9.1	14.7	13.1	16.5	14.4	17.6	11.7	12.2	12.2	12.5	17.7	14.4	10.7	12.8	11.2	15.5
	LAI 20	4.3	5.4	4.2	4.7	4.0	3.9	3.8	3.3	3.6	4.4	3.9	5.1	4.9	5.4	5.4	4.3	4.5	4.3	3.7	4.9	3.4	3.8	3.1	4.5
4	VAR 18	12.6	13.3	20.5	14.0	11.5	13.8	11.7	11.2	10.9	14.7	14.5	16.7	14.4	17.6	15.2	13.5	12.2	14.4	17.7	15.5	12.6	13.5	12.8	15.9
	LAI 12	4.6	5.4	4.3	4.8	4.6	4.1	4.0	3.3	3.8	4.9	4.3	5.3	4.9	5.5	5.5	4.4	4.7	5.0	3.7	5.0	3.9	3.9	3.8	4.7

σ=2.5																									
RANK	HITS																								
1	VAR 24	8.4	8.4	9.3	8.8	7.7	7.8	12.3	8.6	8.3	10.8	12.6	13.2	8.4	11.5	9.3	9.9	8.8	9.3	10.1	10.2	11.6	10.0	12.2	8.9
	LAI 17	6.5	6.3	7.3	6.5	6.4	6.2	6.3	5.9	6.3	6.2	6.2	6.8	7.0	6.5	6.0	6.0	5.9	6.2	6.3	6.8	6.6	6.1	6.5	6.8
2	VAR 23	10.7	10.9	14.3	10.7	9.1	11.8	12.7	9.0	13.1	13.3	13.2	17.7	13.1	14.4	10.8	11.7	12.9	11.9	16.1	12.2	13.0	14.3	13.4	14.9
	LAI 15	6.6	6.7	7.7	7.5	6.4	6.5	6.6	6.0	6.3	6.4	6.6	7.6	7.0	6.8	6.6	6.2	6.3	6.9	6.4	6.9	7.0	6.6	7.1	7.0
3	VAR 19	13.6	14.0	17.9	13.7	11.6	12.6	13.6	9.2	13.6	14.6	14.8	18.0	14.0	18.8	13.2	12.3	12.9	14.0	16.7	15.0	13.0	15.3	13.6	15.8
	LAI 10	7.1	6.8	8.0	7.8	6.9	6.5	6.9	6.4	6.4	6.8	6.9	7.6	7.2	7.0	7.0	7.2	7.2	6.9	7.3	7.0	7.6	7.2	7.3	7.2
4	VAR 14	15.0	14.7	21.3	15.7	12.1	14.2	15.1	11.1	14.5	15.6	15.4	18.7	14.4	19.6	13.9	13.7	13.0	14.3	18.9	16.6	18.9	18.7	19.6	16.8
	LAI 8	7.5	6.8	8.5	7.8	7.0	6.8	7.2	7.1	6.9	6.8	7.5	7.7	7.5	7.0	7.0	7.3	7.3	7.1	7.3	7.0	7.7	7.4	7.4	7.5

σ=3.5																									
RANK	HITS																								
1	VAR 24	10.1	11.2	11.4	12.6	10.1	12.8	11.3	11.2	12.2	12.2	8.2	10.2	8.9	12.0	12.2	11.3	10.9	10.4	10.0	13.2	12.9	13.6	13.3	13.1
	LAI 14	8.9	8.2	7.4	7.6	7.9	8.9	6.6	7.1	6.5	8.6	6.3	6.8	8.3	8.5	9.3	8.2	7.7	7.2	8.0	7.9	7.9	7.9	7.7	7.8
2	VAR 21	11.7	12.8	14.6	12.6	10.8	15.2	11.5	11.3	13.3	12.2	13.4	16.7	12.7	13.1	17.2	14.5	12.7	12.2	17.1	14.3	20.2	15.7	14.2	13.9
	LAI 14	9.0	8.8	7.5	7.7	8.1	9.1	7.0	7.3	7.1	8.7	6.4	7.4	8.4	8.7	9.6	8.4	8.3	7.5	8.2	7.9	8.0	8.0	8.0	8.0
3	VAR 19	13.3	13.7	16.0	14.3	13.0	16.5	12.1	11.6	13.5	16.3	13.5	17.2	14.6	13.8	18.1	15.2	15.4	13.8	17.8	16.5	20.9	16.4	15.0	15.4
	LAI 8	9.5	8.9	7.5	7.7	8.2	9.4	7.0	7.3	7.2	8.8	6.7	7.4	8.5	9.2	9.9	8.8	8.5	8.0	8.3	8.0	8.0	8.1	8.1	8.3
4	VAR 16	15.5	14.6	18.3	14.4	13.1	16.7	13.3	13.6	14.4	17.2	13.8	18.4	16.0	14.4	18.4	15.2	16.4	14.9	18.1	17.3	21.8	19.3	19.3	16.6
	LAI 9	9.5	9.1	7.7	7.9	8.4	9.4	7.0	7.5	7.3	8.8	6.7	7.4	8.9	9.2	10.0	8.8	8.5	8.1	8.4	8.0	8.1	8.5	8.2	8.5

σ=4.5																									
RANK	HITS																								
1	VAR 23	14.1	11.7	14.5	15.3	10.2	8.7	13.9	12.5	16.4	13.3	17.3	16.8	15.0	15.1	13.9	9.5	14.5	14.1	15.7	17.7	17.0	14.6	16.4	
	LAI 9	7.5	8.5	8.4	23.8	8.3	7.9	8.0	8.3	8.0	7.0	8.1	7.8	37.7	8.6	9.3	8.6	8.6	9.0	9.7	8.8	8.4	7.7	37.8	36.5
2	VAR 19	21.4	13.6	14.6	17.0	15.6	14.5	13.3	15.8	16.4	17.9	17.2	18.7	16.9	15.7	19.2	17.1	12.3	14.9	20.5	17.7	19.9	18.5	17.2	18.4
	LAI 9	7.6	8.6	8.8	24.0	8.8	8.1	8.2	8.6	8.1	7.3	8.3	7.9	38.5	9.3	9.7	8.8	9.3	9.1	9.8	9.0	8.9	8.1	38.6	37.0
3	VAR 16	21.7	14.4	18.2	17.3	16.1	15.4	13.9	16.2	17.5	18.2	17.7	18.8	17.2	16.0	19.6	18.2	13.5	18.6	20.9	17.9	19.9	18.9	18.5	18.9
	LAI 9	7.7	8.7	8.9	25.3	9.1	8.3	8.4	8.7	8.4	7.4	8.3	8.3	40.1	9.5	10.1	8.8	9.3	10.1	10.5	9.3	8.9	8.1	40.2	39.0
4	VAR 15	22.7	15.4	18.2	18.0	18.1	15.5	13.9	16.3	17.8	18.5	18.0	19.1	18.5	16.9	20.0	18.6	14.8	18.9	21.1	18.9	20.1	21.8	19.3	19.4
	LAI 10	7.7	8.9	9.1	26.1	9.4	8.5	8.5	8.7	8.5	7.4	8.3	8.5	40.3	9.6	10.4	9.0	9.4	10.1	10.6	9.4	9.2	8.3	40.4	39.1

Σχήμα 17: Πίνακες απόστασης για τις τέσσερις καλύτερες αντιστοιχίσεις για σταδιακά αυξανόμενο θόρυβο περιγράμματος. Συγκρίνονται οι VAR/VAR (VAR στο Σχήμα) και LAII/VAR (LAI στο Σχήμα). Οι βαθμολογίες αντιστοίχισης για την LAII/VAR υποβαθμίζονται με ταχύτερο ρυθμό σε σύγκριση με τις βαθμολογίες της VAR/VAR, γεγονός που αποδεικνύει ότι η VAR/VAR έχει αυξημένη αντίσταση στον θόρυβο.

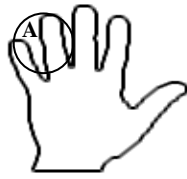
καλύτερο χρόνο εκτέλεσης και καλύτερα αποτελέσματα αναγνώρισης ταυτόχρονα. Ο περιορισμός των πιθανών αρχικών αντιστοιχιών του δυναμικού προγραμματισμού στις διαισθητικές επιλογές παράγει καλύτερα αποτελέσματα αναγνώρισης σε σύγκριση με ευρετικές αναζητήσεις ανεξάρτητες του περιβάλλοντος σχήματος (π.χ. ομοιόμορφη δειγματοληψία). Αυτό συμβαίνει γιατί ο χώρος αναζήτησης μειώθηκε διαισθητικά, μειώνοντας έτσι και την πιθανότητα ψεύτικων θετικών αντιστοιχίσεων. Αυτό θα φανεί καλύτερα στο επόμενο πείραμα της Ενότητας 4.8.4.

Στο δεύτερο πείραμα του Σχ. 17, συγκρίνουμε τη δυνατότητα ανάκτησης της VAR / VAR με αυτή της LAII/ VAR για σταδιακά αυξανόμενες τιμές διασποράς θορύβου. Στο Σχ. 16 του προηγούμενου πειράματος, η σύγκριση πραγματοποιείται για διασπορά ίση με 0,5. Στο δεύτερο πείραμα αυξάνουμε την διασπορά διαταραχής περιγράμματος σταδιακά από 1,5 σε 4,5 με βήμα 1,0. Οι μέθοδοι συγκρίνονται με βάση τον αριθμό των σωστών αντιστοιχίσεων σε κάθε μία από τις τέσσερις πλησιέστερες αντιστοιχίσεις. Η VAR / VAR παρουσιάζει αυξημένη αντοχή στον θόρυβο εφόσον όλες οι πρώτες αντιστοιχίσεις είναι σωστές για κάθε διασπορά θορύβου έως 3,5. Όλες οι άλλες αντιστοιχίσεις είναι επίσης υψηλότερες από εκείνες της LAII/ VAR.

4.8.3. LAII έναντι VAR σε απόδοση, εφαρμοσιμότητα και πολυπλοκότητα

Στην προηγούμενη ενότητα, η πολυπλοκότητα και η απόδοση της φάσης αντιστοίχισης εξετάστηκαν για δύο διαφορετικές εφαρμογές δυναμικού προγραμματισμού. Στην ενότητα αυτή θα συγκρίνουμε την πολυπλοκότητα της φάσης κωδικοποίησης. Ο υπολογισμός της GLT / VAR απαιτεί $n^2/2$ πράξεις (αφού η GLS είναι συμμετρική), ανεξάρτητα από την ποσότητα του θορύβου που ασκείται στο περίγραμμα, όπου n είναι ο αριθμός των σημείων του περιγράμματος. Ο χρόνος εκτέλεσης για ένα περίγραμμα 100 σημείων είναι μικρότερος από ένα δευτερόλεπτο σε μία τυπική διάταξη υπολογιστή. Για την LAII από την άλλη μεριά, αν υποθέσουμε ότι και το περίγραμμα (που εξήχθη) και η αρχική δυαδική εικόνα είναι διαθέσιμα, τότε η πολυπλοκότητα είναι $O(k^2 \times n) = O(n)$, με k να είναι το μέγεθος του κυκλικού φίλτρου και n ο αριθμός των σημείων του περιγράμματος. Για τη μέθοδο VAR, δεν είναι απαραίτητη η ύπαρξη της αρχικής δυαδικής εικόνας, χρειάζεται μόνο το περίγραμμα. Εάν δεν είναι διαθέσιμη η δυαδική εικόνα, τότε η πολυπλοκότητα της μεθόδου LAII αυξάνεται σε $O(k^2 \times m)$, όπου m είναι το μέγεθος (σε εικονοστοιχία) της δυαδικής εικόνας το οποίο συνήθως είναι μεγαλύτερο από n^2 . Αυτό συμβαίνει γιατί η δυαδική εικόνα στη μέθοδο LAII πρέπει να κατασκευαστεί πριν από τη φάση κωδικοποίησης. Εάν μόνο η δυαδική εικόνα υπάρχει αρχικά, τότε πρέπει να πραγματοποιηθεί μία εξαγωγή περιγράμματος και για τις δύο μεθόδους πριν από τη φάση κωδικοποίησης.

Αν και η πολυπλοκότητα των δύο υπό σύγκριση μεθόδων είναι γενικά ίδια, η μέθοδος



Σχήμα 18: Ένα ζήτημα σχετικά με την επιλογή της ακτίνας του κυκλικού δίσκου στην εφαρμογή της μεθόδου LAII παρουσιάζεται εδώ. Ο κυκλικός δίσκος με κέντρο στο σημείο A τέμνει το περίγραμμα σε πολλαπλές περιοχές, συνεπώς, η περιοχή του που βρίσκεται μέσα στο περίγραμμα δεν μπορεί να αναπαραστήσει σωστά την καμπυλότητα στο σημείο A . Αυξάνεται λοιπόν η πολυπλοκότητα σε περιπτώσεις όπου πρέπει να υπολογιστεί μία έγκυρη ακτίνα κυκλικού δίσκου για ολόκληρο το σχήμα ή όπου πρέπει να πραγματοποιηθεί μία δυναμική τροποποίηση της ακτίνας σύμφωνα με τις μεταβολές του σχήματος.

LAII ενδέχεται να ενέχει επιπρόσθετη πολυπλοκότητα στην προσαρμογή της ακτίνας του κυκλικού φίλτρου, ανάλογα με την ποσότητα παραμόρφωσης του θορύβου ή της πολυπλοκότητας⁸ σχήματος του υπό εξέταση περιγράμματος. Αυτό συμβαίνει διότι ένας δίσκος κάποιας ακτίνας ενδέχεται να τέμνει το εσωτερικό του σχήματος σε περισσότερες από μία παρακείμενες περιοχές όπως στο Σχ. 18. Σε μια τέτοια περίπτωση ενδέχεται να είναι απαραίτητη μία δυναμική προσαρμογή της ακτίνας του δίσκου, γεγονός που αυξάνει περαιτέρω την πολυπλοκότητα του αλγορίθμου LAII τουλάχιστον κατά μία τάξη πολυπλοκότητας. Από την άλλη πλευρά, η GLT / VAR υπολογίζεται απευθείας για οποιοδήποτε περίγραμμα και οποιαδήποτε πολυπλοκότητα σχήματος, εφόσον δεν εμπλέκεται παράμετρος κλίμακας. Περαιτέρω, ενώ η μέθοδος LAII είναι επιτυχής στη γενίκευση της έννοιας της καμπυλότητας στην περίπτωση θορυβωδών περιγραμμάτων, σαν περιγραφέας σχήματος έχει παρόμοιους περιορισμούς με τους περιγραφείς καμπυλότητας, αφού είναι ένας τοπικός περιγραφέας. Η σχετική μετατόπιση ενός τμήματος της καμπύλης ως προς την υπόλοιπη καμπύλη για ένα σημείο σε αυτό το τμήμα δεν λαμβάνεται υπόψη στον υπολογισμό της τιμής LAII για το σημείο αυτό. Από την άλλη, ο περιγραφέας VAR ενσωματώνει ολικές και τοπικές περιγραφές ταυτόχρονα για κάθε σημείο της καμπύλης.

4.8.4. GLS/VAR έναντι Shape Contexts

Στο προηγούμενο πείραμα, δείχτηκε ότι η χρήση της VAR για τον περιορισμό της αρχικής αντιστοιχίας των σημείων εκκίνησης του δυναμικού προγραμματισμού στα διαισθητικά σημεία μείωσε την πολυπλοκότητα της φάσης αντιστοίχισης χωρίς να μειώνεται η δυνατό-

⁸Υπό την έννοια της διάστασης Hausdorff.

τητα ανάκτησης. Σε αυτή την ενότητα εξετάζουμε περαιτέρω αυτή την έννοια, εισάγοντας μία νέα μέθοδο αντιστοίχισης χαμηλής πολυπλοκότητας που χρησιμοποιεί την επιφάνεια GLT (GLS) σε συνδυασμό με τη VAR. Ονομάζουμε αυτή τη νέα μέθοδο GLS/ VAR. Η VAR χρησιμοποιείται για τον περιορισμό των πιθανών επιλογών για τα σημεία εκκίνησης της GLS και η απόλυτη κατά σημείο διαφορά των GLS θεωρείται ως απόσταση ανάμεσα στα σχήματα. Ο σκοπός αυτού του πειράματος είναι να εξεταστεί η μείωση στην πολυπλοκότητα και τον χρόνο εκτέλεσης που μπορεί να επιτευχθεί από το προτεινόμενο πλαίσιο σε σχέση με άλλες ευρέως χρησιμοποιούμενες μεθόδους. Για το λόγο αυτό, επιλέγουμε να το συγκρίνουμε με τη μέθοδο Shape Contexts (SC) [167], γιατί είναι μία μέθοδος που χρησιμοποιείται ευρέως και αναγνωρίζεται γενικά για την αντιστοίχιση σχημάτων και για την οποία υπάρχει μία γρήγορη εκδοχή με παρόμοια πολυπλοκότητα και χρόνο εκτέλεσης με της GLS/ VAR. Περαιτέρω, η SC έχει αποτελέσει την βάση για άλλες μεθόδους που χρησιμοποιούν δυναμικό προγραμματισμό και επιτυγχάνουν βαθμολογίες αντιστοίχισης της τρέχουσας στάθμης της τεχνικής που μπορεί να επωφεληθούν από το προτεινόμενο πλαίσιο εργασίας. Για τη σύγκριση, χρησιμοποιούμε την ευρέως δοκιμασμένη βάση δεδομένων MPEG7 CE-Shape-1 [256] η οποία αποτελείται από 1400 εικόνες περιγράμματος από 70 κατηγορίες. Κάθε κατηγορία έχει 20 διαφορετικά σχήματα. Ο ρυθμός αναγνώρισης μετράται από τον αποκαλούμενο *Έλεγχο Bullseye*: Κάθε σχήμα στη βάση δεδομένων, αντιστοιχίζεται με όλα τα άλλα σχήματα και μετρώνται τα 40 εγγύτερα ενδεχόμενα. Το πολύ 20 από τα 40 ενδεχόμενα μπορεί να είναι σωστές επιλογές. Η βαθμολογία προκύπτει σαν ο λόγος των σωστών επιλογών προς τον υψηλότερο δυνατό αριθμό επιτυχιών (ο οποίος είναι 20×1400).

Με βάση το προτεινόμενο πλαίσιο εργασίας, η νέα μέθοδος, που είναι κατάλληλη για προβλήματα μεγάλης κλίμακας, μετρά την απόλυτη κατά σημείο διαφορά ανάμεσα στις GLS. Εάν $S_1(i, j), i, j = 1, 2, \dots, n$ και $S_2(i, j), i, j = 1, 2, \dots, n$ είναι οι διακριτές GLS δύο διαφορετικών σχημάτων, $c(i), c(j)$ μία αντιστοιχία ανάμεσα στις καμπύλες περιγραμμάτων που διατηρεί την τοπικότητα και C_{VAR} το σύνολο όλων αυτών των αντιστοιχιών που είναι 1 προς 1 και επί και απεικονίζουν τα τοπικά ακρότατα των VAR στις δύο καμπύλες ως σημεία εκκίνησης για τον υπολογισμό των GLS, τότε η απόσταση ανάμεσα στα σχήματα υπολογίζεται με βάση τον τύπο:

$$\min_{c \in C_{VAR}} \left\{ \frac{1}{n^2} \sum_{i,j} |S_1(i, j) - S_2(c(i), c(j))| \right\} \quad (35)$$

Θα συγκρίνουμε αυτή τη μέθοδο, που αποκαλούμε GLS/ VAR, με την μέθοδο Shape Contexts. Η Shape Contexts εισήχθη από τους Belongie et al [167]. Περιγράφει τη σχετική κατανομή στο χώρο (απόσταση και προσανατολισμός) των σημείων της καμπύλης. Για n σημεία περιγράμματος p_1, p_2, \dots, p_n ο περιγραφέας Shape Context στο σημείο p_i

ορίζεται ως ένα ιστόγραμμα h_i των σχετικών συντεταγμένων των υπόλοιπων $n - 1$ σημείων

$$h_i(k) = \#\{p_j : j \neq i; p_j - p_i \in \text{bin}(k)\} \quad (36)$$

όπου τα bins διαιρούν ομοιόμορφα τον λογαριθμικό χώρο πολικών συντεταγμένων log-polar space. Η απόσταση ανάμεσα στα δύο ιστογράμματα ορίζεται χρησιμοποιώντας τον στατιστικό τύπο χ^2 . Για τη σύγκριση σχημάτων, οι Belongie et al χρησιμοποιήσαν ένα πλαίσιο εργασίας που συνδυάζει την Shape Contexts και την Thin-Plate-Splines (SC+TPS) [167]. Για τα σημεία σε δύο σχήματα A και B , πρώτα βρίσκονται οι αντιστοιχίες σημείου προς σημείο μέσω σταθμισμένης διμερούς αντιστοίχισης. Έπειτα η TPS χρησιμοποιείται επαναληπτικά για την εκτίμηση του μετασχηματισμού ανάμεσά τους. Τέλος, η ομοιότητα D ανάμεσα στα σχήματα A και B μετράται ως ένας σταθμισμένος συνδυασμός τριών μερών: $D = aD_{ac} + D_{sc} + bD_{be}$ όπου το D_{ac} μετρά τη διαφορά στην εμφάνιση, το D_{be} μετρά την ενέργεια κάμψης και D_{sc} , είναι η απόσταση Shape Context που μετρά την μέση απόσταση ανάμεσα σε κάθε σημείο του A και του πιο όμοιου αντίστοιχου του στο B (υπό την έννοια της σύγκρισης των ιστογραμμάτων τους με τη στατιστική συνάρτηση χ^2) και τα a, b είναι βάρη. Οι Belongie et al απέδειξαν ότι η SC+TPS είναι πολύ αποτελεσματική για εργασίες αντιστοίχισης σχημάτων. Η μέθοδος Shape Context είναι αρκετά δημοφιλής και έχουν προταθεί διάφορες τροποποιήσεις της που ευνοούν την ταχύτητα ή τη δυνατότητα διάκρισης για διαφορετικούς τύπους σχημάτων [257, 258, 259, 260, 261, 262].

Η κατασκευή ενός πίνακα αποστάσεων για το σύνολο σχημάτων MPEG7 CE-Shape-1, απαραίτητο για τον έλεγχο bullseye, απαιτεί τον υπολογισμό $1400 \times 1399/2 = 979300$ αποστάσεων ανα ζεύγη σχημάτων. Εάν για κάθε υπολογισμό απόστασης σχημάτων χρειάζεται ένα δευτερόλεπτο, χρειάζονται 272,3 ώρες ή αλλιώς περίπου 11 ημέρες για να ολοκληρωθεί ολόκληρος ο έλεγχος. Όταν χρησιμοποιείται κάποια μορφή δυναμικού προγραμματισμού για την αντιστοίχιση κάθε ζεύγους σχημάτων, αυτό μεταφράζεται σε μία περίοδο πολλών ημερών για να ολοκληρωθεί ο έλεγχος σε έναν υπολογιστή με συνηθισμένη διάταξη.

Εν όψη αυτών των περιορισμών, επιλέγουμε πειράματα με αυτήν την μεγάλη βάση σχημάτων διότι θα αποκαλύψει καλύτερα τα πλεονεκτήματα ως προς την πολυπλοκότητα και τον χρόνο εκτέλεσης της μεθόδου GLS/VAR. Η SC+TPS [167] χρησιμοποιεί τον Ουγκρικό αλγόριθμο για την αντιστοιχία ο οποίος έχει πολυπλοκότητα $O(n^3)$, ενώ ο δυναμικός προγραμματισμός έχει πολυπλοκότητα $O(n^2)$ (δεδομένου ότι η αναζήτηση για την αρχική αντιστοιχία δεν συνδέεται απευθείας με τον αριθμό των σημείων της καμπύλης), επομένως η SC+TPS είναι ακόμα πιο αργή από τις προσεγγίσεις που χρησιμοποιούν δυναμικό προγραμματισμό για την αντιστοίχιση. Για το λόγο αυτό, εφαρμόζουμε μία απλή εκδοχή της SC που είναι σημαντικά πιο γρήγορη και πιο κοντά στην ταχύτητα της μεθόδου μας. Η απλή εκδοχή χρησιμοποιεί μόνον την D_{SC} για τη φάση αντιστοίχισης.

Αυτή η απλή εκδοχή του Shape Contexts που ονομάζουμε SC/SIMPLE έχει χρησιμοποιηθεί επίσης κι από άλλους συγγραφείς π.χ. [153] (σελίδα 25, τελευταία παράγραφος) ειδικά σε περιπτώσεις τόσο μεγάλης βάσης δεδομένων. Αυτή η απλή εκδοχή είναι επίσης χρήσιμη για τους σκοπούς σύγκρισης γιατί εξαλείφει το αποτέλεσμα της αντιστοίχισης και συμβάλει στην κατανόηση της διαφοράς στους περιγραφείς αυτούς καθεαυτούς. Η κλασική μέθοδος SC+TPS [167] δεν υλοποιήθηκε λόγω της υψηλής πολυπλοκότητάς της και των περιορισμών στο χρόνο εκτέλεσης αλλά η βαθμολογία bullseye της στο σύνολο σχημάτων MPEG7 έχει αναφερθεί από άλλους συγγραφείς [167, 153] και έχει συμπεριληφθεί στα αποτελέσματα του Πίνακα 1. Για τη μέθοδο GLS, παρουσιάζουμε επίσης τις ανάλογες με τις εφαρμογές της προηγούμενης ενότητας σχετικά με τις αντιστοιχίες στην αναζήτηση της καλύτερης αντιστοίχισης ανάμεσα σε δύο σχήματα. Έτσι GLS/ VAR σημαίνει κατά σημείο αντιστοίχιση με τη VAR να επιλέγει τα σημεία εκκίνησης των GLS (τοπικά ακρότατα της VAR), GLS / 5P σημαίνει κατά σημείο αντιστοίχιση GLS με σημεία εκκίνησης των GLS για κάθε 5ο σημείο σε κάθε καμπύλη (σταθερός αριθμός συνολικά $20 \times 20 = 400$ αντιστοιχιών αρχικών σημείων εξετάζονται για κάθε αντιστοίχιση ζεύγους περιγραμμάτων, για περιγράμματα 100 σημείων).

Μέθοδος	Βαθμολογία	Πολυπλοκότητα	Χρόνος Εκτέλεσης
SC/SIMPLE	64.57%	$O(n^2)$	2.7 ώρες
SC+TPS[167]	76.51%	$O(n^3)$	>>4.37 ώρες
GLS/5P	73.57%	$O(n^3)$	4.37 ώρες
GLS/VAR	74.25%	$O(n^2)$	48.70 λεπτά

Πιν. 1: Βαθμολογίες Bullseye, πολυπλοκότητα και χρόνοι εκτέλεσης για διάφορες μεθόδους που χρησιμοποιούν τη βάση MPEG7 CE-Shape-1 Part B.

Στον Πίνακα 1, μπορούμε να δούμε τη βαθμολογία bullseye για τις παραπάνω μεθόδους και τον χρόνο που χρειάστηκε για την ολοκλήρωση κάθε ελέγχου χρησιμοποιώντας τον ίδιο υπολογιστή. Στην περίπτωση της SC+TPS, ο χρόνος καθορίζεται ως προς την πιο αργή από τις άλλες μεθόδους με βάση τη σύγκριση πολυπλοκότητάς τους. Βλέπουμε ότι η GLS/ VAR έχει περίπου την ίδια βαθμολογία με την κλασική SC+TPS αλλά πιο χαμηλή πολυπλοκότητα και σημαντικά μικρότερο χρόνο εκτέλεσης. Ταυτόχρονα, η GLS/ VAR έχει σημαντικά καλύτερη βαθμολογία από την απλή εφαρμογή της SC ενώ ο χρόνος εκτέλεσής της εξακολουθεί να είναι μικρότερος. Αυτή η βαθμολογία είναι σύμφωνη με τη βαθμολογία που αναφέρεται στο [153] για την ίδια μέθοδο. Σε αναλογία προς το αποτέλεσμα της προηγούμενης ενότητας, παρατηρούμε ότι επιτρέποντας στην VAR να επιλέξει τα σημεία εκκίνησης (αντιστοιχίες) μειώνεται ο συνολικός χρόνος και βελτιώνεται η βαθμολογία ταυτόχρονα. Όπως έχουμε ήδη εξηγήσει, αυτό είναι πιθανό γιατί περιορίζουμε τον χώρο αναζήτησης όλων των πιθανών αντιστοιχιών στις διαισθητικές μόνο, όπως αυτές

καθορίζονται από τα τοπικά ακρότατα της VAR. Εδώ φαίνεται η χρησιμότητα ενός διαισθητικά συμβατού περιγραφέα στην αναγνώριση σχημάτων. Άλλες τροποποιήσεις του Shape Contexts που χρησιμοποιούν δυναμικό προγραμματισμό και παράγουν ακόμα καλύτερες βαθμολογίες όταν ελέγχονται με το ίδιο σύνολο δεδομένων έχουν παρουσιαστεί π.χ. η Inner Distance (εσωτερική απόσταση) με χρήση δυναμικού προγραμματισμού για την αντιστοίχιση [153] αλλά η ταχύτητα οποιασδήποτε μεθόδου χρησιμοποιεί δυναμικό προγραμματισμό είναι μικρότερη από αυτή της GLS/5P. Επίσης, για να επιτευχθούν υψηλότερες βαθμολογίες, η εξομάλυνση ως ένα βήμα προεπεξεργασίας απαιτείται για αυτές τις εκδοχές του SC. Για μία δίκαια σύγκριση της GLS/VAR με αυτές τις μεθόδους θα χρειαζόταν: 1) να εγκαταλειφθεί η έννοια μη επεμβατικών περιγραφέων που δεν τροποποιούν το σχήμα και 2) να καθοριστεί μία προσέγγιση δυναμικού προγραμματισμού για την αντιστοίχιση των GLS.

