NATIONAL AND KAPODISTRIAN UNIVERSITY OF ATHENS

SCHOOL OF SCIENCES

Department of Informatics and Telecommunications

Program of Postgraduate Studies

PhD Thesis

Retrieval of 3-Dimensional Rigid and Non-Rigid Objects

Konstantinos A. Sfikas

ATHENS

January 2014

ΕΘΝΙΚΟΝ ΚΑΙ ΚΑΠΟΔΙΣΤΡΙΑΚΟΝ ΠΑΝΕΠΙΣΤΗΜΙΟΝ ΑΘΗΝΩΝ

ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ

Τμήμα Πληροφορικής και Τηλεπικοινωνιών

Πρόγραμμα Μεταπτυχιακών Σπουδών

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

Ανάκτηση Τριδιάστατων Άκαμπτων και Εύκαμπτων Αντικειμένων

Κωνσταντίνος Α. Σφήκας

ΑΘΗΝΑ

Ιανουάριος 2014

PhD Thesis

Retrieval of 3-Dimensional Rigid and Non-Rigid Objects

Konstantinos A. Sfikas

SUPERVISOR PROFESSOR: **Theoharis Theoharis**, Professor UoA

THREE-MEMBER ADVISORY COMMITTEE:
Theoharis Theoharis, Professor UoA
Ioannis Pratikakis, Assistant Professor DUTh
Stavros Perantonis, Research Director NCSR Demokritos

SEVEN-MEMBER EXAMINATION COMMITTEE:

Theoharis Theoharis
Professor UoA, NTNU

Ioannis Pratikakis
Assistant Professor DUTh

Stavros Perantonis
Research Director NCSR Demokritos

Dimitrios Gunopulos
Professor UoA

Ioannis Emiris
Professor UoA

Sergios Theodoridis
Professor UoA

Emmanouil Sangriotis
Associate Professor UoA

Examination Date: December 30, 2013

Διδακτορική Διατριβή

Ανάκτηση Τριδιάστατων Άκαμπτων και Εύκαμπτων Αντικειμένων

Κωνσταντίνος Α. Σφήκας

ΕΠΙΒΛΕΠΩΝ ΚΑΘΗΓΗΤΗΣ: **Θεοχάρης Θεοχάρης**, Καθηγητής ΕΚΠΑ

ΤΡΙΜΕΛΗΣ ΣΥΜΒΟΥΛΕΥΤΙΚΗ ΕΠΙΤΡΟΠΗ:
Θεοχάρης Θεοχάρης, Καθηγητής ΕΚΠΑ
Ιωάννης Πρατικάκης, Επ. Καθηγητής ΔΠΘ
Σταύρος Περαντώνης, Διευθ. Έρευνας ΕΚΕΦΕ Δημόκριτος

ΕΠΤΑΜΕΛΗΣ ΕΞΕΤΑΣΙΚΗ ΕΠΙΤΡΟΠΗ:

Θεοχάρης Θεοχάρης
Καθηγητής ΕΚΠΑ, NTNU

Ιωάννης Πρατικάκης
Επ. Καθηγητής ΔΠΘ

Σταύρος Περαντώνης
Διευθ. Έρευνας ΕΚΕΦΕ Δημόκριτος

Δημήτριος Γουνόπουλος
Καθηγητής ΕΚΠΑ

Ιωάννης Εμίρης
Καθηγητής ΕΚΠΑ

Σέργιος Θεοδωρίδης
Καθηγητής ΕΚΠΑ

Εμμανουήλ Σαγκριώτης
Αν. Καθηγητής ΕΚΠΑ

Ημ/νία Εξέτασης: 30 Δεκεμβρίου 2013

ABSTRACT

In recent years, the increased availability of inexpensive 3D object acquisition hardware and simplified 3D modelling software has resulted in the creation of massive 3D model datasets that are either publicly available or for proprietary use (i.e. industrial or commercial). This increase of information has also created the need for methods that are both effective and efficient in terms of 3D model annotation and search.

This dissertation focuses on the problem of 3D object retrieval from large datasets in a near realtime manner. In order to address this task we focus on three major subproblems of the field: (i) pose normalization of rigid 3D models with applications to 3D object retrieval, (ii) non-rigid 3D model description and (iii) search over rigid 3D model datasets based on 2D image queries.

The first of the three subproblems, 3D model pose normalization, has a wide range of applications. Pose normalization of 3D models is a handy tool used for visualization of large 3D model databases, but also an important preprocessing step in a number of 3D object operations (i.e. retrieval, biometrics etc.). Here, three novel pose normalization methods are presented, based on: (i) 3D Reflective Object Symmetry (ROSy) and (ii, iii) 2D Reflective Object Symmetry computed on Panoramic Views (SymPan and SymPan+). All three are general purpose 3D model pose normalization methods used to enhance the performance of a rigid 3D object retrieval pipeline.

Considering the second subproblem, a non-rigid 3D object retrieval methodology which combines the properties of conformal geometry and graph-based topological information (ConTopo++) has been developed. This method is both robust and efficient in terms of retrieval accuracy and computation speed: while graph-based methods are robust to non-rigid object deformations, they require intensive computations which can be reduced by the use of appropriate representations, addressed through geometry-based methods. In this respect, a 3D object retrieval methodology which combines the above advantages in a unified manner, is presented. Furthermore, a string matching strategy for the comparison of graphs that describe 3D models, is proposed.

Regarding the third subproblem a 3D object retrieval method, based on 2D range image queries that represent partial views of real 3D objects, is presented. The complete 3D models of the database are described by a set of panoramic views and a Bag-of-Visual-Words model is built using SIFT features extracted from them. To address the problem of partial matching, a histogram computation

scheme on the panoramic views (that represents local information by taking into account spatial context), is suggested. Furthermore, a number of optimization techniques are applied throughout the process, for enhancing the retrieval performance.

The methodologies developed and described in this dissertation are evaluated in terms of retrieval accuracy and demonstrated using both quantitative and qualitative measures via an extensive consistent evaluation against state-of-the-art methods on standard datasets.

Subject Area Computer Graphics, Computer Vision, Image Processing, Pattern Recognition, Information Retrieval

Keywords 3D objects, Rotation Normalization, Shape Modelling, Partial Matching, Range Images

ΠΕΡΙΛΗΨΗ

Τα τελευταία χρόνια, η αυξημένη διαθεσιμότητα φθηνού υλικού ψηφιοποίησης τριδιάστατων αντικειμένων και απλοποιημένου λογισμικού τριδιάστατης προσομοίωσης είχε ως αποτέλεσμα τη δημιουργία μαζικών βάσεων δεδομένων τριδιάστατων μοντέλων, οι οποίες είτε είναι διαθέσιμες στο κοινό ή αποκλειστικά για ιδιωτική χρήση (π.χ. βιομηχανική ή εμπορική). Αυτή η αύξηση της πληροφορίας δημιούργησε την ανάγκη για μεθόδους που είναι ταυτόχρονα αποτελεσματικές και αποδοτικές όσον αφορά την αναζήτηση και την κατηγοριοποίηση των τριδιάστατων μοντέλων.

Η παρούσα διδακτορική διατριβή εστιάζει στο πρόβλημα της ανάκτησης τριδιάστατων αντικειμένων από μεγάλες βάσεις δεδομένων σε σχεδόν πραγματικό χρόνο. Προκειμένου να αντιμετωπιστεί αυτό το πρόβλημα, θα επικεντρωθούμε σε τρία βασικά υποπροβλήματα του χώρου: (i) κανονικοποίηση θέσης άκαμπτων τριδιάστατων μοντέλων με εφαρμογές στην ανάκτηση τριδιάστατων αντικειμένων, (ii) περιγραφή εύκαμπτων τριδιάστατων μοντέλων και (iii) αναζήτηση από βάσεις δεδομένων τριδιάστατων μοντέλων βασιζόμενη σε διδιάστατες εικόνες-ερώτησης.

Το πρώτο από τα τρία υποπροβλήματα, η κανονικοποίηση θέσης τριδιάστατων μοντέλων, έχει ένα ευρύ φάσμα εφαρμογών. Η κανονικοποίηση θέσης είναι ένα χρήσιμο εργαλείο για την απεικόνιση μεγάλων βάσεων δεδομένων τριδιάστατων μοντέλων, αλλά και ένα σημαντικό βήμα προεπεξεργασίας σε μια σειρά από εφαρμογές τριδιάστατων αντικειμένων (π.χ. ανάκτηση, αναγνώριση βάσει βιομετρίας, κλπ.). Στην παρούσα διατριβή παρουσιάζονται τρεις νέες μέθοδοι κανονικοποίησης, οι οποίες βασίζονται στις εξής αρχές: (i) Τριδιάστατη Ανακλαστική Συμμετρία Αντικειμένου (ROSy)και (ii, iii) Διδιάστατη Ανακλαστική Συμμετρία Αντικειμένου υπολογιζόμενη επί Πανοραμικών Προβολών (SymPan και SymPan+).Και οι τρεις μέθοδοι είναι γενικού σκοπού μέθοδοι κανονικοποίησης τριδιάστατων μοντέλων και χρησιμοποιούνται για να αυξήσουν την απόδοση μιας διαδικασίας ανάκτησης τριδιάστατων αντικειμένων.

Όσον αφορά το δεύτερο υποπρόβλημα, αναπτύχθηκε μια μεθοδολογία ανάκτησης εύκαμπτων τριδιάστατων αντικειμένων, η οποία συνδυάζει τις ιδιότητες της σύμμορφης γεωμετρίας και της τοπολογικής πληροφορίας βασιζόμενης σε αναπαράσταση γράφων (ConTopo++).Η μέθοδος αυτή είναι ταυτόχρονα εύρωστη και αποτελεσματική όσον αφορά την ακρίβεια και την ταχύτητα ανάκτησης:

ενώ οι μέθοδοι που βασίζονται στην αναπαράσταση γράφων είναι εύρωστες για παραμορφώσεις εύκαμπτων αντικειμένων, απαιτούν εντατικούς υπολογισμούς, οι οποίοι ωστόσο μπορούν να μειωθούν με τη χρήση κατάλληλων αναπαραστάσεων, μέσω μεθόδων που βασίζονται στην γεωμετρία. Σύμφωνα με αυτή την θεώρηση, παρουσιάζεται μια μεθοδολογία ανάκτησης τριδιάστατων αντικειμένων, η οποία συνδυάζει τα παραπάνω πλεονεκτήματα με ενιαίο τρόπο. Επιπλέον, προτείνεται μια στρατηγική συνταιριασμού συμβολοσειρών, για τη σύγκριση των γράφων που αναπαριστούν τριδιάστατα μοντέλα.

Σχετικά με το τρίτο υποπρόβλημα, παρουσιάζεται μια μέθοδος ανάκτησης τριδιάστατων αντικειμένων, βασιζόμενη σε διδιάστατες εικόνες-ερώτησης, οι οποίες αντιπροσωπεύουν προβολές πραγματικών τριδιάστατων αντικειμένων. Τα πλήρη 3Δ μοντέλα της βάσης δεδομένων περιγράφονται από ένα σύνολο πανοραμικών προβολών και ένα μοντέλο Bag-of-Visual-Words δημιουργείται χρησιμοποιώντας τα χαρακτηριστικά SIFT που προέρχονται από αυτά. Για να αντιμετωπιστεί το πρόβλημα της μερικής ταύτισης, προτείνεται ένα σχήμα υπολογισμού ιστογραμμάτων από τις πανοραμικές προβολές (που εκπροσωπεί τοπική πληροφορία λαμβάνοντας υπόψη το χωρικό πλαίσιο). Επιπλέον , ένα σύνολο από τεχνικές βελτιστοποίησης εφαρμόζεται καθ' όλη τη διαδικασία, για την ενίσχυση της απόδοσης ανάκτησης.

Οι μεθοδολογίες που αναπτύχθηκαν και περιγράφονται στην παρούσα διατριβή αξιολογούνται όσον αφορά την ακρίβεια ανάκτησης και παρουσιάζονται κάνοντας χρήση τόσο ποσοτικών όσο και ποιοτικών μέτρων μέσω μιας εκτεταμένης και συνεκτικής αξιολόγησης σε σχέση με μεθόδους τρέχουσας τεχνολογικής στάθμης επάνω σε τυποποιημένες βάσεις δεδομένων.

Θεματική Περιοχή Γραφικά Υπολογιστών, Υπολογιστική Όραση, Επεξεργασία Εικόνας, Αναγνώριση Προτύπων, Ανάκτηση Πληροφορίας

Λέξεις Κλειδιά Τριδιάστατα Αντικείμενα, Κανονικοποίηση Περιστροφής, Μοντελοποίηση Σχήματος, Μερική Ταύτιση, Εικόνες Βάθους

ΕΥΧΑΡΙΣΤΙΕΣ

Για τις μεταπτυχιακές μου σπουδές, καθώς και για την εκπόνηση της παρούσας διδακτορικής διατριβής, έλαβα υποτροφία από Ελληνικό Ίδρυμα Κρατικών Υποτροφιών (ΙΚΥ). Η διάρκεια της χρηματοδότησης ήταν από το 2007 έως και το 2011 και η στήριξη αυτή αναγνωρίζεται δια του παρόντος.

Από το 2012 και έπειτα, η έρευνα που πραγματοποιήθηκε στην παρούσα διδακτορική διατριβή έχει χρηματοδοτηθεί από την Ευρωπαϊκή Ένωση (Ευρωπαϊκό Κοινωνικό Ταμείο - ΕΚΤ) και την Ελλάδα από εθνικούς πόρους μέσω του Επιχειρησιακού Προγράμματος Εκπαίδευση και Δια Βίου Μάθηση του Εθνικού Στρατηγικού Πλαισίου Αναφοράς (ΕΣΠΑ) - Χρηματοδότηση Ερευνητικών Προγραμμάτων: THALES - 3DOR (MIS 379516). Η επένδυση στην κοινωνία της γνώσης μέσω του Ευρωπαϊκού Κοινωνικού Ταμείου.

Η παρούσα διδακτορική διατριβή εκπονήθηκε στο Εργαστήριο Γραφικών Υπολογιστή του Τμήματος Πληροφορικής και Τηλεπικοινωνιών του Εθνικού και Καποδιστριακού Πανεπιστημίου Αθηνών.

Contents

List of Figures

List of Tables

List of Algorithms

Σύνοψη Διατριβής

Εισαγωγή

Ο όρος *Πληροφορία*, συνήθως αναφέρεται σε ένα χρήσιμο μέρος δεδομένων που βρίσκεται ανάμεσα σε μια συλλογή σχετικών οντοτήτων. Οι πρόσφατες εξελίξεις στις τεχνολογίες αποθήκευσης και η ευρεία χρήση του Διαδικτύου, έχουν οδηγήσει σε μεγάλη αύξηση του όγκου των δεδομένων που αποθηκεύονται και διανέμονται από μεγάλες βάσεις δεδομένων. Οποιαδήποτε προσπάθεια για χειροκίνητο χαρακτηρισμό και εξαγωγή πληροφοριών είναι σχεδόν αδύνατη, καθιστώντας ως εκ τούτου υποχρεωτική την ανάγκη αυτοματοποίησης μιας τέτοιας εργασίας.

Η διαδικασία της εξαγωγής χρήσιμων πληροφοριών από μεγάλες ποσότητες δεδομένων, με αυτοματοποιημένο τρόπο και με βάση ένα παράδειγμα ή μία περιγραφική ερώτηση, ονομάζεται ανάκτηση πληροφοριών. Διαδεδομένοι τύποι πληροφοριών που μπορούν να επωφεληθούν από μια τέτοια διαδικασία ανάκτησης είναι: κείμενο, εικόνα ήχος, τριδιάστατα και τετραδιάστατα πλέγματα (η τελευταία κατηγορία αναφέρεται σε τριδιάστατα πλέγματα με χρονική αλληλουχία).

Τα τελευταία χρόνια, μέσω της δημιουργίας φθηνών τριδιάστατων ψηφιοποιητών και της απλούστευσης του λογισμικού τριδιάστατης μοντελοποίησης, έχει δημιουργηθεί ένας μεγάλος όγκος τριδιάστατων δεδομένων, ο οποίος βρίσκεται αποθηκευμένος σε αντίστοιχες επιστημονικές και βιομηχανικές/εμπορικές αποθήκες δεδομένων. Επιπλέον, τα τριδιάστατα δεδομένα μπορούν να υποβάλλονται σε επεξεργασία με διάφορους τρόπους, ανάλογα με την εφαρμογή, και περιστασιακά να συνδυάζονται με δεδομένα άλλων τύπων (π.χ. κείμενα σχολιασμού ή/και μικρογραφίες των τριδιάστατων μοντέλων). Αυτοί οι τύποι δεδομένων μπορούν επιπλέον να χρησιμοποιηθούν ως ερωτήσεις για την ανάκτηση των τριδιάστατων αντικειμένων.

Ορισμένα παραδείγματα εφαρμογών που εκμεταλλεύονται τις ιδιότητες των τριδιάστατων

μοντέλων και μπορούν να ωφεληθούν σημαντικά από μια διαδικασία ανάκτησης είναι τα εξής: στην ιατρική, μεγάλα διαγνωστικά τριδιάστατα δεδομένα συγκρίνονται και ερευνώνται, προκειμένου να βοηθηθεί η διαδικασία λήψης ιατρικών αποφάσεων. Στην βιομετρία, το τριδιάστατο μοντέλο του προσώπου ενός ατόμου αναζητείται σε αντίστοιχες βάσεις δεδομένων με σκοπό την ταυτοποίηση. Η βιομηχανία ανάπτυξης ψηφιακών παιχνιδιών χρησιμοποιεί την ανάκτηση και την επαναχρησιμοποίηση των τριδιάστατων μοντέλων προκειμένου να ελαχιστοποιηθούν οι χρόνοι παραγωγής και να μειωθεί το μέγεθος του τελικού προϊόντος. Άλλοι τομείς εφαρμογής περιλαμβάνουν την μηχανολογία και την αρχαιολογία. Συνεπώς, μπορεί εύκολα να εξαχθεί το συμπέρασμα, ότι η ανάκτηση τριδιάστατων αντικειμένων είναι μια βασική διαδικασία, αν και σε γενικές γραμμές είναι πολύπλοκη και εξαρτάται ιδιαίτερα από την εφαρμογή.

Πλαίσιο Εργασίας και Διατύπωση του Προβλήματος

Οι εφαρμογές ανάκτησης τριδιάστατων αντικειμένων μπορούν να ταξινομηθούν σε δύο μεγάλες κατηγορίες: διακατηγοριακή (inter-class) και ενδοκατηγοριακή (intra-class) . Η inter-class ανάκτηση εστιάζει στο γενικό τομέα των τριδιάστατων αντικειμένων και στοχεύει στην εύρεση της καλύτερης ταύτισης ανάμεσα σε μια σειρά από τριδιάστατα μοντέλα που ανήκουν σε ένα ευρύ φάσμα διαφορετικών κλάσεων. Σε αυτή την περίπτωση, συνήθως δεν υπάρχει προηγούμενη γνώση σχετικά με τα χαρακτηριστικά ή τη φύση των τριδιάστατων αντικειμένων. Η intra-class ανάκτηση στοχεύει σε μια συγκεκριμένη κατηγορία των τριδιάστατων αντικειμένων (π.χ. τριδιάστατα πρόσωπα, εύκαμπτα τριδιάστατα μοντέλα, μοντέλα αναπαράστασης ανθρώπινων δράσεων, μοντέλα μηχανολογίας κλπ.), όπου αναζητείται μια ταύτιση μεταξύ τριδιάστατων μοντέλων που ανήκουν στην ίδια κλάση, αλλά έχουν τα ιδιαίτερα χαρακτηριστικά τους οριζόμενα διαφορετικά. Οι μέθοδοι ανάκτησης intra-class εκμεταλλεύονται συνήθως γνώσεις του τομέα και τα χαρακτηριστικά μορφής των τριδιάστατων μοντέλων, προκειμένου να επιτύχουν υψηλότερες επιδόσεις.

Και για τις δύο περιπτώσεις, το γενικό πλαίσιο ενός συστήματος ανάκτησης τριδιάστατων αντικειμένων μπορεί να περιγραφεί ως εξής: προεπεξεργασία, κανονικοποίηση θέσης, υπολογισμός περιγραφέα σχήματος, συνταιριασμός χαρακτηριστικών.

Κατά την *προεπεξεργασία*τα τριδιάστατα μοντέλα καθαρίζονται από τυχόν ανωμαλίες που μπορεί να παρουσιάζουν λόγω της διαδικασίας ψηφιοποίησης, παραδείγματος χάριν διπλές

ή αντεστραμμένες επιφάνειες, κενά στην δομή τους κλπ. Το βήμα αυτό εξαρτάται σε μεγάλο βαθμό από την μέθοδο/εξοπλισμό που χρησιμοποιείται για τη δημιουργία των τριδιάστατων μοντέλων και μπορεί να διαφέρει σημαντικά από τη μία εφαρμογή στην άλλη.

Η*κανονικοποίηση θέσης* διασφαλίζει ότι οι γεωμετρικές ιδιότητες των τριδιάστατων μοντέλων ορίζονται με ομοιόμορφο τρόπο. Η ποικιλομορφία των πηγών απόκτησης των τριδιάστατων αντικειμένων συνεπάγεται ότι τα τριδιάστατα μοντέλα που μπορεί ακόμη και να είναι μέρος του ίδιου συνόλου δεδομένων, δύνανται να έχουν τις γεωμετρικές ιδιότητές τους αυθαίρετα ορισμένες. Ως εκ τούτου, πριν από την οποιαδήποτε επεξεργασία, πρέπει να εξασφαλίζεται ότι τα τριδιάστατα αντικείμενα έχουν κανονικοποιηθεί ως προς την μεταφορά, την περιστροφή και την κλιμάκωση (στο Σχήμα 1 παρουσιάζεται ένα παράδειγμα κανονικοποίησης περιστροφής). Η κανονικοποίηση θέσης των τριδιάστατων αντικειμένων είναι ένα κοινό βήμα προεπεξεργασίας σε διάφορες εφαρμογές γραφικών υπολογιστή [21, 110, 121, 141]. Η οπτικοποίηση, η ανασυγκρότηση σπασμένων θραυσμάτων, η βιομετρία και ανάκτηση τριδιάστατων αντικειμένων είναι μόνο μερικά παραδείγματα εφαρμογών που επωφελούνται από την διαδικασία κανονικοποίησης θέσης.

Το βασικό βήμα ενός συστήματος ανάκτησης τριδιάστατων αντικειμένων είναι ο *υπολογισμός του περιγραφέα σχήματος* δηλαδή ενός συνόλου χαρακτηριστικών που περιγράφουν το σχήμα του αντικειμένου με συνοπτικό και ακριβή τρόπο. Στο στάδιο αυτό, τα δομικά ή/και άλλα ειδικά χαρακτηριστικά ενός τριδιάστατου αντικειμένου μοντελοποιούνται και δημιουργείται ένας περιγραφέας σχήματος που κωδικοποιεί πιστά το σχήμα του τριδιάστατου μοντέλου.Η επιλογή των χαρακτηριστικών είναι στενά συνδεδεμένη με την αντίστοιχη εφαρμογή και μπορεί να ποικίλει σε μεγάλο βαθμό σε κάθε σύστημα ανάκτησης τριδιάστατων αντικειμένων (π.χ. η intra-class ανάκτηση εκμεταλλεύεται τα χαρακτηριστικά που είναι πιο διακριτικά μέσα σε ένα συγκεκριμένο τομέα, ενώ η inter-class ανάκτηση χρησιμοποιεί πιο γενικά χαρακτηριστικά).