Το δεύτερο θα οδηγούσε σε αύξηση του χρόνου εκτέλεσης της μεθόδου μας σε συγκρίσιμα επίπεδα με τις άλλες τεχνικές που αναφέρθηκαν. Επιπλέον, όλα τα παραπάνω πειράματα υποδηλώνουν ότι η μεμονωμένη χρήση της VAR, για να περιοριστούν οι πιθανές επιλογές των σημείων εκκίνησης στον δυναμικό προγραμματισμό, μπορεί να βελτιώσει την πολυπλοκότητα, τον χρόνο εκτέλεσης και την ακρίβεια όλων των μεθόδων που χρησιμοποιούν τον δυναμικό προγραμματισμό για την αντιστοίχιση σημείων περιγράμματος. Το συμπέρασμα αυτό καθιστά τη VAR σημαντική για τη μείωση της πολυπλοκότητας σε άλλες μεθόδους της τρέχουσας στάθμης της τεχνικής. Στον Πίνακα 2, συγκρίνονται οι δύο εφαρμογές της GLS ως προς τον τρόπο επιλογής των σημείων εκκίνησης στη φάση αντιστοίχισης. Μπορούμε να δούμε τον συνολικό αριθμό των σωστών αντιστοιχιών για όλες τις πιθανές 1400 αντιστοιχίσεις για τις 10 σειρές με την καλύτερη αντιστοιχία του συνόλου Σχημάτων της βάσης MPEG7.

Ο αριθμός των επιτυχιών που εμφανίζονται στην δεύτερη και την τρίτη στήλη υποδηλώνουν ότι η πιο αργή από τις δύο εφαρμογές, για κάθε 5ο σημείο αντιστοιχιών σημείων εκκίνησης, έχει χειρότερες βαθμολογίες ανάκτησης παρά το γεγονός ότι εξετάζει περισσότερους συνδυασμούς από την ταχύτερη εφαρμογή η οποία εξετάζει τις αντιστοιχίες των σημείων εκκίνησης ανάμεσα σε δύο καμπύλες μόνον στα τοπικά ακρότατα των VAR τους.

4.8.5. GLS/VAR έναντι SC ως προς την πολυπλοκότητα και την ταχύτητα εκτέλεσης

Σε αυτή την ενότητα, εξετάζουμε την πολυπλοκότητα και τον χρόνο εκτέλεσης για τη φάση κωδικοποίησης και τη φάση αντιστοίχισης. Η πολυπλοκότητα και ο χρόνος εκτέλεσης της φάσης κωδικοποίησης είναι λιγότερο σημαντικά εφόσον αυξάνονται γραμμικά με τον αριθμό των σχημάτων στη βάση δεδομένων. Από την άλλη μεριά η πολυπλοκότητα και ο χρόνος εκτέλεσης για την φάση αντιστοίχισης ενός ζεύγους σχημάτων είναι σημαντικά γιατί αυξάνονται τετραγωνικά με τον αριθμό των σχημάτων στη βάση δεδομένων (αφού όλα

τα σχήματα πρέπει να συγκριθούν μεταξύ τους). Για το λόγο αυτό, χωρίσαμε αυτές τις δύο φάσεις στο πείραμα και επιχειρήσαμε να μετακινήσουμε όσες περισσότερες εργασίες γινόταν στην φάση κωδικοποίησης. Οι GLS, τα τοπικά ακρότητα της VAR, τα ιστογράμματα SC για όλα τα απαραίτητα σημεία και παραμετροποιήσεις υπολογίστηκαν, λοιπόν, για κάθε σχήμα στη φάση κωδικοποίησης.

ΤΑΞΗ	GLS/VAR	GLS/5P
1	1338	1331
2	1287	1277
3	1244	1207
4	1146	1130
5	1080	1065
6	1024	1028
7	1009	963
8	975	927
9	923	923
10	870	856

Πιν. 2: Παρουσιάζεται ο αριθμός των σωστών αντιστοιχίσεων από όλες τις πιθανές 1400 αντιστοιχίσεις σύμφωνα με το πείραμα της ενότητας 4.8.4. Στην πρώτη στήλη υπάρχει η σειρά της αντιστοιχίας (1η πιο κοντινή, 2η πιο κοντινή κ.λ.π.), στη δεύτερη και τρίτη στήλη υπάρχει ο αριθμός των σωστών αντιστοιχιών για κάθε διαβάθμιση για τις μεθόδους GLS/VAR και GLS/5P αντίστοιχα. Σημειώνουμε ότι η επιλεκτική – με βάση τις διαισθητικές ιδιότητες του σχήματος – μείωση του χώρου αναζήτησης για αντιστοιχίσεις των σημείων εκκίνησης μέσω της VAR δεν μειώνει μόνον την πολυπλοκότητα και βελτιώνει τον χρόνο εκτέλεσης (Πίνακας 1) αλλά οδηγεί επίσης και σε καλύτερες βαθμολογίες αντιστοιχίσης. Το αποτέλεσμα αυτό συμφωνεί με το πείραμα του Σχ. 16.

Η πολυπλοκότητα στον υπολογισμό των απαραίτητων για επακόλουθη αντιστοίχιση GLS και των τοπικών ακροτάτων της VAR για ένα συγκεκριμένο σχήμα έχει εξεταστεί στην υποενότητα 4.8.3 και για κάθε GLS και VAR είναι της τάξης $O(n^2)$, όπου n ο αριθμός των σημείων του περιγράμματος. Συνεπώς για τις k GLS που αντιστοιχούν με τα αντίστοιχα k τοπικά ακρότητα της VAR, η πολυπλοκότητα είναι $O(kn^2)$. Ο υπολογισμός των ιστογραμμάτων SC για ένα δεδομένο σχήμα είναι επίσης της τάξης $O(n^2)$ αφού για κάθε σημείο στο σχήμα, υπολογίζεται ένα ιστόγραμμα από τη σχετική θέση όλων των υπόλοιπων σημείων. Παρόλο που ο υπολογισμός των GLS και των ιστογραμμάτων SC έχει τελικά την ίδια πολυπλοκότητα (εάν χρησιμοποιηθεί η VAR για αντιστοιχίες σημείων εκκίνησης για τη μέθοδο GLS), για τον υπολογισμό των ιστογραμμάτων απαιτούνται περισσότερες πράξεις και συνεπώς χρειάζεται μεγαλύτερος χρόνος εκτέλεσης, αφού η κατασκευή των ιστογραμμάτων περιλαμβάνει τον μετασχηματισμό σε λογαριθμικό χώρο πολικών συντε-

ταγμένων log polar space και τον υπολογισμό όχι μόνον των σχετικών αποστάσεων (όπως συμβαίνει στην GLS) αλλά και των γωνιών. Αυτή η διαφορά στον χρόνο εκτέλεσης δεν είναι ωστόσο κρίσιμη στη φάση κωδικοποίησης εφόσον αυξάνεται γραμμικά με τον αριθμό των σχημάτων στη βάση δεδομένων.

Τώρα, ως προς τη φάση αντιστοίχισης, η πολυπλοκότητα της αντιστοίχισης ενός ζεύγους σχημάτων με τη μέθοδο GLS/VAR είναι $O(kn^2) = O(n^2)$, n ο αριθμός των σημείων περιγράμματος. Για να βρεθεί η απόσταση σχήματος ανάμεσα σε ένα ζεύγος σχημάτων, οι GLS πρέπει να αφαιρεθούν σημειακά για όλους τους συνδυασμούς (k) των τοπικών ακροτάτων VAR και στα δύο σχήματα. Συνεπώς οι λειτουργίες που χρειάζονται είναι kn^2 προσθέσεις, όπου k είναι κατά μέσο όρο ίσο με 25 (5 τοπικά ακρότατα για κάθε VAR σε κάθε ζεύγος) και n ο αριθμός των σημείων περιγράμματος. Για την SC/SIMPLE απόσταση ανάμεσα σε ένα ζευγάρι σχημάτων, η πολυπλοκότητα είναι πάλι $O(n^2)$ αφού πρέπει να αντιστοιχιστούν n^2 συνδυασμοί ιστογραμμάτων. Εάν χρησιμοποιηθούν οι τυπικές παράμετροι των 96 bins (8 bins για τις αποστάσεις και 12 bins για τις γωνίες) για κάθε ιστογράμμα, οι λειτουργίες που χρειάζονται για την αντιστοίχιση ενός ζεύγους ιστογραμμάτων με την στατιστική συνάρτηση χ^2 είναι 96 δυνάμεις + 96 × 2 προσθέσεις + 96 διαιρέσεις, έτσι για το σύνολο των n^2 αντιστοιχίσεων ιστογράμματος, θα χρειαστούν $n^2(96$ δυνάμεις + 96 × 2 προσθέσεις + 96 διαιρέσεις). Αυτό σε σύγκριση με τις kn^2 προσθέσεις που αναλαμβάνει η GLS/ VAR για την ίδια εργασία είναι σημαντικά πιο αργό γιατί το k είναι κατά μέσο όρο περίπου 25 (5 τοπικά ακρότατα για κάθε σχήμα). Η κατάσταση με τις άλλες μεθόδους που ελέγχθηκαν και δεν χρησιμοποιούν τη VAR για την επιλογή των αντιστοιχιών των σημείων εκκίνησης είναι ακόμα χειρότερη αφού η πολυπλοκότητα αυξάνεται, όπως εξηγήσαμε πριν, σε $O(n/5 \times n^2) = O(n^3)$ (για αντιστοιχίες κάθε 5ου σημείου). Η διαφορά στον χρόνο εκτέλεσης είναι κρίσιμη στη φάση σύγκρισης εφόσον αυξάνεται τετραγωνικά με τον αριθμό των σχημάτων στη βάση δεδομένων (1400^2 στην περίπτωσή μας). Έτσι εξηγείται η διαφορά στην ταχύτητα των μεθόδων του Πίνακα 1 και υποδηλώνει το μεγάλο πλεονέκτημα της GLS/ VAR ως προς την πολυπλοκότητα και τον χρόνο εκτέλεσης για τη δεδομένη βαθμολογία bullseye που επιτυγχάνει.

Για τη φάση αντιστοίχισης του παραπάνω πειράματος με τη βάση MPEG, χρησιμοποιήθηκε ένα τετραπύρρηνο μηχάνημα με 4GB μνήμης μόνο για αυτό το σκοπό. Ο αλγόριθμος αντιστοίχισης γράφηκε χρησιμοποιώντας πρότυπο ANSI C και συντάχθηκε με το πρόγραμμα gcc compiler. Όλα τα παραπάνω προϋποθέτουν την εκτέλεση ακολουθιακών αλγορίθμων. Εάν θεωρήσουμε διανυσματική/πολυνηματική εκτέλεση τότε τα παραπάνω συμπεράσματα σχετικά με τον ακολουθιακό χρόνο μεταφέρονται σε απαιτήσεις μνήμης ή/και απαιτήσεις υλικού. Εάν για παράδειγμα, είναι δυνατή η χρήση παράλληλων υπολογισμών, τότε ο υπολογισμός των κατά σημείο διαφορών για όλα τα ζεύγη GLS μπορεί να εκτελεστεί ταυτόχρονα και επομένως θα χρειαστούν k διεργασίες/νήματα για την αντιστοίχιση ενός ζεύγους σχημάτων. Για την παράλληλη εκτέλεση του SC/SIMPLE, εάν

υποθέσουμε ότι όλες οι στατιστικές συγκρίσεις χ^2 ανάμεσα σε όλα τα ζεύγη ιστογραμμάτων πραγματοποιηθούν ταυτόχρονα, θα χρειαζόταν n^2 διεργασίες/νήματα για τον υπολογισμό της απόστασης ανάμεσα σε ένα ζεύγος σχημάτων και επομένως περισσότεροι υπολογιστικοί πόροι. Άλλες μέθοδοι που χρησιμοποιούν δυναμικό προγραμματισμό δεν έχουν εμφανή παραλληλισμό, αφού η βέλτιστη διαδρομή πρέπει να δημιουργηθεί ακολουθιακά.

4.9. Κωδικοποίηση Ολικών Μορφομετρικών Χαρακτηριστικών στα Τοπικά Μέτρα της Ολοτοπικής Επιφάνειας

Σε αυτήν την παράγραφο θα δειχθεί ότι τοπικοί περιγραφείς της ολοτοπικής επιφάνειας αντιστοιχούν σε ολικούς μορφομετρικούς περιγραφείς του αντίστοιχου περιγράμματος και πως είναι δυνατόν μέσω των τοπικών περιγραφών να εντοπισθούν και να μετρηθούν οι μορφομετρικές αυτές περιοχές.

Τώρα θα δειχθεί ότι τα τοπικά ακρότατα της ολοτοπικής επιφάνειας αντιστοιχούν σε μορφομετρικές περιοχές του περιγράμματος και συγκεκριμένα σε 'διατάσεις' και 'συρρικνώσεις' της μορφής του περιγράμματος. Ορίζονται αυτά τα κρίσιμα σημεία και επεκτείνοντας την έννοια των κρίσιμων σημείων σε αυτήν των *κρίσιμων περιοχών* παρέχεται ένας τρόπος μέτρησης αυτών των μορφομετρικών χαρακτηριστικών του περιγράμματος, μέσω των τοπικών μέτρων της ολοτοπικής επιφάνειας. Για την γενική περίπτωση των καμπύλων m διαστάσεων, ο ολοτοπικός μετασχηματισμός επεκτείνεται ως εξής:

Ας είναι $[0, \lambda] \subset \mathbb{R}$ και $\alpha : [0, \lambda] \rightarrow \mathbb{R}^m$ μία συνεχής 1-1 και άπειρες φορές διαφορίσιμη *καμπύλη* στο \mathbb{R}^m . Ας είναι επίσης $\alpha(0) = \alpha(\lambda)$, έτσι ώστε α μία κλειστή καμπύλη στο \mathbb{R}^m .

Εφόσον α συνεχής, 1-1 και επί του $\alpha([0, \lambda]) \subset \mathbb{R}^m$ είναι ισομορφισμός και μεταφέρει την ολική διάταξη του $[0, \lambda] \subset \mathbb{R}$ στο $\text{Im}_\alpha([0, \lambda]) \subset \mathbb{R}^m$. Θεωρήστε το \mathbb{R}^m εφοδιασμένο με την συνήθη μετρική που προκύπτει από την συνήθη νόρμα $\|\cdot\|_m$, $\text{Im}_\alpha([0, \lambda]) := I_\alpha$ είναι το γράφημα της κλειστής καμπύλης α και η διάμετρος του στο \mathbb{R}^m είναι

$$\Delta = \sup_{x, y \in I_\alpha} \|x - y\|_m$$

Θεωρήστε και το γράφημα εφοδιασμένο με την συνήθη μετρική που προκύπτει από την νόρμα $\|\cdot\|_1 = |\cdot|$ (απόλυτη τιμή) στο ισομορφικό του $[0, \lambda]$.

Ας είναι τώρα x_0 ένα σημείο του γραφήματος της καμπύλης α και ας θεωρήσουμε την συνάρτηση v'_{x_0} ως: $v'_{x_0} : I_\alpha \rightarrow [0, \Delta], x \mapsto v_{x_0}(x) := \|x - x_0\|_m$.

Η v'_{x_0} είναι καλά ορισμένη για x_0 σαν τον περιορισμό της d_m στο γράφημα της α , και άρα συνεχής σαν τέτοια. Ορίζεται τώρα σαν v_{x_0} , η σύνθεση της α και της v'_{x_0} ως εξής:

Ορισμός 4 $v_{x_0} := (v'_{x_0} \circ \alpha)$, συνεπώς,

$$v_{x_0} : [0, \lambda] \rightarrow [0, \Delta], x \mapsto v_{x_0}(x) := v'_{x_0}(\alpha(x))$$

Η v_{x_0} καλείται *συνάρτηση άποψης* η απλά *άποψη* του γραφήματος I_α από το σημείο x_0 .

Η v_{x_0} είναι καλά ορισμένη για κάθε x_0 . Είναι επίσης συνεχής σαν σύνθεση της συνεχούς καμπύλης α και της συνεχούς μετρικής v'_{x_0} .

Έστω $t \in [0, \lambda]$ και $\alpha(t) = x$.

Ορισμός 5 Έστω $[0, \lambda] \subset \mathbb{R}$ και $\alpha : [0, \lambda] \rightarrow \mathbb{R}^m$ μία συνεχής και άπειρες φορές διαφορίσιμη καμπύλη του \mathbb{R}^m . Έστω επίσης ότι $\alpha(0) = \alpha(\lambda)$, έτσι ώστε α είναι μία κλειστή καμπύλη του \mathbb{R}^m . Για κάθε $t_1 \in [0, \lambda]$ οι συναρτήσεις απόψεων είναι οι $v_{\alpha(t_1)}(t_2)$. Η συνάρτηση $S_\alpha(t_1, t_2) = v_{\alpha(t_1)}(t_2)$, $t_1, t_2 \in [0, \lambda]$ είναι ο *ολοτοπικός μετασχηματισμός* της καμπύλης α . Η S_α ορίζει μία επιφάνεια $\{t_1, t_2, S_\alpha(t_1, t_2)\}$ η οποία παράγεται από τις συναρτήσεις άποψης $v_{\alpha(t)}$, $t \in [0, \lambda]$:

$$S_\alpha = \bigcup_{t \in [0, \lambda]} v_{\alpha(t)}$$

Η επιφάνεια αυτή καλείται *ολοτοπική επιφάνεια* της κλειστής καμπύλης α και είναι προφανές ότι

$$S_\alpha \subset [0, \lambda] \times [0, \lambda] \times [0, \Delta].$$

Έστω $\alpha \in C_c^\infty([0, \lambda], \mathbb{R}^m)$, μια κλειστή, συνεχής, άπειρες φορές διαφορίσιμη καμπύλη του \mathbb{R}^m , $v_x(t)$ οι απόψεις, $S_\alpha(x, y)$ η ολοτοπική επιφάνεια και $H_\alpha = \begin{vmatrix} \frac{\partial S_\alpha^2}{\partial x^2} & \frac{\partial S_\alpha^2}{\partial x \partial y} \\ \frac{\partial S_\alpha^2}{\partial y \partial x} & \frac{\partial S_\alpha^2}{\partial y^2} \end{vmatrix}$ ο Hessian της S_α . Έστω επίσης $t_1, t_2 \in [0, \lambda]$ και $x_1 = \alpha(t_1)$, $x_2 = \alpha(t_2)$, $x_1, x_2 \in I_\alpha$.

Ορισμός 6 Το σημείο (t_1, t_2) καλείται *ϵ - κρίσιμο* αν και μόνον αν

$$\left| \frac{\partial S_\alpha}{\partial t_1}(t_1, t_2) \right| \leq \epsilon, \epsilon \geq 0$$

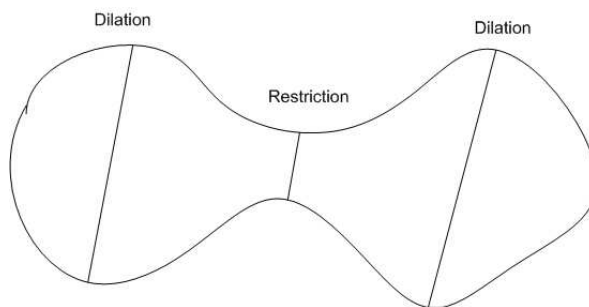
και

$$\left| \frac{\partial S_\alpha}{\partial t_2}(t_1, t_2) \right| \leq \epsilon, \epsilon \geq 0$$

και

$$H_\alpha(t_1, t_2) > 0$$

Εδώ επεκτείνεται η έννοια των τοπικών μεγίστων και ελαχίστων της ολοτοπικής επιφάνειας με τρόπο ώστε τα κρίσιμα σημεία δεν είναι μόνον τα σημεία τοπικών μεγίστων και ελαχίστων



Σχήμα 19: Ένα σχήμα με μία συρρίκνωση και δύο διατάσεις.

αλλά και όλα τα σημεία σε μία περιοχή γύρω τους, το εύρος της οποίας καθορίζεται από το ϵ .

Αν είναι τώρα (t_1, t_2) ένα ϵ κρίσιμο σημείο, το ζεύγος (x_1, x_2) με $x_1 = \alpha(t_1)$ και $x_2 = \alpha(t_2)$ είναι επίσης ϵ κρίσιμα και ορίζουν μία μορφομετρική περιοχή του περιγράμματος, συγκεκριμένα μία *διάταση* ή μία *συρρίκνωση* της μορφής του περιγράμματος. Στο σχήμα 19 φαίνεται ένα περίγραμμα με μία συρρίκνωση και δύο διατάσεις. Τα τοπικά μέτρα λοιπόν της ολοτοπικής επιφάνειας ορίζουν ολικούς μορφομετρικούς περιγραφείς πάνω στο περίγραμμα. Για να αντιστοιχηθούν αριθμητικά μεγέθη στους περιγραφείς αυτούς και έτσι να μετρηθούν τα μορφώματα του περιγράμματος, χρησιμοποιούνται κατάλληλα εμβαδά που οριοθετούνται από την ολοτοπική επιφάνεια στις περιοχές των ϵ -κρίσιμων σημείων.

Ορισμός 7 Έστω $C_\epsilon(\alpha)$ το σύνολο όλων των ϵ -κρίσιμων σημείων του περιγράμματος α και x ένα σημείο του περιγράμματος. Το ολοτοπικό μέτρο του περιγράμματος στο σημείο x είναι:

$$GL_\alpha(x) = \int_{(x,y) \in C_\epsilon(\alpha)} k(x,y) \|x-y\|_m dy$$

όπου $k(x,y)$ κατάλληλη συνάρτηση που εξασφαλίζει ότι η ποσότητα $\|x-y\|_m$ προσμετράται στον υπολογισμό του ολοτοπικού μέτρου μόνον εφόσον το ευθύγραμμο τμήμα που ενώνει τα x και y κείται στο εσωτερικό του περιγράμματος. Αυτό είναι απαραίτητο για να είναι το ολοτοπικό μέτρο αντιληπτικά συμβατό μετρώντας τα κυρτά μέρη των μορφών που σχηματοποιούνται από τα τμήματα του περιγράμματος.

Μέσω του ολοτοπικού μέτρου ολικές μορφομετρικές περιγραφές του περιγράμματος μετρώνται και κωδικοποιούνται στα σημεία των αντιστοίχων τμημάτων του περιγράμματος που μετέχουν στις μορφές αυτές. Όπως έχουμε ήδη αναφέρει αυτή η αντιστοιχία ανάμεσα στην μορφολογία του περιγράμματος και των σημείων του περιγράμματος που μετέχουν στην μορφολογία αυτή είναι σημαντική διότι εκτός από την αντιληπτική συμβατότητα των μετρήσεων που αντιστοιχούν στα σημεία αυτά, επιτυγχάνεται η *μεταφορά* μορφομετρικών χαρακτηριστικών από το εσωτερικό στο περίγραμμα της μορφής, που εξαιτίας της ολικής διάταξης των σημείων του επιτρέπει με μεθόδους σχετικά χαμηλής

πολυπλοκότητας την ανίχνευση βέλτιστης ανταπόκρισης με σημεία άλλου περιγράμματος και τον ορισμό μορφομετρικά συμβατής απόστασης μεταξύ τους. Στα πειραματικά αποτελέσματα της επόμενης παραγράφου χρησιμοποιούνται τεχνικές δυναμικού προγραμματισμού για την εύρεση της απόστασης ανάμεσα σε διαφορετικά περιγράμματα και φαίνεται ότι το ολοτοπικό μέτρο επιτυγχάνει εξαιρετικά αποτελέσματα στην αναγνώριση σχήματος.

4.9.1. Ταξινόμηση Σχημάτων με το Ολοτοπικό Μέτρο

Σε αυτήν την παράγραφο χρησιμοποιείται μία προσέγγιση δυναμικού προγραμματισμού παρόμοια με αυτήν του [251] για να μετρηθεί η απόσταση ανάμεσα σε σχήματα από την βάση σιλουετών του Kimia χρησιμοποιώντας το ολοτοπικό μέτρο για την κατά ζεύγος αντιστοίχιση των σημείων. Ας είναι C_1 και C_2 δύο περιγράμματα. Δεδομένης μίας αντιστοίχισης c ανάμεσα στα σημεία του C_1 και του C_2 έτσι ώστε $c(C_1(t_1)) = C_2(t_2)$ και C το σύνολο όλων των δυνατών τέτοιων αντιστοιχήσεων, για να μετρηθεί η απόσταση μεταξύ δύο περιγραμμάτων πρέπει να ελαχιστοποιηθεί το ενεργειακό συναρτησοειδές:









































































$$E(C_1, C_2, c) = \int_{x \in C_1} |GL_{C_1}(x) - GL_{C_2}(c(x))| dx$$

σε σχέση με την αντιστοίχιση c . Η *ολοτοπική απόσταση* συνεπώς ανάμεσα στα περιγράμματα C_1 και C_2 ορίζεται από:

$$\|C_1 - C_2\|_{GL} = \min_{c \in C} E(C_1, C_2, c)$$

Τα περιγράμματα ψηφιοποιούνται παίρνοντας 100 ισο-απέχοντα σημεία στο κάθε ένα. Στην συνέχεια βρίσκεται μία αρχική αντιστοίχιση ανάμεσα σε ένα ζεύγος σημείων των περιγραμμάτων εκτιμώντας την διάμετρο του εσωτερικού χώρου που περικλείουν τα περιγράμματα. Η πολυπλοκότητα πάντως του προβλήματος επιτρέπει και εξωντωτικό έλεγχο για την βέλτιστη αρχική αντιστοίχιση, εάν τα αποτελέσματα μέσω της διαμέτρου δεν είναι ικανοποιητικά.

Στην συνέχεια η τεχνική δυναμικού προγραμματισμού εξετάζει όλες τις δυνατές κατά ζεύγη αντιστοιχίσεις ανάμεσα στα σημεία των περιγραμμάτων, η τήρηση της ολικής διάταξης των σημείων στα περιγράμματα κρατά την πολυπλοκότητα εφικτή καθώς η μέθοδος βρίσκει την βέλτιστη αντιστοίχιση ανάμεσα στα δύο περιγράμματα. Η επιλογή του ϵ στον υπολογισμό των ολοτοπικών μέτρων είναι καθοριστική στην ισορροπία ανάμεσα στην ταχύτητα του αλγορίθμου και την ακρίβεια των αποτελεσμάτων. Μία χαμηλή τιμή θα οδηγήσει σε λίγα κρίσιμα σημεία που ενδεχομένως δεν θα περιγράψουν επαρκώς τις διαφορές των σχημάτων, ενώ μία υψηλή τιμή θα επιβραδύνει σημαντικά τον υπολογισμό των ολοτοπικών μέτρων. Στα πειραματικά αποτελέσματα που παρουσιάζονται χρησιμοποιήθηκε η τιμή

		1st	2nd	3d		1st	2nd	3d
Distance	0	0.45	0.48	0.51	0	0.58	0.6	0.6
Image #	1	2	42	40	11	15	4	5
Image Pic								
Distance	0	0.45	0.58	0.59	0	0.58	0.64	0.64
Image #	2	1	41	40	15	11	8	9
Image Pic								
Distance	0	0.12	0.6	0.7	0	0.15	0.24	0.25
Image #	4	5	11	6	21	22	23	16
Image Pic								
Distance	0	0.12	0.6	0.71	0	0.39	0.49	0.52
Image #	5	4	11	9	31	33	34	32
Image Pic								
Distance	0	0.35	0.58	0.65	0	0.5	0.58	0.58
Image #	6	7	9	11	41	40	2	39
Image Pic								
Distance	0	0.35	0.6	0.75	0	0.24	0.32	0.36
Image #	7	6	9	5	51	44	52	50
Image Pic								
Distance	0	0.64	0.68	0.76	0	0.24	0.3	0.34
Image #	8	15	9	6	61	67	63	65
Image Pic								
Distance	0	0.58	0.6	0.6	0	0.35	0.38	0.45
Image #	9	6	7	11	91	96	90	89
Image Pic								
Distance	0	0.52	0.67	0.69	0	0.63	0.64	0.67
Image #	10	82	27	75	81	84	80	58
Image Pic								

Σχήμα 20: Σχήματα από την βάση του Kimia και τα τρία πιο κοντινά τους σύμφωνα με το ολοτοπικό μέτρο

0.008 η οποία φαίνεται να είναι ικανοποιητική για το συγκεκριμένο σύνολο σχημάτων. Στον πίνακα του σχήματος 20 παρουσιάζονται 18 σιλουέτες από την βάση δεδομένων του Kimia και συμπεριλαμβάνονται σχήματα από όλες τις κλάσεις. Για κάθε σιλουέτα βλέπουμε τις τρεις πιο κοντινές της στην βάση δεδομένων. Η πιο κοντινή φαίνεται υπό τον τίτλο 1st η δεύτερη πιο κοντινή με 2nd και η τρίτη πιο κοντινή με 3d. Αυτές είναι οι τρεις σιλουέτες που ανάμεσα από όλες τις 99 σιλουέτες της βάσης δεδομένων του Kimia έχουν τις τρεις μικρότερες ολοτοπικές αποστάσεις από την αντίστοιχη αρχική σιλουέτα. Φαίνεται ότι για όλες τις κλάσεις σιλουετών οι τρεις εγγύτερες είναι πάντα σωστές. Δεδομένου ότι οι σιλουέτες που παρουσιάζονται έχουν διαφορετικά μορφολογικά χαρακτηριστικά τα πειραματικά αποτελέσματά επιδεικνύουν ότι ο νέος ολοτοπικός μετασχηματισμός επιτυγχάνει εντυπωσιακά αποτελέσματα στην γενική αναγνώριση σχήματος.

5. ΜΑΡΚΟΒΙΑΝΗ ΣΗΜΑΣΙΟΛΟΓΙΚΗ ΤΑΞΙΝΟΜΗΣΗ

Αν και οι άνθρωποι τείνουν να συνδέουν τις εικόνες με υψηλού επιπέδου έννοιες, οι τρέχουσες τεχνικές τεχνητής όρασης εξάγουν συνήθως χαμηλού επιπέδου χαρακτηριστικά γνωρίσματα εικόνων και η σύνδεση μεταξύ των χαμηλού επιπέδου χαρακτηριστικών γνωρισμάτων και της υψηλού επιπέδου σημασιολογίας του περιεχομένου της εικόνας χάνονται. Ούτε ένα μοναδικό αλλά ούτε και ένας συνδυασμός χαμηλού επιπέδου χαρακτηριστικών γνωρισμάτων δεν έχει γενικά ρητή σημασιολογική έννοια. Επιπλέον, τα μέτρα ομοιότητας μεταξύ των οπτικών χαρακτηριστικών γνωρισμάτων απαραίτητως δεν ταυριάζουν με την ανθρώπινη αντίληψη [263] και, επομένως, τα αποτελέσματα ανάκτησης από μεθόδους χαμηλού επιπέδου δεν είναι γενικά ικανοποιητικά και συχνά απρόβλεπτα.

Για τον λόγο αυτό εδώ προτείνουμε ένα σύστημα αλληλεπίδρασης ανθρώπων-μηχανής-εικόνων που έχει την δυνατότητα ομαδικής εξόρυξης σημασιολογικών συσχετίσεων και διασύνδεσης τους με σκοπό την ομαδοποίηση χρηστών η την ταξινόμηση και ανάκτηση εικόνων. Στην επόμενη Ενότητα 5.1 πρώτα παρουσιάζεται το σύστημα για ομαδοποίηση χρηστών. Το ίδιο μοντέλο χρησιμοποιείται στην συνέχεια στην Ενότητα 5.2 για αυτόματη περιγραφή, ταξινόμηση και ανάκτηση εικόνων.

5.1. Ταξινόμηση Χρηστών Πληροφοριακών Συστημάτων

Με την χρήση ενός στοχαστικού μοντέλου αναπαριστούνται και κατηγοριοποιούνται οι παράμετροι χρηστών ενός συστήματος πρόσβασης πληροφοριών. Σε αυτό το μοντέλο, το γενικό καθώς και το ατομικό για κάθε χρήστη διάγραμμα μετάβασης συλλαμβάνεται υπό μορφή Μαρκοβιανής ανέλιξης όπου οι καταστάσεις είναι οι λέξεις κλειδιά που ρωτούνται προς το σύστημα από τους χρήστες και μια μετάβαση από μία κατάσταση σε μία άλλη αντιστοιχεί στην σειρά που εμφανίζονται οι λέξεις κλειδιά στις ερωτήσεις των χρηστών. Κάτω από αυτό το μοντέλο η πιθανολογική εγγύτητα των καταστάσεων της Μαρκοβιανής ανέλιξης μεταφράζεται σε σημασιολογική εγγύτητα των αντίστοιχων λέξεων κλειδιών με τρόπο ώστε μια ομάδα από μεταβατικά εγγύς καταστάσεις αντιστοιχεί σε μια σημασιολογική κατηγορία λέξεων κλειδιών. Δεδομένου ότι οι καταστάσεις αντιπροσωπεύουν τις λέξεις κλειδιά που ρωτούνται από τους χρήστες, ο χώρος καταστάσεων μπορεί να γίνει πολύ μεγάλος, αλλά συγχρόνως χωρίζεται σε υποσύνολα με ισχυρές αλληλεπιδράσεις μεταξύ των καταστάσεων των υποσυνόλων και αδύνατες αλληλεπιδράσεις μεταξύ των υποσυνόλων. Με την χρήση αυτής της δομής επιτυγχάνεται αποτελεσματική κατηγοριοποίηση του μεγάλου χώρου καταστάσεων και εν συνεχεία αποκάλυψη των αντίστοιχων σημασιολογικών συστάδων των λέξεων κλειδιών.

Θεωρώντας την ακολουθία των λέξεων κλειδιών, που ρωτώνται από τους χρήστες προς το πληροφοριακό σύστημα, ως μεταβάσεις μεταξύ Μαρκοβιανών καταστάσεων και υποθέτοντας σημασιολογική εξάρτηση για τις λέξεις κλειδιά στην ίδια ερώτηση, αναπτύσσεται ένα στοχαστικό μοντέλο πιθανότητας που συλλαμβάνει σημασιολογικές εξαρτήσεις μεταξύ των

λέξεων κλειδιών. Χτίζεται μια στοχαστική ανάλιξη όπου στις καταστάσεις αντιστοιχίζονται λέξεις κλειδιά που ρωτιούνται από τους χρήστες και μια μετάβαση μεταξύ καταστάσεων αντιστοιχεί στην ακολουθία αυτών των λέξεων κλειδιών στις ερωτήσεις. Αυτήν την ανάλιξη καλείται γενική ανάλιξη συμπεριφοράς. Οι τεχνικές που χρησιμοποιούνται στην συνέχεια είναι βασισμένες στην παροδική συμπεριφορά των *σχεδόν πλήρως διαχωρισμένων Μαρκοβιανών ανελίξεων* που διαμερίζουν τον χώρο καταστάσεων (λέξεις κλειδιά) στις ομάδες υψηλής ενδοομαδικής και χαμηλής διαομαδικής συνδεσιμότητας. Οι τεχνικές που χρησιμοποιούνται είναι αποτελεσματικές ακόμη και για πολύ μεγάλους χώρους καταστάσεων που συνήθως προκύπτουν από αυτά τα μοντέλα.

Αυτή η ομαδοποίηση τακτοποιεί τις λέξεις κλειδιά σε ομάδες κατά τρόπο ιεραρχικό και γι αυτό μέσω αυτής της ομαδοποίησης οι λέξεις κλειδιά κανονίζονται σε ομάδες με έναν τρόπο ώστε οι λέξεις κλειδιά μέσα σε μια ομάδα είναι πιο σχετικές η μία με την άλλη από ότι οι λέξεις κλειδιά που ανήκουν σε διαφορετικές ομάδες. Κατά συνέπεια αυτή η συγκέντρωση επιβάλλει μια αυτοπροσαρμοζόμενη αυτόματη οντολογία στον χώρο των λέξεων κλειδιών.

5.1.1. Το Μαρκοβιανό Μοντέλο για τις Λέξεις Κλειδιά

Η συμπεριφορά των χρηστών θα κατασκευαστεί χρησιμοποιώντας τις ερωτήσεις που ρωτούν προς το πληροφοριακό σύστημα. Το περιεχόμενο μιας ερώτησης είναι ο κατάλογος των λέξεων κλειδιών που περιέχει. Η συμπεριφορά ενός συγκεκριμένου χρήστη καθώς επίσης και η συμπεριφορά όλων των χρηστών συλλογικά θα αναπαρασταθούν μέσω διαγραμμάτων μεταβάσεων. Στο προτεινόμενο μοντέλο κάθε χρήστης αντιπροσωπεύεται με μια Μαρκοβιανή ανάλιξη που κατασκευάζεται από τις ερωτήσεις του χρήστη. Η συνολική συμπεριφορά όλων των χρηστών αντιπροσωπεύεται με μία άλλη Μαρκοβιανή ανάλιξη που είναι η *υπέρθυση* των ανελίξεων των επί μέρους χρηστών, όπου η υπέρθεση ορίζεται να είναι μία καινούργια ανάλιξη που έχει χώρο καταστάσεων την ένωση των χώρων καταστάσεων των επί μέρους ανελίξεων και πίνακα μετάβασης τον *κατά σημείο* μέσο όρο των πινάκων μετάβασης των επί μέρους ανελίξεων, αφού αυτοί προσαρμοστούν στην διάσταση της πληθυκότητας του χώρου καταστάσεων της υπέρθεσης με προσθήκη μηδενικών στα κατάλληλα σημεία. Η συνολική ανάλιξη χρησιμοποιείται για να σταθμίσει την απόσταση μεταξύ των λέξεων κλειδιών. Η Μαρκοβιανή ανάλιξη του κάθε χρήστη και η συνολική ανάλιξη κατασκευάζονται ταυτόχρονα με τρόπο παρόμοιο. Κάθε λέξη κλειδί αντιστοιχεί σε μία κατάσταση. Οι πιθανότητες μετάβασης μεταξύ των καταστάσεων μετρώνται αναπαριστώντας μια ερώτηση με αυτόν τον τρόπο.

Κάθε φορά που εμφανίζεται μια λέξη κλειδί σε μια ερώτηση, ο μετρητής της κατάστασής του αυξάνεται, και εάν υπάρχει μετάβαση μεταξύ καταστάσεων αυξάνεται και ο αντίστοιχος διακαταστατικός μετρητής συνδέσεων. Μετρώνται με αυτόν τον τρόπο και οι εμφανίσεις των

λέξεων κλειδιών αλλά και οι αλληλουχίες αυτών των εμφανίσεων. Γίνεται μαζική επεξεργασία πλήθους g λέξεων κλειδιών, αυξάνονται οι μετρητές και ενημερώνονται οι πιθανότητες ως ανωτέρω. Προτού γίνει η επεξεργασία του επόμενου συνόλου g λέξεων κλειδιών, οι μετρητές καθαρίζονται. Εάν ο χρήστης u ρωτά τη λέξη κλειδί k_j αμέσως μετά από τη λέξη κλειδί k_i και αυτό εμφανίστηκε k φορές η πιθανότητα μετάβασης p_{ij} ενημερώνεται ως εξής: Εάν p_{ij} είναι η τρέχουσα πιθανότητα (πριν από την αναπροσαρμογή) βασισμένη σε m λέξεις κλειδιά, η νέα πιθανότητα (βασισμένη σε $m + g$ λέξεις κλειδιά) είναι:

$$p_{ij} = \frac{mp_{ij} + k}{g + m}$$

Η αναδρομική αυτή σχέση ορίζει μία Μαρκοβιανή ανέλιξη. Το σύστημα επεξεργάζεται ξεχωριστά τις ερωτήσεις για κάθε χρήστη έτσι ώστε, εκτός από την γενική ανέλιξη να κατασκευαστεί από τον ίδιο αναδρομικό τύπο και μία ανέλιξη για κάθε χρήστη του συστήματος ξεχωριστά. Το διάνυσμα κατάστασης ισορροπίας της Μαρκοβιανής ανέλιξης που προκύπτει για κάθε χρήστη u_i συμβολίζουμε με p_{i_i} και θα αναπαριστά τον χρήστη. Στην Ενότητα 5.1.6 θα χρησιμοποιηθεί αυτή η αναπαράσταση για να ορισθεί η απόσταση ανάμεσα στους χρήστες σαν μία γενίκευση της απόστασης ανάμεσα στις λέξεις κλειδιά.

Η αξία της αναπαράστασης μιας ερώτησης ως μετάβαση μεταξύ καταστάσεων και της συμπεριφοράς των χρηστών ως Μαρκοβιανές ανελιξεις θα γίνει σαφής μόλις ορισθεί η σημασιολογική απόσταση μεταξύ των λέξεων κλειδιών που υπονοείται από την συνολική Μαρκοβιανή ανέλιξη.

Στο επόμενο τμήμα παρουσιάζονται βασικές έννοιες της θεωρίας Μαρκοβιανών ανελιξεων, κάποιοι συμβολισμοί εισάγονται επίσης. Σε όλο αυτό το έγγραφο τα έντονα γράμματα είτε κεφαλαία είτε μικρά είναι διανύσματα, τα κεφαλαία αλλά όχι έντονα γράμματα είναι πίνακες ενώ το N και τα κανονικά μικρά γράμματα είναι απλοί αριθμοί. Ένα έντονο γράμμα με δείκτη για παράδειγμα είναι ένα διάνυσμα στήλη όπου η θέση της στήλης μέσα στον πίνακα υπονοείται μέσω του δείκτη. Όλα τα διανύσματα είναι διανύσματα στήλες εκτός από το διάνυσμα ισορροπίας που παραδοσιακά θεωρείται διάνυσμα γραμμή.

5.1.2. Η Παροδική Συμπεριφορά Σύγκλισης

Ας είναι $P = \{p_{ij}\}$ ο πίνακας μεταβάσεων μιας Μαρκοβιανής ανέλιξης N καταστάσεων. Μία αλυσίδα είναι ένα είδος συλλογικής παγίδας, ένα σύνολο καταστάσεων όπου η ανέλιξη παγιδεύεται εκεί όταν εισέλθει σε κάποια από τις καταστάσεις της ομάδας. Μία μεταβατική (transient) κατάσταση δεν μπορεί να είναι μέλος μιας αλυσίδας αλλά μπορεί να συνδεθεί με μια εάν είναι δυνατό η ανέλιξη να εισαχθεί σε εκείνη την αλυσίδα από εκείνη την μεταβατική κατάσταση. Όλες οι καταστάσεις μιάς αλυσίδας είναι λοιπόν έμμονες (recurrent). Ο πίνακας μετάβασης ανωτέρας n -τάξης συμβολίζεται με $\Phi(n) = \{\phi_{ij}(n)\}$ και είναι $\Phi(n) =$

$$P^n = \underbrace{P * P * \dots * P}_n.$$

Η i -οστή γραμμή του $\Phi(n)$ που συμβολίζεται με $\phi_i(n)$ αντιστοιχεί στο διάνυσμα πιθανότητας του να βρίσκεται η ανέλιξη σε κάθε μία από τις καταστάσεις μετά από n βήματα εάν άρχισε από την κατάσταση i ή ισοδύναμα από την αρχική κατάσταση e_i , όπου e_i το i -οστό, ορθοκανονικό διάνυσμα N διαστάσεων που έχει 1 στη i -οστή θέση και μηδενικά παντού αλλού (ένα παράδειγμα για τρεις διαστάσεις είναι: $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$, $e_3 = (0, 0, 1)$). Καλούμε το διάνυσμα $\phi_i(n)$ απόκριση n τάξης της ανέλιξης.

Καθώς το n πηγαίνει στο άπειρο ο πίνακας $\Phi(n)$ συγκλίνει στον πίνακα όριο $P^\infty = \Phi$. Για όλες τις καταστάσεις που ανήκουν στην ίδια αλυσίδα αντιστοιχούν όμοιες γραμμές στον πίνακα Φ , κάθε αλυσίδα μπορεί λοιπόν αυτοτελώς να θεωρηθεί ως μονοδεσμική Μαρκοβιανή διαδικασία.

Στις μεταβατικές καταστάσεις που συνδέονται με περισσότερες από μια αλυσίδες θα αντιστοιχεί μια γραμμή στον πίνακα Φ που είναι το άθροισμα των γραμμών που αντιστοιχούν σε κάθε αλυσίδα με την οποία συνδέονται, αφού πρώτα η κάθε γραμμή πολλαπλασιαστεί με την πιθανότητα εισόδου από την μεταβατική κατάσταση στην αντίστοιχη αλυσίδα. Στη γενική περίπτωση λοιπόν, μια Μαρκοβιανή ανέλιξη μπορεί να θεωρηθεί ως συλλογή των αλυσίδων της μαζί με τις πιθανές μεταβατικές καταστάσεις που συνδέονται με αυτές.

Μια Μαρκοβιανή ανέλιξη περισσότερων από μιας αλυσίδων δεν θα συγκλίνει σε μοναδικό διάνυσμα πιθανότητας ισορροπίας. Ανάλογα με την αρχική κατάσταση το διάνυσμα ισορροπίας θα είναι αυτό στο οποίο συγκλίνει η αντίστοιχη αλυσίδα στην οποία ανήκει η αρχική κατάσταση. Έτσι οι κόμβοι που συνδέονται η ανήκουν στην ίδια αλυσίδα μπορούν να προσδιοριστούν από την αντίστοιχη ομάδα όμοιων σειρών του πίνακα Φ .

Η διαδικασία σύγκλισης σε μια αλυσίδα έχει ενιαίο σημείο ισορροπίας ανεξάρτητο από την αρχική συνθήκη. Ο ρυθμός αυτής της σύγκλισης είναι όπως θα δούμε αναλυτικά παρακάτω ανάλογος με το γινόμενο των ιδιοτιμών του πίνακα μετάβασης, δεδομένου ότι ο ρυθμός σύγκλισης όταν αυτή προβληθεί στα ιδιοδιανύσματα του πίνακα μετάβασης είναι ίσος με την αντίστοιχη ιδιοτιμή. Το 1 σαν ιδιοτιμή αντιστοιχεί σε μη σύγκλιση κατά μήκος της κατεύθυνσης του αντίστοιχου ιδιοδιανύσματος, η πολλαπλότητα αυτών των ιδιοτιμών είναι ίσος με τον αριθμό των αλυσίδων της ανέλιξης.

Ο πίνακας $T(n) = \Phi(n) - \Phi$ καλείται μεταβατικός πίνακας και περιέχει τα μεταβατικά συστατικά της ανέλιξης. Το άθροισμα αυτών των πινάκων μέχρι το βήμα n καλείται n -αθροιστικός μεταβατικός πίνακας $T_\Sigma(n) = \sum_{k=0}^n T(k)$. Ο μετασχηματισμός ζήτα του πίνακα $T(n)$ υπολογισμένος στην τιμή 1 ισούται ως γνωστόν με το άπειρο άθροισμα των $T(n)$ που είναι επίσης πίνακας και καλείται αθροιστικός πίνακας μεταβάσεων και συμβολίζεται με $T_g(1)$. Ο πίνακας αυτός περιέχει το άπειρο άθροισμα των μεταβατικών συστατικών της

Μαρκοβιανής ανέλιξης και μπορεί να υπολογιστεί από :

$$T_g(1) = \sum_{n=0}^{\infty} (P^n - \Phi) = [I - P + \Phi]^{-1} - \Phi \quad (37)$$

Είναι σαφές ότι ο $T_g(1)$ είναι το όριο των $T_{\Sigma}(n)$. Οι στήλες του $T_{\Sigma}(n)$, $T(n)$ και $T_g(1)$ θα συμβολίζονται $\tau_i(n)$, $t_i(n)$ και $t_g(i)$ αντίστοιχα και οι συντεταγμένες τους με $\tau_{ij}(n)$, $t_{ij}(n)$ και $t_g(ij)$ αντίστοιχα.

Το διάνυσμα των μέσων αριθμών επισκέψεων στο χρόνο n όταν η ανέλιξη ξεκίνησε από την κατάσταση i δίδεται από :

$$\bar{\eta}_i(n) = \sum_{k=0}^n \phi_i(k) \quad (38)$$

5.1.3. Η Διαδικασία Σύγκλισης σαν Βάση για Ομαδοποίηση των Καταστάσεων της Μαρκοβιανής Ανέλιξης

Σε αυτό το κεφάλαιο παρέχονται κάποιες διαπιστώσεις σε σχέση με την δυναμική της σύγκλισης της Μαρκοβιανής ανέλιξης, δίνονται γεωμετρικές ερμηνείες των γνωστών πιθανοθεωρητικών μέτρων της ανέλιξης και εγκαθίσταται η γνώση ότι ο τρόπος σύγκλισης της ανέλιξης από όλες τις αρχικές καταστάσεις στα αντίστοιχα διανύσματα ισορροπίας αποκαλύπτει μία ομαδοποίηση των καταστάσεων της. Αρχικά δείχνεται ότι η σύγκλιση προς την κατάσταση ισορροπίας γίνεται στην κατεύθυνση των ιδιοδιανυσμάτων του ανάστροφου πίνακα μετάβασης πρώτης τάξης με ρυθμό την αντίστοιχη ιδιοτιμή.

Ας είναι P ο πίνακας μετάβασης πρώτης τάξης και λ_i, v_i τα ζεύγη ιδιοτιμών, αριστερών ιδιοδιανυσμάτων του P , όπου $|\lambda_i| \leq 1, \forall i$ και v_i τα αντίστοιχα κανονικοποιημένα⁹ αριστερά ιδιοδιανύσματα. Ας είναι V ο πίνακας που έχει στήλες τα v_i , V^T ο ανάστροφος πίνακας του V και Λ ο διαγώνιος πίνακας των ιδιοτιμών του P που αντιστοιχεί στον V . Έστω επίσης ότι $|\lambda_i| \geq |\lambda_j|$ όταν $i < j$.

Είναι $V^T P = \Lambda V^T \Leftrightarrow (V^T)^{-1} V^T P = (V^T)^{-1} \Lambda V^T \Leftrightarrow P = (V^T)^{-1} \Lambda V^T \Leftrightarrow P (V^T)^{-1} = (V^T)^{-1} \Lambda$, άρα όταν ο V είναι ο πίνακας που έχει στήλες τα αριστερά ιδιοδιανύσματα του P ο $(V^T)^{-1}$ είναι ο πίνακας που έχει στήλες τα δεξιά ιδιοδιανύσματα του P και τον συμβολίζουμε με $V' = \{v'_{ij}\}$

Από την προφανή σχέση $V' V^T = I$ προκύπτει ότι η i γραμμή του V' έχει ως στοιχεία της συντεταγμένες της αρχικής συνθήκης e_i όταν αυτή προβληθεί επάνω στα ιδιοδιανύσματα στήλες του V .

Έτσι ας είναι λοιπόν i, j δύο τυχαίες καταστάσεις της ανέλιξης και $e_i = \phi_i(0)$, $e_j = \phi_j(0)$ τα διανύσματα των αντίστοιχων αρχικών καταστάσεων. Οι αρχικές καταστάσεις μπορούν να γραφούν σαν γραμμικός συνδυασμός των αριστερών ιδιοδιανυσμάτων του P :

⁹Σε σχέση με την l_1 νόρμα. Όλα τα στοχαστικά διανύσματα είναι κανονικοποιημένα ως προς l_1 .

$$\begin{aligned} e_i &= \phi_i(0) = v'_{i1} \mathbf{v}_1 + v'_{i2} \mathbf{v}_2 + \cdots + v'_{iN} \mathbf{v}_N \\ e_j &= \phi_j(0) = v'_{j1} \mathbf{v}_1 + v'_{j2} \mathbf{v}_2 + \cdots + v'_{jN} \mathbf{v}_N \end{aligned} \quad (39)$$

Εδώ η διάταξη των ιδιοδιανυσμάτων θεωρείται αντίστοιχη με αυτήν των ιδιοτιμών που τους αντιστοιχούν, άρα \mathbf{v}_1 είναι το διάνυσμα ισορροπίας της ανέλιξης που αντιστοιχεί στην ιδιοτιμή $\lambda_1 = 1$.

Έστω ότι οι καταστάσεις i και j ανήκουν στην ίδια αλυσίδα, τότε η απόκριση n -τάξης της ανέλιξης εάν το σύστημα ξεκίνησε από την κατάσταση i και j δίδεται από :

$$\begin{aligned} \phi_i(n) &= \mathbf{v}_1 + \lambda_2^n v'_{i2} \mathbf{v}_2 + \cdots + \lambda_N^n v'_{iN} \mathbf{v}_N = \mathbf{v}_1 + \mathbf{t}_i(n) \\ \phi_j(n) &= \mathbf{v}_1 + \lambda_2^n v'_{j2} \mathbf{v}_2 + \cdots + \lambda_N^n v'_{jN} \mathbf{v}_N = \mathbf{v}_1 + \mathbf{t}_j(n) \end{aligned} \quad (40)$$

Εδώ οι καταστάσεις θεωρούνται ότι ανήκουν σε μία μοναδική αλυσίδα κι έτσι οι δείκτες θεωρούνται προσαρμοσμένοι με τρόπο ώστε \mathbf{v}_1 είναι το μοναδικό διάνυσμα ισορροπίας, $\lambda_i < 1, \forall i > 1$ και N είναι η πληθυσμότητα της αλυσίδας. Η σημασία της παραπάνω έκφρασης είναι ότι αποκαλύπτει τον μηχανισμό της σύγκλισης και εκφράζει τα συστατικά των μεταβατικών μεγεθών της ανέλιξης με γεωμετρικούς όρους.

Το διάνυσμα των μέσων αριθμών επισκέψεων στο χρόνο n (n -τάξης) όταν η ανέλιξη ξεκίνησε από την κατάσταση i μπορεί να γραφεί σαν :

$$\begin{aligned} \bar{\eta}_i(n) &= \sum_{k=0}^n \phi_i(k) = (n+1) \mathbf{v}_1 + v'_{i2} \mathbf{v}_2 \sum_{k=0}^n \lambda_2^k + \cdots + v'_{iN} \mathbf{v}_N \sum_{k=0}^n \lambda_N^k = \\ &= (n+1) \mathbf{v}_1 + \sum_{p=2}^N \frac{1 - \lambda_p^n}{1 - \lambda_p} v'_{ip} \mathbf{v}_p = (n+1) \mathbf{v}_1 + \sum_{k=0}^n \mathbf{t}_i(k) = (n+1) \mathbf{v}_1 + \boldsymbol{\tau}_i(n) \end{aligned} \quad (41)$$

και το διάνυσμα των αναμενόμενων ποσοστών επισκεψιμότητας n -τάξης ξεκινώντας από i είναι :

$$\mathbf{f}_i(n) = \frac{\bar{\eta}_i(n)}{n+1} = \mathbf{v}_1 + \frac{1}{n+1} \sum_{p=2}^N \frac{1 - \lambda_p^n}{1 - \lambda_p} v'_{ip} \mathbf{v}_p = \mathbf{v}_1 + \frac{\boldsymbol{\tau}_i(n)}{n+1} \quad (42)$$

Τα στοιχεία λοιπόν του n -αθροιστικού μεταβατικού πίνακα εκφράζονται με γεωμετρικούς όρους στον υπόχωρο που έχει βάση τα ιδιοδιανύσματα του P .

Τώρα θα εξετασθούν οι διαφορές στην μεταβατική συμπεριφορά της ανέλιξης όταν αυτή αρχίζει απο διαφορετικές αρχικές συνθήκες. Η διαφορά στις αναμενόμενες επισκεψιμότητες των κόμβων μετά από n βήματα όταν η ανέλιξη ξεκινάει από την κατάσταση i αντί της j είναι :

$$\bar{\eta}_i(n) - \bar{\eta}_j(n) = \sum_{k=0}^n \phi_i(k) - \sum_{k=0}^n \phi_j(k) = \sum_{k=0}^n (\phi_i(k) - \phi_j(k))$$

που από την 40 και 41 γίνεται:

$$\bar{\eta}_i(n) - \bar{\eta}_j(n) = \sum_{p=2}^N \sum_{k=0}^n \lambda_p^k (v'_{ip} - v'_{jp}) \mathbf{v}_p = \sum_{p=2}^N \frac{1 - \lambda_p^n}{1 - \lambda_p} (v'_{ip} - v'_{jp}) \mathbf{v}_p = \boldsymbol{\tau}_i(n) - \boldsymbol{\tau}_j(n) \quad (43)$$

Εδώ εκφράζεται με γεωμετρικούς όρους η διαφορά των μεταβατικών αθροισμάτων της ανέλιξης. Παρατηρήστε ότι επειδή τα i και j ανήκουν στην ίδια αλυσίδα, ο όρος $(n+1)\mathbf{v}_1$ απαλείφεται και η διαφορά συγκλίνει. Στην γενική περίπτωση όμως τα i και j μπορεί να ανήκουν σε διαφορετικές αλυσίδες με διανύσματα ισορροπίας τα v_1 και u_1 οπότε ο αντίστοιχος όρος τώρα γίνεται $(n+1)(v_1 - u_1)$ και η διαφορά πηγαίνει στο άπειρο.