Τέλος, κάθε περιγραφέας σχήματος χρησιμοποιείται ως μια *υπογραφή* κατά την διαδικασία *συνταιριασμού*. Σε αυτό το βήμα, οι υπογραφές των τριδιάστατων μοντέλων, που βρίσκονται αποθηκευμένες στη βάση δεδομένων, συγκρίνονται με τις αντίστοιχες υπογραφές του τριδιάστατου μοντέλου ερώτησης, χρησιμοποιώντας μια μετρική σύγκρισης. Η επιλογή της μετρικής επίσης από τα επιλεχθέντα χαρακτηριστικά και την αντίστοιχη εφαρμογή. Η απόκριση του συστήματος ανάκτησης τριδιάστατων αντικειμένων αποτελείται από το σύνολο των τριδιάσ-

τατων αντικειμένων που αντιστοιχούν στα πλησιέστερα ταίρια της συγκεκριμένης ερώτησης.

Συνεισφορά

Η παρούσα διατριβή έχει κάνει τις ακόλουθες συνεισφορές στην έρευνα στον τομέα της ανάκτησης τριδιάστατων αντικειμένων: (i-iii) τρεις νέες μέθοδοι κανονικοποίησης θέσης τριδιάστατων μοντέλων, (iv) μια μεθοδολογία ανάκτησης για εύκαμπτα τριδιάστατα αντικείμενα και (v) έναν αλγόριθμο ανάκτησης τριδιάστατων αντικειμένων, με βάση ερωτήσεις αποτελούμενες από εικόνες βάθους. Πιο αναλυτικά:

i ROSy: Μέθοδος κανονικοποίησης θέσης γενικού σκοπού, η οποία βασίζεται στην ανακλαστική συμμετρία των τριδιάστατων αντικειμένων. Αρχικά, το ελάχιστο πλαίσιο οριοθέτησης ενός άκαμπτου τριδιάστατου αντικειμένου - ευθυγραμμισμένου με τους πρωτεύοντες άξονες του χώρου - τροποποιείται με την απαίτηση το τριδιάστατο μοντέλο να βρίσκεται επίσης σε ελάχιστη γωνιακή διαφορά σε σχέση με τα κανονικά διανύσματα των εδρών του πλαισίου οριοθέτησής του. Για να εκτιμηθεί το τροποποιημένο - ευθυγραμμισμένο πλαίσιο οριοθέτησης, χρησιμοποιείται ένα σύνολο προκαθορισμένων επιπέδων συμμετρίας και υπολογίζεται μια συνδυαστική χωρική και γωνιακή απόσταση μεταξύ του τριδιάστατου μοντέλου και του συμμετρικού μοντέλου του. Με την ελαχιστοποίηση της συνδυαστικής απόστασης, το τριδιάστατο μοντέλο τοποθετείται μέσα στο τροποποιημένο - ευθυγραμμισμένο πλαίσιο οριοθέτησης και κατά συνέπεια επιτυγχάνεται η ευθυγράμμιση με το σύστημα συντεταγμένων [105].

ii SymPan: Μέθοδος κανονικοποίησης θέσης, με βάση τις πανοραμικές προβολές και την ανακλαστική συμμετρία. Αρχικά, η επιφάνεια ενός τριδιάστατου μοντέλου προβάλλεται επάνω στην πλευρική επιφάνεια ενός περιγεγραμμένου κυλίνδρου, ευθυγραμμισμένου με τον πρωτεύοντα κύριο άξονα του χώρου. Με βάση αυτή την κυλινδρική προβολή, εξάγεται ένας χάρτης απόκλισης κανονικών διανυσμάτων και χρησιμοποιώντας μια στρατηγική αναζήτησης βασιζόμενης στα οκταδικά δένδρα, υπολογίζεται εκείνη η περιστροφή που ευθυγραμμίζει τον πρωτεύοντα κύριο άξονα του τριδιάστατου μοντέλου και τον άξονα του κυλίνδρου. Κατόπιν, οι δευτερεύοντες άξονες του τριδιάστατου μοντέλου ευθυγραμμίζονται με τον δευτερεύοντα κύριο άξονα του χώρου με παρόμοιο τρόπο [104].

iii SymPan+: Μέθοδος βασιζόμενη στην μεθοδολογία κανονικοποίησης θέσης SymPan. Η ποιοτική και πειραματική έρευνα επάνω στα τριδιάστατα σύνολα δεδομένων, μας οδήγησε στο συμπέρασμα πως τα περισσότερα αντικείμενα διαθέτουν τουλάχιστον ένα επίπεδο συμμετρίας. Η προτεινόμενη μεθοδολογία καθοδηγείται από αυτή την παρατήρηση. Αρχικά, μέσω μιας βηματικής διαδικασίας, εκτιμάται το επίπεδο συμμετρίας ενός τριδιάστατου μοντέλου. Κατά συνέπεια, υπολογίζεται και ο πρώτος άξονας του μοντέλου. Αυτό επιτυγχάνεται μέσω της περιστροφής του τριδιάστατου μοντέλου και του υπολογισμού του επιπέδου ανακλαστικής συμμετρίας επάνω σε πανοραμικές προβολές. Οι εναπομείναντες άξονες συμμετρίας του τριδιάστατου μοντέλου εκτιμώνται μέσω του υπολογισμού της διακύμανσης των πανοραμικών προβολών του τριδιάστατου μοντέλου [108].

iv ConTopo: Συνδυάζοντας τις ιδιότητες της σύμμορφης γεωμετρίας και των γράφων με βάση την τοπολογική πληροφορία, προτείνεται μια μεθοδολογία ανάκτησης για εύκαμπτα τριδιάστατα αντικείμενα, η οποία είναι ταυτόχρονα εύρωστη και αποτελεσματική σχετικά με την ακρίβεια ανάκτησης και την ταχύτητα υπολογισμού. Ενώ οι μέθοδοι που βασίζονται στους γράφους είναι ανεκτικοί σε παραμορφώσεις εύκαμπτων αντικειμένων, απαιτούν εντατικούς υπολογισμούς, οι οποίοι ωστόσο δύναται να μειωθούν με τη χρήση κατάλληλων αναπαραστάσεων μέσω των μεθόδων που βασίζονται στην γεωμετρία. Σύμφωνα με αυτή την θεώρηση, παρουσιάζεται μια μεθοδολογία ανάκτησης τριδιάστατων αντικειμένων, η οποία συνδυάζει τα παραπάνω πλεονεκτήματα με ενιαίο τρόπο. Επιπλέον, προτείνεται μια στρατηγική συνταιριασμού σειρών για τη σύγκριση των γράφων που περιγράφουν τα τριδιάστατα αντικείμενα [106].

v Αλγόριθμος ανάκτησης τριδιάστατων αντικειμένων, με βάση τις εικόνες-ερωτήσεις που αντιπροσωπεύουν τμηματικές προβολές πραγματικών τριδιάστατων αντικειμένων. Τα πλήρη τριδιάστατα μοντέλα της βάσης δεδομένων περιγράφονται από ένα σύνολο πανοραμικών προβολών και ένα Bag-of-Visual-Words μοντέλο δημιουργείται από χαρακτηριστικά SIFT, τα οποία υπολογίζονται σε αυτές. Για να αντιμετωπιστεί το πρόβλημα του μερικού συνταιριασμού, προτείνεται ένα σχήμα υπολογισμού ιστογραμμάτων, υπολογιζόμενων στις πανοραμικές προβολές. Τα ιστογράμματα αντιπροσωπεύουν τοπική πληροφορία λαμβάνοντας υπόψιν το χωρικό πλαίσιο [107].

Πειραματικά Αποτελέσματα

Μετρικές Αξιολόγησης

Η πειραματική αξιολόγηση των προτεινόμενων μεθόδων βασίζεται στις καμπύλες Ακρίβειας-Ανάκλησης (Precision-Recall) καθώς και σε πέντε ποσοτικά μέτρα: τον Κοντινότερο Γείτονα (Nearest Neighbor), την Πρώτη Τάξη (First Tier), την Δεύτερη Τάξη (Second Trier), το Μέτρο Ε (E-measure) και το Προεξοφλούμενο Συσσωρευμένο Κέρδος (Discounted Cumulative Gain) [110] για τις κλάσεις του εκάστοτε συνόλου δεδομένων. Για κάθε μοντέλο ερώτηση που ανήκει στην κλάση C, η Ανάκληση δηλώνει το ποσοστό των μοντέλων της κλάσης C που ανακτώνται και η Ακρίβεια δηλώνει το ποσοστό των μοντέλων που ανήκουν στην κλάση C ως προς τον συνολικό αριθμό των ανακτημένων μοντέλων. Το καλύτερο αποτέλεσμα είναι 100% και για τις δύο ποσότητες. Το μέτρο του Πλησιέστερου Γείτονα δείχνει το ποσοστό των ερωτημάτων, όπου η πιο κοντινή ταύτιση ανήκει στην κλάση της ερώτησης. Οι στατιστικές Πρώτης και Δεύτερης Τάξης, μετρούν την τιμή ανάκλησης για τις $(D-1)$ και $2(D-1)$ πλησιέστερες ανακτήσεις, αντίστοιχα, όπου D είναι η πληθάριθμος της κλάσης της ερώτησης. Το Μέτρο Ε συνδυάζει την ακρίβεια και την ανάκληση σε ένα ενιαίο αριθμό και το Προεξοφλούμενο Συσσωρευμένο Κέρδος βαρύνει τα σωστά αποτελέσματα που βρίσκονται πιο κοντά στην κορυφή της λίστας ανάκλησης περισσότερο από τα σωστά αποτελέσματα που βρίσκονται σε μετέπειτα θέσεις της λίστας, με βάση την παραδοχή ότι ο χρήστης είναι λιγότερο πιθανό να εξετάσει τα στοιχεία κοντά στο τέλος της λίστας [59, 110].

Βάσεις Δοκιμών

Για κάθε μέθοδο που παρουσιάζεται, οι αξιολογήσεις έγιναν σε ευρέως χρησιμοποιούμενα σύνολα δεδομένων, επιτρέποντας έτσι την σύγκριση των αποτελεσμάτων τόσο με μεθοδολογίες της τρέχουσας τεχνολογικής στάθμης όσο και με μελλοντικές μεθόδους.

Τα σύνολα δεδομένων, επάνω στα οποία διεξήχθησαν τα πειράματα αξιολόγησης της μεθόδου ROSy, είναι τα εξής: τα σύνολα εκπαίδευσης και ελέγχου της βάσης Princeton Shape Benchmark (PSB) [110], τα κατηγοριοποιημένα αντικείμενα της βάσης δεδομένων National Taiwan University database (NTU) [28], το σύνολο δεδομένων MPEG-7 [129], το σύνολο δεδομένων Engineering Shape Benchmark dataset (ESB) [60] και η βάση δεδομένων National Institute of Standards and Technology dataset (NIST), η οποία περιλαμβάνει τριδιάστατα

αντικείμενα, προερχόμενα από τον διαγωνισμό SHREC 2009 [43], και των οποίων το σχήμα έχει κανονικοποιηθεί και έχουν κατηγοριοποιηθεί οπτικά. Τέλος, πειράματα διεξήχθησαν στα αρθρωτά και μη αρθρωτά αντικείμενα του συνόλου δεδομένων McGill [144]. Από το σύνολο δεδομένων NTU, χρησιμοποιήθηκαν μόνο τα ταξινομημένα αντικείμενα, καθώς τα μη ταξινομημένα αντικείμενα δεν θα παρείχαν ακριβή αποτελέσματα ανάκτησης. Ο Πίνακας 2 (Table 2) δείχνει τον αριθμό των κατηγοριών και τον συνολικό αριθμό των αντικειμένων σε κάθε σύνολο δεδομένων που χρησιμοποιήθηκε για τα πειράματα. Το σύνολο των δεδομένων, που χρησιμοποιήθηκε για την αξιολόγηση της μεθόδου SymPan είναι το υποσύνολο δοκιμής της βάσης Princeton Shape Benchmark (PSB) [110]. Αυτό το σύνολο δεδομένων αποτελείται από 907 τριδιάστατα μοντέλα που κατατάσσονται σε 92 κλάσεις. Αντίστοιχα, για την μέθοδο κανονικοποίησης θέσης SymPan+ εκτός από το υποσύνολο δοκιμής της βάσης PSB, χρησιμοποιήθηκαν επίσης τα σύνολα δεδομένων MPEG-7 [129], ESB [60] και NIST.

Τα σύνολα δεδομένων που χρησιμοποιούνται για την πειραματική αξιολόγηση της μεθοδολογίας ανάκτησης εύκαμπτων τριδιάστατων αντικειμένων ConTopo είναι τα εξής: (i) TOSCA [19, 18], (ii) SHREC'07 *Watertight Models* [50], (iii) SHREC'10 *Non-rigid 3D Models* [71] και (iv) SHREC'11 *Non-rigid 3D Watertight Meshes* [70].

Τέλος, τα σύνολα δεδομένων που χρησιμοποιήθηκαν για την πειραματική αξιολόγηση της προτεινόμενης μεθόδου ανάκτησης τριδιάστατων αντικειμένων, με βάση εικόνες-ερωτήσεις είναι τα εξής: (i) SHREC'09 *Querying with Partial Models* [39], (ii) SHREC'10 *Range Scan Retrieval* [40] και (iii) SHREC'11 *Shape Retrieval Contest of Range Scans* [113]. Το υποσύνολο στόχος των χρησιμοποιούμενων συνόλων δεδομένων βασίζεται στο γενικό σύνολο δεδομένων ελέγχου που κατασκευάστηκε στο NIST [43].

Μέθοδοι Κανονικοποίησης Θέσης ROSy, SymPan και SymPan+

Για την εκτίμηση της απόδοσης της μεθόδου κανονικοποίησης θέσης ROSy, επιλέχθηκε μια μεθοδολογία ανάκτησης τριδιάστατων αντικειμένων, τρέχουσας τεχνολογικής στάθμης, από τους Papadakis et al. [88], σαν όχημα αποτίμησης.

Το σύστημα ανάκτησης τριδιάστατων αντικειμένων του Παπαδάκη, στην αρχική του μορφή, χρησιμοποιεί ένα συνδυασμό δύο μεθόδων κανονικοποίησης θέσης για να επιτύχει την βέλτιστη κανονικοποίηση ενός τριδιάστατου αντικειμένου. Η προτεινόμενη μέθοδος παρουσιάζει όμοιες επιδόσεις με τις δύο υφιστάμενες, ωστόσο ο συνδυασμός των τριών μεθόδων κανον-

ικοποίησης θέσης (ονομαζόμενος ROSy+), παρέχει μια σημαντική ώθηση στην διακριτική ικανότητα του συστήματος ανάκτησης, ξεπερνώντας την αρχική προσέγγιση.

Στο Σχήμα 35 διαφαίνεται πως το σύστημα ανάκτησης τριδιάστατων αντικειμένων, κάνοντας χρήση της συνδυασμένης μεθόδου κανονικοποίησης θέσης ROSy+ , ξεπερνάει δύο πρόσφατες μεθόδους κανονικοποίησης: DLA [27] και GSMD+SHD+R [72], τον περιγραφέα PANORAMA καθώς και τρεις κλασικές μεθοδολογίες ανάκτησης τριδιάστατων αντικειμένων: Lightfield [28], SH-GEDT [65] και DESIRE [132]. Η προτεινόμενη μέθοδος εξετάσθηκε επίσης στο σύνολο των διαθέσιμων βάσεων δοκιμών. Τα αντίστοιχα αριθμητικά αποτελέσματα παρουσιάζονται στον Πίνακα 4 (Table 4). Τα αποτελέσματα δείχνουν πως η προτεινόμενη μέθοδος είναι σταθερή σε όλες τις βάσεις δεδομένων.

Σε όλα τα 8 σύνολα δεδομένων, η προτεινόμενη μέθοδος είναι σε θέση να επιτύχει μια μέση αύξηση επιδόσεων της τάξης του 3% σχέση με την αρχική υβριδική προσέγγιση του Παπαδάκη (μέσος όρος των ποσοτικών μέτρων αξιολόγησης). Αυτό το κέρδος είναι σημαντικό, διότι επιτυγχάνεται αποκλειστικά και μόνο με την ενίσχυση της διαδικασίας κανονικοποίησης θέσης και όχι του πυρήνα του αλγορίθμου ανάκτησης. Επιπλέον, καθίσταται σαφές ότι η προτεινόμενη μεθοδολογία αποδίδει καλύτερα σε σχέση με τις υπόλοιπες μεθόδους ανάκτησης τριδιάστατων αντικειμένων κατά μέσο όρο 2% - 5% (μέσος όρος των ποσοτικών μέτρων αξιολόγησης).

Παρόμοια με τον τρόπο που πραγματοποιήθηκε η ποσοτική αξιολόγηση της μεθοδολογίας κανονικοποίησης θέσης ROSy, για τους αλγορίθμους κανονικοποίησης θέσης SymPan και SymPan+ επιλέχθηκε το σύστημα ανάκτησης τριδιάστατων αντικειμένων, τρέχουσας τεχνολογικής στάθμης, PANORAMA [89] ως όχημα αξιολόγησης. Η προτεινόμενες μέθοδοι αντικαθιστούν την μέθοδο κανονικοποίησης θέσης NPCA, στο υπάρχον υβριδικό σύστημα.

Το άμεσο αποτέλεσμα των προτεινόμενων μεθόδων ευθυγράμμισης μπορεί να αξιολογηθεί συγκρίνοντας τις με την αρχική απόδοση της μεθόδου PANORAMA. Όσον αφορά την απόδοση της ανάκτησης αντικειμένων, τα ολοκληρωμένα συστήματα συγκρίθηκαν με τις αντίστοιχες μεθοδολογίες DLA [27], GSMD+SHD+R [72], Rosy+ [105], Lightfield [28], SH-GEDT [65] και DESIRE [132].

Στο Σχήμα 40, χρησιμοποιώντας τα πειραματικά αποτελέσματα που δίνονται στο [105, 104], απεικονίζονται τα διαγράμματα ακρίβειας-ανάκλησης για το υποσύνολο δοκιμής του συνόλου δεδομένων PSB για το σύστημα ανάκτησης τριδιάστατων αντικειμένων PANORAMA,

όπως αυτό έχει ενισχυθεί τις προτεινόμενες μεθόδους, αντίστοιχα. Ο Πίνακας 5 (Table 5) δείχνει ποσοτικά μέτρα για τις ίδιες μεθόδους. Ο Πίνακας 6 (Table 6) παρουσιάζει τα ποσοτικά αποτελέσματα για τις επιπλέον βάσεις δοκιμής στις οποίες ελέγχθηκε η μέθοδος SymPan+. Τα αποτελέσματα δείχνουν ότι το προτεινόμενο σύστημα υπερτερεί μεθόδων τρέχουσας τεχνολογικής στάθμης και αυξάνει σημαντικά την απόδοση του συστήματος ανάκτησης τριδιάστατων αντικειμένων σε σύγκριση με την αρχική υλοποίηση του.

Στο Σχήμα 41 παρουσιάζονται συγκριτικές ευθυγραμμίσεις μεταξύ των μεθόδων κανονικοποίησης θέσης SymPan+ και CPCA, NPCA για διάφορα τριδιάστατα μοντέλα από το σύνολο δεδομένων PSB. Οι ευθυγραμμίσεις αυτές δείχνουν ότι η προτεινόμενη μέθοδος είναι σε θέση να παράγει ακριβή αποτελέσματα ευθυγράμμισης που, ανεξάρτητα από την κλάση προέλευσης ή τη μορφολογία των αντικειμένων εισόδου, είναι συνεπή και σταθερά.

Σημείωση: Για τα αντίστοιχα αριθμητικά αποτελέσματα, ανατρέξτε στους πίνακες των Παραρτημάτων Α και Β.

Μέθοδος Ανάκτησης Εύκαμπτων Τριδιάστατων Αντικειμένων ConTopo

Στην συνέχεια, πραγματοποιήθηκε σύγκριση μεταξύ της μεθόδου ανάκτησης εύκαμπτων τριδιάστατων αντικειμένων ConTopo και άλλων αντίστοιχων μεθόδων τρέχουσας τεχνολογικής στάθμης σε πρότυπα σύνολα δεδομένων.

Στο Σχήμα 42, σύμφωνα με τα πειραματικά αποτελέσματα που εμφανίζονται στα [11] και [102], απεικονίζονται τα διαγράμματα ακρίβειας-ανάκλησης για το σύνολο των δεδομένων της βάσης TOSCA για την προτεινόμενη μεθοδολογία ανάκτησης εύκαμπτων τριδιάστατων αντικειμένων, την μεθοδολογία Διακριτών Σύμμορφων Συντελεστών (Discrete Conformal Factors - Cfact) και δύο βασικούς περιγραφείς σχήματος για τριδιάστατα αντικείμενα: τον περιγραφέα Πεδίου Φωτός (LightField - LF) και τον περιγραφέα που βασίζεται στις σφαιρικές αρμονικές (Spherical Harmonics - SH). Τα διαγράμματα ακρίβειας-ανάκλησης των παραπάνω μεθόδων δείχνουν σαφώς την υψηλή ακρίβεια της προτεινόμενης μεθόδου.

Στο Σχήμα 43 απεικονίζονται τα διαγράμματα ακρίβειας-ανάκλησης για το πλήρες σύνολο δεδομένων SHREC'07 *Watertight Models*. Η προτεινόμενη μέθοδος αποδίδει καλύτερα από το σύνολο των απεικονιζόμενων περιγραφέων.

Στα Σχήματα 44 και 45 απεικονίζονται τα αποτελέσματα ακρίβειας-ανάκλησης της προ-

τεινόμενης μεθόδου σε σχέση με τα δημοσιευμένα αποτελέσματα των συνόλων δεδομένων SHREC'10 *Non-rigid 3D Models* και SHREC'11 *Non-rigid 3D Watertight Meshes*. Τα αντίστοιχα ποσοτικά μεγέθη εμφανίζονται στους Πίνακες 8 και 9 (Tables 8, 9).

Από τα παραπάνω, είναι σαφές ότι η προτεινόμενη μέθοδος υπερτερεί των αντίστοιχων μεθόδων των διαγωνισμών, παρόλο που αυτές παρουσίαζαν ήδη πολύ υψηλές επιδόσεις.

Όμοια συμπεράσματα μπορούν επίσης να εξαχθούν από τα ποσοτικά μέτρα του Πίνακα 8 (Table 8), όπου η προτεινόμενη μέθοδος έχει τις υψηλότερες βαθμολογίες στα τρία από τα πέντε μέτρα (και είναι επίσης η δεύτερη καλύτερη για τα υπόλοιπα δύο μέτρα, με μικρή διαφορά). Στο σύνολο δεδομένων SHREC'11 *Non-rigid 3D Watertight Meshes*, η προτεινόμενη μέθοδος αποδίδει καλύτερα ή παρουσιάζει ανάμεικτα αποτελέσματα σε σχέση με περίπου το 65% των αποτελεσμάτων των μεθόδων που παρουσιάστηκαν στον διαγωνισμό. Αυτό απεικονίζεται επίσης στο Σχήμα 45 και στον Πίνακα 9 (Table 9).

Σημείωση: Για τα αντίστοιχα αριθμητικά αποτελέσματα, ανατρέξτε στους πίνακες του Παραρτήματος C.

Μέθοδος Ανάκτησης Τριδιάστατων Αντικειμένων βάσει Εικόνων-Ερώτησης

Η προτεινόμενη μέθοδος ανάκτησης άκαμπτων τριδιάστατων αντικειμένων βάσει εικόνων-ερώτησης συγκρίθηκε σε δύο σύνολα δεδομένων, το SHREC'09 *Querying with Partial Models* και το SHREC'10 *Range Scan Retrieval*, σε σχέση με τα υπάρχοντα αποτελέσματα των μεθόδων που συμμετείχαν αντίστοιχα.

Πιο συγκεκριμένα, στο σύνολο δεδομένων SHREC'09 *Querying with Partial Models*, πραγματοποιήθηκε σύγκριση με τις παραλλαγές του περιγραφέα CMVD (Compact Multi-View Descriptor) των Δάρας και Αξενόπουλος [34] και τις μεθόδους BF-SIFT και BF-GridSIFT των Furuya και Ohbuchi. Η αυξημένη απόδοση της προτεινόμενης μεθόδου παρουσιάζεται τόσο στο διάγραμμα ακρίβειας-ανάκλησης του Σχήματος 46 όσο και στα ποσοτικά μέτρα του Πίνακα 10 (Table 10).

Στο σύνολο δεδομένων SHREC'10 *Range Scan Retrieval* πραγματοποιήθηκε σύγκριση με τις παραλλαγές της μεθόδου BF-DSIFT-E των Ohbuchi και Furuya [83] και με τις παραλλαγές της μεθόδου SURFLET, όπως αυτή προτάθηκε από τους Hillebrand et al. [133]. Ο Πίνακας 11 (Table 11) δείχνει πως η προτεινόμενη μέθοδος παρουσιάζει την υψηλότερη

βαθμολογία στις τρεις από τις πέντε μετρικές. Τα διαγράμματα ακρίβειας-ανάκλησης του Σχήματος 47, δείχνουν πως η προτεινόμενη μέθοδος ξεπερνάει το σύνολο των διαγωνιζόμενων.

Σημείωση: Για τα αντίστοιχα αριθμητικά αποτελέσματα, ανατρέξτε στους πίνακες του Παραρτήματος D.

Συμπεράσματα

Στην παρούσα διδακτορική διατριβή παρουσιάστηκαν οι μεθοδολογίες που αναπτύχθηκαν για την αντιμετώπιση των προβλημάτων της κανονικοποίησης θέσης τριδιάστατων μοντέλων, της ανάκτησης τριδιάστατων αντικειμένων με εφαρμογές σε άκαμπτα και εύκαμπτα μοντέλα, καθώς και της ανάκτησης τριδιάστατων αντικειμένων με βάση εικόνες-ερωτήσεις. Στον τομέα της κανονικοποίησης θέσης τριδιάστατων μοντέλων, προτάθηκαν τρεις νέες μέθοδοι με βάση τις ιδιότητες της ανακλαστικής συμμετρίας των τριδιάστατων αντικειμένων. Ο αλγόριθμος κανονικοποίησης θέσης ROSy συμπληρώνει με επιτυχία τις αντίστοιχες μεθόδους CPCA και NPCA ως ένα βήμα προεπεξεργασίας για ένα σύστημα ανάκτησης τριδιάστατων αντικειμένων. Η προσθήκη της προτεινόμενης μεθόδου αυξάνει την διακριτική ικανότητα του συστήματος κατά περίπου 3% πάνω από την προηγούμενη καλύτερη προσέγγιση. Η μεθοδολογία κανονικοποίησης θέσης SymPan βασίζεται σε έναν Χάρτη Απόκλισης Κανονικών Διανυσμάτων της επιφάνειας του τριδιάστατου μοντέλου καθώς και στις ιδιότητες ανακλαστικής συμμετρίας του αντικειμένου. Βασιζόμενοι στις ίδιες αρχές όπως στην μέθοδο κανονικοποίησης θέσης τριδιάστατων μοντέλων SymPan, καθώς και οδηγούμενοι από την πεποίθηση πως τα περισσότερα τριδιάστατα αντικείμενα έχουν τουλάχιστον ένα επίπεδο συμμετρίας, στην μέθοδο SymPan+ επιχειρούμε να προσδιορίσουμε αυτό το επίπεδο και στην συνέχεια να το ευθυγραμμίσουμε με ένα από τα κύρια επίπεδα του χώρου. Στην συνέχεια, οι κύριοι άξονες του τριδιάστατου μοντέλου εκτιμώνται μέσω του υπολογισμού της διακύμανσης των τιμών των εικονοστοιχείων των πανοραμικών προβολών. Οι προτεινόμενες μέθοδοι κανονικοποίησης θέσης χρησιμοποιούν πληροφορίες που προέρχονται από την προβολή των τριδιάστατων μοντέλων επάνω στον περιγεγραμμένο κύλινδρό τους. Η ποιότητα των μεθόδων ευθυγράμμισης αποδεικνύεται τόσο οπτικά όσο και μέσα από τις επιδόσεις ενός συστήματος ανάκτησης τριδιάστατων αντικειμένων τρέχουσας τεχνολογικής στάθμης. Η μέθοδος SymPan βελτιώνει

τη διακριτική ικανότητα του συστήματος ανάκτησης τριδιάστατων αντικειμένων PANORAMA κατά μέσο όρο 6% πάνω από την αρχική προσέγγιση. Η μέθοδος SymPan+ επιπλέον βελτιώνει την διακριτική ικανότητα του συστήματος ανάκλησης PANORAMA (κάνοντας χρήση της μεθοδολογίας κανονικοποίησης θέσης SymPan) κατά μέσο όρο 1%, προσφέροντας έτσι μια συνολική αύξηση της τάξης του 7% σε σχέση με την αρχική προσέγγιση. Και οι δύο μέθοδοι παρουσιάζουν βελτιωμένη απόδοση σε σχέση με την μεθοδολογία ROSY κατά μέσο όρο 2 - 3%.

Και οι τρεις προτεινόμενες μέθοδοι είναι σε θέση να παράγουν υψηλής ποιότητας ευθυγραμμίσεις των τριδιάστατων αντικειμένων, ανεξαρτήτως κλάσης προέλευσης ή μορφολογίας. Οι ευθυγραμμίσεις αυτές είναι ταυτόχρονα σταθερές και συνεπείς.

Για την αντιμετώπιση του προβλήματος της ανάκτησης εύκαμπτων τριδιάστατων αντικειμένων, προτάθηκε ο περιγραφέας ConTopo++. Αυτός ο περιγραφέας ενσωματώνει τόσο γεωμετρικά όσο και τοπολογικά χαρακτηριστικά σε μια ενιαία διαδικασία εξαγωγής περιγραφέα. Επιπλέον, εισάγεται μια τεχνική συνταιριασμού γράφων, βασιζόμενη στην αντιστοίχιση σειρών. Αυτή η βελτιωμένη μεθοδολογία ανάκτησης τριδιάστατων αντικειμένων, αποτιμήθηκε όχι μόνο σε σχέση με την βασική μορφή του αλγορίθμου (ConTopo) και τα αντίστοιχα σύνολα δεδομένων που εμφανίζονται στο [102] αλλά και σε σχέση με τυποποιημένα σύνολα δεδομένων από τα SHREC'10 *Non-rigid 3D Models* και SHREC'11 *Non-rigid 3D Watertight Meshes* και τους αντίστοιχους τριδιάστατους περιγραφείς σχήματος τρέχουσας τεχνολογικής στάθμης. Σε κάθε περίπτωση, η προτεινόμενη μεθοδολογία ανάκτησης για εύκαμπτα τριδιάστατα αντικείμενα είναι σε θέση να επιτύχει υψηλά επίπεδα ακρίβειας ανάκτησης, ξεπερνώντας πολλούς από τους ανταγωνιστικούς περιγραφείς, με χαμηλό υπολογιστικό κόστος. Το Σχήμα 48 απεικονίζει μερικά παραδείγματα ανάκτησης από το σύνολο δεδομένων SHREC'10 *Non-rigid 3D Models*.

Στον τομέα της ανάκτησης τριδιάστατων αντικειμένων με βάση εικόνες-ερώτησης, προτείναμε μια στρατηγική χωρικών ιστογραμμάτων σε ένα Bag-of-Visual-Words πλαίσιο, η οποία προσαρμόζει την πληροφορία των πανοραμικών προβολών στο έργο της μερικής ταύτισης. Ιδιαίτερη προσοχή έχει δοθεί στο στάδιο προεπεξεργασίας των εικόνων-ερώτησης, όπου μια σειρά από διαδοχικά φίλτρα εφαρμόζονται στις εικόνες, προκειμένου να αμβλυνθούν τα προβλήματα που μπορεί να έχει εισάγει η διαδικασία ψηφιοποίησης. Αυτή η βελτιωμένη μεθοδολογία ανάκτησης τριδιάστατων αντικειμένων, εκτιμήθηκε όχι μόνο σε σχέση με την

προηγούμενη προσέγγισή μας [103] και το αντίστοιχο σύνολο δεδομένων SHREC'10 *Range Scan Retrieval*, αλλά και σε τυποποιημένα σύνολα δεδομένων από το SHREC'09 *Querying with Partial Models* με τις αντίστοιχες μεθοδολογίες ανάκτησης τριδιάστατων αντικειμένων, και το SHREC'11 *Shape Retrieval Contest of Range Scans*. Σε κάθε περίπτωση, η προτεινόμενη μέθοδος ανάκτησης 3Δ αντικειμένου υπερτερεί των ανταγωνιστικών περιγραφέων.

Η παρούσα διατριβή αντιμετωπίζει με επιτυχία μια σειρά από προβλήματα του τομέα της ανάκτησης τριδιάστατων αντικειμένων. Η ανάκτηση τριδιάστατων αντικειμένων μπορεί να θεωρηθεί ως ένα σχετικά νέο πεδίο που εμπίπτει στις επιστημονικές περιοχές της υπολογιστικής όρασης, των γραφικών υπολογιστή και της ανάκτησης πληροφορίας με βάση το περιεχόμενο. Οι περιγραφόμενες μέθοδοι έχουν αποδειχθεί εύρωστοι όσον αφορά την ακρίβεια ανάκτησης και ξεπέρασαν προηγούμενες μεθόδους τρέχουσας τεχνολογικής στάθμης στις αντίστοιχες δοκιμές αξιολόγησης. Αυτές οι δοκιμές διεξήχθησαν σε δημόσια διαθέσιμες βάσεις δεδομένων.

Preface

Although the pursuit for knowledge is a laborious and long-winded process, its results are of utmost value and can have the deepest effects on oneself. For giving me the opportunity not only to understand but to attempt and create new knowledge in the field of Computer Science, I would like to express my gratitude to my mentors, Prof. Theoharis Theoharis (Department of Informatics and Telecommunications, University of Athens) and Assistant Prof. Ioannis Pratikakis (Department of Electrical and Computer Engineering, Democritus University of Thrace). Their trust and unlimited support date long before the beginning of my pursuit for the doctorate degree and have been vital for its completion.

My thanks also go to my co-advisor, Dr. Stavros Perantonis (Research Director, NCSR Demokritos) for the scientific advice he has offered me, and also for his belief in me for many years, even since when I was still in pursuit of my B.Sc. and M.Sc. degrees.

I would also like to thank the members of the examination committee for honoring me by their participation in the scientific evaluation of this dissertation.

I am also grateful to all my colleagues and friends at the Computer Graphics Laboratory of the Informatics and Telecommunications Department of the University of Athens, for listening my problems and sharing their thoughts. Their collaboration is invaluable. I would like to especially thank Dr. Panagiotis Perakis for his help, especially at the beginning of this study.

My parents, Diamantis and Syrmatoula, receive my deepest gratitude for their love and for caring and supporting me from the first moment of my life. Everything I have achieved

so far is because they have provided me with all the necessary means.

Finally, I would like to dedicate this dissertation to my companion, Vicky and my friend Maria, who stood with me and encouraged me to pursue my dreams. They always had a solution to my problems (related and unrelated to this thesis), allowing me to see further than I thought I could.

Konstantinos A. Sfikas

– Athens, January 2014

1 Introduction

Well begun is half done

– Aristotle

Information, commonly refers to a useful portion of data located among a collection of related entities. Recent advances in storage technologies and the widespread use of the Internet, have resulted in a vast increase of the amount of data stored in and distributed from large databases. Any attempt for manual annotation and information extraction is almost impossible, therefore rendering the need for an automated procedure, mandatory.

The process of extracting useful information from large amounts of data, in an automated manner and based on an example or descriptive query, is called information retrieval. Common types of information that can benefit from such a retrieval process are: textual, visual, audio and video data and most recently, 3D and 4D (3D over time) data.

In recent years, through the creation of inexpensive 3D scanners and the simplification of 3D modelling software, a large volume of 3D data has been created and stored in corresponding scientific and industrial/commercial repositories. Furthermore, 3D data can be processed in various, application dependent, ways and occasionally be combined with data of other types and modalities (e.g. textual annotation and/or thumbnails of 3D models). These data types can further be used as queries for the retrieval of 3D objects.

Some example applications that exploit the properties of 3D models and could greatly benefit from a retrieval process follow: in medicine large diagnostic 3D data are compared and researched in order to assist the process of making medical decisions. In biometrics a person's 3D facial model is searched over corresponding databases for identification

purposes. Game development utilizes retrieval and reusability of 3D models in order to minimize production times and reduce the size of the final product. Other example application areas include engineering and archaeology. It can therefore be easily deduced, that 3D object retrieval is a key process, although in general it is complex and highly depended on the application.

1.1 Framework and problem statement

3D object retrieval applications can be classified into two major categories: inter-class and intra-class retrieval. Inter-class retrieval focuses on a generic domain of 3D objects and aims at finding the closest match among a set of 3D models that belong to a broad range of different classes. In this case, there is usually no prior knowledge regarding the characteristics or the nature of the 3D objects. Intra-class retrieval targets a specific 3D object domain (e.g. 3D faces, non-rigid 3D models, human action models, engineering models etc), where a match is sought between 3D models that belong to the same class but have their special characteristics defined differently. Intra-class 3D object retrieval methods usually exploit domain knowledge and shape characteristics of the 3D models, in order to attain higher performance.

For both categories, the generic framework of a 3D object retrieval system can be outlined as follows: preprocessing, pose normalization, shape descriptor extraction, feature matching.

Initially, let us define that, if not otherwise stated, a triangular 3D polygonal object is composed of a set of vertices $v_i = [v_{ix}, v_{iy}, v_{iz}]^T, i \in \{1, 2, ..., N_v\}$, where N_v denotes the total number of vertices, and a set of triangular faces $f_j = \{v_{j1}, v_{j2}, v_{j3}\}$, where $j = 1, 2, ..., N_p$ and N_p denotes the total number of triangular faces.

At the first step of the 3D object retrieval pipeline, 3D models are preprocessed. In this step, the 3D models are cleaned up of any inconsistencies present due to the digitization process, i.e. double or reversed faces, structural gaps, etc. This step is highly dependent on the method/equipment used for the creation of the 3D models and may differ greatly from one application to another.

After basic preprocessing, *Pose Normalization* ensures that the geometric properties of the 3D models are defined in a uniform manner. The diversity of 3D object acquisition

sources implies that 3D objects which may even be part of the same dataset, have their geometrical properties arbitrarily defined. Therefore, before any kind of processing is carried out, it must be ensured that the 3D objects have been normalized in terms of position, scaling and rotation (Fig 1 shows an example rotation normalization). Pose normalization of 3D objects is a common preprocessing step in various computer graphics applications [21, 110, 121, 141]. Visualization, broken fragment reconstruction, biometrics and 3D object retrieval are only a few examples of applications that benefit from a pose normalization procedure. To achieve pose normalization, for every 3D object, a corresponding set of normalization transformations in 3D space must be defined.

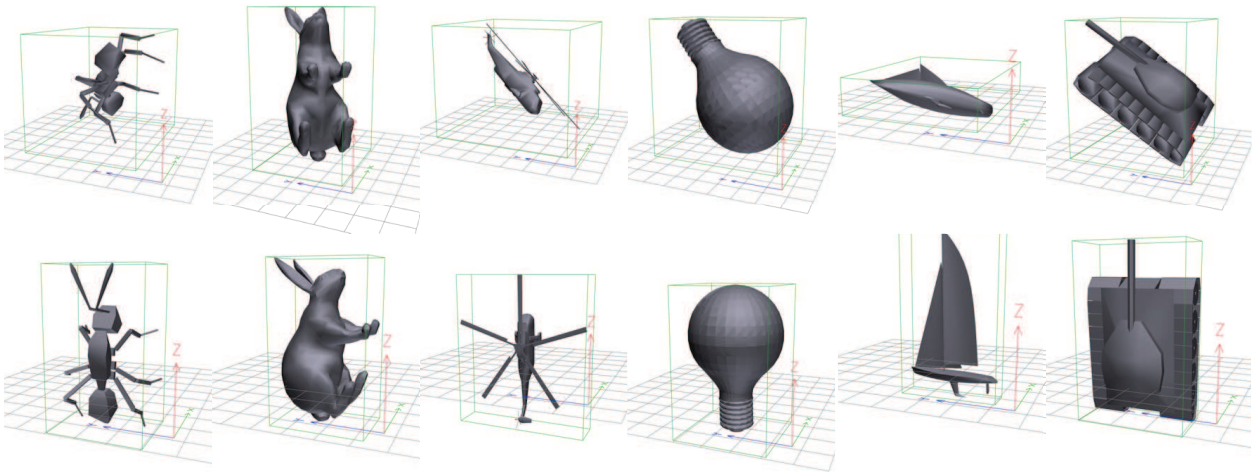


Figure 1: Examples of non-aligned objects (top-row) and the corresponding rotation normalizations (bottom row).

The main step of a 3D object retrieval system is the computation of a feature set for each 3D model. In this step, the structural and/or other special characteristics of a 3D object are modelled and a shape descriptor that faithfully encodes the shape of the 3D model, in an efficient manner, is created. Feature selection is tightly connected to the corresponding application and can vary greatly for each 3D object retrieval system (e.g. intra-class retrieval exploits features that are more distinguishing within a specific domain, whereas inter-class retrieval uses more generic characteristics).

Finally, each 3D object's shape descriptor is used as a signature during the matching procedure. At this step, the signatures of the 3D models, stored in the database, are compared to the corresponding signatures of the query 3D model(s), using a specified

metric. The selected metric is also dependent on both the features selected and the corresponding application. Finally, the response of the 3D object retrieval system is the set of 3D object(s) that correspond to the closest match(es) of the given user query.

1.2 Contributions

This thesis has made the following research contributions in the area of 3D object retrieval: (i-iii) three new 3D model pose normalization methods, (iv) a non-rigid 3D object retrieval methodology and (v) a 3D object retrieval algorithm, based on range image queries. In detail, the contributions of this dissertation are the following:

i ROSy: A general purpose global pose normalization method based on 3D object reflective symmetry. Initially, the axis-aligned minimum bounding box of a rigid 3D model is modified by requiring that the 3D model is also in minimum angular difference with respect to the normals to the faces of its bounding box. To estimate the modified axis-aligned bounding box, a set of predefined planes of symmetry are used and a combined spatial and angular distance, between the 3D model and its symmetric model, is calculated. By minimizing the combined distance, the 3D model fits inside its modified axis-aligned bounding box and alignment with the coordinate system is achieved [105].

ii SymPan: A pose normalization method, based on panoramic views and reflective symmetry, is presented. Initially, the surface of a 3D model is projected onto the lateral surface of a circumscribed cylinder, aligned with the primary principal axis of space. Based on this cylindrical projection, a normals' deviation map is extracted and using an octree-based search strategy, the rotation which optimally aligns the primary principal axis of the 3D model and the cylinder's axis is computed. The 3D model's secondary principal axis is then aligned with the secondary principal axis of space in a similar manner [104].

iii SymPan+: A pose normalization method based on similar principles as the SymPan methodology. Qualitative and experimental investigation in 3D data-sets has led us to the observation that most objects possess a single plane of symmetry. Our approach

is thus guided by this observation. Initially, through an iterative procedure, the symmetry plane of a 3D model is estimated, thus computing the first axis of the model. This is achieved by rotating the 3D model and computing reflective symmetry scores on panoramic view images. The other principal axes of the 3D model are estimated by computing the variance of the 3D model's panoramic views [108].

iv ConTopo: Combining the properties of conformal geometry and graph-based topological information, a non-rigid 3D object retrieval methodology is proposed, which is both robust and efficient in terms of retrieval accuracy and computation speed. While graph-based methods are robust to nonrigid object deformations, they require intensive computation which can be reduced by the use of appropriate representations, addressed through geometry-based methods. In this respect, a 3D object retrieval methodology, which combines the above advantages in a unified manner, is presented. Furthermore, a string matching strategy for the comparison of graphs which describe 3D objects, is proposed [106].

v A 3D object retrieval method, based on range image queries that represent partial views of real 3D objects, is presented. The complete 3D models of the database are described by a set of panoramic views and a Bag-of-Visual-Words model is built using SIFT features extracted from them. To address the problem of partial matching, a histogram computation scheme, on the panoramic views, that represents local information by taking into account spatial context, is suggested. Furthermore, a number of optimization techniques are applied throughout the process, for enhancing the retrieval performance [107].

1.3 Overview of this thesis

The organization of this thesis is as follows:

In **Chapter** 2, state-of-the-art methods in the different fields of 3D object retrieval are reviewed. More specifically, recent and noteworthy works on rigid and non-rigid 3D shape descriptors, pose normalization and 3D object retrieval based on 2D image queries, are presented.

In **Chapter** 3, the proposed pose normalization methods, as well as corresponding key concepts used, are discussed in detail.

In **Chapter** 4, the proposed non-rigid 3D object retrieval method is presented. Its basic components, namely, conformal mappings and topological graphs, along with the proposed string-based graph matching technique, are discussed in detail.

In **Chapter** 5, the proposed methodology for 3D object retrieval based on 2D range image queries, is presented. Key concepts, such as Panoramic Views, SIFT, Bag-of-Visual-Words modelling and Spatial Histograms are discussed.

Chapter 6 presents the evaluation methodology and illustrates the extended experimentation results of the methods presented in this dissertation, against state-of-the-art works on standard datasets. These results are discussed in detail.