Η διαφορά όμως των αναμενόμενων ποσοστών επισκεψιμότητας εξ αιτίας της διαφορετικής αρχικής συνθήκης πάντα συγκλίνει. Για αρχικές συνθήκες της ίδιας αλυσίδας είναι:

$$\mathbf{f}_i(n) - \mathbf{f}_j(n) = \frac{\bar{\eta}_i(n) - \bar{\eta}_j(n)}{n+1} = \frac{1}{n+1} \sum_{p=2}^N \frac{1 - \lambda_p^n}{1 - \lambda_p} (v'_{pi} - v'_{pj}) \mathbf{v}_p = \frac{\boldsymbol{\tau}_i(n) - \boldsymbol{\tau}_j(n)}{n+1} \quad (44)$$

και εκφράζεται και αυτή με γεωμετρικούς όρους σαν γραμμικός συνδυασμός των ιδιοδιανυσμάτων του P .

Καθώς το n τείνει στο άπειρο η 43 γίνεται:

$$\bar{\eta}_i(\infty) - \bar{\eta}_j(\infty) = \sum_{p=2}^N \frac{v'_{ip} - v'_{jp}}{1 - \lambda_p} \mathbf{v}_p = \mathbf{t}_g(i) - \mathbf{t}_g(j) \quad (45)$$

που είναι η διαφορά των μερικών αθροισμάτων της ανέλιξης εκφρασμένη με γεωμετρικούς όρους στις κατευθύνσεις της σύγκλισης. Η σημασία αυτής της έκφρασης βρίσκεται στην συσχέτιση του μηχανισμού της σύγκλισης με τα πραγματικά στατιστικά μέτρα της ανέλιξης όπως αυτά εκφράζονται μέσω των στοιχείων του αθροιστικού πίνακα μεταβάσεων της σχέσης 37 και η απόδοση γεωμετρικής ερμηνείας σε αυτά.

Τώρα παρουσιάζεται ένα βασικό μέτρο συνδεσιμότητας ανάμεσα στις καταστάσεις μιας αλυσίδας. Ας είναι i, j δύο καταστάσεις της ίδιας αλυσίδας με j έμμονη. Ο αναμενόμενος χρόνος πρώτης μετάβασης $\bar{\theta}_{ij}$ από την κατάσταση i στην j είναι ο αναμενόμενος αριθμός μεταβάσεων που θα χρειαστεί η ανέλιξη για να φτάσει για πρώτη φορά στην κατάσταση j εάν ξεκινήσει από την κατάσταση i και δίδεται από:

$$\bar{\theta}_{ij} = \frac{t_g(jj) - t_g(ij)}{\pi_j} \quad (46)$$

Η σχέση 45 δείχνει ότι ο αναμενόμενος χρόνος πρώτης μετάβασης από i σε j συσχετίζεται με τα στοιχεία του V' και συγκεκριμένα με τις συντεταγμένες των αρχικών συνθηκών e_i και e_j όταν τα ιδιοδιανύσματα του P αποτελούν την βάση του χώρου \mathbb{R}^N . Ειδικότερα από

την 45 και την 46 έχουμε:

$$\pi_j \bar{\theta}_{ij} = \frac{v'_{j2} - v'_{i2}}{1 - \lambda_2} v_{j2} + \dots + \frac{v'_{jN} - v'_{iN}}{1 - \lambda_N} v_{jN} \quad (47)$$

που δείχνει ότι ο αναμενόμενος χρόνος πρώτης μετάβασης από το i στο j δίνεται από το άθροισμα των ενισχυμένων κατά την αντίστοιχη ιδιοτιμή διαφορών των ομωνύμων συντεταγμένων των αντιστοιχών αρχικών συνθηκών όταν αυτές προβάλλονται στα ιδιοδιανύσματα του P και η αξία αυτής της έκφρασης είναι ότι συνδέει τον αναμενόμενο χρόνο πρώτης μετάβασης με τον τρόπο σύγκλισης της Μαρκοβιανής διαδικασίας.

Μαρκοβιανές ανελίξεις ανακλύπτουν συχνά στην προσομοίωση πραγματικών φαινομένων και στις περισσότερες περιπτώσεις η αξία τους βρίσκεται στον προσδιορισμό του αναμενόμενου ποσοστού του χρόνου που η διαδικασία ξοδεύει σε κάθε μία από τις καταστάσεις. Η εύρεση αυτού του σημείου ισορροπίας για μεγάλους χώρους καταστάσεων που προκύπτουν συνήθως σε αυτές τις διαδικασίες δεν είναι ένας εύκολος στόχος και πολλές τεχνικές έχουν προταθεί για αυτόν τον λόγο. Η κατάσταση γίνεται περαιτέρω περίπλοκη όταν η Μαρκοβιανή ανέλιξη είναι σχεδόν πλήρως διαχωρίσιμη που σημαίνει ότι υπάρχουν ομάδες καταστάσεων με υψηλή συνδεσιμότητα μεταξύ των καταστάσεων κάθε ομάδας και σχετικά χαμηλή συνδεσιμότητα μεταξύ των καταστάσεων των διαφορετικών ομάδων ή των ίδιων των ομάδων μεταξύ τους. Από την άποψη της σύγκλισης αυτό σημαίνει ότι η μετάβαση από τις αρχικές καταστάσεις στο διάνυσμα πιθανότητας ισορροπίας για ολόκληρη τη ανέλιξη είναι πολύ αργή ή ισοδύναμα από την άποψη των ιδιοδιανυσμάτων ότι υπάρχει τουλάχιστον μία ιδιοτιμή πολύ κοντά στο 1 που επιβραδύνει τη σύγκλιση κατά μήκος του αντίστοιχου ιδιοδιανύσματος αλλά συγχρόνως η μερική σύγκλιση μεταξύ των καταστάσεων μιας ομάδας υψηλής ενδοσυνδεσιμότητας συμβαίνει γρήγορα δεδομένου ότι οι αντίστοιχες ιδιοτιμές που την περιγράφουν είναι πολύ κοντά στο μηδέν.

Από την άποψη της σύγκλισης λοιπόν η κατάσταση με τις σχεδόν πλήρως διαχωριζόμενες Μαρκοβιανές ανελίξεις μπορεί να περιγραφεί μέσω των ιδιοτιμών του πίνακα μετάβασης πρώτης τάξης (που είναι οι ίδιες με αυτές του ανάστροφου του) ότι υπάρχει ένα σύνολο από πολύ μικρές κατά απόλυτη τιμή ιδιοτιμές κοντά στο μηδέν αρμόδιο για τη γρήγορη σύγκλιση μέσα στις διάφορες υψηλά ενδοσυνδεόμενες ομάδες καταστάσεων και τουλάχιστον μία ιδιοτιμή πολύ κοντά στο ένα αρμόδια για την αργή σύγκλιση ολόκληρης της ανέλιξης. Αυτή η ιδιοτιμή καλείται συνήθως ρίζα Perron της ανέλιξης. Η ύπαρξη μιας ρίζας Perron λ_2 με μέτρο κοντά στο ένα σημαίνει ότι ανεξάρτητα από την κατάσταση εκκίνησης της ανέλιξης, έστω i , σε έναν κατάλληλα μεγάλο χρόνο n το διάνυσμα κατάστασης της ανέλιξης θα είναι ίσο με $v_1 + \lambda_2^n v'_{i2} v_2$ επειδή οι άλλοι όροι θα έχουν εξαφανιστεί, έτσι εκτός αν το v'_{i2} είναι μηδέν η ανέλιξη θα κινηθεί προς την ισορροπία με το ρυθμό που η ιδιοτιμή λ_2 συρρικνώνει το ιδιοδιάνυσμα v_2 που είναι αργός εάν η λ_2 έχει μέτρο κοντά στο ένα. Εάν η ανέλιξη άρχιζε από j αντί από i έπειτα από κατάλληλα μεγάλο n το διάνυσμα κατάστασης της θα

ήταν $v_1 + \lambda_2^n v'_{j_2} v_2$ και η διαφορά εξ αιτίας της έναρξης της ανέλιξης από την κατάσταση j αντί της i , των διανυσμάτων κατάστασης μετά από n βήματα θα ήταν στην ποσότητα $|v'_{i_2} - v'_{j_2}|$. Αυτό προτείνει ότι στην περίπτωση μιας σχεδόν πλήρους διαχωρίσιμης Μαρκοβιανής ανέλιξης όπως αυτή, μια ομαδοποίηση των στοιχείων της στήλης του πίνακα V' που αντιστοιχεί στην κατά απόλυτη τιμή μεγαλύτερη μικρότερη από ένα ιδιοτιμή του πίνακα μετάβασης πρώτης τάξης αντιστοιχεί σε μια ομαδοποίηση των καταστάσεων όσον αφορά την μεταβατική σύγκλιση της ανέλιξης και κατά συνέπεια την μεταξύ τους συνδεσιμότητα. Στην περίπτωση όπου άλλες ιδιοτιμές κοντά στο ένα υπάρχουν μια ομαδοποίηση των στοιχείων των αντίστοιχων στηλών του V' θα περιέγραφε παράπλευρες ομαδοποιήσεις στη διαδικασία σύγκλισης που μεταφράζονται σε ιεραρχικά επίπεδα ομαδοποιήσεων των καταστάσεων της ανέλιξης με βάση την συνδεσιμότητά τους. Στην πραγματικότητα η ομαδοποίηση των στοιχείων μιας στήλης του πίνακα V' είναι τόσο σημαντική όσο πιο κοντά στο ένα είναι η αντίστοιχη ιδιοτιμή. Μα ο πίνακας V' έχει ως στήλες τα δεξιά ιδιοδιανύσματα του P . Έτσι ενώ τα αριστερά ιδιοδιανύσματα του P περιγράφουν τις διευθύνσεις της σύγκλισης, τα δεξιά ιδιοδιανύσματα του P περιγράφουν τον τρόπο της σύγκλισης για κάθε μία από τις αρχικές καταστάσεις της ανέλιξης σε διάφορες χρονικές κλίμακες που καθορίζονται από τις αντίστοιχες ιδιοτιμές. Φαίνεται λοιπόν ότι στον υπολογισμό του αναμενόμενου χρόνου μετάβασης μεταξύ δύο καταστάσεων η διαφορά στις συντεταγμένες των δεξιών ιδιοδιανυσμάτων του P που αντιστοιχούν σε αυτές τις καταστάσεις βαρύνουν ανάλογα με το πόσο κοντά στο 1 είναι η αντίστοιχη ιδιοτιμή. Έτσι οι διαφορές στις συντεταγμένες του v'_N έχουν το μέγιστο βάρος και αυτές του v'_1 το ελάχιστο.

5.1.4. Στατιστικά Μέτρα για τις Αρχικές Συνθήκες

Σε αυτό το κεφάλαιο εξετάζονται οι καταστάσεις της Μαρκοβιανής ανέλιξης με έναν πιο 'στατιστικό' τρόπο. Συνδέονται τα στατιστικά με τα στοχαστικά μέτρα και δίδονται γεωμετρικές ερμηνείες. Ας θεωρηθούν λοιπόν τα μέχρι την στιγμή n ποσοστά επισκεψιμότητας στις καταστάσεις της ανέλιξης ως μια σειρά N παρατηρήσεων της αντίστοιχης τυχαίας μεταβλητής που συμβολίζεται με $S_i(n)$ μιας και είναι προφανές ότι εξαρτάται από την αρχική συνθήκη i και τον χρόνο n , είναι επίσης προφανές ότι ο μέσος των παρατηρήσεων είναι πάντοτε $\frac{1}{N}$ ενώ η διασπορά είναι αντίστροφα ανάλογη με τον πλουραλισμό των επισκέψεων της ανέλιξης μέχρι και το βήμα n . Το στατιστικό μέτρο που εξετάζεται στην συνέχεια είναι η διασπορά της τ.μ. $S_i(n) - S_j(n)$, δηλαδή της διαφοράς των τ.μ. που αντιστοιχούν σε δύο διαφορετικές αρχικές συνθήκες αλλά στον ίδιο χρόνο n διότι αυτή η διασπορά συνδέει τα στοχαστικά με τα στατιστικά μέτρα αλλά και τον μηχανισμό σύγκλισης της στοχαστικής διαδικασίας δίδοντας γεωμετρικές ερμηνείες.

Αρχικά λοιπόν είναι:

$$var(S_i(n) - S_j(n)) = var(S_i(n)) + var(S_j(n)) - 2cov(S_i(n), S_j(n)) \quad (48)$$

όπου η διασπορά της διαφοράς συνδέεται με την συνδιακύμανση των ίδιων τ.μ. Η συνδιακύμανση είναι σημαντική εδώ διότι μέσω αυτής θα γίνει δυνατή η εφαρμογή της θεωρίας των Principal Components σε σχέση με τα ιδιοδιανύσματα του πίνακα μετάβασης και έτσι θα ενοποιηθούν τα στοχαστικά με τα στατιστικά συμπεράσματα για την κατασκευή της απόστασης μεταξύ ανελιξων. Στην συνέχεια μιας και πάντοτε οι τ.μ $S_i(n)$ και $S_j(n)$ θα θεωρούνται στον ίδιο χρόνο n για εργονομία συμβολισμού το n θα μεταφέρεται στο αντίστοιχο στατιστικό μέτρο. Έτσι η 48 γίνεται:

$$var_n(S_i - S_j) = var_n(S_i) + var_n(S_j) - 2cov_n(S_i, S_j) \quad (49)$$

Όμως η διαφορά έχει πάντα μέσο όρο μηδέν και έτσι η διασπορά της μπορεί να γραφεί και σαν:

$$var_n(S_i - S_j) = \frac{\|\mathbf{f}_i(n) - \mathbf{f}_j(n)\|^2}{N} \quad (50)$$

που δίνει μια γεωμετρική ερμηνεία σε σχέση με τα διανύσματα των ποσοστών επισκεψιμότητας της ανέλιξης. Όμως από την 44 προκύπτει μία έκφραση του $\mathbf{f}_i(n) - \mathbf{f}_j(n)$ σαν γραμμικού συνδυασμού των ιδιοδιανυσμάτων του πίνακα μετάβασης, δηλαδή δίδονται οι συντεταγμένες του σε μια άλλη βάση και όπως είναι γνωστό μπορεί να υπολογισθεί η νόρμα του με την χρήση των συντεταγμένων της άλλης βάσης ως εξής:

$$\begin{aligned} \|\mathbf{f}_i(n) - \mathbf{f}_j(n)\|^2 &= \langle \mathbf{f}_i(n) - \mathbf{f}_j(n), \mathbf{f}_i(n) - \mathbf{f}_j(n) \rangle \\ &= (\boldsymbol{\lambda}(n) \square (\mathbf{v}'_i - \mathbf{v}'_j)) \cdot U \cdot ((\mathbf{v}'_i - \mathbf{v}'_j)^T \square \boldsymbol{\lambda}(n)^T) \\ &= \boldsymbol{\lambda}(n) \cdot (A \square U \square A^T) \cdot \boldsymbol{\lambda}(n)^T \end{aligned} \quad (51)$$

όπου

$$\boldsymbol{\lambda}(n) = \frac{1}{n+1} \left(\frac{1 - \lambda_1^n}{1 - \lambda_1}, \dots, \frac{1 - \lambda_N^n}{1 - \lambda_N} \right),$$

$$A = \begin{pmatrix} v'_{i1} - v'_{j1} & \dots & v'_{i1} - v'_{j1} \\ v'_{i2} - v'_{j2} & \dots & v'_{i2} - v'_{j2} \\ \dots & \dots & \dots \\ v'_{iN} - v'_{jN} & \dots & v'_{iN} - v'_{jN} \end{pmatrix},$$

$$U = \begin{pmatrix} \langle \mathbf{v}_1, \mathbf{v}_1 \rangle & \dots & \langle \mathbf{v}_1, \mathbf{v}_N \rangle \\ \langle \mathbf{v}_2, \mathbf{v}_1 \rangle & \dots & \langle \mathbf{v}_2, \mathbf{v}_N \rangle \\ \dots & \dots & \dots \\ \langle \mathbf{v}_N, \mathbf{v}_1 \rangle & \dots & \langle \mathbf{v}_N, \mathbf{v}_N \rangle \end{pmatrix}$$

με v_i τα ιδιοδιανύσματα του πίνακα μετάβασης πρώτης τάξης και \square ο κατά σημείο πολλ/σμός μεταξύ ομοδιάστατων πινάκων. Οι ιδιότητες αυτού του πολλ/σμού μελετούνται στο [264].

Οι παραπάνω συνδιακυμάνσεις για όλα τα ζεύγη των καταστάσεων της ανέλιξης την στιγμή n προκύπτουν σαν τιμές του πίνακα συνδιακυμάνσεων του δείγματος $F(n)^T = (\frac{1}{n+1} \sum_{k=0}^n P^k)^T$. Παρατηρήστε εδώ ότι ο πίνακας $F(n)$ είναι ένα πολυώνυμο του πίνακα P και άρα ως γνωστόν έχει τα ίδια ιδιοδιανύσματα με αυτόν. Οι ιδιοτιμές τους δεν είναι ίδιες αλλά η διάταξη των ιδιοδιανυσμάτων σε σχέση με την αντιστοιχία τους στις αντίστοιχες ιδιοτιμές ως προς το μέτρο τους είναι επίσης η ίδια μιας και οι ιδιοτιμές του $F(n)$ είναι οι εικόνες αυτών του P μέσα από την πολυωνυμική συνάρτηση. Παρατηρήστε επίσης ότι τα δεξιά(αριστερά) ιδιοδιανύσματα του $F(n)^T$ είναι τα αριστερά(δεξιά) του $F(n)$ όπως είναι προφανές και άρα τα αριστερά ιδιοδιανύσματα του $F(n)^T$ είναι τα δεξιά ιδιοδιανύσματα του P δηλαδή τα ιδιοδιανύσματα για την σημασία των οποίων στην ομαδοποίηση των καταστάσεων της ανέλιξης έγινε λόγος στο τέλος της παραγράφου 5.1.3.

5.1.5. Η Απόσταση Ανάμεσα στις Λέξεις Κλειδιά

Είναι γνωστό από την στατιστική θεωρία ότι η διεύθυνση του κύριου ιδιοδιανύσματος του πίνακα συνδιακύμανσης ενός στατιστικού δείγματος παρέχει την βέλτιστη γραμμική προσέγγιση της χωρικής κατανομής του δείγματος αυτού με την έννοια ότι ελαχιστοποιείται η μέση απόσταση του δείγματος από την ευθεία αυτήν. Την ίδια στιγμή έχει δειχθεί από τα παραπάνω ότι εάν θεωρήσουμε ως στατιστικό δείγμα τον παροδικό πίνακα του αριθμού των αναμενόμενων επισκέψεων της Μαρκοβιανής διαδικασίας τότε η διεύθυνση που προσεγγίζει βέλτιστα την κατανομή αυτού του δείγματος στον χώρο, λόγω των μηχανισμών της σύγκλισης που δείχθηκαν παραπάνω, είναι αυτή του ιδιοδιανύσματος του πίνακα απλής μετάβασης που αντιστοιχεί στην Perron ιδιοτιμή. Άρα το ιδιοδιάνυσμα του πίνακα απλής μετάβασης της ανέλιξης είναι το κύριο ιδιοδιάνυσμα του πίνακα συνδιακύμανσης του παροδικού πίνακα αναμενόμενων επισκέψεων της ανέλιξης. Αυτή είναι η ιδέα που χρησιμοποιείται εδώ για την κατασκευή μιάς βέλτιστης απόστασης μεταξύ των λέξεων κλειδιών όταν οι συμπεριφορές των χρηστών αντιπροσωπεύονται με Μαρκοβιανές ανελιξεις. Η απόσταση αυτή λαμβάνει υπόψη της την γενική συμπεριφορά όλων των χρηστών συγκεντρωτικά. Η προτεινόμενη απόστασή θα μετρήσει την απόσταση μεταξύ των λέξεων κλειδιών όπως αυτή υπαγορεύεται από τη γενική συμπεριφορά.

Σε κάθε περίπτωση από τώρα και στο εξής όταν αναφέρεται μία ανέλιξη θα υπονοείται η μεταβατική απόκριση n -βημάτων του πίνακα του αναμενόμενου αριθμού επισκέψεων της αντίστοιχης ανέλιξης. Αυτό γίνεται διότι η χρήση της Μαρκοβιανής ανέλιξης γίνεται για να αποκαλυφθεί η ομαδοποίηση των λέξεων κλειδιών που χρησιμοποιούνται σε αυτήν την ανέλιξη και δεν ενδιαφέρει η κατάσταση ισορροπίας. Οι σημαντικές πληροφορίες

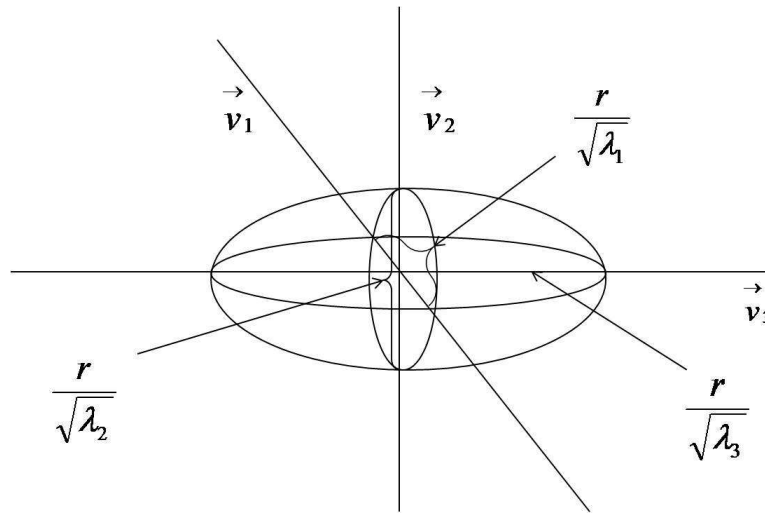
για το προτεινόμενο μοντέλο βρίσκονται στην μεταβατική συμπεριφορά της ανέλιξης και ειδικότερα στις γραμμές του μεταβατικού πίνακα αναμενόμενων αριθμών επισκέψεων δεδομένου ότι μια ομαδοποίηση αυτών των γραμμών αποκαλύπτει την ομαδοποίηση των καταστάσεων όπως παρουσιάστηκε στο προηγούμενο τμήμα. Για να διαμορφωθούν οι ομάδες των λέξεων κλειδιών υπολογίζεται ο πίνακας των αναμενόμενων επισκέψεων στις λέξεις κλειδιά μετά από χρόνο n , συμβολίζοντας αυτόν τον πίνακα με $F(n)$, έτσι είναι $F(n) = \sum_{k=0}^n P^k$ όπου P είναι ο πίνακας μετάβασης πρώτης τάξης. Εδώ πρέπει να ορισθεί ένα κριτήριο διακοπής δεδομένου ότι είναι επιθυμητό ένα τέτοιο n που να σταματήσει τη σειρά στο σημείο όπου η αργή σύγκλιση έχει αναλάβει αλλά όχι πριν να ολοκληρωθεί η γρήγορη σύγκλιση. Δεδομένου ότι η γρήγορη σύγκλιση θα διαμορφώσει συστάδες στις γραμμές του $F(n)$ όσον αφορά την Ευκλείδεια απόστασή τους και δεδομένου ότι η αργή σύγκλιση δεν θα αλλάξει σημαντικά αυτήν την διαμόρφωση σε μια επανάληψη, κριτήριο διακοπής της σειράς μπορεί να αποτελέσει η σύγκριση διαδοχικών διανυσματικών διασπορών των γραμμών του $F(n)$. Εάν η αλλαγή στις διαδοχικές διασπορές είναι κάτω από ένα κατώτατο όριο διακόπεται ο υπολογισμός και εν πάση περιπτώσει διακόπεται εάν το n υπερβεί ένα όριο που θεωρείται το μέγιστο για την πολυπλοκότητα του αλγορίθμου. Άλλο κριτήριο διακοπής θα μπορούσε να αποτελέσει η ορίζουσα του $F(n)$ μιας και η συσταδοποίηση των γραμμών του θα τείνει να ριζώνει την τάξη του αντίστοιχου πίνακα. Από τώρα και στο εξής όποτε γίνεται χρήση του συμβόλου F θα εννοείται $F(n)$ που έχει υπολογιστεί στο επιθυμητό n .

Η απόσταση που ορίζεται παρακάτω λειτουργεί με στατιστικούς όρους στο στοχαστικό μοντέλο των χρηστών. Σύμφωνα με αυτούς τους όρους ο πίνακας συνδιακύμανσης ενός συνόλου παρατηρήσεων, όταν κάθε παρατήρηση είναι ένα διάνυσμα των χαρακτηριστικών της γνωρισμάτων είναι το μέγεθος που γίνεται ευρύτατα αποδεκτό και που χρησιμοποιείται για να δώσει τις περισσότερες πληροφορίες για την έλλειψη ομοιογένειας και ισορροπίας του χώρου χαρακτηριστικών γνωρισμάτων όπως διαταράσσεται από το πρότυπο του δείγματος [265]. Σύμφωνα με την ιδέα πίσω από την προτεινόμενη απόσταση, οι παρατηρήσεις είναι οι σειρές του πίνακα F_G όπου F_G είναι ο πίνακας των αναμενόμενων επισκέψεων που έχει υπολογιστεί σε επιθυμητό n όπως παραπάνω και που αντιστοιχεί στη γενική συμπεριφορά όλων των χρηστών. Ο πίνακας γενικής συμπεριφοράς F_G έχει τον πίνακα συνδιακύμανσης $\Sigma(F_G)$, τα στοιχεία του οποίου είναι οι συνδιακυμάνσεις των ζευγαριών των κόμβων της ανέλιξης, όπως εξηγήθηκε ανωτέρω.

Ορίζουμε λοιπόν την απόσταση μεταξύ \mathbf{k}_i και \mathbf{k}_j και συμβολίζουμε με $\delta_{i,j}$ να είναι:

$$\delta_{i,j} = (\mathbf{k}_i - \mathbf{k}_j)^T \Sigma(F_G) (\mathbf{k}_i - \mathbf{k}_j) \quad (52)$$

όπου η διαστασιμότητα των $\mathbf{k}_i, \mathbf{k}_j$ έχει, με προσθήκη μηδενικών στα κατάλληλα σημεία, επεκταθεί σε αυτήν του $\Sigma(F_G)$.



Σχήμα 21: Γεωμετρική αναπαράσταση της r -ισοεπιφάνειας της προτεινόμενης MSI απόστασης στις τρεις διαστάσεις. Το τρισδιάστατο ελλειψοειδές καταδεικνύει ως ισοαπέχοντα τα αντισυμμετρικά ζεύγη σημείων στους άξονές του, συνεπώς οι αποστάσεις στην διεύθυνση του ιδιοδιανύσματος v_1 βαραίνουν περισσότερο και αυτές του v_3 λιγότερο.

Οι ισο-επιφάνειες που ορίζονται από την τετραγωνική αυτή μορφή αναπαριστούν ελλειψοειδές στον \mathbb{R}^N με κέντρο την αρχή των αξόνων και άξονες στις διευθύνσεις των ιδιοδιανυσμάτων του $\Sigma(F_G)$ με μήκος αντιστρόφως ανάλογο των αντίστοιχων ιδιοτιμών (Σχ. 21). Ο υπολογισμός της απόστασης παραπάνω περιγράφει το μήκος του $k_i - k_j$ αφού όμως πρώτα εφαρμοσθεί ένας μη ομοιόμορφος μετασχηματισμός που μεταβάλλει τις προβολές του στην διεύθυνση των ιδιοδιανυσμάτων του πίνακα συνδιακύμανσης του F_G ανάλογα με την αντίστοιχη ιδιοτιμή. Τα ιδιοδιανύσματα του $\Sigma(F_G)$ είναι κάθετα μεταξύ τους και παράγουν τον \mathbb{R}^N . Στην κατεύθυνση του κάθε ιδιοδιανύσματος του $\Sigma(F_G)$ οι γραμμές του F_G προβάλλονται με διασπορά ίση με την αντίστοιχη ιδιοτιμή. Η διεύθυνση όπου οι γραμμές του F_G προβάλλονται με μέγιστη διασπορά είναι αυτή του πρωτεύοντος ιδιοδιανύσματος του $\Sigma(F_G)$, που όπως δείχθηκε στα προηγούμενα, είναι και αυτή που διαχωρίζει βέλτιστα τις καταστάσεις τις ανέλιξης ως προς την συνδεσιμότητά τους στο αντίστοιχο Μαρκοβιανό μοντέλο. Το μήκος του $k_i - k_j$, που είναι η απόσταση των δύο λέξεων κλειδιών, ρυθμίζεται ανάλογα με την διεύθυνσή της στον N -διάστασο χώρο με μέγιστο μήκος (άρα και διαφορά ανάμεσα στις δύο λέξεις κλειδιά) στην διεύθυνση του πρωτεύοντος ιδιοδιανύσματος του $\Sigma(F_G)$ διότι αυτό σημαίνει ότι οι συγκεκριμένες λέξεις κλειδιά δεν έχουν ισχυρή σημασιολογική σύνδεση μιας και όπως καταδεικνύει η γενική συμπεριφορά οι χρήστες δεν ρωτούσαν μαζί αυτές τις λέξεις κλειδιά. Η παραπάνω απόσταση είναι προφανώς καλά ορισμένη και επιτυγχάνει να μετρήσει την διαφορά μεταξύ των λέξεων κλειδιών σε σχέση με την γενική συμπεριφορά όλων των χρηστών και μάλιστα να μεγιστοποιεί την απόσταση ανάμεσα σε λέξεις κλειδιά με την ελάχιστη στην γενική συμπεριφορά συνδεσιμότητα σαν καταστάσεις του διαγράμματος μετάβασης.