Finally, in **Chapter** 7 the contributions of this dissertation are summarized and conclusions that were drawn are discussed.

1.4 List of Publications

Work from this dissertation has appeared in the following co-authored publications (Citations are according to the "APA" style):

Journals

- [105] **Sfikas, K.**, Theoharis, T., & Pratikakis, I. (2011). ROSy+: 3D object pose normalization based on PCA and reflective object symmetry with application in 3D object retrieval. *International Journal of Computer Vision*, vol. 91, no. 3, pp. 262-279.

- [106] **Sfikas, K.**, Theoharis, T., & Pratikakis, I. (2012). Non-rigid 3D object retrieval using topological information guided by conformal factors. *The Visual Computer*, vol. 28, no. 9, pp. 943-955.

- [107] **Sfikas, K.**, Theoharis, T., & Pratikakis, I. (2013). 3D Object Retrieval via Range Image Queries in a Bag-of-Visual-Words Context. *The Visual Computer*, vol. 29, no. 12, pp. 1351-1361.

- [108] **Sfikas, K.**, Theoharis, T., & Pratikakis, I. (2014). Pose Normalization of 3D Models via Reflective Symmetry on Panoramic Views. *The Visual Computer*, **submitted to**.

Conferences

- [102] **Sfikas, K.**, Pratikakis, I., & Theoharis, T. (2011). ConTopo: Non-Rigid 3D Object Retrieval using Topological Information guided by Conformal Factors. in *3DOR*, H. Laga, T. Schreck, A. Ferreira, A. Godil, I. Pratikakis, and R. C. Veltkamp, Eds. Eurographics Association, pp. 25-32.

- [103] **Sfikas, K.**, Pratikakis, I., & Theoharis, T. (2012). 3D object retrieval via range image queries based on SIFT descriptors on panoramic views. in *3DOR*, M. Spagnuolo, M. M. Bronstein, A. M. Bronstein, and A. Ferreira, Eds. Eurographics Association, pp. 9-15.

- [104] **Sfikas, K.**, Pratikakis, I., & Theoharis, T. (2013). SymPan: 3D Model Pose Normalization via Panoramic Views and Reflective Symmetry. in *3DOR*, U. Castellani, T. Schreck, S. Biasotti, I. Pratikakis, A. Godil, and R. C. Veltkamp, Eds. Eurographics Association, pp. 41-48.

2 Related Work

If I have seen further
it is by standing on the shoulders of giants

– Isaac Newton,
letter to Robert Hooke,
1676

A common top-level categorization of 3D object retrieval methods is based on the structure of the 3D models that they use: rigid vs non-rigid. This categorization has prevailed, mainly due to the different processing required to search over each domain. The former case has shown many significant works in recent years, while the latter one is evolving rapidly and constantly important novelties are being presented. We also survey a relatively new category of 3D object retrieval which is based on the use of 2D (instead of 3D) queries. Finally, we deal separately with pose normalization techniques.

2.1 Rigid 3D Object Retrieval

Content-based rigid 3D object retrieval methods can be classified into three major categories according to the spatial dimensionality of the information used, namely 2D, 3D and their combination. According to this categorization, a brief overview of the related work in the area of rigid 3D shape descriptors for generic 3D object retrieval is presented.

2.1.1 2D-based 3D Object Retrieval

In the first category, shape descriptors are computed on projection images, which can be contours, silhouettes, depth buffers and/or other types of 2D representations. Therefore, similarity is measured using 2D matching techniques.

One of the most acknowledged methods for 3D object retrieval, based on the extraction of features from 2D representations of the 3D objects, was the Light Field descriptor, proposed by Chen et al. [28]. This descriptor is comprised of Zernike moments and Fourier coefficients computed on a set of projections taken at the vertices of a dodecahedron.

Lian et al. [72] proposed an enhancement to the Light Field descriptor, by computing the same features on projections taken from the vertices of geodesic spheres (triangulated solid approximating the real sphere) generated by the regular unit octahedron.

Vranic [131] proposed a shape descriptor where features are extracted from depth buffers produced by six projections of the object, one for each side of a cube which encloses the object. In the same work, the Silhouette-based descriptor is proposed which uses the silhouettes produced by the three projections taken from the Cartesian planes.

Zarpalas et al. [142] introduced a 3D shape descriptor called the spherical trace transform, which is the generalization of the 2D trace transform proposed by Kadyrov and Petrou [62]. In this method, a variety of 2D features are computed for a set of planes intersecting the volume of a 3D object.

In [91], Passalis et al. proposed PTK, a depth buffer based descriptor which uses parallel projections to capture the object's thickness and an alignment scheme that is based on symmetry. The PTK descriptor encompasses symmetry, eigenvalue-related weighting and an object thickness related measure.

Shih et al. [109] proposed the elevation descriptor, where six depth buffers (elevations) are computed from the faces of the 3D model's bounding box and each buffer is described by a set of concentric circular areas that give the sum of pixel values within the corresponding areas.

Ohbuchi et al. [84] proposed the Multiple Orientation Depth Fourier Transform (MODFT) descriptor where the model is projected from 42 viewpoints to cover all possible view aspects. Each depth buffer is then transformed to the $r-\theta$ domain and the Fourier transform is applied. To compare two models, all possible pairs of coefficients are compared which

inevitably increases comparison time.

A newly proposed method is the depth line descriptor proposed by Chaouch and Verroust - Blondet [27, 25] where a 3D object is projected to the faces of its bounding box giving six depth buffers. Each depth buffer is then decomposed into a set of horizontal and vertical depth lines that are converted to state sequences which describe the change in depth at neighboring pixels.

Papadakis et al. in [89] proposed PANORAMA, a 3D shape descriptor that uses a set of panoramic views of a 3D object which describe the position and orientation of the object's surface in 3D space. For each view the corresponding 2D Discrete Fourier Transform and the 2D Discrete Wavelet Transform are computed.

The method proposed by Alizadeh et al. [5] uses 60 different 2D silhouettes, which are automatically extracted from different view-angles of 3D models. Solving the Poisson equation [53] for each Silhouette assigns a number to each pixel as the pixel's signature. Counting and accumulating these pixel signatures generates a histogram-based signature for each silhouette.

2.1.2 3D-based 3D Object Retrieval

In the second major category of 3D object retrieval techniques, shape descriptors are extracted from 3D shape representations. A set of subcategories can be identified here, namely, statistical, graph-based and spherical function based descriptors.

In the shape histogram descriptor proposed by Ankerst et al. [7], 3D space is divided into concentric shells, sectors, or both and for each part, the object's shape distribution is computed giving a sum of histograms bins.

The shape distributions descriptor proposed by Osada et al. [86] measures a set of shape characteristics for a random set of points belonging to the object, using appropriate shape functions, e.g. the D2 function which measures the distance between two random surface points.

Zaharia and Petreux [140] presented the 3D shape spectrum descriptor which is the histogram that describes the angular representation of the first and second principal curvature along the surface of the 3D object.

Xiang et al. in [135] propose a rigid transformation insensitive descriptor, called the

Poisson shape histogram descriptor, extracted by a voxelized representation of the 3D objects.

In [137] a two-step descriptor called Sorted Extended Gaussian Image (SEGI) is presented. Based on Extended Gaussian Image and Shell histograms, SEGI initially performs approximate 3D object retrieval based on the sorted histogram bins and them refines the results by recording the relations between the bins.

Ohbuchi et al. [82] proposed enhanced shape functions, namely the (absolute) angle distance histogram for inconsistently oriented meshes, which are extensions of the D2 shape distribution.

Hilaga et al. [57] introduced the multi-resolution Reeb graph, which represents a 3D object's topology and skeletal structure at various levels of detail.

Yu et al. [136] used spherical functions to describe the topology and concavity of the surface of a 3D object and the amount of effort required to transform it to its bounding sphere.

In Zhang et al. [144] consider the use of medial surfaces to compute an equivalent directed acyclic graph of an object.

In the work of Sundar et al. [119], the 3D object passes through a thinning process producing a set of skeletal points, which finally form a directed acyclic graph by applying the minimum spanning tree algorithm.

Cornea et al. [32] propose the use of curve skeletons produced by the application of the generalized distance field to the volume of the 3D object and similarity is measured using the earth mover's distance [97].

The P3DS descriptor developed by Kim et al. [67] uses an attributed relational graph whose nodes correspond to parts of the object that are represented using ellipsoids and the similarity is computed by employing the earth mover's distance.

In [66, 64] Kazhdan et al. proposed planar reflective symmetry descriptor (PRSD), a collection of spherical functions that describes the measure of a model's rotational and reflective symmetry with respect to every axis passing through the center of mass. This descriptor is used for determining the existence or not of a 3D object's principal symmetry axes. Extending this work to every possible plane Podolak et al. presented the planar reflective symmetry transformation (PRST) in [94].

2.1.3 2D/3D Hybrid 3D Object Retrieval

Besides the previous categories, combinations of different methods have been considered in order to enhance the overall performance.

Vranic [131] proposed the Ray-based descriptor which characterizes a 3D object by a spherical extent function capturing the furthest intersection points of the object's surface with rays emanating from the origin. Spherical harmonics or moments can be used to represent the spherical extent function. A generalization of the previous approach uses several spherical extent functions of different radii.

The GEDT descriptor proposed by Kazhdan et al. [65] is a volumetric representation of the Gaussian Euclidean Distance Transform of a 3D object, expressed by norms of spherical harmonic frequencies.

In Papadakis et al. [87], the CRSP descriptor was proposed which uses the Continuous PCA (CPCA) along with Normals PCA (NPCA) to alleviate the rotation invariance problem and describes a 3D object using a volumetric spherical-function based representation expressed by spherical harmonics.

Generalizing from 2D to 3D, Novotni and Klein [81] presented the 3D Zernike descriptor, Daras et al. [35] introduced the generalized Radon transform and Ricard et al. [96] developed the 3D ART descriptor by generalizing the 2D angular radial transform. Zaharia and Preteux [139] proposed the C3DHTD descriptor by generalizing the 2D Hough Transform.

The approach of Bustos et al. [21] assumes that the classification of a particular dataset is given, in order to estimate the expected performance of the individual shape descriptors for the submitted query and automatically weigh the contribution of each method. However, in the general case, the classification of a 3D model dataset is not fixed since the content of a 3D model dataset is not static.

In the context of partial shape matching, Funkhouser and Shilane [46] use the predicted distinction performance of a set of descriptors based on a preceding training stage and perform a priority driven search in the space of feature correspondences to determine the best match of features between a pair of models. The disadvantages of this approach is its time complexity which is prohibitive for online interaction as well as the storage requirements for the descriptors of all the models in the database.

Based on the idea of combining features obtained from 2D and 3D representations, Song and Golshani [115] developed a descriptor that described an object by obtaining a set of orthogonal projections from different viewpoints and by measuring the curvature of the object's surface.

Vranic [132] developed a hybrid descriptor called DESIRE, that consists of the Silhouette, Ray and Depth buffer based descriptors, which are combined linearly by fixed weights.

Papadakis et al. [88] proposed a hybrid descriptor formed by combining features extracted from a depth-buffer and spherical function based representation, with enhanced translation and rotation invariance properties. The advantage of this method over similar approaches is the top discriminative power along with minimum space and time requirements.

2.1.4 Rigid 3D Object Retrieval Databases

In recent years, aiming at prototyping the research and to provide tools for the experimentation and the comparison of the corresponding results, an active set of rigid 3D object databases has been created. In this chapter we will consider the most prevalent of these databases, which have been used to conduct experiments by research groups and form a commonly accepted benchmark.

In [110], the authors describe a widespread rigid 3D object database, known as the Princeton Shape Benchmark (PSB). The PSB is publicly available and consisted of 3D models, software tools and a standardized set of benchmarks for comparison between 3D model retrieval algorithms. The dataset contains 1,814 polygonal models collected from the World Wide Web and which have been categorized by humans depending on their function and form. Included is also a set from hierarchical classifications, discrete sets for training and testing, metadata for each object and a suite of software tools for testing, analysis and visualization of retrieval results. Fig. 2 shows examples of models from the PSB dataset.

In [138] the NTU 3D object database is presented. This database comprises a benchmark dataset for 3D object retrieval methods, consisted of 1,833 3D models. The NTU dataset includes 3D models from 3DCafe, retrieved in December 2001. For the categoriza-

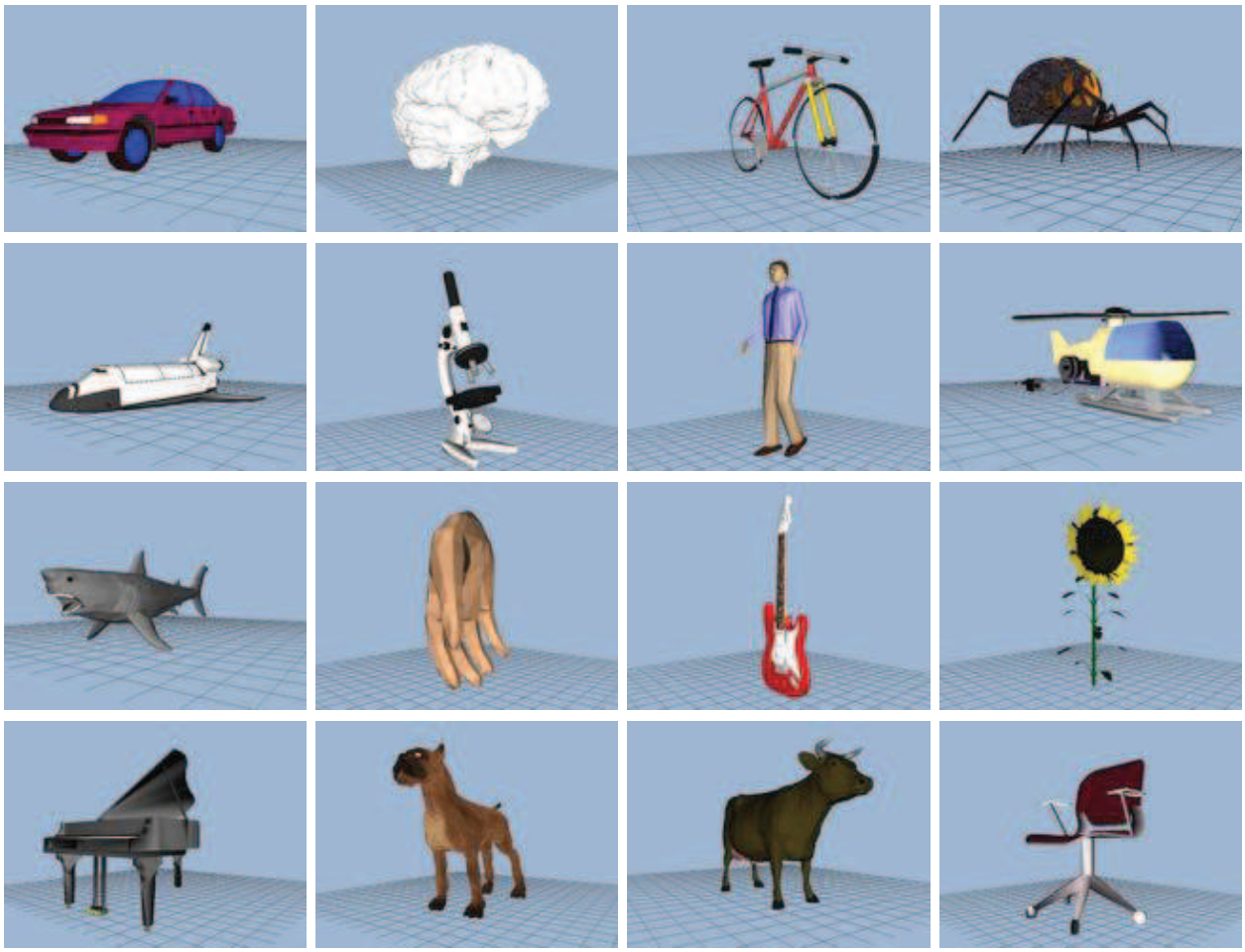


Figure 2: 3D model examples from the PSB dataset.

tion of the models a user was used and was based on the functionality of the objects. The dataset has been divided into 47 classes, including 549 3D models, mainly vehicles and household items (e.g., airplanes, cars, chairs, tables, etc.), and 1,289 additional models, which were categorized as 'Miscellaneous'.

In [144], the McGill database is presented. The specific database includes 3D models produced by 3D models of the PSB dataset. The dataset includes a total of 320 samples that have been created by several primary objects (hands, humans, tables, chairs, glasses, airplanes, birds, dolphins, etc.). Some of these models are illustrated in Fig. 3. This database provides categorizations classified as both rigid and non-rigid.

The ESB database is presented in [60]. This database is consisted of 3D models derived from a variety of sources, including the National Design Repository and industry. Furthermore, this specific database includes many CAD models, which have been created by students at undergraduate design course at the University of Purdue. The authors

Figure 3: Sample 3D models from the McGill dataset. Illustrated are models that exhibit significant articulation, as well as moderate or no articulation.

showed spacial attention so that the specific database includes only common industrial objects that are entirely consistent with the corresponding distribution of industrial objects of the real world. The particular dataset consists of 801 models, divided into 42 classes of similar parts, as Discs, T-shaped parts and Bracket-like parts, and 66 models divided into three classes of other common shapes.

Finally, a special reference must be made for the SHREC databases, which set a

reference point for many 3D object retrieval methods and become annually updated with new data. The original dataset of SHREC [43], was created using two human evaluators that carried out the classification. In the case of disagreement, a third evaluator was used for resolving. This categorization was carried out based solely on geometric and topological similarity between the 3D models. Each class consists of 20 models and the complete database includes 40 classes of objects commonly encountered in everyday life.

For all the aforementioned databases, Table 1 show a summary of the categories and the number of models they are consisted of.

Table 1: Categories and cardinalities of the rigid 3D object benchmark databases

3D object database	# of Categories	# of objects
PSB training	90	907
PSB test	92	907
NTU	47	549
McGill Articulated	10	254
McGill Non Articulated	9	202
ESB	45	867
SHREC(NIST)	40	800

2.2 Non-Rigid 3D Object Retrieval

Over the past few years, a lot of research effort has successfully addressed the problem of rigid 3D model retrieval, by exploiting inter-class variability. However, in the case of intra-class variability, non-rigid 3D shape descriptors are more effective since the objects of the class can assume a variety of transformations, including non-rigid deformations.

Here, non-rigid 3D object retrieval methods are categorized, based on two common classes of features, which have proven to be very discriminative for the specific domain: geometric and topological. A third category is formed by methods exploiting both types of features.

2.2.1 Methods based on Geometric Features

A large number of methods are based on the discrete Laplace-Beltrami operator.

Reuter at al. [95] compare two triangulated surfaces by computing the distance between two isometry-invariant feature vectors, given by the first *n* eigenvalues (the spectra) of the Laplace-Beltrami operator.

Similarly, Rustamov [100] uses the eigenvectors of the Laplace-Beltrami operator.

Zaharia and Preteux [140] presented the 3D shape spectrum descriptor which is the histogram that describes the angular representation of the first and second principal curvature of the 3D object surface.

Xiang et al. [135] use the histogram of the solution to the volumetric Poisson equation ($\nabla^2 U = -1$) as a pose invariant shape descriptor.

In a similar manner, Ben-Chen and Gotsman [11], create a descriptor that maps the local curvature characteristics of the 3D object. The histogram of the solution to the conformal factor equation $\nabla^2 \phi$ is used as the descriptor. The same principles were also used by Wang et al. [134] for face recognition using 3D conformal maps.

Some non-rigid shape descriptors are derived from geodesic distances on the mesh, which are invariant to isometric transformations.

Elad and Kimmel [42] proposed a canonical representation for triangulated surfaces: a surface in $\mathbb{R}^3$ is transformed into canonical coordinates in the Euclidean space $\mathbb{R}^m$ by applying multi-dimensional scaling. In this canonical representation the geodesic distances on the original surface are approximated by the corresponding Euclidean distances. The matching problem of non-rigid and deformed objects is then reduced to the problem of matching rigid objects embedded in $\mathbb{R}^m$, which can be approached with well-known algorithms.

Jain and Zhang [58] compare non-rigid objects by matching spectral embeddings that are derived from the eigenvectors of affinity matrices, computed by considering geodesic distances.

A shape descriptor for non-rigid 3D objects based on histograms of surface functions is presented by Gal et al. [48], where two scalar functions are used on the mesh, the *local-diameter* function which measures the diameter of the 3D shape in the neighborhood of each vertex and the *centricity* function which measures the average geodesic distance from a vertex to all other vertices on the mesh.

Carlsson et al. [22] compared barcode descriptors of point clouds computed by using

the persistence homology theory.

Dey et al. [37] compared noisy point clouds, by matching signatures extracted from segmented parts of the point sets by making use of Morse theory.

In [75], Mademlis et al. proposed a novel shape descriptor based on the impact that the 3D objects have when they are exposed to a specific type of force field (i.e. the Newtonian or the Coulombian field). The 3D objects are initially voxelized and subsequently the histograms of the field factors are compared.

2.2.2 Methods based on Topological Features

Another category of non-rigid 3D object descriptors comprises methods that use graphs for object representation.

Siddiqi et al. [114] based on a singularity theory, derive skeletal descriptors from shape models. These descriptors (shocks) are organized into directed acyclic graphs.

Based on the same idea of shock graphs Sebastian et al. [101] match two shapes by measuring the effort needed for the deformation of one shape into the other.

Zhang et al. [144] consider the use of medial surfaces to compute an equivalent directed acyclic graph of an object.

In the work of Sundar et al. [119], the 3D object passes through a thinning process producing a set of skeletal points, which form a directed acyclic graph by applying the minimum spanning tree algorithm.

The P3DS descriptor developed by Kim et al. [67] uses an attributed relational graph whose nodes correspond to parts of the object that are represented using ellipsoids and the similarity is computed by employing the earth mover's distance.

Hilaga et al. [57] presented a technique to match the topology of triangulated models, by comparing Multiresolution Reeb Graphs (MRGs). Their algorithm for matching two MRGs is a coarse-to-fine strategy, which searches the node pairs providing the largest value of similarity while maintaining topological consistency. Similarly to [57], Hamza and Krim [55] considered a discrete approximation of the global squared geodesic distance function. The dissimilarity between two objects was calculated by computing the Jensen-Shannon divergence between the corresponding statistical shape descriptors.

Tung and Schmitt [127] used geodesic distances for building an MRG and merge the

graph geometrical and visual information to improve matching and calculation of shape similarity between models.

2.2.3 Hybrid Methods

In recent research works, Tierny et al. [125] compare 3D models by extracting partial signatures from disk or annulus-like charts using Reeb graph topology.

Bronstein et al. [20] instead of using geodesic distances for extracting shape signatures, they exploit the properties of diffusion distances within the Gromov-Hausdorff framework.

Agathos et al. [2] create a one-level Attributed Relation Graph (ARG) of 3D objects by using a segmentation algorithm. Furthermore, the authors propose a graph matching algorithm based on the Earth Mover's Distance (EMD). The attributes applied on the graph are described in [88] and [67].

Biasotti et al. [14] described a method for the characterization of shapes by using a set of patches, which are automatically tiled and stitched, in order to approximate the original shape. Reeb graphs are used for the definition of the main shape features that drive the approximation process.

2.2.4 Non-Rigid 3D Object Retrieval Databases

Similar to the rigid 3D object retrieval case, non-rigid 3D object retrieval databases have also grown both in number and data size during the last few years. The benchmark datasets, as well as the testing tools are common to the corresponding rigid ones, since most of them derived from the latter. In this chapter we will consider the most prevalent of the non-rigid 3D object retrieval benchmark datasets.

The TOSCA dataset, publicly available, contains three-dimensional non-rigid shapes in a variety of poses for non-rigid shape similarity and correspondence experiments. The database contains a total of 148 objects, including 9 cats, 11 dogs, 3 wolves, 17 horses, 15 lions, 21 gorillas, 1 shark, 24 female figures, and two different male figures, containing 15 and 20 poses. The database also contains 6 centaurs, and 6 seahorses for partial similarity experiments. Each object contains approximately 3000 vertices. Fig 4 depicts some of the 3D models included in the dataset.

Figure 4: Sample models of various classes and poses from the TOSCA collection.

The Aim@Shape Watertight database, presented in has been used for various retrieval contests. This dataset is composed of 400 watertight 3D models, divided to 20 classes, each containing 20 3D models. The Aim@Shape database has been also known as the SHREC'07 *Watertight Models Track dataset* [50], one of the first SHREC tracks, aimed specifically for the retrieval of watertight 3D models.

The McGill Articulated Shape Benchmark database [144] consists of 255 non-rigid 3D models which are classified into 10 categories. The maximum number of the objects in a class is 31, while the minimum number is 20. Based on this dataset, the SHREC'10 *Non-rigid 3D Shape Retrieval database* [71] has been created. More specifically, 200 models are selected (or modified) to generate the test database ensuring that every class contains equal number of models.

Following the same spirit, the SHREC'11 *Non-rigid 3D Watertight Meshes* [70] has been one of the most recent benchmark databases, aimed at non-rigid 3D object retrieval. This large-scale database consists of 600 watertight triangle meshes that are derived from 30 original models, among which 25 objects are collected from several freely accessible repositories (e.g. PSB database, McGill database, TOSCA shapes etc.), while the remaining four models (i.e. *lamp, paper, scissor,* and *twoballs* are created using *Autodesk 3D studio MAX*. The 600 non-rigid models of the dataset have been equally classified into 30 categories

(see Fig. 5).

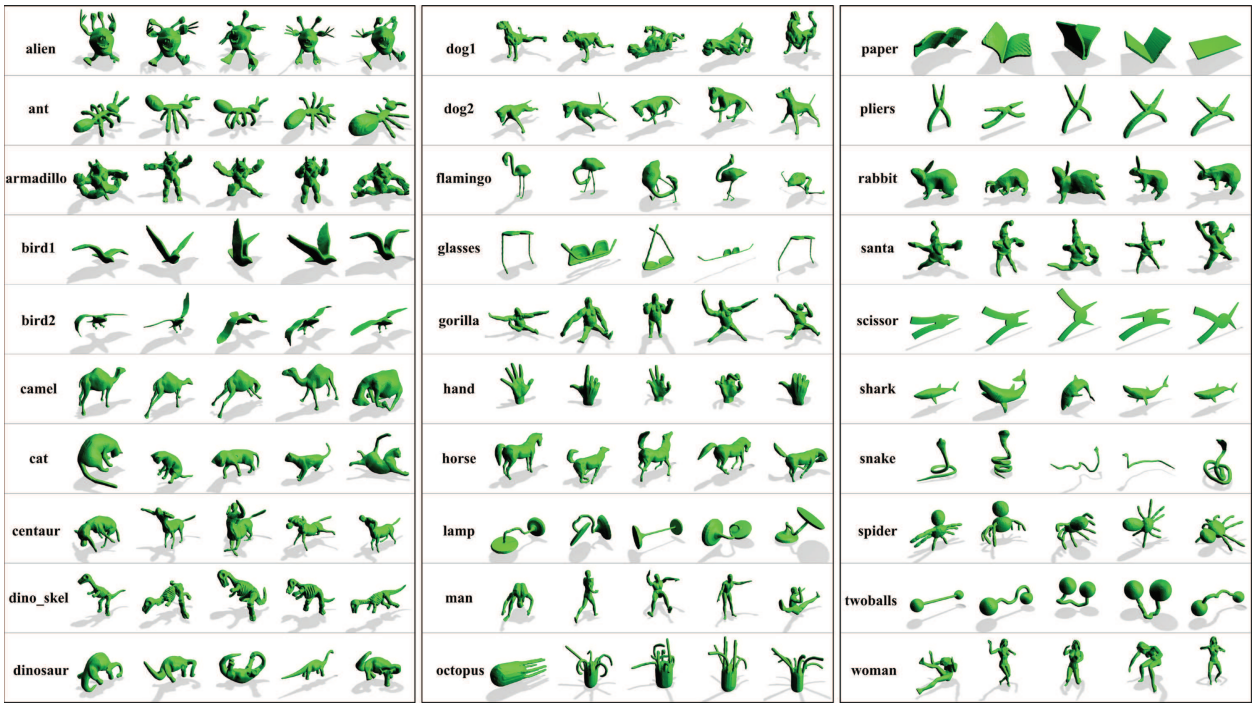


Figure 5: Examples of models in the SHREC'11 *Non-rigid 3D Watertight Meshes* database that are classified into 30 categories.

2.3 Pose Normalization

Common to both rigid and non-rigid 3D object retrieval methods, a pose normalization preprocessing step is usually mandatory, so that the geometric properties of the 3D models are directly comparable by the subsequent shape descriptor extraction and matching steps.

We define *pose normalization* to include translation, scaling and rotation normalization; whereas *alignment* will refer to rotation normalization only. In most cases, translation and scale normalization can be achieved by standard techniques. Alignment, however, is the most difficult part and still under investigation [27, 138, 63, 90, 99, 130]. Alignment methods can be divided into three major categories: (i) methods that are based on principal component analysis of the 3D models; (ii) methods that exploit symmetry characteristics of the 3D models and (iii) methods that utilize invariant shape descriptors for the 3D models. Based on this categorization, an overview of pose normalization methods follows.

2.3.1 PCA-based Alignment Methods

The best-known approach for computing the alignment of 3D objects is Principal Component Analysis (PCA) or Karhunen - Loeve transformation [90, 110, 123, 130, 141].

The PCA algorithm, based on the computation of 3D object moments, estimates the principal axes of a 3D object that are used to determine its orientation. In its original form, PCA has a number of disadvantages: it can be imprecise and often the principal axes of 3D objects that belong to the same class produce poor alignments [28].

To alleviate these problems, Vranic introduced an improvement to the original method, the Continuous PCA (CPCA) algorithm [131, 130, 132]. CPCA computes the principal axes of a 3D object based on the continuous triangle set.

Similar to the CPCA method, Papadakis et al. proposed the Normal PCA (NPCA) algorithm [87, 88], which computes the principal axes of the 3D object based on the surface normal set.

Related to PCA is the use of Singular Value Decomposition (SVD) for alignment [123]. In [41, 86], the SVD of the covariance matrix of the 3D object is computed and the unitary matrix is applied to the 3D object for rotation normalization.

The PCA-based pose normalization methods and especially the CPCA and NPCA variants, although general and well performing in most cases can fail to capture some specific characteristics of 3D objects such as symmetries and large planar or bumpy surfaces.

2.3.2 Symmetry-based Alignment Methods

Another major category of normalization methods exploits symmetry characteristics found in a large number of 3D objects.

Kazhdan et al. [64] define a reflective symmetry descriptor that represents a measure of reflective symmetry for an arbitrary 3D voxel object, for all planes through the object's center of mass. This descriptor is used for finding the main axes of symmetry or to determine that none of them exist in a 3D object. The descriptor is defined on a regular parametrization (the unit sphere) and describes the global characteristics of a 3D shape.

In [94], Podolak et al. extended this work and introduced a Planar Reflective Symmetry Transform (PRST) that computes a measure of the reflective symmetry of a 3D shape with respect to all possible planes. This measure is used to define the center of symmetry and

the principal symmetry axes of the global coordinate system.

Rustamov improved this approach with the augmented symmetry transform in [99].

Minovic et al. [78] compute symmetries of a 3D object, based on the computation of a principal octree aligned with the principal axes. Then the degree of symmetry is computed, based on the number of distinct eigenvalues associated with the principal axes. The algorithm is based on the traversal of the octree, which is aligned with the principal axes of space. An object can be both randomly positioned and have an arbitrary rotation within the space of the octree. A wide range of symmetries can be recognized, represented by a set of regular and irregular rotations.

Martinet et al. [76] use generalized moments to detect perfect symmetries in 3D shapes. The authors perform an analysis of the extrema values, as well as the components of the spherical harmonics and recover the parameters of the symmetries that characterize an object. The algorithm presented operates incrementally, allowing the determination of symmetries in larger objects, based on existing symmetries of their parts.

Mitra et al. [80] compute partial and approximate symmetries in 3D objects. Their method is based on the matching of simple local characteristics, in pairs, and the use of these matchings for the augmentation of information about the existence of symmetries in the corresponding space transformations. A segmentation step extracts potential significant symmetries of the object.

Sun and Sherrah [118] convert the symmetry detection problem to the correlation of the Gaussian image. Once the 3D model has been transformed to its Gaussian image, the principal axes are calculated and the existence of significant symmetries is examined near them. The calculations are performed on the rotational histograms of the images.

Using both PCA-alignment and planar reflective symmetry, Chaouch and Verroust - Blondet [27] compute a 3D object's alignment axes and then, using a Local Translational Invariance Cost (LTIC), make a selection of the most suitable ones.

Using a rectilinearity measure, Lian et al. [72] attempt to find a 3D object's best rotation by estimating the maximum ratio of its surface area to the sum of its three orthogonal projected areas. Similar to the previous approach, [27], a selection between the proposed and a PCA-based alignment is made. Most of the methods that exploit symmetry characteristics for achieving pose normalization, seem to perform quite well in most cases.

However, a major problem related to the symmetry-based techniques is that symmetry detection either focuses on small fragments or larger abstract areas of the 3D objects, and thus it is unable to handle 3D objects that present complex, multilevel (global and local) symmetry in their structure.

In [9] Axenopoulos et al. combine the properties of plane reflection symmetry and rectilinearity for achieving alignment. In this paper both CPCA and principal plane reflective symmetry are used, in order to achieve alignment. Rectilinearity is utilized to improve the alignment results.

2.3.3 Rotation Invariant Rigid Shape Descriptors

A third category of methods achieves rotation invariance by the definition of the shape descriptor. In general the price to pay for this rotational invariance, is reduced discriminative power.

Descriptors based on spherical harmonics [65, 121], Zernike moments [81, 121] and shell histograms [7, 135, 137] are examples of representation methods able to achieve rotation invariance by definition.

Kazhdan et al. in [65] introduce the Spherical Harmonic Representation, a general method for obtaining a rotation invariant representation of spherical shape descriptors that describes them in terms of the distribution of energies across different frequencies. The same authors extended this method with symmetry information to provide a more discriminating representation in [66].

Novotni and Klein in [81] use 3D Zernike invariants as descriptors for 3D shape retrieval and Ankerst et al. [7] proposed the Shape Histograms descriptor, where 3D space is divided into concentric shells, sectors, or both and for each part, the object's shape distribution is computed giving a sum of histogram bins.

Finally, Chen and Ouhyoung [138] use a region based 2D shape descriptor to recover the affine transformation between two 3D objects and thus achieve normalization between them.

The majority of methods that achieve rotation invariance by the definition of shape descriptors, perform best on specific 3D object classes that are composed of 3D objects with similar structure. Due to this explicit behavior, these methods are unable to han-

dle general 3D objects originating from different classes, or with significant structural differences. Also, these methods generally result in descriptors with relatively low discriminating power.

A careful review of the works presented shows that in order to achieve better results, most recent studies attempt to combine techniques from the same or different categories; these often include variations of PCA and/or exploitation of 3D object symmetry characteristics. However, although it seems that most of these methods perform exceptionally well, a major problem is that they usually combine results blindly without taking into account any complementarities involved.

2.4 3D Object Retrieval based on 2D Range Image Queries

In the past few years, the increasing availability of low-cost 3D scanners has resulted in the creation of large 3D model repositories, thus making content-based retrieval a key operation. 3D model retrieval has considerably matured and a number of very accurate and robust descriptors have been proposed [89, 106, 28, 65, 132]. These methodologies use a 3D model query to search a database of 3D models in a content-based manner. However, in practical situations, it is often difficult to come up with a suitable 3D model query in the first place: this has either to be found or built, a random and time-consuming action, respectively.

Nowadays, 3D scanners that typically produce *range images* (also called *range scans* and/or *depth buffers*) from real world 3D objects are becoming common and cheap, e.g. Microsoft Kinect [112]. It would thus, be beneficial, to use the range scans of real objects as queries on 3D model repositories.

However, a number of challenges exist. First, a range image represents only a partial object. Thus, it is not straightforward to effectively match such data against a complete 3D model representation, since an important part of it may be missing. Second, range images can be rough and noisy. Third, it is not straightforward how to bridge the gap between the 3D model representation and the range image, i.e. how to produce descriptors that can be relatively invariant to these two representations. The representation gap makes it difficult to extract a signature that will be (at least partially) similar when presented with a complete, clean 3D model and when presented with a partial and noisy range image of

a similar query object.

Over the past few years, the number of works addressing the problems of multimodal 3D object retrieval and recognition (and particularly those based on range image queries), have increased significantly. Although this task still remains non-trivial, the quality of existing works shows that very important steps have been made in the field. Common retrieval scenarios deal with two different query image types: (i) directly captured range images from real 3D objects (i.e. using a 3D range scanner) and (ii) artificially produced range images via depth buffer capturing of complete 3D models. The first query image type is closer to real-world applications and is being increasingly adopted as 3D scanning becomes common place and corresponding datasets are created.

Hetzel et al. [56] explore a view based approach for the recognition of free-form objects in range images. They combine a set of local features (pixel depth, surface normal and curvature metrics) in a multidimensional histogram in order to achieve classification.

Johnson and Hebert [61], use a spin image representation scheme in order to achieve simultaneous recognition of multiple 3D object in cluttered scenes. The spin image representation is used for matching surface points.

Chen and Bhanu [29] introduce a local surface descriptor for 3D model recognition. This descriptor is computed on feature points of a 3D surface, where large shape variations occur. The local surface descriptor is characterized by its centroid, its local surface type and a 2D histogram. The latter shows the frequency of occurrence of shape index values (calculated from principal curvatures) vs the angles between the normal of the reference feature point and those of its neighbors.

Ruiz-Correa et al. [98] propose a method for recognizing 3D objects in real range image scenes. Initially, shape class components are learnt and extracted from range images and then the spatial relationships among the extracted components are used to form a model that consists of a three-level hierarchy of classifiers.

Adan et al. [1] explore the use of Depth Gradient Image (DGI) models for the recognition of 3D models. The DGI representation synthesizes both surface and contour information, for a specific viewpoint, by mapping the distance between each contour point and the edge of the viewpoint image in terms of internal and external object pixels. This measure is computed for the entire model, taken from the nodes of a tessellated sphere.

Frome et al. [45] introduced two regional shape descriptors, the 3D generalization of the 2D shape context descriptor and the harmonic shape descriptor. The authors evaluate the performance of the proposed descriptors in recognizing similar objects in scenes with noise or clutter.

Ohbuchi et al. [84] proposed the Multiple Orientation Depth Fourier Transform (MODFT) descriptor where the model is projected from 42 viewpoints to cover all possible view aspects. Each depth buffer is then transformed to the $r-\theta$ domain and the Fourier transform is applied. To compare two models, all possible pairs of coefficients are compared which inevitably increases comparison time.

Stavropoulos et al. [116] present a retrieval method based on the matching of salient features between the 3D models and query range images. Salient points are extracted from vertices that exhibit local maxima in terms of protrusion mapping for a specific window on the surface of the model. A hierarchical matching based scheme is used for matching. The authors experimented on range images acquired from the SHape REtrieval Contest 2007 (SHREC'07) *Watertight models* [50] and the Princeton Shape Benchmark (PSB) standard [110] datasets.

Chaouch and Verroust-Blondet [24] present a 2D/3D shape descriptor which is based on either silhouette or depth-buffer images. For each 3D model a set of six projections in calculated for both silhouette and depth-buffers. The 2D Fourier transform is then computed on the projection. Furthermore, they compute a relevance index measure which indicates the density of information contained in each 2D view. The same authors in [26] propose a method where a 3D model is projected to the faces of its bounding box, resulting in 6 depth buffers. Each depth buffer is then decomposed into a set of horizontal and vertical depth lines that are converted to state sequences which describe the change in depth at neighboring pixels. Experimentations were conducted on range images artificially acquired from the PSB dataset.

Shih et al. [109] proposed the elevation descriptor where six depth buffers (elevations) are computed from the faces of the 3D model's bounding box and each buffer is described by a set of concentric circular areas that give the sum of pixel values within the corresponding areas. The models were selected from the standard PSB dataset.

Experimenting on the SHREC'09 *Querying with Partial Models* [39] dataset, Daras and

Axenopoulos in [34] present a view-based approach for 3D model retrieval. The 3D model is initially pose normalized and a set of binary (silhouette) and range images are extracted from predefined views on a 32-hedron. The set of features computed on the views are the Polar-Fourier transform, Zernike moments and Krawtchouk moments. Each query image is compared to all the extracted views of each model of the dataset.

Ohbuchi et al. [85] extract features from 2D range images of the model viewed from uniformly sampled locations on a view sphere. For every range image a set of multi-scale 2D visual features are computed using the Scale Invariant Feature Transform (SIFT) [74]. Finally, the features are integrated into a histogram using the Bag-of-Features approach [49]. The same authors enhanced their approach by pre-processing the range images, in order to minimize interfere caused by any existing occlusions, and also and by refining the positioning of SIFT interest points, so that higher resolution images are favored [47, 83]. Their works have experimented on and have participated on both corresponding SHREC'09 *Querying with Partial Models* and SHREC'10 *Range Scan Retrieval* [40] contests.

Wahl et al. [133] propose a four-dimensional feature that parameterizes the intrinsic geometrical relation of an oriented surface point pair (surflets). For a 3D model a set of surflet pairs is computed over a number of uniformly sampled viewing directions on the surrounding sphere. This work was one of the two contestants of the SHREC'10 *Range Scan Retrieval* track.

2.4.1 3D Object Retrieval on 2D Range Image Queries Databases

In this subcategory of problems, only a few dedicated benchmark databases exist, mainly due to two reasons: It is only recently that inexpensive and relatively accurate 3D object acquisition hardware became available, and a common methodology for testing the proposed methodologies was to artificially create partial (range image) queries from known complete 3D object retrieval databases.

Apart from the last case, where many authors experiment on many of the rigid and non-rigid 3D object retrieval databases that we have already presented in previous chapters, the SHape REtrieval Contest has released a series of datasets targeted especially for this task.

Beginning with the SHREC'09 *Querying with Partial Models* dataset [39], the authors presented a database aimed at both partial 3D object retrieval and retrieval of complete 3D objects based on query views. The target database is a subset of the shape benchmark constructed at NIST, described in [43]. It contains 720 complete 3D models, which are categorized into 40 classes. The classes are defined with respect to their semantic categories and each contains 18 3D models. The query set is composed of 20 range images, which are obtained by capturing range data of 20 objects from arbitrary view directions (see Fig 6). The range images were captured using the NextEngine desktop 3D scanner.

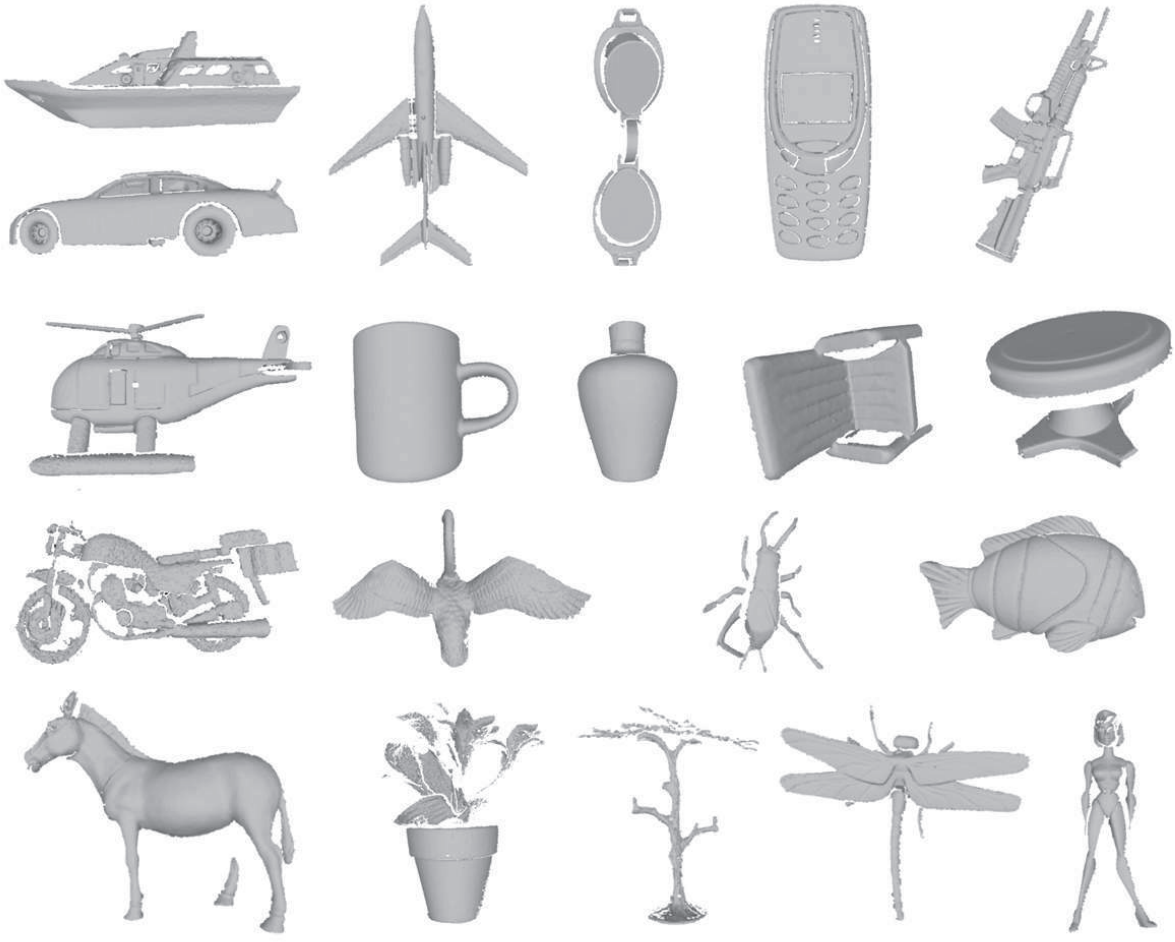


Figure 6: The SHREC'09 *Querying with Partial Models* dataset query views

In SHREC'10 *Range Scan Retrieval*, [40] the same authors extended the previous database and focused specifically on the range image query retrieval. The target database is again the generic shape benchmark constructed at NIST, containing 800 complete 3D models, which are categorized into 40 classes. The classes are defined with respect to

their semantic categories, and each contains 20 complete 3D models. The query set is composed of 120 range images, which are acquired by capturing 3 range scans of 40 real objects from arbitrary view directions. These objects correspond to the 23 classes in the target database, with one exception. The 3 scans of that object were excluded from the evaluation; hence the evaluation is based on 117 scans. The range images are captured using a Minolta Laser Scanner.