5.1.6. Η απόσταση ανάμεσα στους χρήστες

Εδώ γενικεύουμε τα αποτελέσματα της προηγούμενης Ενότητας για να ορίσουμε μια απόσταση ανάμεσα στους χρήστες του πληροφοριακού συστήματος. Από την Ενότητα 5.1.1 έχουμε μια αναπαράσταση για τον κάθε χρήστη στην μορφή ενός διανύσματος ισορροπίας της αντίστοιχης Μαρκοβιανής ανέλιξης συμπεριφοράς του. Κατ'αντιστοιχία λοιπόν με τα προηγούμενα έχουμε:

Ορισμός 8 Έστω \mathbf{u}_i και \mathbf{u}_j δύο χρήστες που αναπαρίστανται από τα αντίστοιχα διανύσματα ισορροπίας¹⁰ $\boldsymbol{\pi}_i$ και $\boldsymbol{\pi}_j$. Ας είναι επίσης $\Sigma(F_G^T)$ ο πίνακας συνδιακύμανσης του ανάστροφου πίνακα αναμενόμενων ποσοστών επισκεψιμότητας n -τάξης της Μαρκοβιανής αλυσίδας γενικής συμπεριφοράς χρηστών, που έχει υπολογιστεί όπως παραπάνω στο επιθυμητό n και έχει μηδενικό μέσο. Τότε ορίζεται μία απόσταση ανάμεσα στους χρήστες \mathbf{u}_i και \mathbf{u}_j ως:

$$\begin{aligned} d(\mathbf{u}_i, \mathbf{u}_j) &= (\boldsymbol{\pi}_i - \boldsymbol{\pi}_j)\Sigma(F_G^T)(\boldsymbol{\pi}_i - \boldsymbol{\pi}_j)^T \\ &= \boldsymbol{\delta}_{ij}\Sigma(F_G^T)\boldsymbol{\delta}_{ij}^T \end{aligned} \quad (53)$$

που καλείται *MSI* απόσταση χρηστών. \square

Οι διαστάσεις του $\boldsymbol{\pi}_i$ και του $\boldsymbol{\pi}_j$ παραπάνω έχουν επεκταθεί σε αυτές του $\Sigma(F_G^T)$ συμπληρώνοντας με μηδενικά στις επιπλέον θέσεις. Η απόσταση d είναι καλώς ορισμένη σαν γενικευμένη Ευκλείδεια συνάρτηση απόστασης που χρησιμοποιεί πίνακα συνδιακύμανσης που είναι πάντοτε θετικά ορισμένος.

Η *MSI* απόσταση είναι μία γενίκευση της απόστασης ανάμεσα στις λέξεις κλειδιά που ορίστηκε παραπάνω και συνεπώς ισχύουν οι ίδιες παρατηρήσεις σε σχέση με την γεωμετρική ερμηνεία των ισοεπιφανειών που ορίζει, σαν ελλειψοειδή του \mathbb{R}^N . Η διεύθυνση όπου οι γραμμές του F_G προβάλλονται με μέγιστη διασπορά είναι αυτή του πρωτεύοντος ιδιοδιανύσματος του $\Sigma(F_G)$, που όπως δείχθηκε στα προηγούμενα, είναι και αυτή που διαχωρίζει βέλτιστα τις καταστάσεις τις ανέλιξης ως προς την συνδεσιμότητά τους στο αντίστοιχο Μαρκοβιανό μοντέλο. Η παραπάνω απόσταση επιτυγχάνει να μετρήσει την διαφορά μεταξύ των χρηστών σε σχέση με την γενική συμπεριφορά όλων των χρηστών και μάλιστα να μεγιστοποιεί την απόσταση ανάμεσα σε χρήστες με την ελάχιστη στην γενική συμπεριφορά συνδεσιμότητα των καταστάσεων στις οποίες διαφέρουν. Στην συνέχεια θα χρησιμοποιηθεί η ίδια τεχνική για να ορισθεί μία απόσταση ανάμεσα στις εικόνες που επιλέγουν οι χρήστες στα πλαίσια ενός πληροφοριακού συστήματος ανάκτησης εικόνων. Πειραματικά αποτελέσματα θα παρουσιαστούν στο τέλος αυτού του κεφαλαίου συγκρίνοντας τις δυνατότητες *MSI* σε σχέση με τις αντίστοιχες μεθόδους *LSI*.

¹⁰Θεωρούμενα ως διανύσματα γραμμές

5.2. Αυτόματη Περιγραφή και Ταξινόμηση Εικόνων μέσω Διάδρασης με τους Χρήστες Πληροφοριακών Συστημάτων

Αυτό που καλείται *σημσιολογικό χάσμα* είναι η έλλειψη σύμπτωσης μεταξύ των πληροφοριών που κάποια μέθοδος μπορεί να εξαγάγει από τα οπτικά στοιχεία και την ερμηνεία που τα ίδια τα στοιχεία έχουν για έναν χρήστη σε μια δεδομένη κατάσταση. Εντούτοις, η διαδικασία ανάκτησης αποτυγχάνει επίσης λόγω του *αισθητήριου χάσματος*: το χάσμα μεταξύ του αντικειμένου στον κόσμο και των πληροφοριών σε μια (υπολογιστική) περιγραφή που ορίζεται σε μια καταγραφή αυτού του αντικειμένου. Ενώ το προηγούμενο χάσμα θέτει το ζήτημα των ερμηνειών των χρηστών των εικόνων και πώς είναι εγγενώς δύσκολο να συλληφθούν στο οπτικό περιεχόμενο, το τελευταίο χάσμα κάνει την αναγνώριση εικόνας μία πρόκληση λόγω των περιορισμών στις ικανότητες καταγραφής και περιγραφής. Αυτήν την περίοδο, μόνο 10% των αρχείων εικόνας έχουν μια επαγγελματική περιγραφή (σχολιασμός). Κατά συνέπεια, οι μηχανές αναζήτησης εικόνας είναι μόνο ικανές να επιτύχουν μία ακρίβεια περίπου 42% και ανάκληση περίπου 12% [266], ενώ 60% των επισκεπτών μηχανών αναζήτησης χρησιμοποιούν τουλάχιστον δύο διαφορετικές μηχανές αναζήτησης δεδομένου ότι δεν ικανοποιούνται από το ανακτημένο περιεχόμενο. Η πιο κοινή καταγγελία είναι ότι οι μηχανές αναζήτησης δεν αναγνωρίζουν την σημασιολογία περιεχομένου. Επιπλέον περίπου 77% των ερευνητών αλλάζουν τις λέξεις κλειδιά περισσότερο από μία φορά επειδή δεν μπορούν να ανακτήσουν το περιεχόμενο ενδιαφέροντος τους [267, 268].

Τα συστήματα ανάκτησης εικόνων που βασίζονται σε κειμενικές περιγραφές, (στο εξής θα αναφέρονται ως ABIR) είναι μια προσπάθεια να ενσωματωθεί σημασιολογικό περιεχόμενο τόσο στις κειμενικές ερωτήσεις των χρηστών όσο και στους κειμενικούς περιγραφείς τίτλων εικόνας (όπως η Αναζήτηση εικόνας στο Google, η το Yahoo!). Οι τεχνικές *Latent Semantic Indexing (LSI)* που αρχικά εφαρμόστηκαν με αυξανόμενη επιτυχία στην ταξινόμηση εγγράφων ενσωματώθηκαν στα συστήματα ABIR για να καλύψουν την ανάγκη για μια πιο αξιόπιστη ενσωμάτωση εννοιών. Εντούτοις, το επίπεδο επιτυχίας σε αυτές τις προσπάθειες είναι αμφισβητήσιμο ένας λόγος για αυτό βρίσκεται στην έλλειψη επαρκών λέξεων κλειδιών ανά-εικόνα σε σύγκριση με τον αριθμό λέξεων κλειδιών που ορίζονται συνήθως ανά έγγραφο.

Η *Μαρκοβιανή σημασιολογική ταξινόμηση* (στο εξής MSI), είναι μια νέα μέθοδος με σκοπό τον αυτόματο σχολιασμό εικόνων με κειμενικές περιγραφές και βάσει αυτού του σχολιασμού ανάκτηση εικόνων. Οι ιδιότητες του MSI το καθιστούν ιδιαίτερα κατάλληλο για εργασίες ABIR ειδικά όταν οι ανά εικόνα κειμενικές περιγραφές είναι περιορισμένες. Τα χαρακτηριστικά της μεθόδου την καθιστούν επίσης καλύτερα εφαρμόσιμη στα πλαίσια πληροφοριακών συστημάτων πραγματικού χρόνου με διάδραση χρηστών.

5.2.1. Σχετικές Εργασίες

Η ανάπτυξη των συστημάτων ανάκτησης είναι αδιαφιλονίκητα ταχεία. Ήδη πρόσφατα, περισσότερα από 200 συστήματα ανάκτησης περιεχομένου αναπτύχθηκαν [269], η πλειοψηφία των οποίων είναι βασισμένη σε χαμηλού επιπέδου χαρακτηριστικά. Πιο συγκεκριμένα, τα συστήματα αυτά μπορούν να ταξινομηθούν σε δύο κύριες κατηγορίες: (α) εκείνα που εκτελούν εξόρυξη σημασιολογίας βασισμένη στην ανάλυση κειμενικών πληροφοριών που συνδέονται με τις εικόνες, όπως σχολιασμοί, λέξεις κλειδιά, τίτλοι, εναλλακτικό κείμενο στις σελίδες HTML ή το περιβάλλον κείμενο και (β) εκείνα που βασίζονται στην εξαγωγή χαμηλού επιπέδου οπτικών χαρακτηριστικών γνωρισμάτων όπως το χρώμα, ή υφή κλπ, προκειμένου να εκτελεσθεί ταξινόμηση, κατηγοριοποίηση, περιληπτική παρουσίαση της πληροφορίας κ.λπ. σε συλλογές εικόνων. Οι μέθοδοι της πρώτης κατηγορίας εξαρτώνται από τον επίμοχθο σχολιασμό, ενώ οι τελευταίες μέθοδοι δεν μπορούν συνήθως να συλλάβουν αποτελεσματικά τη σημασιολογία. Επιπλέον, μερικές άλλες τεχνικές χρησιμοποιούν συνδυασμό χαμηλού επιπέδου χαρακτηριστικών υπό μορφή οπτικών λέξεων-κλειδιών [270] και σχολιασμό κειμένων για να εκτελέσουν ανάκτηση με βάση το περιεχόμενο όμως απαιτούν συνήθως τη ρητή συμμετοχή των χρηστών για το γλωσσικό σχολιασμό των εικόνων [271, 272].

Τα συστήματα ανάκτησης που βασίζονται σε κειμενικές περιγραφές εικόνων (ABIR) ενσωματώνουν σημασιολογικό περιεχόμενο βασισμένο στο κείμενο των ερωτήσεων και στους τίτλους της εικόνας. Μια άμεση συνέπεια είναι ότι μέθοδοι που αρχικά αναπτύχθηκαν για την ανάκτηση εγγράφων μπορεί να είναι κατάλληλες και για τα συστήματα ABIR. Η μέθοδος (LSI) [273] αναπτύχθηκε αρχικά για την ανάκτηση εγγράφων. Ο Hofmann, βασισμένος σε αμιγώς στατιστική βάση [274] εισήγαγε την μέθοδο (pLSI) [275] ως εναλλακτική της (LSI) για κατηγοριοποίηση και ανάκτηση εγγράφων. Μία επίσης στατιστική μέθοδος, η (LDA) προτάθηκε [276] για να εξετάσει τους περιορισμούς της pLSI σχετικά με τη γενίκευση, ενώ οι Griffiths και Steyvers ενσωμάτωσαν στην μέθοδο LDA μια τεχνική Μόντε Κάρλο Μαρκοβιανής αλυσίδας [277]. Οι M. Steyvers et al εισήγαγαν ένα νέο πιθανολογικό πρότυπο που αντιπροσωπεύει και τους συντάκτες και τα θέματα στην ανάκτηση εγγράφων, ενσωματώνοντας μια δειγματοληψία Gibbs [278].

Προσπάθειες εφαρμογής των LSI/pLSI τεχνικών για μια πιο αξιόπιστη ενσωμάτωση σημασιολογίας στα συστήματα ABIR έχουν ήδη αναφερθεί [279, 280, 281, 282]. Υβριδικά μοντέλα που συνδυάζουν οπτικά χαρακτηριστικά και κείμενο έχουν επίσης αναφερθεί στην βιβλιογραφία της τεχνητής όρασης [283]. Ένα άλλο μοντέλο που βασίζεται στην κατά Bayes πιθανοθεωρητική επαγωγή προτάθηκε από τους J.Li et al [284], ενώ οι J.Fan et al, πρότειναν ένα σχέδιο βασισμένο στα τοπικά και ολικά οπτικά χαρακτηριστικά γνωρίσματα [285]. Η προτεινόμενη εν τω παρόντι προσέγγιση, ενώ είναι στοχαστική στη φύση της, εν τούτοις προάγει μια ποιοτική πτυχή συλλογισμού των πιθανοτήτων, δεδομένου ότι κα-

θορίζει ρητούς συνδέσμους σχετικότητας μεταξύ των λέξεων κλειδιών.

Η χρησιμοποίηση δικτύων για τη σύλληψη της σημασιολογίας είναι συνηθισμένη τεχνική στα συστήματα ΑΙ. Στο πλαίσιο της θεωρίας πιθανοτήτων καλούνται Bayesian networks, causal nets, η influence diagrams [286]· στην θεωρία Dempster-Shafer καλούνται galleries [287], Qualitative Markov Networks [288], η Constraint Networks [289]. Πολλά μοντέλα του ανθρώπινου συλλογισμού απεικονίζονται συχνά μέσω παραστάσεων γράφων (π.χ., τα σημασιολογικά δίκτυα [290], constraint networks [289], inference networks [291], conceptual dependencies [292], και conceptual structures [293]) και σε πολλές περιπτώσεις έχει δειχθεί ότι απεικονίζουν καλύτερα την ανθρώπινη ευλογοφάνεια [286].

5.2.2. Συνεισφορά της μεθόδου MSI σε εργασίες ABIR

Με τον όρο *πρόβλημα της μηδενικής συχνότητας* αναφερόμαστε στο γεγονός ότι η πιθανότητα να βρεθούν κοινές λέξεις κλειδιά στις περιγραφές ακόμη και στενά συνδεδεμένων εικόνων είναι εξαιρετικά μικρή διότι οι εικόνες δεν περιγράφονται με *ακριβώς* τις ίδιες λέξεις κλειδιά. Η μεθοδολογία που προτείνεται εδώ είναι μία νέα (εναλλακτική) στοχαστική προσέγγιση για συστήματα (ABIR) και θα δείξουμε ότι σε σύγκριση με τις μεθόδους LSI αντιμετωπίζει καλύτερα το πρόβλημα μηδενικής συχνότητας κατασκευάζοντας ένα ρητό σύνδεσμο συσχέτισης μεταξύ όλων των λέξεων-κλειδιών που φέρει ένα πιθανολογικό βάρος. Δείχνουμε ότι ορίζοντας λογικές συνδέσεις μεταξύ των λέξεων-κλειδιών με τη βοήθεια ενός μερικώς διευκρινισμένου Μαρκοβιανού μοντέλου, επιτυγχάνουμε καλύτερη γενίκευση στις περιπτώσεις αραιά σχολιασμένων σημαντικών περιοχών ως εκ τούτου η προτεινόμενη προσέγγιση εγείρει την ποιοτική δίπλα στην ποσοτική έννοια των πιθανοτήτων. Η βασική ιδέα πίσω από την προσέγγιση είναι να αντισταθμιστούν τα αραιά στοιχεία με την ενσωμάτωση μιας διαδικασίας σχολιασμών πιθανολογικού ποιοτικού συλλογισμού που θα διαδώσει τις μερικές πεποιθήσεις σχετικά με τις συνδέσεις μεταξύ των λέξεων κλειδιών. Ένας μηχανισμός που κερδίζει σε απόδοση εξάγοντας τη δομή των υπαρχόντων στοιχείων χωρίς απαραίτητα την ενσωμάτωση νέων στοιχείων, όπως συμβαίνει με τα παραδοσιακά μοντέλα, εισάγεται ως εκ τούτου.

Επιπλέον, ακόμα κι αν, βασισμένα στις κειμενικές περιγραφές / σχολιασμό, αυτόματα συστήματα ανάκτησης εικόνας έχουν παρουσιαστεί στη βιβλιογραφία, το προτεινόμενο σύστημα είναι νέο ως προς τον τρόπο που ενοποιεί αυτούς τους δύο στόχους. Πράγματι και ο αυτόματος τρόπος κειμενικής περιγραφής / σχολιασμών και η διαδικασία ανάκτησης θεωρούνται σε ένα ενιαίο πλαίσιο διάδρασης με τους χρήστες του συστήματος δυναμικά εξάγοντας σημασιολογία με σκοπό τον καθορισμό ενός στοχαστικού μοντέλου ποιοτικής διασύνδεσης. Αυτό έχει επιπτώσεις όσον αφορά την στοχοθέτηση δεδομένου ότι η εκπαίδευση της λογικής του συστήματος εκτελείται δυναμικά από τους ίδιους τους χρήστες που χρησιμοποιούν πραγματικά το σύστημα. Η ενοποιημένη Μαρκοβιανή στοχαστική

οργάνωση πίσω από το προτεινόμενο σύστημα επιτρέπει στην τεχνική ανάκτησης να επωφελείται από την ελλοχεύουσα δομή των στοιχείων κειμενικών περιγραφών - σχολιασμών ενώ συγχρόνως τα στοιχεία σχολιασμών αποκτούν συγκεκριμένη στοχαστική ερμηνεία. Μια σύγκριση με την μέθοδο LSI στην εφαρμογή των ABIR συστημάτων αποκαλύπτει ότι η προτεινόμενη προσέγγιση επιτυγχάνει καλύτερα αποτελέσματα ανάκτησης.

5.2.3. Η προτεινόμενη προσέγγιση

Στο [294], η χρήση ενός πίνακα συνδιακύμανσης για την αντιπροσώπευση των χαρακτηριστικών της εικόνας έκανε απαραίτητη την μεταφορά σε μία πολλαπλότητα Riemann προτού γίνει δυνατός ο καθορισμός μίας απόστασης. Αυτό το παράδειγμα δείχνει πόσο κρίσιμο είναι, ο ορισμός της απόστασης και η αντιπροσώπευση χαρακτηριστικών να γίνουν με ενοποιημένη προσέγγιση. Στην περίπτωση μας, οι εικόνες αντιπροσωπεύονται με διανύσματα πιθανότητας που ορίζονται αυτόματα από τη διαδικασία σχολιασμών, όπως θα εξηγηθεί στις επόμενες παραγράφους. Ένα διανυσματικός χώρος είναι επομένως διαθέσιμος βάσει του οποίου θα προτείνουμε μια απόσταση που έχει συγκεκριμένη στοχαστική ερμηνεία.

Η προτεινόμενη προσέγγιση θα παρουσιαστεί στο πλαίσιο ενός διαδραστικού συστήματος ανάκτησης εικόνων (παρόμοιου με το σύστημα Google Image) όπου οι χρήστες ψάχνουν για εικόνες με την υποβολή ερωτήσεων που αποτελούνται από λέξεις κλειδιά. Το σύστημα αποκρίνεται στις ερωτήσεις επιστρέφοντας διάφορες εικόνες. Ο χρήστης ερευνά τις επιστρεφόμενες εικόνες και μεταφορτώνει τις σχετικές με τις ανάγκες / προτιμήσεις του διαφορετικά, αγνοεί τα αποτελέσματα και υποβάλλει μια άλλη ερώτηση για περαιτέρω επεξεργασία. Αυτή η διαδραστική διαδικασία έχει υπονοούμενες συνέπειες που εκμεταλλευόμαστε μία προς μία σε μια βαθμιαία κατασκευή του προτεινόμενου συστήματος:

Βήμα 1:

Ο χρήστης σιωπηρά συσχετίζει τις ανακτημένους (μεταφορτωμένες) εικόνες με την ερώτησή του (λέξεις κλειδιά). Υποθέτοντας Μαρκοβιανές μεταβάσεις στις λέξεις κλειδιά ο στόχος της προτεινόμενης προσέγγισης είναι να ποσοτικοποιηθούν οι λογικές συνδέσεις μεταξύ των λέξεων κλειδιών. Εάν κάποιος χρήστης συσχετίζει την εικόνα I_i με την ερώτησή του q_i , όπου η λέξη κλειδί k_2 ακολουθεί την λέξη κλειδί k_1 και αυτό συμβαίνει m φορές, τότε η πιθανότητα μετάβασης πρώτης τάξης $p_i(k_1, k_2)$ ανάμεσα σε αυτές τις λέξεις κλειδιά ενημερώνεται ως εξής:

Εάν $p_i(k_1, k_2)$ είναι η τρέχουσα τιμή πιθανότητας (πριν την ενημέρωση) υπολογισμένη από M λέξεις κλειδιά τότε η καινούργια τιμή της πιθανότητας (υπολογισμένη σε $M + m$ λέξεις κλειδιά) υπολογίζεται από την αναδρομική εξίσωση:

$$p_i(k_1, k_2) = \frac{Mp_i(k_1, k_2) + m}{M + m} \quad (54)$$

Αυτή η διαδικασία κατασκευάζει μια Μαρκοβιανή αλυσίδα όπου κάθε λέξη κλειδί αντιστοιχεί σε μία κατάσταση. Κάθε φορά που εμφανίζεται μια λέξη κλειδί σε μια ερώτηση, ο μετρητής της κατάστασης της αυξάνεται, εάν μια άλλη λέξη κλειδί ακολουθεί στην ίδια ερώτηση, ο διακαταστατικός μετρητής των συνδέσεων τους επίσης αυξάνεται. Οι εμφανίσεις των λέξεων κλειδιών αλλά και η αλληλουχία αυτών των εμφανίσεων μετριοούνται και τα δύο με αυτόν τον τρόπο. Οι ερωτήσεις σχετικά με μια εικόνα επεξεργάζονται ομαδικά για αυτήν την εικόνα, οι μετρητές αυξάνονται, οι πιθανότητες ενημερώνονται ως ανωτέρω. Προτού υποβληθεί σε επεξεργασία το επόμενο σύνολο ερωτήσεων, οι μετρητές μηδενίζονται. Το διάνυσμα της κατάστασης ισορροπίας της Μαρκοβιανής αλυσίδας που κατασκευάστηκε με αυτόν τον τρόπο, για κάθε εικόνα I_i , θα συμβολίζεται με π_i και θα αντιπροσωπεύσει την εικόνα από τώρα και στο εξής.

Αυτή η προσέγγιση μοντελοποίησης δικαιολογείται: (α) από τον ποιοτικό χαρακτήρα της εννοιολογικής προσέγγισης που υιοθετείται σε αυτήν την εργασία, και (β) τους σκοπούς στοχοθέτησης του *MSI*, δηλαδή την ανάγκη να συλληφθούν οι συγκεκριμένες, αυτών των χρηστών αντιλήψεις για τις εικόνες. Πράγματι, το γεγονός ότι κάθε ακολουθία λέξεων κλειδιών (ερώτηση) προέρχεται από έναν συγκεκριμένο χρήστη, που φιλτράρεται μέσω της αντίληψής του για την επιλεγμένη εικόνα, δικαιολογεί την ανάθεση μιας λογικής σύνδεσης μεταξύ αυτών των λέξεων κλειδιών, πέρα από τις μεμονωμένες συνδέσεις μεταξύ κάθε λέξης κλειδιού και της επιλεγμένης εικόνας. Είναι αυτή η λογική σύνδεση που ευνοείται από την προτεινόμενη εδώ προσέγγιση αντί για τον υπολογισμό της κατανομής των εικόνων ως προς τις λέξεις κλειδιά με τον παραδοσιακό στατιστικό τρόπο.

Βήμα 2: Μια προσπάθεια για άμεση σύγκριση ανάμεσα στα διανύσματα πιθανότητας π_i και π_j που υπολογίστηκαν στο προηγούμενο βήμα και αντιπροσωπεύουν τις δύο εικόνες, θα προσέκρουε, όπως αναφέρθηκε παραπάνω, στο πρόβλημα της μηδενικής συχνότητας. Από το ίδιο το γεγονός, ότι ένας χρήστης βάζει ορισμένες λέξεις κλειδιά μαζί, (σε μια ερώτηση) δηλώνει σιωπηρά ότι οι αυτές λέξεις κλειδιά είναι κατά κάποιο τρόπο σχετικές μεταξύ τους ανεξάρτητα από τις εικόνες που θα επιλεγθούν η δεν θα επιλεγθούν από αυτόν τον χρήστη. Προτείνουμε λοιπόν να χρησιμοποιήσουμε αυτό το γεγονός και να αντιμετωπίσουμε το πρόβλημα της μηδενικής συχνότητας στην βάση ύπαρξης συστάδων από *παρόμοιες* λέξεις κλειδιά. Για τον σκοπό αυτό, η συνολική Μαρκοβιανή αλυσίδα (*AMC*) όλων των ερωτήσεων που ρωτώνται από όλους τους χρήστες, ανεξάρτητα από τις επιλεγμένες εικόνες, κατασκευάζεται σε αυτό το βήμα. Ο πίνακας πρώτης τάξης αυτής της διαδικασίας που συμβολίζεται με P_G , υπολογίζεται με έναν παρόμοιο με τον προηγούμενο τρόπο βημάτων από τον αναδρομικό τύπο της εξίσωσης (54). Ο P_G , αν και είναι Μαρκοβιανός πίνακας θα χρησιμοποιηθεί για να ορισθούν συστάδες στον χώρο των λέξεων κλειδιών και όχι για τον υπολογισμό μιας κατανομής πιθανότητας, όπως συνήθως συμβαίνει με τα Μαρκοβιανά μοντέλα. Ως εκ τούτου ο σκοπός του *AMC* είναι να ορισθούν βάρη σχετικότητας (συνδεσιμότητας) ανάμεσα στις λέξεις κλειδιά μέσω των συστάδων τους.

Βήμα 3:

Βήμα βελτιστοποίησης. Ο πίνακας AMC θα χρησιμοποιηθεί εδώ για να καθοριστούν συστάδες στις λέξεις κλειδιά και ως εκ τούτου να ορισθούν ρητές συνδέσεις σχετικότητας μεταξύ των λέξεων κλειδιών. Αυτός ο στόχος συγκέντρωσης συνδέεται με τα χαρακτηριστικά σύγκλισης της Μαρκοβιανής αλυσίδας AMC όπως αυτή καθορίζεται στην σειρά $F_G(n) = \sum_{k=0}^n P_G^k$ όπου P_G είναι ο πίνακας μετάβασης πρώτης τάξης AMC . Ο $F_G(n)$ είναι ο πίνακας αναμενόμενων ποσοστών επισκεψιμότητας n -τάξης (Ενότητα 5.1.3). Μία κατάλληλη συνθήκη τερματισμού σταματά την σειρά στο επιθυμητό n όπου η αργή σύγκλιση έχει αναλάβει, αλλά όχι προτού τελειώσει η γρήγορη σύγκλιση. Η τιμή της ορίζουσας του $F_G(n)$ χρησιμοποιείται ως συνθήκη τερματισμού δεδομένου ότι οι συστάδες στις σειρές του $F_G(n)$ θα ριξουν την τάξη του και τότε η ορίζουσα θα γίνει σχεδόν μηδέν. Μία διαδικασία βελτιστοποίησης εμπλέκεται σε αυτήν την διαδικασία σύγκλισης και αφορά τη συνολική διασπορά των στηλών του $F_G(n)$, όταν προβάλλονται στην κατεύθυνση των ιδιοδιανυσμάτων του πίνακα P_G . Περαιτέρω ανάλυση σε αυτήν την έννοια δίδεται σε σχετικές μας δημοσιεύσεις. Από τώρα και στο εξής, θα χρησιμοποιήσουμε ακριβώς F_G για να υπονοούμε $\frac{1}{n+1}F_G(n)$, τον n -τάξης πίνακα αναμενόμενων ποσοστών επισκεψιμότητας που έχει υπολογισθεί στο επιθυμητό n .