Following the same scheme, in the SHREC'11 *Shape Retrieval Contest of Range Scans*, [113] the cardinalities of both the target and query datasets has been raised. The target database is composed of 1000 complete 3D models, which are categorized into 50 classes, each containing 20 models. The query set is composed of 150 range images, which are acquired by capturing 3 range scans of 50 models from arbitrary view directions. The range images are captured using a Minolta Laser Scanner.

3 Pose Normalization

Nature does nothing without purpose

– Aristotle

In this chapter, three novel pose normalization methods, ROSy, SymPan and SymPan+, are discussed in detail.

ROSy consists a general purpose global pose normalization method based on 3D object reflective symmetry. Initially, the axis-aligned minimum bounding box of a rigid 3D model is modified by requiring that the 3D model is also in minimum angular difference with respect to the normals to the faces of its bounding box. To estimate the modified axis-aligned bounding box, a set of predefined planes of symmetry are used and a combined spatial and angular distance, between the 3D model and its symmetric model, is calculated. By minimizing the combined distance, the 3D model fits inside its modified axis-aligned bounding box and alignment with the coordinate system is achieved [105]. SymPan is a pose normalization method, based on panoramic views and reflective symmetry. Initially, the surface of a 3D model is projected onto the lateral surface of a circumscribed cylinder, aligned with the primary principal axis of space. Based on this cylindrical projection, a normals' deviation map is extracted and using an octree-based search strategy, the rotation which optimally aligns the primary principal axis of the 3D model and the cylinder's axis is computed. The 3D model's secondary principal axis is then aligned with the secondary principal axis of space in a similar manner [104]. An extension to the SymPan method, SymPan+, estimates the symmetry plane (and consecutively the corresponding principal axis) using an iterative process. This is achieved by rotating the 3D model and

computing reflective symmetry scores on panoramic view images. The remaining principal axes of the 3D model are estimated by computing the pixel variance of the 3D model's panoramic view images [108].

All three pose normalization methods are integrated in a rigid 3D object retrieval pipeline for quantitative evaluation purposes. The results are discussed in Chapter 6.

3.1 ROSy Pose Normalization

In the ROSy method, the problem of pose normalization is described through the Surface-Oriented Minimum Bounding Box (SoMBB), a modified version of the Axis-Aligned Bounding Box (AABB) which is commonly used in collision detection techniques [128, 51]. The motivation behind the proposed method is to minimize the SoMBB of a 3D object so that the latter becomes aligned with its SoMBB and consecutively with the principal axes of space. Furthermore, to ensure that the 3D object's large planar areas are also in alignment with the principal planes of space, it is required that the average normal to the object's large planar areas become parallel to the box's face normals. Using the properties of the reflective symmetry transformation the problem will be mathematically formulated and the foundation for the proposed solution will be set.

3.1.1 Surface-Oriented Minimum Bounding Box

Pose normalization is, by definition, a complex procedure, highly dependent on the target application. For that reason, various definitions have been proposed in the literature [66, 121, 131]. We next attempt to provide the definition of a 3D object's SoMBB and formulate the problem of pose normalization based on the notion of the SoMBB. A 3D object's *Surface-Oriented Minimum Bounding Box* (SoMBB) is an axis-aligned bounding box that has the minimum possible volume while simultaneously the normals to its faces are in minimum angular difference with the majority of the contained 3D object's face normals (Figure 7). Pose normalization is the procedure of finding a set of homogeneous transformations (translation, scaling, rotation) that fit a 3D object into its SoMBB.

Constraining the face normals of a 3D object to be in *minimum angular difference* with the faces of its SoMBB can be interpreted as making the average normal to the object's

large planar areas parallel to the box's face normals [54]. Example 3D object classes that explicitly define large planar areas are buildings, airplanes, ships, tables, billboards, etc (Figure 8a). Object classes that define large planar areas as the average normal of many small triangles with similar orientation are quadruped animals, hands, human bodies, etc (Figure 8b).

If a 3D object is not aligned with the coordinate system axes but arbitrarily positioned in space, the calculation of the corresponding SoMBB is not an easy task. However, let us suppose that the SoMBB is already precalculated, although the object has an arbitrary rotation in space. If the object is fit inside its SoMBB, then it becomes aligned with the coordinate system. Translation and scaling of the 3D object, to fit inside the SoMBB can be adequately solved by one of the common techniques, however, the rotation of the 3D object, so as to satisfy the two SoMBB conditions (minimum volume and minimum angular difference between the normals to its faces and the 3D object's face normals) remains a

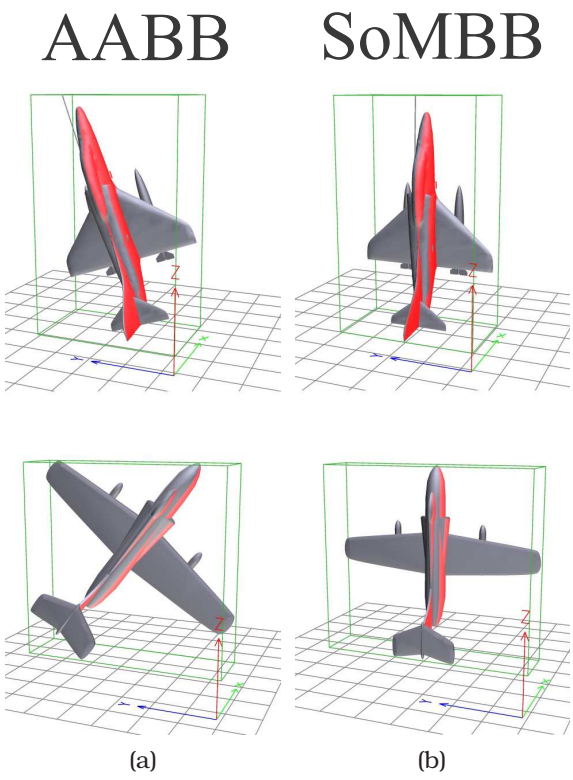
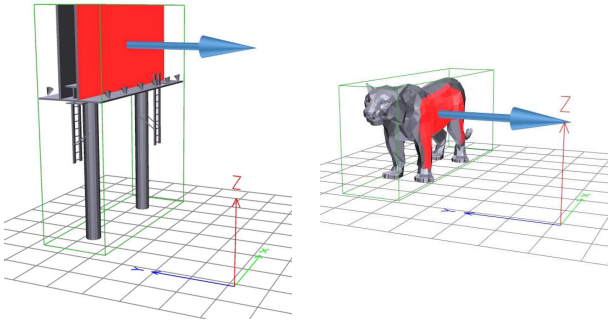


Figure 7: Difference between the AABB and SoMBB on objects of the same class. On column 7a two 3D objects are enclosed inside their AABB, while on column 7b, the same 3D objects are enclosed inside their SoMBB. Marked faces indicate faces whose normals are parallel to the SoMBB's face normals.



(a) (b)

Figure 8: 3D objects enclosed in their SoMBBs. Marked faces illustrate the components of a 3D objects' planar surface and the arrows show the direction of its averaged normal axis (parallel to the normals to the left - right faces of the SoMBB). 8a illustrates a 3D object that has a planar surface defined by structure and 8b illustrates a 3D object that defines a planar surface through many small triangles with similar orientation.

hard task. To solve this problem, properties of the reflective symmetry transformation will be taken into account.

3.1.2 Reflective Symmetry Transformation

Assume that a 3D object M, represented by a set of m vertices $P \in \mathbb{R}^{m\times 3}$, and a corresponding set of normals $N \in \mathbb{R}^{m\times 3}$, exists in the Euclidean $\mathbb{R}^3$ space. The reflective symmetry transformation (or reflection) [124] is a linear transformation which computes the symmetric object M_{Π}^{-1} about a candidate plane of symmetry Π ($\Pi : ax+by+cz+d = 0$). The planes of symmetry used, are the three principal planes of the Cartesian coordinate system (i.e. XY, XZ, YZ). Figure 9 illustrates example reflective symmetry transformations against the principal planes of the Cartesian coordinate system.

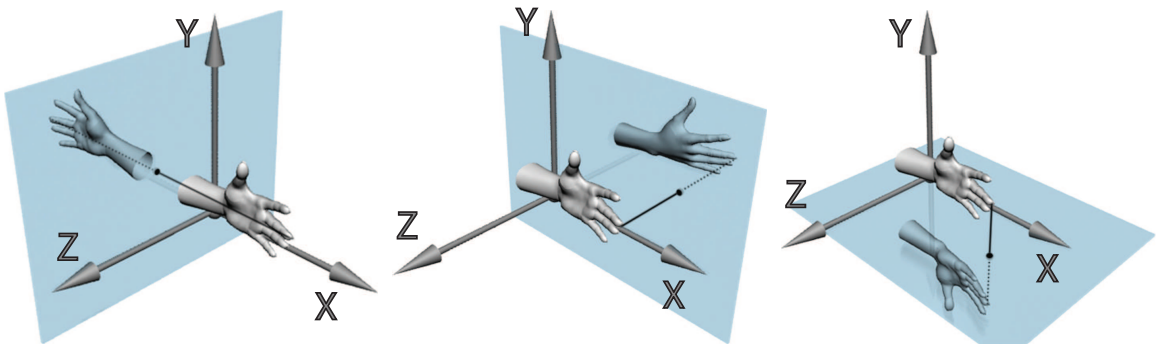


Figure 9: Examples of the reflective symmetry transformation on planes YZ, XY and XZ, respectively.

3.1.3 Euler Angles

We also need to introduce Euler angles which will prove useful in the sequel. According to Euler's Rotation Theorem, to reach any target frame, a specific sequence of three rotations, that are described by three angles is required. If the three rotations are written in terms of rotation matrices, then the general rotation can be written as the product of these rotation matrices. The three angles giving the corresponding rotation matrices are called the Euler angles [79, 52]. There are six possible conventions regarding the Euler angles ($X-Y-X$, $X-Z-X$, $Y-X-Y$, $Y-Z-Y$, $Z-X-Z$ and $Z-Y-Z$), depending on the axes about which the rotations occur. The first two rotations establish a common principal rotation axis between the source and target frames (also known as the 'line of nodes'). The third rotation, about the principal rotation axis, aligns the remaining axes of the reference and target frames. Different conventions result in a different axis ordering of the target frame.

3.1.4 The Complete Method

In this section, a minimization criterion that measures the fitness of a 3D object into its SoMBB will be defined and the complete pose normalization method will be described.

According to [23] if an arbitrary model is reoriented in such way that the areas of the bounding boxes of its projections onto the three principal planes are minimum, then the volume of the bounding box of the reoriented model is also minimum. Ahn et al. in [4, 3] showed that finding the minimum convex hull of the union of two convex sets[1] is equivalent to finding maximum overlap[2] between them. Therefore, defining the maximum overlap between a 3D object and its symmetric object, with regard to a given symmetry plane Π is equivalent to defining the two objects' minimum convex hull projected onto Π. However, since 3D objects M and M_{Π}^{-1} are symmetric, their projected convex hulls are identical and will be minimized simultaneously.

To define the maximum overlap between 3D objects M and M_{Π}^{-1} we use a measure based on the distance between their corresponding vertices. The distance between each vertex p_i of the original 3D object M and the corresponding vertex $p_{i,\Pi}^{-1}$ of the symmetric object M_{Π}^{-1} is twice the distance between p_i and the plane of symmetry Π. The mean

[1] E.g. the convex hulls of our 3D objects.

[2] The maximum overlap is achieved when the union of the two convex sets occupies the minimum total space.

distance between all the corresponding vertices of objects M and M_{Π}^{-1} can be used as a measure of the total distance between the original 3D object and the plane of symmetry (eq. 1). When this distance becomes minimal, with regard to the dimension defined by the normal to the plane of symmetry, the volume of the 3D object's SoMBB is minimized as well (by maximizing the overlap between the two objects). During the minimization process (see eq. 6) the 3D object is transformed while the planes of symmetry remain fixed.

$$Dist(M, M_{\Pi}^{-1}) = \frac{1}{m} \sum_{i=1}^{m} \left| p_i - p_{i,\Pi}^{-1} \right| \tag{1}$$

Using eq. 1 as the minimization function, we achieve the results shown in Figure 7a. As has been described before, our motivation is to achieve an alignment which is intuitively shown in Figure 7b. To this end, we enrich the functional with an additional term that expresses the minimum angular difference measure based on the angular distance between the corresponding face normals of 3D objects M and M_{Π}^{-1}. The angle $\theta_{i,\Pi}$ between a surface normal n_i of the original 3D object M and the corresponding surface normal $n_{i,\Pi}^{-1}$ of the symmetric object M_{Π}^{-1} is supplementary to the angle between the surface normal n_i and the normal to the plane of symmetry (eq. 2). When $\theta_{i,\Pi}$ is minimized, surface normals n_i and $n_{i,\Pi}^{-1}$ become perpendicular to the normal to the plane of symmetry, whereas when $\theta_{i,\Pi}$ is maximized surface normals n_i and $n_{i,\Pi}^{-1}$ become parallel to the normal to the plane of symmetry.

$$\theta_{i,\Pi} = \left[\cos^{-1} \left(\frac{n_i \cdot n_{i,\Pi}^{-1}}{|n_i| \left| n_{i,\Pi}^{-1} \right|} \right) \right] \tag{2}$$

The mean angle between all the corresponding normals of 3D objects M and M_{Π}^{-1} can be used to define the angular difference between the normal to the plane of symmetry and the face normals of the 3D object (eq. 3).

$$Ang(M, M_{\Pi}^{-1}) = \frac{1}{m} \sum_{i=1}^{m} \frac{1}{\pi} \theta_{i,\Pi} \tag{3}$$

Since the parallelization of the corresponding face normals of a 3D object and its symmetric object is not always perfect, $Ang(M, M_{\Pi}^{-1})$ could be relaxed so that narrower angles are rewarded over wider angles (eq. 4). Figure 10 illustrates the difference between

equations (3) and (4).

$$Ang_{tanh}(M, M_{\Pi}^{-1}) = \frac{1}{m}\sum_{i=1}^{m}\tanh\left(\frac{3}{2}\left(\theta_{i,\Pi} - \pi\right)\right) + 1 \tag{4}$$

When $Ang_{tanh}(M, M_{\Pi}^{-1})$ is minimized, the mean angle between the face normals of the 3D object and the normals to the SoMBB's faces that are parallel to the plane of symmetry is also minimized.

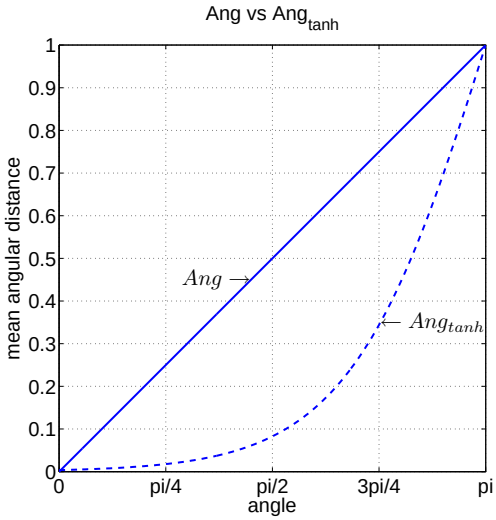


Figure 10: Graphical representation of $Ang(M, M_{\Pi}^{-1})$ and $Ang_{tanh}(M, M_{\Pi}^{-1})$, respectively.

Equations (1) and (4) aim to minimize the volume of the SoMBB and to have the 3D object's face normals as parallel as possible to the normals to the SoMBB's faces. We use them together (eq. 5), with equal weights, to create our minimization criterion.

$$k_{\Pi} = \arg\min\left\{Dist(M, M_{\Pi}^{-1}) + Ang_{tanh}(M, M_{\Pi}^{-1})\right\} \tag{5}$$

In equation (5), $Dist(M, M_{\Pi}^{-1})$ is dependent on the scaling of the 3D objects, the distance between their centroids and the distance between the corresponding vertices, while $Ang_{tanh}(M, M_{\Pi}^{-1})$ depends only on the angle between the normals and lies in the range [0, 1]. To normalize equation (5) and give equal weights to the two factors, 3D objects M and M_{Π}^{-1} need to be centered at the origin and properly scaled so as to fit inside the unit sphere. Once translation and scale normalization are performed, $Dist(M, M_{\Pi}^{-1})$ ranges in the interval [0, 2]. The final form of the minimization criterion is expressed by equation (6), where the ranges of $Dist(M, M_{\Pi}^{-1})$ and $Ang_{tanh}(M, M_{\Pi}^{-1})$ are equalized

through multiplication of $Dist(M, M_{\Pi}^{-1})$ with a normalization factor of 0.5.

$$\begin{aligned} k_{\Pi} =& \arg\min \left\{ \frac{1}{2} Dist(M, M_{\Pi}^{-1}) + Ang_{tanh}(M, M_{\Pi}^{-1}) \right\} \\ =& \arg\min \left\{ \frac{1}{2m} \sum_{i=1}^{m} \left| p_i - p_{i,\Pi}^{-1} \right| + \right. \\ & \left. \frac{1}{m} \sum_{i=1}^{m} \tanh\left(\frac{3}{2} \left(\theta_{i,\Pi} - \pi \right) \right) + 1 \right\} \end{aligned} \tag{6}$$

$Dist(M, M_{\Pi}^{-1})$ and $Ang_{tanh}(M, M_{\Pi}^{-1})$ contribute equally to the computation of k_{Π} and, as will be shown in the evaluation section, experimental results support this choice.

Next, the use of the minimization criterion will be expanded on all three principal planes of the Euclidean space. The complete method is called 3D object pose normalization based on Reflective Object Symmetry (ROSy) and aligns an arbitrary 3D object with a reference coordinate system. The corresponding algorithm is outlined in Algorithm 1.

A preliminary step of the proposed method is the resampling of the input 3D object M (Algorithm 1: step 2) to ensure that any deficiencies of the digitization process are eliminated. Unwanted conditions like irregular distribution of vertices on the 3D object's surface could potentially result in a poor alignment. Object resampling is achieved by redistributing the vertices of the 3D object along its surface triangles, based on a ratio of $f_i = \frac{t_i^{area}}{t_{sum}^{area}}$, where t_i^{area} denotes the area of triangle i and t_{sum}^{area} denotes the total area of the object's surface. Note that at least one vertex will accrue from every initial triangle and so, it is possible that the final number of 3D object vertices is greater than the original.

Standard translation and scale normalizations are then performed (Algorithm 1: steps 3 - 4). Translation invariance is achieved by using the Continuous Principal Component Analysis (CPCA). The centroid of the object is computed using CPCA and then the whole object is translated, so that the centroid coincides with the coordinate origin [87, 131, 130]. Scale invariance is achieved through the scaling of M so that it fits exactly into the unit sphere. Translation and scale normalization, position the object 'inside' its SoMBB and limit the range of $Dist(M, M_{\Pi}^{-1})$ in the interval [0, 2].

At this point, although 3D object M is positioned at the center of its SoMBB, it is uncertain if it complies with the two SoMBB constraints: minimum bounding box volume

Algorithm 1 Object pose normalization based on Reflective Object Symmetry (ROSy)

1: Read input 3D object M;
2: Resampling of M;
3: Translation normalization of the centroid of M to the origin of the coordinate system;
4: Scale normalization of M to the unit sphere;
5: **for** ROT_AXIS in (Z, Y, Z) **do**
6: **if** ROT_AXIS $= Y$ **then**
7: $\Pi \leftarrow YZ$;
8: **else if** ROT_AXIS $= Z$ **then**
9: $\Pi \leftarrow XZ$;
10: **end if**
11: $k_{best} \leftarrow \infty$;
12: $M_{\Pi}^{-1} \leftarrow$ Reflection_Transform(M, Π);
13: **for** ANGLE = 0° to 180° step 2° **do**
14: $M \leftarrow$ Rotation_Transform(M, ANGLE, ROT_AXIS);
15: $M_{\Pi}^{-1} \leftarrow$ Rotation_Transform(M_{Π}^{-1}, -ANGLE, ROT_AXIS);
16: $k_{\Pi} = \frac{1}{2} Dist(M, M_{\Pi}^{-1}) + Ang_{tanh}(M, M_{\Pi}^{-1})$;
17: **if** $k_{\Pi} < k_{best}$ **then**
18: $k_{best} \leftarrow k_{\Pi}$;
19: $M_{best} \leftarrow M$;
20: **end if**
21: **end for**
22: **end for**
23: **for** Π in (XY, YZ, ZX) **do**
24: $M_{\Pi}^{-1} \leftarrow$ Reflection_Transform(M_{best},Π);
25: $d_{\Pi} = \frac{1}{2} Dist(M_{best}, M_{\Pi}^{-1})$;
26: **end for**
27: Label axes in ascending plane distance order $(XY < YZ < ZX)$; **return** M_{best};

and minimum angular difference between the 3D object's face normals and the normals to the bounding box's faces. To ensure this, criterion k_{Π} must be minimized on all three orthogonal planes of symmetry (Algorithm 1: steps 5 - 22).

In the proposed methodology, following the concept of Euler angles, the target is to establish the common principal rotation axis, by aligning the 3D object with two faces of the corresponding SoMBB. Then, the input object will be rotated about this common principal rotation axis, so as to further align it with the third face of the SoMBB and thus with the coordinate system. Criterion k_{Π} will ensure that the normals to the faces of the SoMBB will attain minimum angular difference to the 3D object face normals and, simultaneously, that the SoMBB will attain minimum volume. Note that since the ordering of axes is not important, the $Z - Y - Z$ convention will be arbitrarily selected, for the

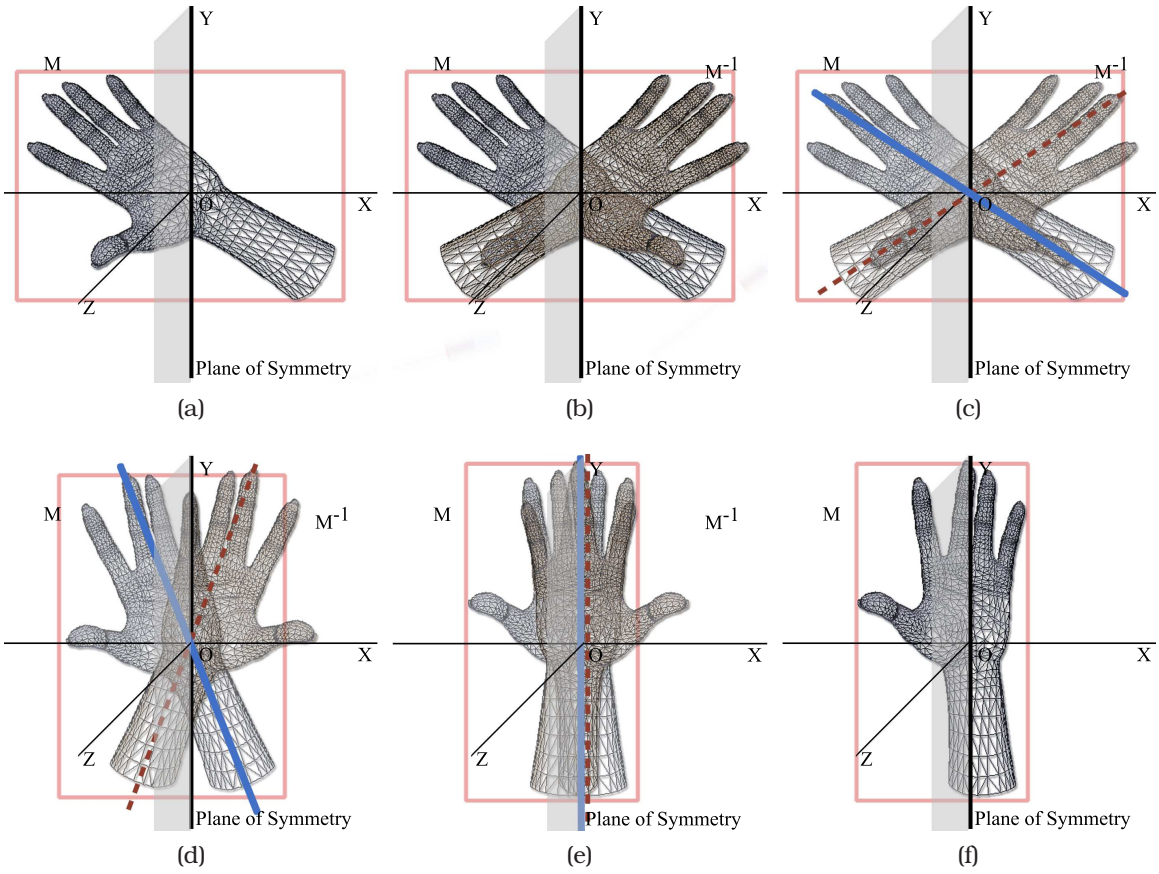


Figure 11: Illustration of the alignment procedure for a single rotation step on the YZ plane of symmetry. The bounding rectangle illustrates the SoMBB of the 3D object's projection on plane XY. 11a Original 3D object orientation. 11b Original (M) and its symmetric (M_{Π}^{-1}) 3D objects. 11c Selected principal axes projection of the original (dashed axis) and the symmetric (solid axis) 3D object, on principal plane YZ. 11d - 11f Stepwise minimization of the distance between the original and its symmetric 3D object.

alignment procedure.

The selected plane of symmetry for each iteration step must fulfill two conditions: (i) its normal must be perpendicular to the target SoMBB face normal and (ii) the axis about which the rotation occurs must be perpendicular to the plane's normal. On each iteration, 3D object M and its symmetric object M_{Π}^{-1} are rotated by 180 degrees in opposing directions with a step of 2 degrees, until k_{Π} is minimized (Figure 11). Since vertex and normal cardinality is fixed for each model, the time required for the alignment process to complete is linear in the number of iteration steps. We have selected a step of 2 degrees which results in good alignments while preserving acceptable processing speed. Exhaustive search is performed here as k_{Π} is not necessarily a monotonic function. When the rotation normalization procedure is complete, object M will be aligned with the Cartesian

coordinate system.

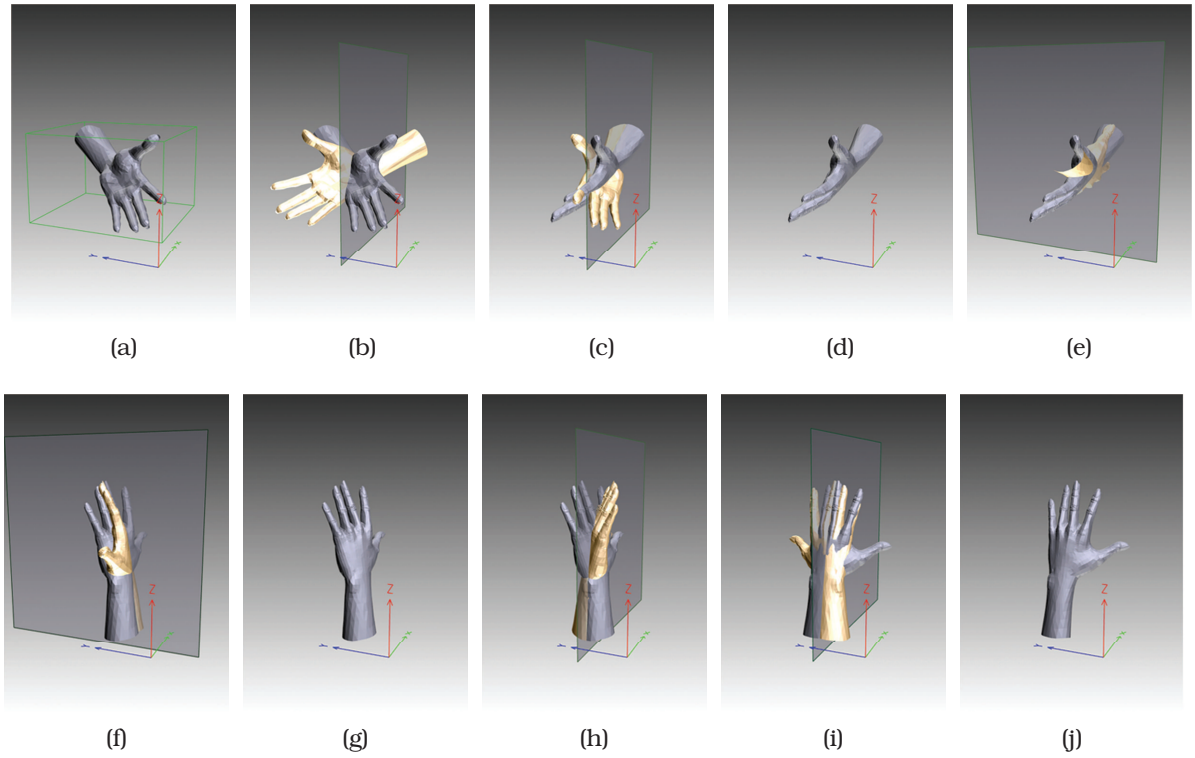


Figure 12: Overview of all the steps of the alignment procedure. 12a initial 3D object orientation. 12b symmetric 3D object against plane XZ. 12c - 12d minimum k_{XZ} point after rotation of the 3D object and its symmetric object about axis z. 12e symmetric object against plane YZ. 12f - 12g minimum k_{YZ} point after rotation of the 3D object and its symmetric object about axis y. 12h symmetric 3D object against plane XZ. 12i minimum k_{XZ} point after rotation of the 3D object and its symmetric object about axis z. 12j final aligned 3D object.

In detail, let us assume that the initial orientation of the 3D object is arbitrary. The first rotation about axis z, given minimization criterion k_{XZ}, aligns the 3D object with the first selected plane of symmetry, XZ. The direction of the 3D object's orientation becomes constrained by the first plane of symmetry. The second rotation, about axis y, given minimization criterion k_{YZ} (symmetry plane YZ), further aligns the input 3D object M with axis z of the coordinate system (the 3D object is aligned with both planes XZ and YZ). The direction of the 3D object's orientation thus becomes constrained by axis z of the coordinate system, which is the intersection of the first and the second planes of symmetry, XZ and YZ, respectively. Once the common principal rotation axis has been established, rotation of the 3D object about this axis, around symmetry plane XZ, given minimization criterion k_{XZ}, results in the alignment of the 3D object with the remaining

axes x and y of the coordinate system (Figure 12).

The symmetric object M_{Π}^{-1} of the input 3D object M is created by the reflective symmetry transformation function (Algorithm 1: step 12) that takes as input the vertices and surface normals of a 3D object M and a plane of symmetry. The rotation transformation function (Algorithm 1: steps 14 - 15) takes as input a 3D object, an angle of rotation and the axis about which the rotation occurs and returns the rotated 3D object. The rotation function is iteratively used with opposing angles for the 3D object and its symmetric object (to maintain the symmetry property) and the transformation that results from the best k_{Π} value, is kept.

The aforementioned pose normalization procedure is able to orient the principal axes of a 3D object with the Cartesian coordinate system axes. However, although the object is correctly aligned with the reference coordinate system, the ordering of the dimensions is not defined, yet. The final step of the method is to label the principal axes of the aligned object by computing the mean distance of its vertices from each coordinate system axis (Algorithm 1:steps 23 - 27). Although the direct calculation of the mean distance between the vertices of the 3D object and each of the three coordinate system axes is the simplest method, structural specificities like symmetries or density variations of the 3D object's surface could lead to inaccurate results. To overcome this problem, symmetric objects can also be used for the labeling of the 3D object's axes.

This procedure is similar to using the Manhattan distance for the calculation of the mean distance between the 3D object vertices and the coordinate system axes, and the results derive from a 2-step calculation, therefore distinguishing better similar dimensions of the 3D object. If the mean distance between the vertices of the original and symmetric 3D objects is small, against a specific principal plane, then most of the 3D object's vertices lie close to that plane, which possibly contains the primary and secondary principal axes of the 3D object. If the vertex distance is large, then most of the 3D object's vertices lie far from the principal plane and thus, this plane cannot contain the principal axis of the 3D object.

The primary principal plane, which has the smallest vertex distance between the 3D object and its symmetric 3D object, is assumed to contain the 3D object principal axis, while the tertiary principal plane (with the largest vertex distance) is assumed to contain

the tertiary principal axis. Since the principal axes are perpendicular, they are defined as the two non-common axes of the primary and the tertiary principal planes and the second principal axis of the 3D object is defined as their common axis.

3.2 SymPan, SymPan+ Pose Normalization

In the sequel, two novel pose normalization methods that produce state-of-the-art results, in both qualitative and quantitative terms, are presented.

The motivation for both methods is that the use of reflective symmetry as a feature for pose normalization and 3D object retrieval seems to enhance the results [66], as most of the 3D objects exhibit symmetrical properties to some degree. These properties tend to be distinct between different classes and similar between objects of the same class, therefore enhancing the discrimination achieved by other commonly used characteristics, such as the spatial distribution and/or surface orientation of the 3D models.

In the SymPan methodology, the surface of the 3D model is initially projected on the lateral surface of a circumscribed cylinder, aligned with the primary principal axis in space. For each point on the cylinder's surface, the angular difference of its normal against a corresponding surface point of the 3D model is computed, thus creating a Normals' Deviation Map. The maximization of the mean value of the NDM will result in the minimization of the angular difference between the 3D model's surface normals and those of its circumscribed cylinder, therefore achieving alignment between the primary principal axis of the 3D model and the corresponding axis of the cylinder. The alignment process is further enhanced by a measure based on the 3D model's symmetry characteristics using an octree-based search strategy. In particular, the NDM image is scanned for a column of vertical symmetry that results in the polar coordinate at which the model exhibits the greatest degree of reflective symmetry. Following the alignment of the 3D model's primary principal axis, a similar search methodology is used for its secondary principal axis. The insight for the proposed approach is that the principal axis of a model often coincides with its axis of symmetry.

The SymPan+ method is based on the same principle as for SymPan, plus on the observation, that many common 3D objects possess at least one plane of symmetry. Detecting this symmetry plane is, therefore, the first step of our method and gives the

first principal axis of the model. The surface of the 3D model is initially projected on the lateral surface of a selected circumscribed cylinder, aligned with one of the principal axes of space. The 3D model is iteratively rotated accordingly and the reflective symmetry scores of its projections are computed. The axis of the projection cylinder, along with the axis of maximum reflective symmetry, as this is computed on the set of panoramic projections, define the 3D model's plane of symmetry, the normal of which, defines one of the 3D model's principal axes. The other principal axes of the 3D model are estimated via the computation of the variance on the pixel values of the panoramic views.

3.2.1 Panoramic Views Representation

Let us define a projection cylinder as a cylinder whose axis is aligned with one of the principal axes of space (e.g. the z axis), as described by Papadakis et al. [89]. To obtain a panoramic view for a 3D model, we project the 3D model onto the lateral surface of a cylinder of radius R and height $H = 2R$, centered at the origin, with its axis parallel to one of the principal axes of space, see Fig. 13a. We set the value of R to $2 * d_{max}$ where d_{max} is the maximum distance of the model's surface from its centroid.

In the following, we parameterize the lateral surface of the cylinder using a set of points $s(\phi, y)$ where $\phi \in [0, 2\pi]$ is the angle in the XY plane, $y \in [0, H]$ and we sample the ϕ and y coordinates at rates $2B$ and B, respectively (we set $B = 180$). We sample the ϕ dimension at twice the rate of the y dimension to account for the difference in length between the perimeter of the cylinder's lateral surface and its height. Thus we obtain the set of points $s(\phi_u, y_v)$ where $\phi_u = u * 2\pi/(2B)$, $y_v = v * H/B$, $u \in [0, 2B - 1]$ and $v \in [0, B - 1]$. These points are shown in Fig. 13b.

We shall next determine the value at each point $s(\phi_u, y_v)$ of the panoramic view. The computation is carried out iteratively for $v = 0, 1, ..., B - 1$, each time considering the set of coplanar $s(\phi_u, y_v)$ points, i.e. a cross section v of the cylinder at height y_v and for each cross section we cast rays from its center c_v in the ϕ_u directions.

The cylindrical projections are used to capture two different characteristics of a 3D model's surface; (i) the position of the model's surface in 3D space, (referred to as **Spatial Distribution Map** or SDM), and (ii) the orientation of the model's surface, (referred to as **Normals' Deviation Map** or NDM). To capture these characteristics we use two kinds of

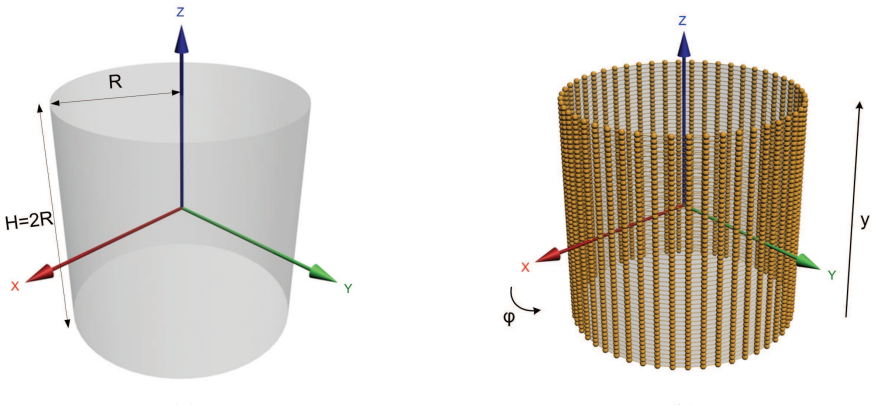


Figure 13: (a) A projection cylinder for the acquisition of a 3D model's panoramic view and (b) the corresponding discretization of its lateral surface to the set of points $s(\phi_u, y_v)$

cylindrical projections $s_1(\phi_u, y_v)$ and $s_2(\phi_u, y_v)$.

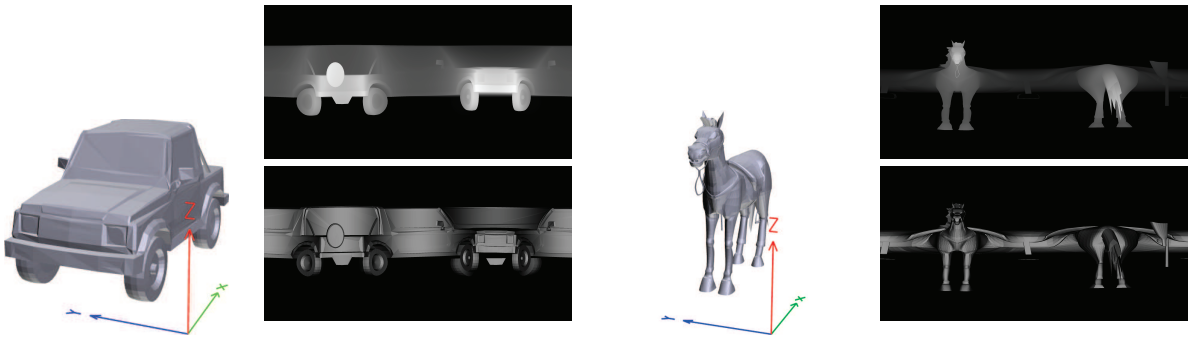
To capture the position of the model's surface, for each cross section at height y_v we compute the distances from c_v of the intersections of the model's surface with the rays at each direction ϕ_u. Let $pos(\phi_u, y_v)$ denote the distance of the furthest from c_v point of intersection between the ray emanating from c_v in the ϕ_u direction and the model's surface; then $s_1(\phi_u, y_v) = pos(\phi_u, y_v)$. The value of a point $s(\phi_u, y_v)$ lies in the interval $[0, R]$, where R denotes the radius of the cylinder.

To capture the orientation of the model's surface, for each cross section at height y_v we compute the intersections of the model's surface with the rays at each direction ϕ_u and measure the angle between a ray and the normal vector of the triangle that is intersected. To determine the value of a point $s_2(\phi_u, y_v)$ we use the cosine of the angle between the ray and the normal vector of the furthest from c_v intersected triangle of the model's surface. If $ang(\phi_u, y_v)$ denotes the aforementioned angle, then the values of the $s(\phi_u, y_v)$ points are given by $s_2(\phi_u, y_v) = |\cos(ang(\phi_u, y_v))|^n$.

We take the nth power of $|\cos(ang(\phi_u, y_v))|$, where $n \geq 2$, since this setting enhances the contrast of the produced cylindrical projection. We have experimentally found that setting n to a value in the range $[4, 6]$ gives the best results. Also, taking the absolute value of the cosine is necessary to deal with inconsistently oriented triangles along the model's surface.

A cylindrical projection can be viewed as a 2D gray-scale image where pixels correspond to the $s_k(\phi_u, y_v)$ intersection points in a manner reminiscent of cylindrical texture mapping

and their values are mapped to the $[0, 1]$ space. In Figs. 14a and 14c, we show two example 3D models aligned with the z axis. In Figs. 14b and 14d the unfolded visual representations of its corresponding cylindrical projections $s_1(\phi_u, y_v)$ and $s_2(\phi_u, y_v)$ are given, respectively.



(a) (b) (c) (d)

Figure 14: (a),(c) Two example 3D models and (b),(d) their corresponding cylindrical projections capturing the position of the surface (top row) and the orientation of the surface (bottom row), on the z axis.

3.2.2 Image Reflective Symmetry Estimation

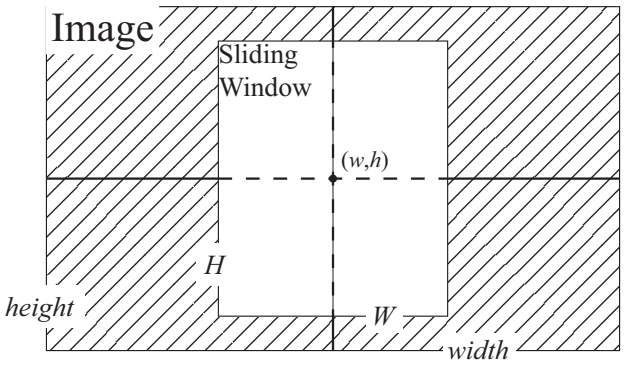


Figure 15: Illustration of the sliding window parameters used for the estimation of the reflective symmetry found in a 2D image

To measure the reflective symmetry of a 2D grayscale image I, a method similar to the one proposed in [143] is followed. A sliding window of width W and height H is defined. The sliding window is initially positioned at the center of the image, see Fig. 15. At each window position, the reflective symmetry measured at its central column w, is:

$$Sym(w) = 1 - \frac{1}{2m} \sum_{h=\frac{height}{2}-m}^{\frac{height}{2}+m} SymDiff(w,h) \tag{7}$$

$$SymDiff(w,h) = \frac{1}{n} \sum_{x=1}^{n} |(w-x,h) - (w+x,h)| \tag{8}$$

where (w,h) denotes the image pixel located at column w and row h. We set $n = 10\%$ of the image's *width* and $m = 40\%$ of the image's *height*, resulting in a sliding window of dimensions $W = 0.20 * \textit{width}$ and $H = 0.80 * \textit{height}$.

The process is repeated for every sliding window position on the image and the maximum $Sym(w)$ value is stored as the image's symmetry score. Fig. 16 illustrates two example panoramic view images with the symmetry columns indicated and the corresponding symmetry score graphs, as extracted by the proposed method. We thus define the maximum symmetry score for a 2D image I and the image column where this score is attained as:

$$S(I) = \max\{Sym(w) | w \in 1 : width\} \tag{9}$$

$$S_{index}(I) = \arg\max\{Sym(w) | w \in 1 : width\} \tag{10}$$

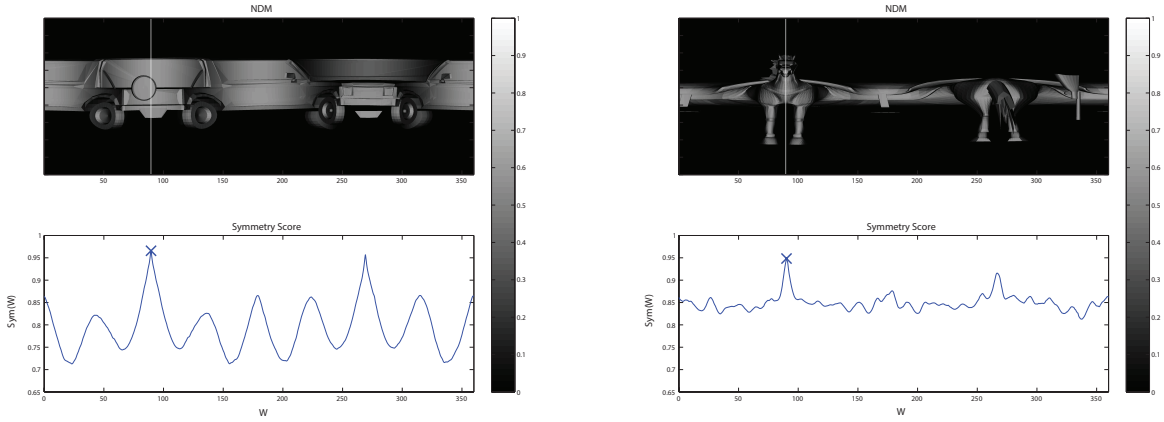


Figure 16: Top row: example panoramic view images with symmetry column indicated and, bottom row: the corresponding symmetry score graphs as extracted by the proposed method.