Βήμα 4:

Μπορούμε σε αυτό το σημείο να δώσουμε τον ορισμό της απόστασης MSI:

Ορισμός 9 Έστω i και j δύο εικόνες που αναπαρίστανται από τα αντίστοιχα διανύσματα ισορροπίας¹¹ π_i και π_j . Ας είναι επίσης $\Sigma(F_G^T)$ ο πίνακας συνδιακύμανσης του ανάστροφου πίνακα αναμενόμενων ποσοστών επισκεψιμότητας n -τάξης της Μαρκοβιανής αλυσίδας (AMC), που έχει υπολογιστεί στο επιθυμητό n και έχει μηδενικό μέσο. Τότε η απόσταση MSI ανάμεσα στις εικόνες i και j ορίζεται ως:

$$\begin{aligned} d(i, j) &= (\pi_i - \pi_j) \Sigma(F_G^T) (\pi_i - \pi_j)^T \\ &= \delta_{ij} \Sigma(F_G^T) \delta_{ij}^T \end{aligned} \quad (55)$$

όπου οι διαστάσεις του π_i και του π_j έχουν επεκταθεί σε αυτές του $\Sigma(F_G^T)$ συμπληρώνοντας με μηδενικά στις επιπλέον θέσεις. □

Η MSI απόσταση είναι καλώς ορισμένη σαν γενικευμένη Ευκλείδεια συνάρτηση απόστασης που χρησιμοποιεί πίνακα συνδιακύμανσης.

5.2.4. Σύγκριση με LSI

Συγκρίνουμε εδώ την προτεινόμενη μέθοδο MSI με την LSI σε εφαρμογές ABIR.

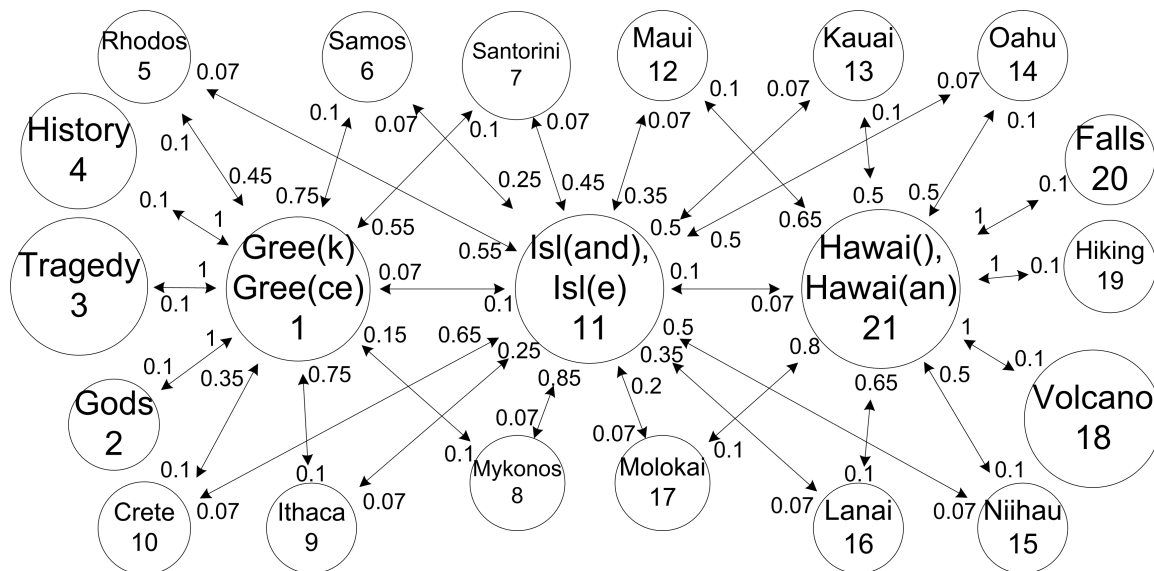
¹¹Θεωρούμενα ως διανύσματα γραμμές

	K1	K2	K3	K4	K5	K6	K7	K8	K9	K10	K11	K12	K13	K14	K15	K16	K17	K18	K19	K20	K21	
	GRE	GOD	TRA	HIS	RHO	SAM	SAN	MYK	ITH	CRE	ISL	MAU	KAU	OAH	NIH	LAN	MOL	VOL	HIK	FAL	HAW	
GRE	-	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	-	-	-	-	-	-	-	-	-	-	-
GOD	1.0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
TRA	1.0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
HIS	1.0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
RHO	0.45	-	-	-	-	-	-	-	-	-	0.55	-	-	-	-	-	-	-	-	-	-	-
SAM	0.75	-	-	-	-	-	-	-	-	-	0.25	-	-	-	-	-	-	-	-	-	-	-
SAN	0.55	-	-	-	-	-	-	-	-	-	0.45	-	-	-	-	-	-	-	-	-	-	-
MYK	0.15	-	-	-	-	-	-	-	-	-	0.85	-	-	-	-	-	-	-	-	-	-	-
ITH	0.75	-	-	-	-	-	-	-	-	-	0.25	-	-	-	-	-	-	-	-	-	-	-
CRE	0.35	-	-	-	-	-	-	-	-	-	0.65	-	-	-	-	-	-	-	-	-	-	-
ISL	0.07	-	-	-	0.07	0.07	0.07	0.07	0.07	0.07	-	0.07	0.07	0.07	0.07	0.07	0.07	-	-	-	-	0.07
MAU	-	-	-	-	-	-	-	-	-	-	0.35	-	-	-	-	-	-	-	-	-	-	0.65
KAU	-	-	-	-	-	-	-	-	-	-	0.5	-	-	-	-	-	-	-	-	-	-	0.5
OAH	-	-	-	-	-	-	-	-	-	-	0.5	-	-	-	-	-	-	-	-	-	-	0.5
NIH	-	-	-	-	-	-	-	-	-	-	0.5	-	-	-	-	-	-	-	-	-	-	0.5
LAN	-	-	-	-	-	-	-	-	-	-	0.35	-	-	-	-	-	-	-	-	-	-	0.65
MOL	-	-	-	-	-	-	-	-	-	-	0.2	-	-	-	-	-	-	-	-	-	-	0.8
VOL	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1.0
HIK	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1.0
FAL	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1.0
HAW	-	-	-	-	-	-	-	-	-	-	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	-

Πιν. 3: Ο πίνακας μετάβασης πρώτης τάξης της Μαρκοβιανής αλυσίδας του Σχ. 22.

64 εικόνες που διαμορφώνουν δύο διαισθητικές κατηγορίες χρησιμοποιήθηκαν για αυτό το πείραμα. 32 εικόνες είναι σχετικές με τον όρο *Greek* και θεωρούνται ότι να ανήκουν στην πρώτη κατηγορία και 32 εικόνες είναι σχετικές με τον όρο *Hawaii* και θεωρούνται ότι ανήκουν στην δεύτερη κατηγορία. Αρχικά, η απόσταση των 64 εικόνων από την ερώτηση "*Greek Islands*" υπολογίζεται και ταξινομείται και για τις δύο μεθόδους και τα αποτελέσματα εξετάζονται. Κατόπιν, ένας πλήρης πίνακας απόστασης χιτίζεται για όλες τις κατά ζεύγη αποστάσεις αυτών των 64 εικόνων χρησιμοποιώντας και τις δύο μεθόδους. Η σύγκριση γίνεται με το διάγραμμα ακρίβειας / ανάκλησης.

Ζητήσαμε από μια ομάδα χρηστών να μας αφήσει να καταγράψουμε τις ερωτήσεις και τα στοιχεία τους (διεύθυνση ip, ερώτηση q , επιστρεφόμενη ταξινόμηση και εικόνες r και κλικς c) κατά την αλληλεπίδρασή τους με την μηχανή αναζήτησης εικόνας Google Image Search που χρησιμοποιεί την Αγγλική γλώσσα. Οι χρήστες κλήθηκαν σε διαφορετικές περιόδους να μιμηθούν διαφορετικές ομάδες-στόχους με συγκεκριμένα ενδιαφέροντα (τουρίστες, φίλαθλοι, οπαδοί καλλιτέχνες κ.λπ.) δεδομένου ότι η προσαρμοστικότητα του προτεινόμενου συστήματος στις προτιμήσεις του χρήστη έπρεπε να εξεταστεί. Ένα απλό αρχείο καταγραφής χρησιμοποιήθηκε για την κράτηση των ερωτήσεων-ταυτότητα μηχανήματος, URL, εικόνες και αρχείο κλικς. Από τις καταγραμμένες συνόδους, για κάθε



Σχήμα 22: Η συνολική Μαρκοβιανή διαδικασία του πειράματος της Ενότητας 5.2.4. Μόνο το μέρος του δικτύου που διαμορφώνει συστάδες γύρω από τις έννοιες *Greek*, *Hawaiian* και *Islands* παρουσιάζεται γραφικά. Οι πιθανότητες μετάβασης που οδηγούν έξω από αυτό το μέρος του δικτύου έχουν τεθεί μηδέν και το βάρος τους διανεμήθηκε ομοιόμορφα στις αντίστοιχες συνδέσεις εντός του εμφανούς δικτύου. Οι πιθανότητες μετάβασης μικρότερες από 0.05 παρελήφθησαν.

επιλεγμένη εικόνα, η μεμονωμένη Μαρκοβιανή αλυσίδα της χτίστηκε, με βάση τους όρους ερώτησης (λέξεις κλειδιά), όπως εξηγείται ανωτέρω. Μόνο οι επιλεγμένες εικόνες μεταξύ των εικόνων που επιστράφηκαν από το Google Image Search εξετάστηκαν. Το διάνυσμα ισορροπίας της αντίστοιχης Μαρκοβιανής αλυσίδας χρησιμοποιήθηκε μετέπειτα για να περιγράψει αυτόματα κάθε εικόνα. Κατά τρόπο παρόμοιο, από το συνολικό αριθμό των ερωτήσεων από όλους τους χρήστες, ανεξάρτητα από τις επιλεγμένες εικόνες, μια Μαρκοβιανή αλυσίδα, που αντιπροσωπεύει τη συνολική Μαρκοβιανή αλυσίδα, κατασκευάστηκε, χρησιμεύοντας ως μια δυναμικά προσαρμοζόμενη ταξινόμηση για τις λέξεις κλειδιά που έχει δει το σύστημα στις ερωτήσεις χρηστών. Ένα υποσύνολο του δικτύου αυτού, που διαμόρφωσε μια συστάδα γύρω από τις λέξεις κλειδιά *Greek* και *Hawaiian* παρουσιάζεται γραφικά στο Σχ. 22. Οι πιθανότητες μετάβασης χαμηλότερες του 0.05 έχουν παραλειφθεί.

Στον πίνακα (4) μπορούμε να δούμε μια ταξινόμηση των 64 εικόνων σύμφωνα με τη σημασιολογική απόστασή τους από την ερώτηση *Greek Islands* για τις δύο υπο σύγκριση μεθόδους. Η πρώτη στήλη κρατά την αρχική απαρίθμηση για κάθε εικόνα, οι στήλες 4-24 κρατούν το διάνυσμα ισορροπίας της Μαρκοβιανής αλυσίδας που αντιπροσωπεύει κάθε εικόνα, οι δεύτερη και τρίτη στήλες αυτού του πίνακα κρατούν την ταξινόμηση όσον αφορά την απόσταση από την ερώτηση *Greek Islands* για τις μεθόδους MSI και LSI αντίστοιχα.

Η ερώτηση *Greek Islands*, αντιπροσωπεύεται από το διάνυσμα $[Greek(0.5), Islands(0.5)]$. Όπως μπορεί να δει κανείς στην ταξινόμηση των εικόνων που επιστρέφονται από το προ-

τεινόμενο σύστημα (MSI) στον πίνακα (4), η σημασιολογική εγγύτητα των λέξεων κλειδιών, όπως αυτή υπονοείται από την συνδεσιμότητά τους στο συνολικό Μαρκοβιανό δίκτυο, επιτρέπει ευφυή ανάκτηση βάσει εξαρτήσεων πέρα από τα απλά ομο-περιστατικά λέξεων κλειδιών. Ενώ, π.χ., οι εικόνες (1), (2), και (6) είναι οι τρεις πιο κοντινές στην ερώτηση, όπως αναμένεται δεδομένου ότι μοιράζονται τις ίδιες λέξεις κλειδιά, οι εικόνες (3), (4) και (5) βρίσκονται στις 29η, 25η και 17η θέσεις αντίστοιχα, παρά το ότι μοιράζονται τις ίδιες ακριβώς λέξεις κλειδιά με τις εικόνες (1), (2) και (6). Πριν από τις (3), (4) και (5), συναντάμε στις θέσεις από 4 έως 16 πολλές εικόνες που περιλαμβάνουν πρόσθετες λέξεις κλειδιά. Αυτό συμβαίνει επειδή, όπως μπορεί να δει κανείς στην συνολική Μαρκοβιανή αλυσίδα του Σχ. 22, αρχικοποιώντας την Μαρκοβιανή διαδικασία στις αρχικές καταστάσεις που αντιπροσωπεύει κάθε μια από αυτές τις εικόνες (π.χ., (15), (14), (7), (8), (11), (9) κ.λπ.), το δίκτυο είναι πιο κοντά, υπό την έννοια του αναμενόμενου αριθμού μεταβάσεων, να φθάσει σε μία κατάσταση που αντιπροσωπεύεται από την ερώτηση, από ότι είναι να φθάσει σε κατάσταση που αντιπροσωπεύεται από τους αρχικούς όρους που αντιστοιχούν στις εικόνες (3), (4) ή (5).

Μια παρόμοια κατάσταση παρατηρείται πιο κάτω στην ταξινόμηση, όπου, π.χ., η εικόνα (14) προηγείται πολλών εικόνων που περιλαμβάνουν και τις δύο λέξεις κλειδιά ερώτησης, παρά το ότι περιλαμβάνει μόνο μια από τις δύο λέξεις κλειδιά ερώτησης. Αυτό εξηγείται επίσης από τη συνολική διαδικασία του Σχ. 22, όπου έχει καταγραφεί ότι η πλειοψηφία των χρηστών είναι πιθανότερο να ρωτήσει οποιοδήποτε εκ των λέξεων κλειδιών ερώτησης *Island* η *Greek*, αφότου ρωτήσουν *Santorini*, από ότι είναι αφότου ρωτήσει οποιοδήποτε από τα άλλα ονόματα νησιών.

Βλέπουμε επίσης σε αυτήν την ταξινόμηση, ότι όλες οι εικόνες που περιλαμβάνουν *Greek* προηγούνται των εικόνων που περιλαμβάνουν *Hawaii* όπως θα πρέπει να συμβεί και απεικονίζεται σαφώς στην Μαρκοβιανή αλυσίδα *AMC*. Αυτό είναι πράγματι έτσι, δεδομένου ότι όλα τα Ελληνικά νησιά έχουν υψηλότερη πιθανότητα να οδηγήσουν και στους δύο όρους ερώτησης από ότι έχουν τα νησιά της Χαβάης. Επίσης παρατηρούμε ότι οι εικόνες που περιλαμβάνουν οποιοδήποτε από τους τρεις όρους (*Gods*, *Tragedy*, *History*) που συνδέονται μόνο με *Greek* αλλά όχι με *Island* έρχονται μετά από τις εικόνες που περιλαμβάνουν τα Ελληνικά νησιά, αλλά πριν από τις εικόνες που περιλαμβάνουν τα νησιά της Χαβάης, όπως πρέπει να είναι επιθυμητό, διότι ακόμα κι αν συνδέονται μόνο με μια λέξη κλειδί, όλες οι λέξεις κλειδιά που περιλαμβάνουν τα νησιά της Χαβάης οδηγούν, με υψηλή πιθανότητα, μακριά από τουλάχιστον μια από τις λέξεις κλειδιά της ερώτησης.

Όσον αφορά τις *LSI* ταξινομήσεις παρατηρούμε ότι ενώ αρχικά επιστρέφουν πιο διαισθητικά αποτελέσματα, όπως παραδείγματος χάριν τις εικόνες (1) - (6) που είναι οι πιο κοντινές στην ερώτηση δεδομένου ότι μοιράζονται ακριβώς τις ίδιες λέξεις κλειδιά, πιο κάτω στην ταξινόμηση βλέπουμε μη-διαισθητικά αποτελέσματα όπως, π.χ., στη 10η θέση που συναντάμε την εικόνα 40, ένα νησί της Χαβάης, να είναι πιο κοντά στην ερώτηση

The table displays a distance matrix between 64 images, grouped into two categories: Greek (G1-G15) and Hawaii (H33-H47). Each cell in the matrix contains two numerical values: the top value represents the Mean Similarity Index (MSI) and the bottom value represents the Lower Similarity Index (LSI). The diagonal elements are all 1.00, indicating zero distance from self-comparisons. The matrix is symmetric, with the lower triangle mirroring the upper triangle. The images are ordered as G1, G2, G3, G4, G5, G6, G7, G8, G9, G10, G11, G12, G13, G14, G15, H33, H34, H35, H36, H37, H38, H39, H40, H41, H42, H43, H44, H45, H46, H47.

Σχήμα 23: Ο πίνακας απόστασης μεταξύ 30 εικόνων δύο κατηγοριών. Οι εικόνες G1–G15 είναι οι πρώτες 15 εικόνες του πίνακα 4 και ανήκουν στην κατηγορία *Greek* ενώ οι εικόνες H33–H47 αντιπροσωπεύονται από τις αντίστοιχες 15 σειρές του ίδιου πίνακα και ανήκουν στην κατηγορία *Hawaii*. Η προτεινόμενη απόσταση *MSI* (άνω αριθμός) και η απόσταση *LSI* (κάτω αριθμός) συγκρίνονται. Οι εννέα καλύτερες αντιστοιχίες είναι μαρκαρισμένες με τους αντίστοιχους εκθέτες. Η *MSI* επιτυγχάνει καλύτερα αποτελέσματα δεδομένου ότι όλες οι πρώτες εννέα αντιστοιχίες είναι σωστές (οι εικόνες ταιριάζουν με τις εικόνες στην ίδια κατηγορία), ενώ πολλές από τις πρώτες εννέα αντιστοιχίες δεν είναι σωστές για την *LSI*. Το διάγραμμα ακρίβειας / ανάκλησης για αυτό το πείραμα, για όλες τις 64 εικόνες του πίνακα 4, παρουσιάζεται στο Σχ. 24.

Greek Island από πολλές εικόνες που περιλαμβάνουν Ελληνικά νησιά. Η συνέπεια που παρατηρήσαμε με το *MSI*, ότι δηλαδή όλα τα Ελληνικά νησιά προηγήθηκαν των νησιών της Χαβάης, χάνεται στη *LSI* ταξινόμηση. Για να εξετάσουμε περαιτέρω την συμπεριφορά των δύο μεθόδων κατασκευάσαμε έναν πίνακα απόστασης μεταξύ όλων των ζευγαριών των 64 εικόνων και θεωρώντας ότι οι εικόνες (1) - (32) ανήκουν στην κατηγορία *Greek* και οι εικόνες (33) - (64) ανήκουν στην κατηγορία *Hawaiian*, ένα διάγραμμα ακρίβειας / ανάκλησης κατασκευάστηκε για τις δύο μεθόδους.

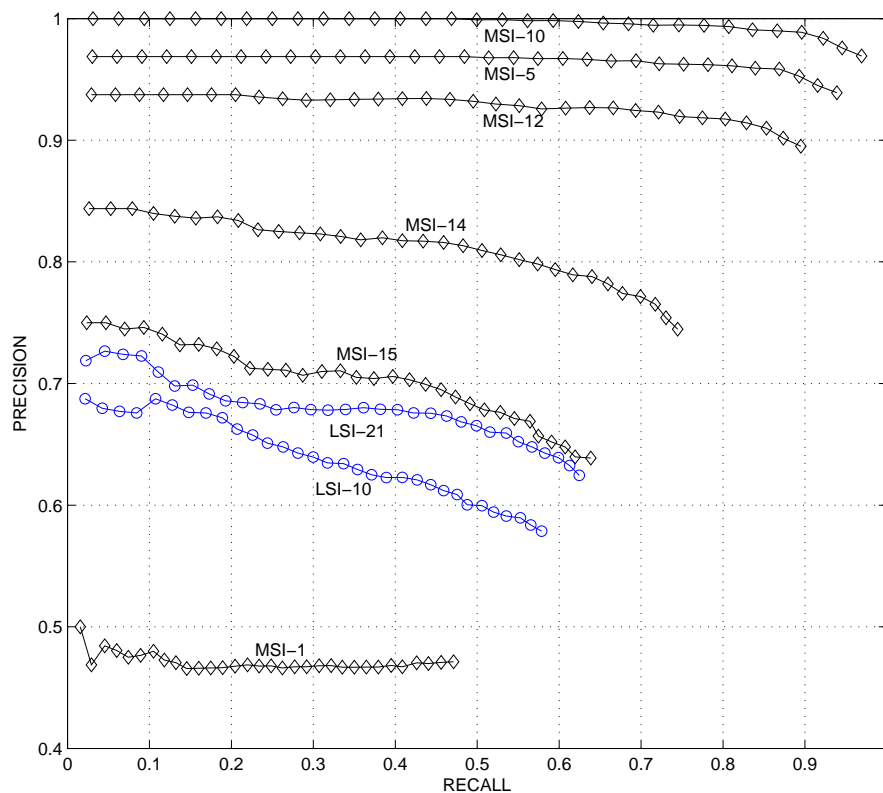
Στο Σχ. 24 μπορούμε να δούμε την ακρίβεια σε σχέση με την ανάκληση για την μέθοδο *MSI* για $n = 1, 5, 10, 12, 14, 15$ και *LSI* για $k = 21$ (καμία μείωση διαστατικότητας) και $k = 10$. Είναι εμφανές ότι η *LSI* δεν μπορεί να συλλάβει τα λανθάνοντα χαρακτηριστικά γνωρίσματα που συλλαμβάνονται με τη βοήθεια του *AMC* και χρησιμοποιεί η *MSI*, αφού το καλύτερο *LSI* αποτέλεσμα, που πραγματοποιείται όπως αναμενόταν όταν δεν εφαρμόζεται καμία μείωση στις διαστάσεις, είναι πολύ φτωχό έναντι των αποτελεσμάτων *MSI*. Για την *MSI* βλέπουμε πώς η επιλογή του n είναι κρίσιμη για την ακρίβεια της μεθόδου, δεδομένου ότι όσο το n αυξάνει βελτιώνεται και η ακρίβεια έως ότου επιτυγχάνεται ένα μέγιστο για $n = 10$ όπου το αποτέλεσμα *MSI* ταξινομεί σχεδόν όλες τις εικόνες στη σωστή κατηγορία. Για $n > 10$ η ακρίβεια μειώνεται αργά και είναι συγκρίσιμη με το καλύτερο *LSI* αποτέλεσμα μόνο για $n > 15$.

Στο σχήμα 23 μπορούμε να δούμε τον πίνακα απόστασης μεταξύ των πρώτων 15 Ελληνικών εικόνων¹² και των πρώτων 15 εικόνων της Χαβάης¹³ του πίνακα 4, όπου παρατηρούμε το τέλειο αποτέλεσμα του *MSI*, για $n = 10$, όπου όλες οι πρώτες 9 ταξινομήσεις για όλες τις 30 εικόνες είναι στη σωστή κατηγορία σε αντιδιαστολή με τις *LSI* ταξινομήσεις όπου πολλές εικόνες αντιστοιχούνται στην λανθασμένη εκ των δύο κατηγοριών.

Αυτά τα αποτελέσματα δικαιώνουν την προσπάθειά μας για μια στοχαστική ποιοτική προσέγγιση στο σημασιολογικό ζήτημα αραιών περιοχών. Η ακρίβεια /ανάκληση επιτυγχάνει σχεδόν τέλεια αποτελέσματα για ορισμένες τιμές μίξης του προτεινόμενου Μαρκοβιανού μοντέλου, ενώ την ίδια στιγμή οι περιγραφές των εικόνων έχουν εκτελεσθεί σωτηρά από τα στοιχεία ερωτήσεων των ίδιων των χρηστών, επομένως βλέπουμε το προτεινόμενο σύστημα να προσαρμόζεται δυναμικά στη συγκεκριμένη ομάδα χρηστών που χρησιμοποιούν πραγματικά το σύστημα. Η ανάγκη για εξωτερικά συστήματα ταξινόμησης προκειμένου να ορισθούν μετρικές σχετικότητας μεταξύ των λέξεων κλειδιών, και, επομένως, η επίλυση ενός νέου προβλήματος, της συμβατότητας μεταξύ εκείνων των συστημάτων και της σημασιολογίας πίσω από τους πραγματικούς χρήστες που χρησιμοποιούν το σύστημα, εξαλείφεται με την χρήση του *MSI*. Ο μηχανισμός πίσω από το *LSI* συσχετίζει τις λέξεις κλειδιά που εμφανίζονται στις ίδιες εικόνες και τις εικόνες που περιγράφονται με τις ίδιες λέξεις κλειδιά. Ξεκινώντας από ένα πίνακα συχνότητας, λέξης κλειδιού / εικόνας

¹²Απαριθμημένες ως Γ1-Γ15.

¹³Απαριθμημένες ως H33-H47, κατά την αντιστοιχία στους δείκτες του πίνακα 4.



Σχήμα 24: Σύγκριση ακρίβειας / ανάκλησης μεταξύ *LSI* και του προτεινόμενου συστήματος *MSI* για διάφορες παραμέτρους των δύο αλγορίθμων. Το πείραμα εκτελείται στις 64 εικόνες του πίνακα 4. Οι πρώτες 32 εικόνες ανήκουν στην κατηγορία Greek και οι επόμενες 32 στην κατηγορία Hawaiian. Οι γραφικές παραστάσεις διαμαντιών αντιπροσωπεύουν τα αποτελέσματα της προτεινόμενης απόστασης (που υπονοείται από την γενική Μαρκοβιανή διαδικασία του πίνακα 3) για τις περιπτώσεις $n = 1, n = 5, n = 10, n = 12, n = 14, n = 15$. Οι γραφικές παραστάσεις κύκλων αντιπροσωπεύουν τα αποτελέσματα όταν χρησιμοποιείται η *LSI* απόσταση με τον πίνακα συχνότητας εικόνας/λέξης κλειδιού του πίνακα 4 για τις παραμέτρους $k = 10$ και $k = 21$. Περισσότερα στο κείμενο.

μια *Ανάλυση Ιδιαζουσών Τιμών* παρέχει μια συμπαγέστερη αντιπροσώπευση των λέξεων κλειδιών και των εικόνων σε ένα χώρο λιγότερων διαστάσεων.

Στην *MSI* προσέγγισή μας από την άλλη πλευρά χρειαζόμαστε ένα τετραγωνικό πίνακα μετάβασης πρώτης τάξης της μορφής λέξη κλειδί / λέξη κλειδί. Ο πίνακας αυτός είναι ο (*AMC*) απο τον οποίο προκύπτουν οι σχέσεις μεταξύ των λέξεων κλειδιών. Έπειτα χρησιμοποιούμε τον πίνακα συχνότητας λέξης κλειδιού/εικόνας για να αντιπροσωπεύσουμε τα διανύσματα εικόνας στη διαφορά των οποίων προβάλλονται οι στήλες του πίνακα συνδιακύμανσης.

Ο μηχανισμός πίσω από το *MSI* αποκαλύπτεται στη συμπεριφορά του συστήματος όσον αφορά τη δύναμη n . Βλέπουμε ότι καθώς το n αυξάνεται τα αποτελέσματα βελτιώνονται έως ότου επιτυγχάνεται το καλύτερο αποτέλεσμα ανάκλησης ακρίβειας για $n = 10$. Αυτή η συμπεριφορά δείχνει ότι κατορθώσαμε να βελτιώσουμε την ανάκτηση χωρίς προσθήκη παρατηρήσεων στο σύστημα αλλά με τη διάδοση των συνδέσεων σχετικότητας που καθορίστηκαν στις υπάρχουσες παρατηρήσεις. Την στιγμή που οι μέθοδοι *LSI* ή *pLSI* θα χρειάζονταν περισσότερες παρατηρήσεις για να βελτιώσει τα αποτελέσματα, το προτεινόμενο σύστημα χρησιμοποιεί τη σύγκλιση της Μαρκοβιανής αλυσίδας (με την αύξηση του *AMC* στην κατάλληλη δύναμη) για να εκμεταλλευτεί περαιτέρω τις υπάρχουσες παρατηρήσεις εξάγοντας βαθύτερα, ποιοτικά συμπεράσματα μεταξύ των λέξεων κλειδιών.

IMG #	RANK		Λέξεις-Κλειδιά			
	MSI	LSI				
1	1	1	GRE(0.5)	-	ISL(0.5)	-
2	2	2	GRE(0.4)	-	ISL(0.6)	-
3	29	6	GRE(0.3)	-	ISL(0.7)	-
4	25	3	GRE(0.8)	-	ISL(0.2)	-
5	17	4	GRE(0.2)	-	ISL(0.8)	-
6	3	5	GRE(0.7)	-	ISL(0.3)	-
7	6	29	GRE(0.2)	-	ISL(0.3)	CRE(0.5)
8	7	25	GRE(0.1)	-	ISL(0.3)	CRE(0.6)
9	9	24	GRE(0.3)	-	ISL(0.2)	CRE(0.5)
10	15	11	GRE(0.2)	-	ISL(0.1)	CRE(0.7)
11	8	31	GRE(0.4)	-	ISL(0.2)	ITH(0.4)
12	20	58	GRE(0.2)	-	ISL(0.3)	ITH(0.5)
13	10	62	GRE(0.1)	-	ISL(0.1)	MYK(0.8)
14	5	63	-	-	ISL(0.1)	MYK(0.9)
15	4	38	GRE(0.3)	-	ISL(0.1)	MYK(0.6)
16	26	61	GRE(0.1)	-	ISL(0.2)	MYK(0.7)
17	16	12	GRE(0.4)	-	ISL(0.1)	MYK(0.5)
18	12	9	GRE(0.4)	-	ISL(0.1)	SAN(0.5)
19	13	7	GRE(0.5)	-	-	SAN(0.5)
20	18	18	-	-	ISL(0.4)	SAN(0.6)
21	11	17	GRE(0.1)	-	ISL(0.1)	SAN(0.8)
22	24	36	GRE(0.3)	-	-	SAN(0.7)
23	27	32	GRE(0.3)	-	ISL(0.1)	SAM(0.6)
24	28	30	GRE(0.4)	-	ISL(0.2)	SAM(0.4)
25	14	19	GRE(0.3)	-	ISL(0.4)	SAM(0.3)
26	19	40	GRE(0.2)	-	ISL(0.1)	RHO(0.7)
27	22	59	-	-	ISL(0.1)	RHO(0.9)
28	21	15	GRE(0.2)	-	-	RHO(0.8)
29	23	23	GRE(0.4)	-	ISL(0.3)	RHO(0.3)
30	31	8	GRE(0.5)	-	-	HIS(0.5)
31	30	20	GRE(0.7)	-	-	TRA(0.3)
32	32	55	GRE(0.5)	-	-	GOD(0.5)
33	58	64	-	HAW(0.4)	-	FAL(0.6)
34	62	41	-	HAW(0.3)	-	HIK(0.7)
35	38	50	-	HAW(0.5)	-	VOL(0.5)
36	61	49	-	HAW(0.1)	ISL(0.5)	MOL(0.4)
37	49	37	-	HAW(0.2)	ISL(0.3)	MOL(0.5)
38	46	16	-	HAW(0.1)	ISL(0.6)	LAN(0.3)
39	63	26	-	HAW(0.5)	-	LAN(0.5)
40	40	10	-	-	ISL(0.5)	LAN(0.5)
41	59	56	-	HAW(0.3)	ISL(0.3)	LAN(0.4)
42	47	22	-	HAW(0.4)	ISL(0.2)	NIH(0.4)
43	36	60	-	HAW(0.4)	ISL(0.1)	NIH(0.5)
44	55	42	-	HAW(0.3)	-	NIH(0.7)
45	64	47	-	HAW(0.6)	ISL(0.2)	OAH(0.2)
46	48	54	-	-	ISL(0.1)	OAH(0.9)
47	50	45	-	HAW(0.3)	ISL(0.2)	OAH(0.5)
48	41	13	-	HAW(0.3)	-	OAH(0.7)
49	56	28	-	HAW(0.2)	ISL(0.3)	OAH(0.5)
50	45	43	-	HAW(0.4)	ISL(0.3)	KAU(0.3)
51	42	52	-	HAW(0.7)	ISL(0.1)	KAU(0.2)
52	53	57	-	HAW(0.2)	ISL(0.1)	KAU(0.7)
53	60	51	-	-	ISL(0.1)	KAU(0.9)
54	52	21	-	HAW(0.3)	ISL(0.2)	MAU(0.5)
55	43	46	-	-	ISL(0.4)	MAU(0.6)
56	54	14	-	-	ISL(0.3)	MAU(0.7)
57	37	53	-	HAW(0.7)	ISL(0.1)	MAU(0.2)
58	44	27	-	-	ISL(0.7)	MAU(0.3)
59	51	48	-	HAW(0.5)	ISL(0.5)	-
60	57	34	-	HAW(0.7)	ISL(0.3)	-
61	39	44	-	HAW(0.4)	ISL(0.6)	-
62	35	33	-	HAW(0.2)	ISL(0.8)	-
63	33	35	-	HAW(0.2)	ISL(0.8)	-
64	34	39	-	HAW(0.6)	ISL(0.4)	-

Πιν. 4: 64 εικόνες τακτοποιημένες όσον αφορά την απόστασή τους από την ερώτηση *Greek Islands*. Κάθε σειρά αντιπροσωπεύει μια εικόνα. Η πρώτη στήλη κρατά τον αριθμό ετικετών εικόνας, οι δεύτερη και τρίτη στήλες κρατούν τον αριθμό ταξινόμησης όσον αφορά τις αποστάσεις *MSI* και *LSI* από την ερώτηση *Greek Islands* αντίστοιχα για αυτήν την εικόνα. Το υπόλοιπο των καταχωρήσεων αντιστοιχεί στις πιθανότητες ισορροπίας της αντίστοιχης Μαρκοβιανής αλυσίδας που αντιπροσωπεύει κάθε εικόνα. Εμφανίζονται μόνο οι μη μηδενικές καταχωρήσεις ομαδοποιημένες κατά κατηγορία για καλύτερη επισκόπηση. Η ταξινόμηση για την απόσταση *MSI* υπονοείται από τον *AMC* του Σχ. (22). 30 από αυτές τις εικόνες παρουσιάζονται στον πίνακα απόστασης του Σχ. (23) για τις σειρές 1-15 και 33-47 αυτού του πίνακα.

6. ΣΥΜΠΕΡΑΣΜΑΤΑ ΚΑΙ ΜΕΛΛΟΝΤΙΚΗ ΕΡΕΥΝΑ

Στον τομέα της συμβατής με το ΑΟΣ υπολογιστικής όρασης παρουσιάστηκε μια βιολογικά συμβατή νευρωνική αρχιτεκτονική για τον υπολογισμό καμπυλότητας και την ενσωμάτωσή της στον μετασχηματισμό σκελετού αυτής της μορφής. Κωδικοποιώντας αυτήν την αναπαράσταση που ονομάστηκε *σκελετός καμπυλότητας* στις τιμές έντασης φωτεινότητας των ασπρόμαυρων εικόνων επιτεύχθηκε βελτίωση στην ικανότητα γενίκευσης των ταξινομητών νευρωνικών δικτύων σε περιπτώσεις παραμόρφωσης του αρχικού σχήματος.

Στον τομέα της υπολογιστικής όρασης επίσης παρουσιάστηκε και αξιολογήθηκε επιτυχώς ο ολοτοπικός μετασχηματισμός *GLT* που αντιστοιχίζει τοπικούς σε ολικούς περιγραφείς μεταξύ ενός κλειστού περιγράμματος και της επιφάνειας μετασχηματισμού. Μέσω του μετασχηματισμού αυτού επιτυγχάνεται τόσο η αναπαράσταση *μορφομετρικών* ολικών χαρακτηριστικών του περικλειομένου από το περίγραμμα σχήματος, όσο και η αναπαράσταση τοπικών και υβριδικών χαρακτηριστικών του περιγράμματος. Ο ολοτοπικός μετασχηματισμός, έχει την ιδιότητα να περιγράφει ιδιότητες μορφής μιας κλειστής καμπύλης ενώ παρέχει ανθεκτικότητα στον θόρυβο όπως και αντιληπτική συμβατότητα. Δείχθηκε ότι η αντιληπτική συμβατότητα που παρέχει είναι σημαντική για την μείωση της πολυπλοκότητας αλλά και την βελτίωση της ικανότητας αναγνώρισης άλλων μεθόδων της τρέχουσας στάθμης της τεχνικής. Οι Ολοτοπικές Εξισώσεις αποκαλύπτουν για πρώτη φορά μία μη προφανή σχέση μεταξύ της τοπικής έννοιας της καμπυλότητας και ολικών, διαισθητικά συμβατών, μέτρων της μορφής του σχήματος. Επιπλέον, η συσχέτιση αυτή γίνεται μέσω της VAR, ενός βαθμωτού περιγραφέα. Καθώς οι τιμές της VAR μετρούν εμβαδά στον χώρο μετασχηματισμού, είναι ανθεκτική στις παραμορφώσεις του περιγράμματος και έτσι η συσχέτιση μέσω της VAR που αποδεικνύεται στις *Ολοτοπικές Εξισώσεις* αποκτά επιπλέον σημασία.

Στο πλαίσιο της πρόσβασης των χρηστών στην οπτική πληροφορία, εξετάστηκαν στοχαστικές τεχνικές που επιτρέπουν την δυναμική συσχέτιση λέξεων κλειδιών και την εν συνεχεία σημασιολογική ομαδοποίηση χρηστών πληροφοριακών συστημάτων που έκαναν χρήση αυτών. Επίσης παρουσιάστηκαν μηχανισμοί μιας δυναμικά προσαρμοζόμενης απόστασης μεταξύ εικόνων που βασίζεται σε ένα Μαρκοβιανό μοντέλο διάδρασης με τους χρήστες του πληροφοριακού συστήματος που επιλέγουν τις εικόνες αυτές. Η Μαρκοβιανή Σημασιολογική Ταξινόμηση (*MSI*) που ορίστηκε για αυτόν τον σκοπό επιτυγχάνει αυτόματη περιγραφή εικόνων και αποδεικνύεται ότι σε σύγκριση με τις τεχνικές *LSI*, δίνει καλύτερα αποτελέσματα ακρίβειας / ανάκλησης σε εργασίες ανάκτησης εικόνων από κειμενικές περιγραφές.

Στην κατεύθυνση μελλοντικής έρευνας θα μπορούσαν να εξεταστούν: Για την μέθοδο του σκελετού καμπυλότητας η ικανότητά του αναγνώρισης σχήματος σε σχέση με άλλους μη βιολογικά συμβατούς ταξινομητές. Σε σχέση με τον Ολοτοπικό Μετασχηματισμό,

υπάρχει δυνατότητα ορισμού περισσότερων περιγραφέν που προκύπτουν από την ολοτοπική επιφάνεια και συνεπώς οι δυνατότητες τους θα μπορούσαν να διερευνηθούν περαιτέρω. Επίσης μέσω πολυωνυμικής προσέγγισης των συναρτήσεων απόψεων και διερεύνηση της άλγεβρας των συντελεστών των προσεγγιστικών αυτών πολυωνύμων, θα μπορούσε κανείς να συνδέσει μορφολογικά χαρακτηριστικά σχήματος με αλγεβρικές δομές πολυωνύμων. Στο πλαίσιο της ταξινόμησης και ανάκτησης εικόνων θα μπορούσαν να εξεταστούν υβριδικά μοντέλα που να συνδυάζουν την Μαρκοβιανή Σημασιολογική Ταξινόμηση με οπτικά χαρακτηριστικά της εικόνας.

7. ΑΝΑΦΟΡΕΣ

- [1] Pavlidis, T.: A review of algorithms for shape analysis. *Computer Graphics and Image Processing* **7** (1978) 243-258
- [2] Pavlidis, T.: Algorithms for shape analysis of contours and waveforms. *IEEE Transactions on PAMI* **7** (1980) 301-312
- [3] Aloimonos, Y.: Visual shape computation. *Proceedings of the IEEE* **76** (1987) 899-916
- [4] L. Uhr, e.: *Pattern Recognition: Theory, Experiments, Computer Simulations and Dynamical Models of Form, Perception and Discovery*. John Wiley (1996)
- [5] Duda, R., Hart, P.: *Pattern Classification and Scene Analysis*. John Wiley (1973)
- [6] Nevatia, R.: *Machine Perception*. Prentice Hall (1982)
- [7] Rosenfeld, A., Kak, A.C.: *Digital Picture Processing*, 2nd edition. Volume 1. Academic Press (1982)
- [8] Rosenfeld, A., Kak, A.C.: *Digital Picture Processing*, 2nd edition. Volume 2. Academic Press (1982)
- [9] Ballard, D.H., Brown, C.M.: *Computer Vision*. Prentice Hall (1982)
- [10] Levine, M.: *Vision in Man and Machine*. McGraw Hill (1985)
- [11] Horn, B.: *Robot Vision*. MIT Press and McGraw Hill (1986)
- [12] Young, I., K. S. Fu, e.: *Handbook of Pattern Recognition and Image Processing*. Academic Press (1986)
- [13] Jain, A.K.: *Fundamentals of Digital Image Processing*. Prentice Hall (1989)
- [14] Schalkoff, R.: *Digital Image Processing and Computer Vision*. John Wiley (1989)
- [15] Haralick, R., Shapiro, L.: *Computer and Robot Vision*. Volume 1. Addison Wesley (1992)
- [16] Haralick, R., Shapiro, L.: *Computer and Robot Vision*. Volume 2. Addison Wesley (1992)
- [17] Biasotti, S., De Floriani, L., Falcidieno, B., Frosini, P., Giorgi, D., Landi, C., Papa-
leo, L., Spagnuolo, M.: Describing shapes by geometrical-topological properties
of real functions. *ACM Comput. Surv.* **40** (2008) 12:1-12:87

- [18] Duda, R., Hart, P.: Descriptions of Line and Shape. John Wiley (1973)
- [19] Young, I., Walker, J., Bowie, J.: An analysis technique for biological shape. *Computer Graphics and Image Processing* **25** (1974) 357-370
- [20] Kashyap, R., Chellappa, R.: Stochastic models for closed boundary analysis: Representation and reconstruction. *IEE Transactions on Information Theory* **27** (1981) 627-637
- [21] Nevatia, R.: Shape Analysis and Recognition. Prentice Hall (1982)
- [22] Davis, L.: Two-dimensional shape representation. Academic Press (1986)
- [23] Kartikeyan, B., Sarkar, A.: Shape description by time series. *IEE Transactions on PAMI* **11** (1986) 977-984
- [24] Taza, A., Suen, C.: Discrimination of planar shapes using shape matrices. *IEE Transactions on SMC* **19** (1989) 1281-1289
- [25] Zahn, C., Roskies, R.: Fourier descriptors for plane closed curves. *Computer Graphics and Image Processing* **21** (1972) 269-281
- [26] Granlund, G.H.: Fourier preprocessing for hand print character recognition. *IEEE Transactions on Computers* **21** (1972) 195-201
- [27] Richards, C.W., Hemami, H.: Identification of three-dimensional objects using fourier descriptors of the boundary curve. *IEEE Transactions on SMC* **4** (1974) 371-378
- [28] Persoon, E., Fu, K.: Shape discrimination using fourier descriptors. *IEEE Transactions on SMC* **7** (1977) 170-179
- [29] Wallace, T., Wintz, P.: An efficient three-dimensional aircraft recognition algorithm using normalized fourier descriptors. *Computer Graphics and Image Processing* **13** (1980) 99-126
- [30] Blum, H.: A transformation for extracting new descriptors of shape. *Models for the Perception of Speech and Visual Forms*. MIT Press (1967)
- [31] Mott-Smith, J.: Medial axis transformations. *Picture Processing and Psychopictorics*. Academic Press (1970)
- [32] Blum, H.: Biological shape and visual science. *Journal of Theoretical Biology* **38** (1973) 205-287

- [33] Blum, H., Nagel, R.: shape description using weighted symmetric axis features. *Pattern Recognition* **10** (1978) 167–180
- [34] Peleg, S., Rosenfeld, A.: A min-max medial axis transformation. *IEEE Transactions on PAMI* **3** (1981) 208–210
- [35] Hu, M.K.: Visual pattern recognition by moment invariants. *IRE Transactions on Information Theory* **8** (1962) 179–187
- [36] Teague, M.R.: Image analysis via the general theory of moments. *Journal of the Optical Society of America* **70** (1980) 920–930
- [37] Belkasim, S.O., Shridhar, M., Ahmadi, M.: Pattern recognition with moment invariants: A comparative study and new results. *Pattern Recognition* **24** (1991) 1117–1138
- [38] Prokop, R.J., Reeves, A.P.: A survey of moment-based techniques for unoccluded object representation and recognition. *CVGIP: Graphical Models and Image Processing* **54** (1992) 438–460
- [39] O'Rourke, J., Badler, N.: Decomposition of three-dimensional objects into spheres. *IEEE Transactions on PAMI* **1** (1979) 295–305
- [40] Shankar, F., Sklansky: Minimal rectangular partition of digitized books. 5th Intl. Conf. Pattern Recognition (1980)
- [41] Kim, H., Park, K., Kim, M.: Shape decomposition by collinearity. *Pattern Recognition Letters* **6** (1987) 335–340
- [42] Jagadish, H.V., Bruckstein, A.M.: On sequential shape descriptions. *Pattern Recognition* **25** (1992) 165–172
- [43] Bennet, J.R., McDonald, J.S.: On the measurement of curvature in a quantized environment. *IEEE Transactions on Computers* **24** (1975) 803–820
- [44] Arkin, E.M., Chew, L.P., Huttenlocher, D.P., Kedem, K., Mitchell, J.S.B.: An efficiently computable metric for comparing polygonal shapes. *IEEE Transactions on PAMI* **13** (1986) 209–216
- [45] Chang, C.C., Hwang, S.M., Buehrer, D.J.: A shape recognition scheme based on relative distances of feature points from the centroid. *Pattern Recognition* **24** (1991) 1053–1063

- [46] Wang, S., Chen, P., Lin, W.: Invariant pattern recognition by moment fourier descriptor. *Pattern Recognition* **27** (1994) 1735-1742
- [47] Huang, T.S.: Coding of two-tone images. *IEEE Transactions on Communications* **25** (1977) 1406-1424
- [48] Dubois, S., Glanz, F.: An autoregressive model approach to two-dimensional shape classification. *IEEE Transactions on PAMI* **8** (1986) 55-66
- [49] Sebestyen, G.S.: *Decision-Making Processes in Pattern Recognition*. Macmillan (1962)
- [50] Devijver, P.A., Kittler, J.: *Pattern Recognition: A Statistical Approach*. Prentice Hall (1982)
- [51] Glendinning, R.H.: Signal plus noise models in shape classification. *Pattern Recognition* **27** (1994) 777-784
- [52] Das, M., Paulik, M.J., Loh, N.K.: A bivariate autoregressiv modeling technique for analysis and classification of planar shapes. *IEEE Transactions on PAMI* **12** (1990) 97-103
- [53] He, Y., Kundu, A.: 2-d shape classification using hidden markov model. *IEEE Transactions on PAMI* **13** (1991) 1172-1184
- [54] Lin, Y., Dou, J., Wang, H.: Contour shape description based on an arch height function. *Pattern Recognition* **25** (1992) 17-23
- [55] Freeman, H.: On the encoding of arbitrary geometric configurations. *IRE Transactions* **10** (1961) 260-268
- [56] Freeman, H.: Computer processing of line-drawing images. *ACM Computing Surveys* **6** (1974) 57-97
- [57] Freeman, H.: Shape description via the use of critical points. *Pattern Recognition* **10** (1978) 159-166
- [58] Parui, S., Majumder, D.: Symmetry analysis by computer. *Pattern Recognition* **16** (1983) 6367
- [59] Fu, K.S.: *Syntactic methods in pattern recognition*. Academic Press (1974)
- [60] Pavlidis, T.: *Structural pattern recognition*. Springer Verlag (1977)

- [61] Fu, K.S.: A general (syntactic-semantic) approach to picture analysis. Springer Verlag (1982)
- [62] Fu, K.S.: Syntactic Pattern Recognition and Applications. Prentice Hall (1982)
- [63] Henderson, T.: Syntactic and structural methods I and II. Cambridge University Press (1983)
- [64] Pavlidis, T., Ali, F.: A hierarchical syntactic shape analyzer. IEEE Transactions on PAMI **1** (1977) 2-9
- [65] Fu, K.S.: Syntactic image modeling using stochastic tree grammars. Computer Graphics and Image Processing **12** (1980) 136-152
- [66] Pavlidis, T.: Algorithms for Graphics and Image Processing. Springer Verlag (1982)
- [67] Pavlidis, T.: Polygonal approximations by newton's method. IEEE Transactions on Computers **26** (1977) 800-807
- [68] Wu, J., Leou, J.: New polygonal approximation schemes for object shape representation. Pattern Recognition **26** (1993) 471-484
- [69] Bengtsson, A., Eklundh, J.: Shape representation by multiscale contour approximation. IEEE Transactions on PAMI **13** (1991) 85-93
- [70] Witkin, A.P.: Scale-space filtering. In: Proceedings of the 8th Int'l Joint Conference on Artificial Intelligence. (1983) 1019-1022
- [71] Ikebe, Y., Miyamoto, S.: Shape design, representation, and restoration with splines. In Fu, K., Kunii, T., eds.: Picture Engineering. Lecture Notes in Computer Science, Berlin Heidelberg New York, Springer-Verlag (1982) 75-95
- [72] Boor, C.D.: A Practical Guide to Splines. Lecture Notes in Computer Science. Springer-Verlag, Berlin Heidelberg New York (1978)
- [73] Atkinson, K.E.: An Introduction to Numerical Analysis. John Wiley (1989)
- [74] Kincaid, D., Cheney, W.: Numerical Analysis. Brooks and Cole (1991)
- [75] Giloi, W.K.: Interactive Computer Graphics. Prentice Hall (1978)
- [76] Mortenson, M.E.: Geometric Modeling. John Wiley (1985)
- [77] Bartels, R., Beatty, J., Barsky, B.: An Introduction to Splines for use in Computer Graphics and Geometric Modeling. Morgan Kaufmann (1987)

- [78] Foley, J.D., van Dam, A., Feiner, S.K., Hughes, J.F.: Computer Graphics: Principles and Practice, 2nd ed. Addison Wesley (1990)
- [79] Cohen, F.S., Huang, Z., Yang, Z.: Invariant matching and identification of curves using b-splines curve representation. *IEEE Transactions on Image Processing* **4** (1995) 1–10
- [80] Chung, P., Tsai, C., Chen, E., Sun, Y.: Polygonal approximation using a competitive hopfield neural network. *Pattern Recognition* **27** (1994) 1505–1512
- [81] Babaud, J., Witkin, A., Baudin, M., Duda, R.: Uniqueness of the gaussian kernel for scale-space filtering. *IEEE Transactions on PAMI* **8** (1986) 26–33
- [82] Asada, H., Brady, M.: The curvature primal sketch. *IEEE Transactions on PAMI* **8** (1986) 2–14
- [83] Mokhtarian, F., Mackworth, A.K.: Scale-based description and recognition of planar curves and two-dimensional shapes. *IEEE Transactions on PAMI* **8** (1986) 34–43
- [84] Mokhtarian, F., Mackworth, A.K.: A theory of multiscale, curvature-based shape representation for planar curves. *IEEE Transactions on PAMI* **14** (1992) 789–805
- [85] Chen, M., Yan, P.: A multiscaling approach based on morphological filtering. *IEEE Transactions on PAMI* **11** (1989) 694–700
- [86] Dilla, A., Levine, M., Noble, P.: Multiple resolution skeletons. *IEEE Transactions on PAMI* **9** (1987) 495–504
- [87] Arcelli, C.: Pattern thinning by contour tracing. *Computer Graphics and Image Processing* **17** (1981) 130–144
- [88] Liu, H., Srinath, M.: Partial shape classification using contour matching in distance transformation. *IEEE Transactions on PAMI* **12** (1990) 1072–1079
- [89] Borgefors, G.: Distance transformations in arbitrary dimensions. *Computer Vision, Graphics, and Image Processing* **27** (1984) 321–345
- [90] Borgefors, G.: Distance transformations in digital images. *Computer Vision, Graphics, and Image Processing* **34** (1986) 344–371
- [91] Weiss, I.: Geometric invariants and object recognition. *International Journal of Computer Vision* **10** (1993) 207–231

- [92] Matheron, G.: Random Sets and Integral Geometry. John Wiley (1975)
- [93] Serra, J.: Image Analysis and Mathematical Morphology. Academic Press (1982)
- [94] Haralick, R., Sternberg, S., Zhuang, X.: Image analysis using mathematical morphology. IEEE Transactions on PAMI **9** (1987) 532-550
- [95] J. Serra, e.: Image Analysis and Mathematical Morphology. Volume 2. Academic Press (1998)
- [96] Giardina, C.R., Daugherty, E.R.: Morphological Methods in Image Processing. Prentice Hall (1988)
- [97] Maragos, P.: Pattern spectrum and multiscale shape representation. IEEE Transactions on PAMI **11** (1989) 701-716
- [98] Dougherty, E.R., Pelz, J.: Morphological granulometric analysis of electrophotographic images-size distribution statistics for process control. Optical Engineering **30** (1991) 438-445
- [99] Dougherty, E.R., Pelz, J., Sand, F., Lent, A.: Morphological image segmentation by local granulometric size distributions. Journal of Electronic Imaging **1** (1992) 46-60
- [100] Shih, F.Y., Pu, C.C.: Morphological shape description using geometric spectrum on multidimensional binary images. Pattern Recognition **25** (1992) 921-927
- [101] Maragos, P.: Optimal morphological approaches to image matching and object recognition. In: Proceedings of the IEEE Int'l Conference on Computer Vision. (1988) 695-699
- [102] Maragos, P.: Morphological correlation and mean absolute error criteria. In: Proceedings of the IEEE Int'l Conference on ASSSP. (1989) 1568-1571
- [103] Shapiro, L., MacDonald, R., Sternberg, S.: Ordered structural shape matching with primitive extraction by mathematical morphology. Pattern Recognition **20** (1987) 75-90
- [104] Calabi, L., Hartnett, W.: Shape recognition, prairie fires, convex deficiencies, and skeletons. The American Mathematical Monthly **75** (1968) 335-341
- [105] Goh, W.B.: Strategies for shape matching using skeletons. Computer Vision and Image Understanding **110** (2008) 326-345

- [106] Gordon, S.: *The Synaptic Organization of the Brain*. Oxford (1990)
- [107] Pfaltz, J., Rosenfeld, A.: Computer representation of planar regions by their skeletons. *Communications of the ACM* **10** (1967) 119–125
- [108] Rosenfeld, A., Pfaltz, J.: Sequential operations in digital picture processing. *Journal of ACM* **13** (1966) 471–494
- [109] Nakagawa, Y., Rosenfeld, A.: A note on the use of local min and max operations in digital picture processing. *IEEE Transactions on SMC* **8** (1978) 632–635
- [110] Goetcheerian, V.: From binary to grey tone image processing using fuzzy logic concepts. *Pattern Recognition* **12** (1980) 7–15
- [111] Duff, M., Watson, D., Fountain, T., Shaw, G.: A cellular logic array for image processing. *Pattern Recognition* **5** (1973) 229–247
- [112] Preston, K., Duff, M., Levialdi, S., Norgren, P., Toriwaki, J.: Basics of cellular logic with some applications in medical image processing. *Proceedings of the IEEE* **67** (1979) 826–859
- [113] Preston, K.: Some notes on cellular logic operators. *IEEE Transactions on PAMI* **3** (1981) 476–481
- [114] Rosenfeld, A.: *Cellular architectures: From automata to hardware*. *Multicomputers and Image Processing Algorithms and Programs* (1982)
- [115] Rosenfeld, A.: Parallel image processing using cellular arrays. *IEEE Computer* **16** (1983) 14–20
- [116] Preston, K.: Cellular logic computers for pattern recognition. *IEEE Computer* **16** (1983) 36–47
- [117] Gerritsen, F., Verbeek, P.: Implementation of cellular-logic operators using 3x3 convolution and table lookup hardware. *Computer Vision, Graphics, and Image Processing* **27** (1984) 115–123
- [118] Preston, K.: *Cellular logic arrays for image processing*. Academic Press (1986)
- [119] Meyer, F.: Digital euclidean skeletons. In: *SPIE-1360 Visual Communications and Image Processing*. (1990) 251–262
- [120] Brooks, R.A.: Symbolic reasoning among 3-D models and 2-D images. *Artificial Intelligence* (1981)

- [121] Brady, M., Asada, H.: Smoothed local symmetries and their implementation. *Int'l Journal of Robotics Research* **3** (1984) 36–61
- [122] Rosenfeld, A.: Axial representations of shape. *Computer Vision, Graphics, and Image Processing* **33** (1986) 156–173
- [123] Ponce, J.: On characterizing ribbons and finding skewed symmetries. *Computer Vision, Graphics, and Image Processing* **52** (1990) 328–340
- [124] Pizer, S., Oliver, W., Bloomberg, S.: Hierarchical shape description via the multiresolution symmetric axis transform. *IEEE Transactions on PAMI* **9** (1987) 505–511
- [125] Rom, H., Medioni, G.: Hierarchical decomposition of axial shape description. *IEEE Transactions on PAMI* **15** (1993) 973–981
- [126] Hoffman, D.D., Richards, W.A.: Parts of recognition. *Cognition* **18** (1984) 65–96
- [127] Leymarie, F., Levine, M.D.: Simulating the grassfire transform using an active contour model. *IEEE Transactions on PAMI* **14** (1992)
- [128] Gauch, J.M., Pizer, S.M.: The intensity axis of symmetry and its application to image segmentation. *IEEE Transactions on PAMI* **15** (1993) 753–770
- [129] Han, C.C., Fan, K.C.: Skeleton generation of engineering drawings via contour matching. *Pattern Recognition* **27** (1994) 261–276
- [130] Shih, F.Y., Pu, C.C.: A skeletonization algorithm by maxima tracking on euclidean distance transform. *Pattern Recognition* **28** (1995) 331–341
- [131] Ogniewicz, R.L., Kuebler, O.: Hierarchic voronoi skeletons. *Pattern Recognition* **28** (1995) 343–359
- [132] Maragos, P., Schafer, R.: Morphological skeleton representation and coding of binary images. *IEEE Transactions on ASSP* **34** (1986) 1228–1244
- [133] Shapiro, L., Haralick, R.: Decomposition of two-dimensional shapes by graph-theoretic clustering. *IEEE Transactions on PAMI* **1** (1979) 10–20
- [134] Shapiro, L.: A structural model of shape. *IEEE Transactions on PAMI* **2** (1980) 111–126
- [135] Bjoklund, C., Pavlidis, T.: Global shape analysis by k-syntactic similarity. *IEEE Transactions on PAMI* **3** (1981) 144–155

- [136] Vanderheydt, L., Dom, F., den Berghe, H.V.: Two-dimensional shape decomposition using fuzzy subset theory applied to automated chromosome analysis. *Pattern Recognition* **13** (1981) 147–157
- [137] Zadeh, L.: Fuzzy sets. *Information and Control* **8** (1965) 338–353
- [138] Zadeh, L.: Outline of a new approach to the analysis of complex systems and decision processes. *IEEE Transactions on SMC* **3** (1973) 28–44
- [139] di Baja, G.S., Thiel, E.: (3,4)-weighted skeleton decomposition for pattern representation and description. *Pattern Recognition* **27** (1994) 1039–1049
- [140] Pitas, I., Venetsanopoulos, A.: Shape decomposition by mathematical morphology. In: *Proceedings of the Int'l Conference on Computer Vision*. (1987) 621–625
- [141] Pitas, I.: Morphological signal decomposition. In: *Proceedings of the Int'l Conference on ASSP*. (1990) 2169–2172
- [142] Pitas, I., Venetsanopoulos, A.: Morphological shape decomposition. *IEEE Transactions on PAMI* **12** (1990) 38–45
- [143] Pitas, I., Maglar, A.: Range image analysis by using morphological signal decomposition. *Pattern Recognition* **24** (1991) 165–181
- [144] Pitas, I., Venetsanopoulos, A.: Morphological shape representation. *Pattern Recognition* **25** (1992) 555–565
- [145] Schonfeld, D., Goutsias, J.: Morphological shape representation. In: *Proceedings of the Int'l Conference on ASSP*. (1990) 2065–2068
- [146] Schonfeld, D., Goutsias, J.: Optimal morphological pattern restoration from noisy binary images. *IEEE Transactions on PAMI* **13** (1991) 14–29
- [147] Goutsias, J., Schonfeld, D.: Morphological representation of discrete and binary images. *IEEE Transactions on Signal Processing* **39** (1991) 1369–1379
- [148] Giblin, P.J., Kimia, B.B.: On the intrinsic reconstruction of shape from its symmetries. *IEEE Transactions on PAMI* **25** (2003) 895–911
- [149] Siddiqi, K., Kimia, B.: A shock grammar for recognition. *Proc. IEEE Conf. Computer Vision and Pattern Recognition* (1996) 507–513
- [150] Aslan, C., Erdem, A., Erdem, E., Tari, S.: Disconnected skeleton: Shape at its absolute scale. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* **30** (2008) 2188–2203

- [151] Daliri, M.R., Torre, V.: Shape recognition based on kernel-edit distance. *Computer Vision and Image Understanding* **114** (2010) 1097–1103
- [152] Ion, A., Artner, N.M., Peyri, G., Kropatsch, W.G., Cohen, L.D.: Matching 2d and 3d articulated shapes using the eccentricity transform. *Computer Vision and Image Understanding* (2011) 817–834
- [153] Ling, H., Jacobs, D.: Shape classification using the inner-distance. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **29** (2007) 286–299
- [154] Lin, W.Y., Boston, N., Hu, Y.H.: Summation invariant and its applications to shape recognition. In: *IEEE International Conference on Acoustics, Speech and Signal Processing*. Volume 5. (2005) 205–208
- [155] Leventon, M.E., Grimson, W.E.L., Faugeras, O.D.: Statistical shape influence in geodesic active contours. In: *CVPR*. (2000) 1316–1323
- [156] Rousson, M., Paragios, N.: Shape priors for level set representations. In: *Proceedings of the 7th European Conference on Computer Vision-Part II. ECCV '02*, London, UK, UK, Springer-Verlag (2002) 78–92
- [157] Cremers, D., Osher, S.J., Soatto, S.: Kernel density estimation and intrinsic alignment for knowledge-driven segmentation: Teaching level sets to walk. *Pattern Recognition* (2004)
- [158] Wang, J., Athitsos, V., Sclaroff, S., Betke, M.: Detecting objects of variable shape structure with hidden state shape models. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* **30** (2008) 477–492
- [159] Cerri, A., Frosini, P.: Advances in multidimensional size theory. *Image Analysis and Stereology* **29** (2010) 19–26
- [160] Biasotti, S., Cerri, A., Frosini, P., Giorgi, D., Landi, C.: Multidimensional size functions for shape comparison. *Journal of Mathematical Imaging and Vision* **32** (2008) 161–179
- [161] d’Amico, M., Frosini, P., Landi, C.: Using matching distance in size theory: a survey. *International Journal of Imaging Systems and Technology* **16** (2006) 154–161
- [162] Felzenszwalb, P.F., Schwartz, J.D.: Hierarchical matching of deformable shapes. In: *Computer Vision and Pattern Recognition 2007. CVPR '07. IEEE Conference on*. (2007)

- [163] Bai, X., Latecki, L.J.: Path similarity skeleton graph matching. *IEEE Transactions on PAMI* **30** (2008) 1282–1292
- [164] Yang, X.W., Bai, X., Latecki, L.J., Tu, Z.W.: Improving shape retrieval by learning graph transduction. In: *Proceedings of the ECCV08*. (2008) 788–801
- [165] Lin, L., Zhu, S.C., Wang, Y.: Layered graph match with graph editing. In: *Computer Vision and Pattern Recognition, IEEE Computer Society Conference on*,. (2007) 1–8
- [166] Morel, J.M., Yu, G.: Asift: A new framework for fully affine invariant image comparison. *SIAM Journal on Imaging Sciences* **2** (2009)
- [167] Belongie, S., Malik, J., Puzicha, J.: Shape matching and object recognition using shape contexts. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* **24** (2002) 509–522
- [168] Mori, G., Belongie, S., Malik, J.: Efficient shape matching using shape contexts. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* **27** (2005) 1832–1837
- [169] Manay, S., Hong, B.W., Yezzi, A.J., Soatto, S.: Integral invariant signatures. In: *ECCV* (4). (2004) 87–99
- [170] Reiss, T.H.: Recognizing planar objects using invariant image features. In: *Lecture Notes in Computer Science. Volume 676*. (1993)
- [171] Manay, S., Cremers, D., Hong, B.W., Yezzi, A., Soatto, S.: Integral invariants for shape matching. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* **28** (2006) 1602–1618
- [172] W. Wathen-Dunn, e.: *Models for the Perception of Speech and Visual Forms*. MIT (1967)
- [173] Zusne, L.: *Visual Perception of Form*. Academic Press (1970)
- [174] Cornsweet, T.N.: *Visual Perception*. Academic Press (1970)
- [175] Lowe, D.G.: *Perceptual Organization and Visual Recognition*. Kluwer Academic (1985)
- [176] Biederman, I.: Human image understanding: Recent research and a theory. *Computer Vision, Graphics and Image Processing* **32** (1985) 29–73

- [177] M. I. Posner, e.: Foundations of Cognitive Science. MIT Press (1989)
- [178] Hake, M.W.: Form discrimination and the invariance of form. John Wiley (1966)
- [179] Zusne, L.: Contemporary Theory of Visual Form Perception: III. The Global Theories. Academic Press (1970)
- [180] Wertheimer, M.: Laws of organization in perceptual forms. Harcourt Brace Jovanovic (1923)
- [181] W. D. Ellis, e.: A Source Book of Gestalt Psychology. Harcourt Brace Jovanovic (1923)
- [182] Koehler, W.: Gestalt psychology. Liverlight (1929)
- [183] Koffka, K.: Principles of Gestalt psychology. Harcourt Brace Jovanovic (1935)
- [184] Hebb, D.O.: The organization of behavior. John Wiley (1949)
- [185] Gibson, J.J.: The perception of the visual world. Houghton (1950)
- [186] Pinker, S.: Visual cognition: An introduction. cognition **18** (1984) 1-63
- [187] Marr, D.: A theory for cerebral neocortex. Proceedings of the Royal Society of London (1970) 161-234
- [188] Marr, D.: Early processing of visual information. Proceedings of the Royal Society of London (1976) 483-519
- [189] Marr, D., Nishihara, H.: Representation and recognition of the spacial organization of three-dimensional shapes. Proceedings of the Royal Society of London (1978) 269-294
- [190] Marr, D., Poggio, T.: A computational theory of human stereo vision. Proceedings of the Royal Society of London (1979) 301-328
- [191] Marr, D., Hildreth, E.: Theory of edge detection. Proceedings of the Royal Society of London (1980) 187-217
- [192] Marr, D., Ullman, S.: Directional selectivity and its use in early visual processing. Proceedings of the Royal Society of London (1981) 151-180
- [193] Marr, D.: Vision. Freeman (1982)
- [194] Horn, B.: Height and gradient from shading. Int'l Journal of Computer Vision **5** (1990) 37-75

- [195] Horaud, R., Brady, M.: On the geometric interpretation of image contours. *Artificial Intelligence* **37** (1988) 333–353
- [196] Ulupinar, F., Nevatia, R.: Inferring shape from contour for curved surfaces. *Proceedings of the Int'l Conference on Pattern Recognition* (1990)
- [197] Hoff, W., Ahuja, N.: Surfaces from stereo: Integrating feature matching, disparity estimation, and contour detection. *IEEE Transactions on PAMI* **11** (1989) 121–136
- [198] Chen, S., Keller, J., Crownover, R.: Shape from fractal geometry. *Artificial Intelligence* **43** (1990) 199–218
- [199] Koenderink, J.: The structure of images. *Biological Cybernetics* **50** (1984) 363–370
- [200] Koenderink, J., Doorn, A.V.: Dynamic shape. *Biological Cybernetics* **53** (1986) 383–396
- [201] Gombrich, H.E.: *Art and Illusion*. Phaidon Press (1960)
- [202] Attneave, F.: Some informational aspects of visual perception. *Psychological Review* **61** (1954) 183–193
- [203] Attneave, F., Arnoult, M.D.: *The quantitative study of shape and pattern perception*. John Wiley (1966)
- [204] Rosenfeld, A., Johnston, E.: Angle detection in digital curves. *IEEE Transactions on Computers* **22** (1973) 875–878
- [205] Pavlidis, T., Horowitz, S.L.: Segmentation of plane curves. *IEEE Transactions on Computers* **23** (1974) 860–870
- [206] Fischler, M.A., Bolles, R.C.: Perceptual organization and curve partitioning. *IEEE Transactions on PAMI* **8** (1986) 100–105
- [207] Teh, C.H., Chin, R.T.: On the detection of dominant points on digital curves. *IEEE Transactions on PAMI* **11** (1989) 859–872
- [208] Wuecher, D.M., Boyer, K.L.: Robust contour decomposition using a constant curvature criterion. *IEEE Transactions on PAMI* **13** (1991) 41–51
- [209] Brault, J., Plamondon, R.: Segmenting handwritten signatures at their perceptually important points. *IEEE Transactions on PAMI* **15** (1993) 953–957

- [210] Katzir, M., Lindenbaum, M., Porat, M.: Curve segmentation under partial occlusion. *IEEE Transactions on PAMI* **16** (1994) 513–519
- [211] Imai, H., Iri, M.: Computational-geometric methods for polygonal approximations of a curve. *Computer Vision, Graphics, Image Processing* **36** (1986) 31–34
- [212] Ghosh, P.: A solution of polygon containment, spatial planning and other related problems using minkowski operations. *Computer Vision, Graphics, Image Processing* **49** (1990) 1–35
- [213] Maes, M.: Polygonal shape recognition using string-matching techniques. *Pattern Recognition* **24** (1991) 433–440
- [214] Xu, J.: Decomposition of convex polygonal morphological structuring elements into neighborhood subsets. *IEEE Transactions on PAMI* **13** (1991) 153–162
- [215] Davis, L.: Understanding shape: Angles and sides. *IEE Transactions on Computers* **26** (1977) 236–242
- [216] Stokely, E.M., Wu, S.Y.: Surface parametrization and curvature measurement of arbitrary 3-d objects: Five practical methods. *IEE Transactions on PAMI* **14** (1992) 833–840
- [217] Richards, W., Hoffman, D.: Codon constraints on closed 2-d shapes. *Computer Vision, Graphics and Image Processing* **31** (1985) 265–281
- [218] Guillemin, V., Pollack, A.: *Differential Topology*. Prentice Hall (1974)
- [219] Leyton, M.: Symmetry-curvature duality. *Computer Vision, Graphics and Image Processing* **38** (1985) 327–341
- [220] Leyton, M.: Inferring causal history from shape. *Cognitive Science* **13** (1989) 357–387
- [221] Pentland, A.: Fractal-based description of natural scenes. *IEEE Transactions on PAMI* **6** (1984) 661–674
- [222] Pentland, A.: Perceptual organization and the representation of natural form. *Artificial Intelligence* **28** (1986) 293–331
- [223] Mandelbrot, B.: *Fractals: form, chance, and dimension*. Freeman (1977)
- [224] Mandelbrot, B.: *The fractal geometry of nature*. Freeman (1983)
- [225] Barnsley, M.: *Fractals Everywhere*. Academic Press (1988)

- [226] Thomson, D.: *On Growth and Form*. Cambridge University Press (1942)
- [227] Stevens, S.: *Patterns in Nature*. Atlantic-Little Brown Books (1974)
- [228] Lowe, D.G.: Three-dimensional object recognition from single two-dimensional images. *Artificial Intelligence* **31** (1987) 355–395
- [229] Goodale, M., Meenan, J., Buelthoff, H., Nicolle, D., Murphy, K., Racicot, C.: Separate neural pathways for the visual analysis of object shape in perception and prehension. *Current Biol.* **4** (1994) 604–610
- [230] Goodale, M., Milner, A., Jakobson, L., Carey, D.: A neurological dissociation between perceiving objects and grasping them. *Nature* **349** (1991) 154–156
- [231] Ungerleider, L.G., Courtney, S.M., Haxby, J.V.: A neural system for human visual working memory. In: *Proc. Natl. Acad. Sci. USA*. Volume 95. (1998) 883–890
- [232] Dobbins, A., Zucker, S.W., Cynader, M.S.: Endstopped neurons in the visual cortex as a substrate for calculating curvature. *Nature* **329** (1987) 438–441
- [233] Hubel, D., Wiesel, T.: Binocular interaction in striate cortex of kittens. *J. Neurophysiol.* **28** (1965) 229–289
- [234] Biederman, I.: Recognition-by-components: A theory of human image understanding. *Psychological Review* **94** (1987) 115–147
- [235] Bookstein, F.: Landmark methods for forms without landmarks: morphometrics of group differences in outline shape. *Med. Im. Anal.* **1** (1996) 225–243
- [236] Berretti, S., Bimbo, A.D., Pala, P.: Retrieval by shape similarity with perceptual distance and effective indexing. *IEEE Trans on Multimedia* **2** (2000) 225–239
- [237] Attneave, F.: Some informational aspects of visual perception. *Psychol. Rev.* **61** (1994) 183–193
- [238] Pomerantz, J.R., Sager, L.C., Stoever, R.J.: Perception of wholes and their component parts: Some configural superiority effects. *J. Exp. Psychol.* **3** (1977) 422–435
- [239] Pavlidis, T.: Algorithms for shape analysis of contours and waveforms. *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-2** (1980) 301–312
- [240] Wang, S.L.Y.P., Lee, K.T.: "multiscale curvature based shape representation using b-spline wavelets. *IEEE Trans. Image Process.* **8** (1999) 1586–1592

- [241] Asada, H., Brady, M.: The curvature primal sketch. *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-8** (1986) 2–14
- [242] Mokhtarian, F.: A theory of multiscale, curvature-based shape representation for planar curves. *IEEE Trans. Pattern Anal. Mach. Intell.* **14** (1992) 789–805
- [243] Mokhtarian, F., Mackworth, A.K.: Scale-based description and recognition of planar curves and two-dimensional shapes. *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-8** (1986) 34–43
- [244] Dudek, G., Tsotsos, J.K.: Shape representation and recognition from multiscale curvature. *Comput. Vis. Image Understand.* **68** (1997) 170–189
- [245] Super, B.: Fast correspondence-based system for shape-retrieval. *Patt. Recog. Lett.* **24** (2004) 217–225
- [246] Hubel, D.: Evolution of ideas on the primary visual cortex. *Nature* (1955–1978)
- [247] Hubel, D., Wiesel, T.: Brain mechanisms of vision. *The Perceptual world* (1990) 3–24
- [248] Hubel, D., Wiesel, T., Stryker, M.: Anatomical demonstration of orientation columns in macaque monkey. *J. Comp. Neurol.* **177** (1978) 361–380
- [249] Wandell, B.A.: *Foundations of vision*. Sinauer Associates, Inc (1995)
- [250] Haykin, S.: *Neural Networks: A Comprehensive Foundation*. Prentice Hall (1998)
- [251] Sebastian, T.B., Klein, P.N., Kimia, B.B.: On aligning curves. *IEEE Transactions on PAMI* **25** (2003) 116–125
- [252] Donatini, P., Frosini, P.: Natural pseudo-distances between closed curves. *Forum Math.* **21** (2009) 981–999
- [253] Groisser, D.: Certain optimal correspondences between plane curves ii. existence, local uniqueness, regularity, and other properties. *Trans. Amer. Math. Soc.* **361** (2009) 3001–3030
- [254] Groisser, D.: Certain optimal correspondences between plane curves i. manifolds of shapes and bimorphisms. *Trans. Amer. Math. Soc.* **361** (2009) 2959–3000
- [255] Michor, P.W., Mumford, D.: Riemannian geometries on spaces of plane curves. *J. Eur. Math. Soc. (JEMS)* **8** (2006) 1–48

- [256] Latecki, L.J., Lakamper, R., Eckhardt, U.: Shape descriptors for non-rigid shapes with a single closed contour. In: IEEE Conf. on Computer Vision and Pattern Recognition I. (2000) 424–429
- [257] Mori, G., Malik, J.: Recognizing objects in adversarial clutter: Breaking a visual captcha. In: IEEE Conf. on Computer Vision and Pattern Recognition I. (2003) 1063–6919
- [258] Mortensen, E.N., Deng, H., Shapiro, L.: A sift descriptor with global context. In: IEEE Conf. on Computer Vision and Pattern Recognition I. (2005) 184–190
- [259] Thayananthan, A., Stenger, B., Torr, P.H.S., Cipolla, R.: Shape context and chamfer matching in cluttered scenes. In: IEEE Conf. on Computer Vision and Pattern Recognition, I. (2003) 127–133
- [260] Tu, Z., Yuille, A.L.: Shape matching and recognition-using generative models and informative features. In: European Conf. on Computer Vision. Number 3 (2004) 195–209
- [261] Zhang, H., Malik, J.: Learning a discriminative classifier using shape context distances. In: IEEE Conf. on Computer Vision and Pattern Recognition. (2003)
- [262] Leibe, B., Schiele, B.: Analyzing appearance and contour based methods for object categorization. In: IEEE Conf. on Computer Vision and Pattern Recognition, II. (2003) 409–415
- [263] Santini, S., Jain, R.: Similarity measures. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **21** (1999)
- [264] Howard, R.: *Dynamic probabilistic systems*. Addison-Wesley John Wiley and Sons Inc. (1971)
- [265] Duda, P., Hart, P.: *Pattern Classification and scene analysis*. Wiley, New York (1973)
- [266] Stevenson, K., Leung, C.: Comparative evaluation of web image search engines for multimedia applications. In: IEEE International Conference on Multimedia and Expo. (2005)
- [267] comScore's Report Article: comscore's qsearch 2.0 service, www.comscore.com. comScore's Report Article (2007)

- [268] Jansen, B.J., Spink, A., Saracevic, T.: Real life, real users, and real needs: A study and analysis of user queries on the web. *Information Processing and Management* **36** (2000) 207–227
- [269] Smeulders, A., Worring, M., Santini, S., Gupta, A., Jain, R.: Content-based image retrieval at the end of the early years. *IEEE Trans. Pattern Analysis and Machine Intelligence* **22** (2000) 1349–1380
- [270] Bhattacharya, A., Ljosa, V., Pan, J.Y., Verardo, M.R., Yang, H., Faloutsos, C., Singh, A.K.: Vivo: Visual vocabulary construction for mining biomedical images. In: *Fifth IEEE International Conference on Data Mining*, Houston, Texas. (2005)
- [271] Li, J., Wang, J.: Real-time computerized annotation of pictures. In: *ACM Multimedia*. (2006)
- [272] Joshi, D., Wang, J.Z., Li, J.: The story picturing engine - a system for automatic text illustration. *ACM Trans. Multimedia Computing, Communications and Applications* **2** (2006) 68–89
- [273] Berry, M. W., D.S.T., O'Brien, G.W.: Using linear algebra for intelligent information retrieval. *SIAM Review* **37** (1995) 573–595
- [274] Hofmann, T.: Probabilistic latent semantic indexing. In: *SIGIR'99*. (1999)
- [275] Hofmann, T.: Unsupervised learning by probabilistic latent semantic analysis. *Machine Learning* **42** (2001) 177–196
- [276] Blei, D.M., Ng, A.Y., Jordan, M.I.: Latent dirichlet allocation. *Journal of Machine Learning Research* (2003) 993–1022
- [277] Griffiths, T.L., Steyvers, M.: Finding scientific topics. In: *Proceedings of the National Academy of Sciences*. Volume 101. (2004) 5228–5235
- [278] Steyvers, M., Smyth, P., Rosen-Zvi, M., Griffiths, T.: Probabilistic author-topic models for information discovery. In: *10th ACM SigKDD Conference on Knowledge Discovery and Data Mining*. (2004)
- [279] Guo, Z., Zhu, S., Chi, Y., Zhang, Z., Gong, Y.: A latent topic model for linked documents. In: *SIGIR*. (2009)
- [280] Pham, T.T., Maillot, N.E., Lim, J.H., Chevallet, J.P.: Latent semantic fusion model for image retrieval and annotation. In: *CIKM*. (2007)

- [281] Ritendra Datta, Dhiraj Joshi, J.L., Wang, J.Z.: Image retrieval: Ideas, influences, and trends of the new age. *ACM Computing Surveys* **40** (2008) 1-60
- [282] Monay, F., Gatica-Perez, D.: On image auto-annotation with latent space models. In: *Proc. ACM Int. Conf. on Multimedia (ACM MM)*. (2003)
- [283] Barnard, K., Forsyth, D.: Learning the semantics of words and pictures. In: *International Conference on Computer Vision*. Volume 2. (2001) 408-415
- [284] Li, L.J., Wang, G., Fei-Fei, L.: Optimol: automatic online picture collection via incremental model learning. *International Journal Of Computer Vision* **88** (2010) 147-168
- [285] Fan, J., Gao, Y., Luo, H.: Integrating concept ontology and multitask learning to achieve more effective classifier training for multilevel image annotation. *IEEE Transactions on image processing* **17** (2008)
- [286] Pearl, J.: *Probabilistic Reasoning in Intelligent Systems*. Morgan Kaufmann (1988)
- [287] Lowrance, L.D., Garvey, T.D., Strat, T.M.: A framework for evidential reasoning systems. In: *Proceedings of the 5th National Conference on AI (AAAI-86)*, Philadelphia (1986) 896-901
- [288] Shafer, G., Shenoy, P.P., Mellouli, K.: Propagating belief functions in qualitative markov trees. *International journal of approximate reasoning* **1** **4** (1987) 394-400
- [289] Montanari, U.: Networks of constraints, fundamental properties and applications to picture processing. *Information science* **7** (1974) 95-132
- [290] Woods, W.: *Representation and understanding*. New York: Academic Press (1975)
- [291] Duda, R.O., Hart, P.E., Nilsson, N.J.: Subjective bayesian methods for rule-based inference systems. In: *Proceedings of the National Computer Conference (AFIPS)*. Volume 45. (1976) 1075-1082
- [292] Schank, R.: Conceptual dependency: A theory of natural language understanding. *Cognitive Psychology* **4** (1972) 552-631
- [293] Sowa, J.F.: *Conceptual structures: Information processing in mind and machine*. Mass: Addison-Wesley (1984)
- [294] O. Tuzel, F.P., Meer, P.: Pedestrian detection via classification on riemannian manifolds. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* **30** (2008) 1713-1727

A. ΣΧΕΤΙΚΕΣ ΔΗΜΟΣΙΕΥΣΕΙΣ ΣΥΓΓΡΑΦΕΑ

A.1. Περιοδικά Ταξινομημένα κατά Science Citation Index (SCI)

- K.A. Raftopoulos and S.D. Kollias: The Global-Local transformation for noise resistant shape representation, *Computer Vision and Image Understanding (CVIU)*, Volume 115, Number 8, Pages 1170-1186, August 2011.
- K. Raftopoulos and N. Papadakis and K. Ntalianis and S. Kollias: A visual pathway for shape-based invariant classification of gray scale images, *Integrated Computer-Aided Engineering*, Volume 14, Number 4, Pages 365-378, December 2007.

A.2. Συνέδρια με Κριτές

- K.A. Raftopoulos and S.D. Kollias: Visual pathways for shape abstraction, *International Conference on Artificial Neural Networks, 2011. ICANN 2011*, 14-17 June 2011, Espoo, Finland.
- P. Tzouveli, K. Raftopoulos, K. Ntalianis and S. Kollias: Semantic Object Watermark Re-Synchronization Based on Skeleton Vertex Corresponds, *ACM Multimedia 2010 Workshop - Multimedia in Forensics, Security and Intelligence (MiFor)*, Florence, Italy, October 29th, 2010.
- K. Raftopoulos, P. Tzouveli, K. Ntalianis, S. Kollias, D.Kalomoirakis, N. Fyssas, G. Foukaneli: Saint Classification in Byzantine Art, *14th International Conference on Virtual Systems and MultiMedia VSMM'08*, Cyprus, 20-26th October, 2008.
- K.A.Raftopoulos and S.D.Kollias: The Global-Local transformation for invariant shape representation, *3d International Symposium on Visual Computing, 2007. ISVC 2007*, 26-28 Nov. 2007, On pages: 224-233.
- K.A.Raftopoulos, N.Papadakis, K.Ntalianis, S.Kollias: Semantic clustering of information systems' users with stochastic techniques, *IEEE International Conference on Semantic Computing, 2007, ICSC 2007* 17-19 Sept. 2007, On page(s): 535-542
- K.A.Raftopoulos, N.Papadakis, K.Ntalianis: Visual pathways for detection of landmark points, *International Conference on Artificial Neural Networks, 2006. ICANN 2006*, 17-19 Sept. 2006, On page(s):728-739

A.3. Σε Στάδιο Κρίσης

- K.A. Raftopoulos, K.S. Ntalianis, D.D. Surlas, S.D. Kollias: Mining User Queries with Markov Chains: Application to Online Image Retrieval, *Knowledge and Data*

Engineering, IEEE Transactions on, η πρώτη έκδοση κατατέθηκε προς κρίση τον Δεκέμβριο του 2009, ακολούθησαν αναθεωρημένες εκδόσεις, μέχρι σήμερα εκρεμεί η τελική απόφαση.