3.2.3 SymPan Pose Normalization

To achieve alignment between a 3D model and a projection cylinder, we compute two equally weighted factors: (i) a measure of how parallel the surface of the 3D model is to the lateral surface of the cylinder, as given by the mean value of the NDM and (ii) the degree of reflective symmetry established by the NDM:

$$D = \overline{NDM} + S(NDM) \tag{11}$$

where $\overline{NDM}$ stands for the mean value of the NDM and S(NDM) measures the reflective symmetry of the NDM.

According to Euler's Rotation Theorem, to reach any target frame, a specific sequence of three rotations, that are described by three angles is required. The first two rotations establish a common principal rotation axis between the source and target frames (also known as the 'line of nodes'). The third rotation, about the principal rotation axis, aligns the remaining axes of the reference and target frames [52, 79].

To this end, we use an octree search strategy for the estimation of the two rotations required for the 3D model's principal axis to become aligned with the cylinder's axis:

Rotate the 3D model around the secondary and tertiary principal axes of space by both θ° and $-\theta^\circ$ and compute the symmetry measure D for the resulting NDM image. Set the 3D model's new orientation to the one which results in the maximum value of D. Set $\theta = \theta/2$ and repeat the search process until $\theta = 0.125^\circ$. Initially $\theta = 90^\circ$. During our experiments the algorithm always converged within 20 steps; we used a maximum of 30 steps to guarantee termination.

After the alignment of the 3D model's principal axis, a search on the 3D model's NDM for its secondary principal axis is carried out, based solely on the 3D model's reflective symmetry characteristics. Again, $Sym(w)$ is computed for every column of the aligned model's NDM and the position of the maximum $Sym(w)$ value is stored as the 3D model's secondary principal axis:

$$S_{index}(NDM) = \arg\max\{Sym(w)|w \in 1 : width\} \tag{12}$$

The model is then rotated around its primary principal axis so that its secondary

principal axis becomes aligned with the corresponding secondary principal axis of space (i.e. the x axis).

3.2.4 SymPan+ Pose Normalization

3.2.4.1 3D Model Symmetry Plane Estimation The estimation of the plane of symmetry of the 3D model is based on the reflective symmetry of its corresponding panoramic view. Once the plane of symmetry is defined, the first principal axis of the model is also defined as the normal to the plane of symmetry.

Initially, a 3D model, having arbitrary pose, is normalized in terms of translation and scaling using the corresponding standard techniques described in ROSy pose normalization method. More specifically, translation normalization is achieved though the definition of the 3D model's centroid and the displacement of this centroid to the coordinate system origin. Consecutively, the 3D model is scaled so that it becomes exactly inscribed inside the unit sphere.

Since translation normalization has been performed, the plane of symmetry of the 3D object passes through the origin of the coordinate system. Our aim is to rotate the symmetry plane so that it includes the z axis; then the pane of symmetry will be detected in the panoramic image.

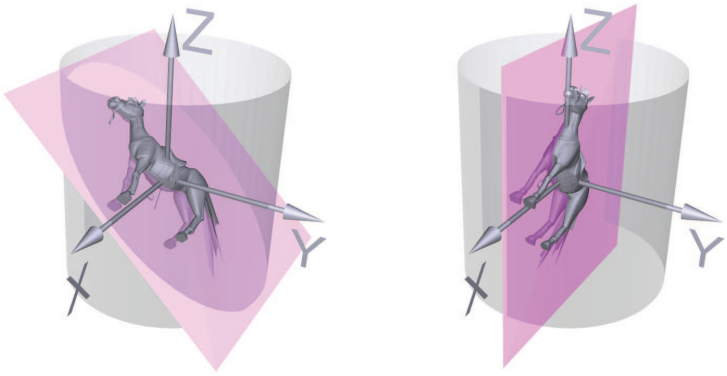


Figure 17: Aligning the symmetry plane normal with the XY plane

Consider the normal $\mathbf{vn}$ to the plane of symmetry. When this normal is included in the XY plane, then the plane of symmetry will include the z axis; a rotation around the x axis by $BestRot$ is sufficient for this to happen (Fig 17):

$$BestRot = \arg\max\{S(I_\theta)|\theta \in 0^\circ : 180^\circ\} \tag{13}$$

Where I_θ is the panoramic image produced by rotating the 3D model by θ degrees around the x axis. $S(I_\theta)$ is the corresponding symmetry score.

$S_{index}(I_\theta)$ along with the z axis then define the plane of symmetry.

In Fig. 18 the best rotation of the 3D model and its symmetry axis is illustrated. The corresponding graph of $S(I_\theta)$ score is shown and $BestRot$ is indicated. Furthermore, the NDM image and its symmetry score graph, resulting from $BestRot$ are also depicted. On both the NDM image and the graph is indicated the S_{index} point.

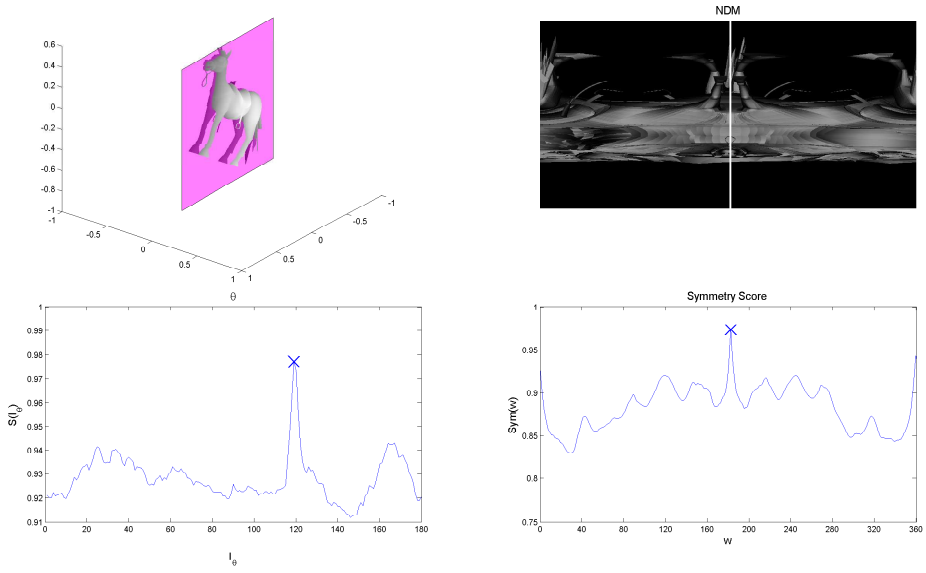


Figure 18: 3D model symmetry plane estimation. (left top) The 3D model with its plane of symmetry and (left bottom) the symmetry scores of the corresponding NDM images, computed under rotation θ. Marked is the rotation (in degrees) giving the maximum symmetry score. (right top) The NDM image of the best rotation, (right bottom) along with its symmetry score plot. Marked is the maximum symmetry score value and the corresponding symmetry column.

Before we proceed to the principal axes estimation, we shall prove that to transform normal vector **vn** so that it is included in the XY plane (or it becomes perpendicular to the normal to the XY plane, the z axis), given any rotation axis R, which is perpendicular to the z axis (i.e. the x and y axes), is dependent only on rotation angle θ.

Let us consider that the normal to the symmetry plane has an arbitrary rotation in space, $\mathbf{vn} = [x', y', z']$. Furthermore, let us define a matrix describing a rotation of angle θ

around an axis perpendicular to the target vector. This is a combination of two principal rotations, a first one around axis z by an angle γ, defining the angular difference between the arbitrary rotation axis and a known axis in space (i.e. the x-axis), and then the actual rotation around the known axis, by angle θ:

$$R = R_z Rx = \begin{bmatrix} \cos(\gamma) & -\sin(\gamma) & 0 \\ \sin(\gamma) & \cos(\gamma) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & -\sin(\theta) \\ 0 & \sin(\theta) & \cos(\theta) \end{bmatrix}$$
$$= \begin{bmatrix} \cos(\gamma) & -\sin(\gamma)\cos(\theta) & \sin(\gamma)\sin(\theta) \\ \sin(\gamma) & \cos(\gamma)\cos(\theta) & -\cos(\gamma)\sin(\theta) \\ 0 & \sin(\theta) & \cos(\theta) \end{bmatrix}$$

Given rotation R, the dot product of $\mathbf{vn}$ and the z axis should equal to zero, in order to be perpendicular:

$$R\mathbf{vn} \cdot z = 0$$

$$\begin{bmatrix} \cos(\gamma) & -\sin(\gamma)\cos(\theta) & \sin(\gamma)\sin(\theta) \\ \sin(\gamma) & \cos(\gamma)\cos(\theta) & -\cos(\gamma)\sin(\theta) \\ 0 & \sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 0 \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} x'\cos(\gamma) - y'\sin(\gamma)\cos(\theta) + z'\sin(\gamma)\sin(\theta) \\ x'\sin(\gamma) + y'\cos(\gamma)\cos(\theta) - z'\cos(\gamma)\sin(\theta) \\ y'\sin(\theta) + z'\cos(\theta) \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 0 \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$z(y'\sin(\theta) + z'\cos(\theta)) = 0 \tag{14}$$

Equation 14 shows that given a rotation of vector $\mathbf{vn}$ around an axis perpendicular to axis z (i.e. the x axis), an angle θ exists, which makes the two vectors perpendicular.

3.2.4.2 3D Model Principal Axes Estimation The normal to the plane of symmetry defines one of the principal axes of the 3D model, but the remaining two principal axes have yet

to be estimated. The 3D model can thus be rotated so that its symmetry plane coincides with one of the principal planes of space (e.g. the XY plane).

To complete the rotation normalization task, the 3D model is projected onto the surface of a projection cylinder whose axis is one of the principal axes of space, perpendicular to the symmetry plane's normal. The 3D model is iteratively rotated around the normal axis to the symmetry plane and the corresponding SDM images are calculated. For each SDM image, the variance of its pixel values is computed and the rotation that minimizes this variance, is defined as the rotation which aligns the principal axis of the 3D model to the axis of the projection cylinder (see Fig. 19).

Since the SDM image represents the distance between the surface if the 3D model and the surface of the projection cylinder, low variation of its pixel values can be interpreted as higher distance between the surface of the 3D model and the surface of the cylinder (i.e. the 3D model is parallel to the surface of the cylinder), whereas high variation denotes that certain areas of the 3D model lie closer to the cylinder than others (i.e. the 3D model is, close to, perpendicular to the cylinder's axis).

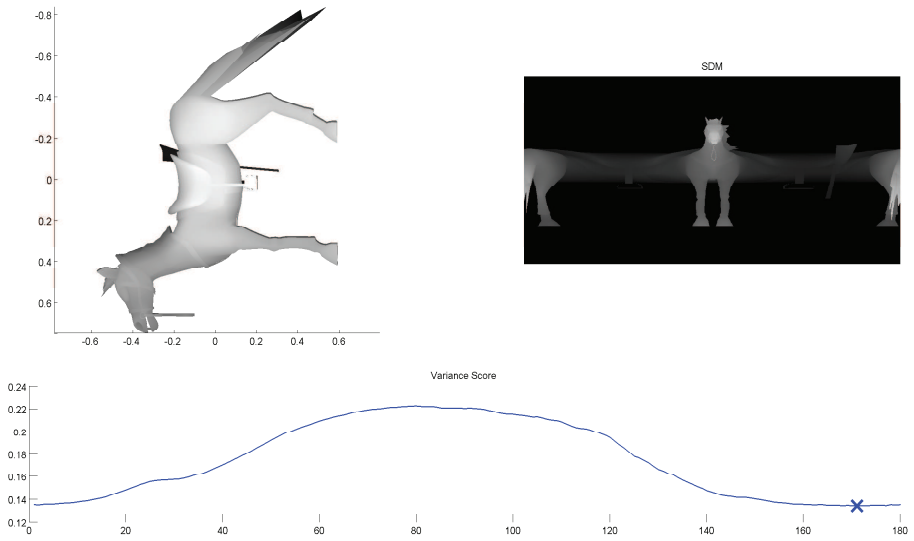


Figure 19: 3D model principal axes estimation, illustrated are the axis aligned 3D model, the corresponding SDM image along with its pixel value variance score. Marked is the minimum variance score.

The complete algorithm for the proposed pose normalization methodology is outlined in Algorithm 2.

Algorithm 2 Object pose normalization based on Reflective Symmetry and Panoramic Views

Input: M; ▷ Input 3D model M

Translation normalization of the centroid of M to the origin of the coordinate system;
Scale normalization of M to the unit sphere;

$BestSym \leftarrow 0$; ▷ Best symmetry value so far
$MinVar \leftarrow -\inf$; ▷ Minimum 2D image pixel values variance so far
$CylAxis \leftarrow (0, 0, 1)$; ▷ Projection cylinder's axis
$RotAxis \leftarrow (1, 0, 0)$; ▷ Selected rotation axis

for $\theta \leftarrow 0^\circ$ to 180°, step 1° **do**
    $M' \leftarrow$ ROTATE$(M, RotAxis, \theta)$; ▷ 3D model rotation transformation (in degrees)
    $NDM \leftarrow$ PROJECTION$(M', CylAxis)$; ▷ 3D model cylindrical projection transformation
    $SymNDM \leftarrow S(NDM)$; ▷ Equation 9
    if $SymNDM \geq BestSym$ **then**
        $BestRot \leftarrow \theta$ ▷ Store θ giving the maximum $Sym(NDM)$ value
        $BestSym \leftarrow SymNDM$; ▷ Equation 10
        $PosSym \leftarrow S_{index}(NDM)$;
    end if
end for

$SymPlane(1) \leftarrow (0, 0, 0)$; ▷ Define the 3D model's plane of symmetry by 3 points in 3D space
$SymPlane(2) \leftarrow CylAxis$;
$SymPlane(3) \leftarrow (0, \cos(PosSym), \sin(PosSym))$;
$M \leftarrow$ ROTATE$(M, RotAxis, BestRot)$ ▷ Based on the best rotation found...
$M \leftarrow$ TRANSFORM_BASIS$((1, 1, 0), SymPlane)$; ▷ ...transform 3D model's plane of symmetry to principal plane

for $\theta \leftarrow 0^\circ$ to 180°, step 1° **do**
    $M' \leftarrow$ ROTATE$(M, CylAxis, \theta)$; ▷ The new rotation axis is perpendicular to the selected principal plane
    $SDM \leftarrow$ PROJECTION$(M', RotAxis)$; ▷ The new projection axis is contained in the selected principal plane
    $VarSDM \leftarrow var(SDM)$; ▷ 2D image pixel values variance
    if $VarSDM \leq MinVar$ **then**
        $MinVar \leftarrow VarSDM$;
        $PosVar \leftarrow \theta$;
    end if
end for
$M_{out} \leftarrow$ ROTATE$(M, CylAxis, PosVar)$;

Output: M_{out}; ▷ Return the pose normalized 3D model

4 Non-Rigid 3D Object Retrieval

The unapparent connection
is more powerful than the apparent one

– Heraclitus

Figure 20: Example models from the class 'Centaur' of the TOSCA dataset, color-coded with the corresponding conformal factors.

In this chapter, the proposed non-rigid 3D object retrieval methodology will be detailed. The motivation behind this methodology, named ConTopo, is to combine, in a unified manner, the properties of both geometrical (i.e. conformal geometry) and topological descriptors (i.e. graph-based topological information), in order to achieve elevated performance. While graph-based methods are robust to nonrigid object deformations, they require intensive computation which can be reduced by the use of appropriate representations, addressed through geometry-based methods. In this respect, a topological structure of a 3D model is constructed with the aid of discrete conformal factors. Further-

more, a string matching strategy for the comparison of graphs which describe 3D objects, is proposed [106].

Initially, let us define a triangular mesh $M = \{V, F, E\}$, represented by the set of vertices V, the set of faces F and the set of edges E connecting neighboring vertices. Optionally, the set of the boundary vertices B could be defined. For a vertex v_i, $V_1(i)$ denotes the 1-ring set of adjacent vertices to v_i and $F_1(i)$ denotes the 1-ring set of adjacent faces to v_i.

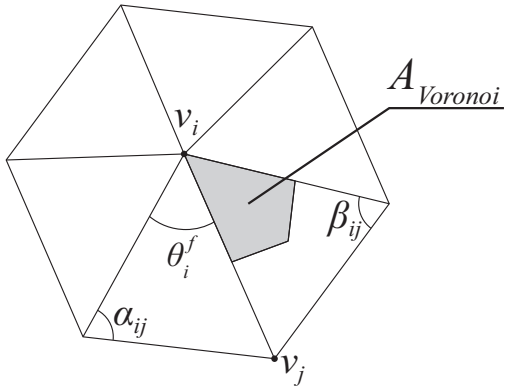


Figure 21: Adjacent faces of vertex v_i at 1-ring neighborhood.

4.1 The Discrete Conformal Factor

Ben-Chen and Gotsman in [11] have introduced the discrete conformal factor for a 3D mesh, which is used as a non-rigid shape descriptor. The conformal factor ϕ_i at a vertex v_i of the triangular 3D mesh M is the solution to the following discrete linear equation (see Fig. 20):

$$\phi_i = \frac{k_i^{targ} - k_i^{orig}}{\mathcal{L}(v_i)} \tag{15}$$

where k_i^{orig} is defined as the discrete Gaussian Curvature at vertex v_i of the triangular 3D mesh:

$$k_i^{orig} = \begin{cases} 2\pi - \sum\limits_{f \in F_1(i)} \theta_i^f, & v_i \in V \backslash B \\ \pi - \sum\limits_{f \in F_1(i)} \theta_i^f, & v_i \in B \end{cases} \tag{16}$$

The first case of equation (16) is used for vertices of the triangular mesh whose 1-ring of adjacent faces is closed, whereas the second case is used for vertices that belong to the boundary of the triangular mesh (if such boundary exists). θ_i^f is the angle near vertex v_i of face f (see Fig. 21).

In (15), k_i^{targ} denotes the uniform Gaussian Curvature:

$$k_i^{targ} = \left(\sum_{j \in V} k_j^{orig} \right) \frac{\sum_{f \in F_1(i)} \frac{1}{3} area(f)}{\sum_{f \in F} area(f)} \tag{17}$$

k_i^{targ} assigns to each vertex a portion of the total curvature of the mesh.

In (15), $\mathcal{L}(v_i)$ denotes the discrete Laplace - Beltrami operator, at vertex v_i, with cotangent weights, defined in [77] as:

$$\mathcal{L}(v_i) = \frac{1}{2A_{Mixed}} \sum_{j \in V_1(i)} (\cot \alpha_{ij} + \cot \beta_{ij}) |v_i - v_j| \tag{18}$$

where A_{Mixed} denotes the mesh surface area in a 1-ring neighborhood around a vertex v_i, which is computed as shown in Algorithm 3 (see also Fig. 21).

Algorithm 3 Pseudo-code for the calculation of the A_{Mixed} surface area of vertex v_i on an arbitrary mesh

```
1: A_Mixed = 0
2: for f ∈ F_1(i) do
3:     if f is non-obtuse then
4:         A_Mixed+ = A_Voronoi(i)
5:     else
6:         if the angle of f at v_i is obtuse then
7:             A_Mixed+ = area(f)/2
8:         else
9:             A_Mixed+ = area(f)/4
10:        end if
11:    end if
12: end forreturn (A_Mixed)
```

$area(f)$ denotes the triangular area of face f, which is computed using Heron's formula. $A_{Voronoi}$ denotes the surface area contribution of a single non-obtuse triangle in $F_1(i)$ (see Fig. 21):

$$A_{Voronoi} = \frac{1}{8} \sum_{j \in F_1(i)} (\cot \alpha_{ij} + \cot \beta_{ij}) \|v_i - v_j\|^2 \tag{19}$$

where α_{ij} and β_{ij} denote the two angles opposite to the common edge $\widehat{v_i v_j}$ (see Fig. 21).

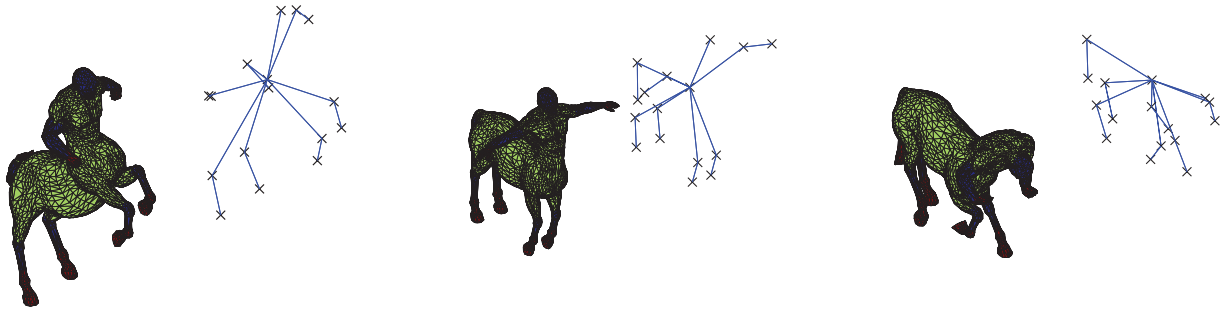


Figure 22: Example partitionings and the corresponding graphs of 3D meshes from the class 'Centaur' using the discrete conformal factor.

4.2 Graph Construction

A graph can be used as a topological map that represents the skeletal structure of an object with arbitrary dimensions. An example of methods for the characterization of 3D mesh topological information is the Reeb Graph [57, 31, 13]. Here, in similar manner, we will define a graph, that captures the topological structure of an arbitrary 3D mesh.

A graph is a representation of a set of objects where some pairs of objects are connected by links. The interconnected objects are represented by mathematical abstractions called nodes (or vertices), and the links that connect some pairs of nodes are called edges [126].

Each node of the graph represents a unique connected component, while each edge of the graph describes the relation between adjacent connected components. Each connected component is composed of 3D mesh faces that have the same label and are also pathwise-connected (i.e. there exists a path that connects each face of the connected component to every other face that belongs to the same connected component). A simple example of a multi-thresholding criterion is the height function [36, 111, 120, 122]. Other options include functions that measure the distance between each vertex on the surface of a 3D mesh and an approximation of its center of mass and/or using other geodesic

properties of the 3D mesh.

In this work we have used discrete conformal factors [11] as a multi-thresholding criterion due to their ability to identify the protrusive parts in a mesh (Fig. 22). The faces of the 3D mesh are partitioned based on a linear multi-thresholding of the values of the discrete conformal factor, thus splitting the mesh into a set of connected components (see also Fig. 23a).

At this point, we note that the computation of the graph is based on the triangle set F of the 3D mesh and not on the corresponding vertex set. Therefore, it is necessary to map the magnitude of the discrete conformal factors from the vertices to the faces of the triangulated mesh. To achieve this, for each triangular face we aggregate the conformal factor values of the three vertices ϕ_i^f, weighted by their barycentric coordinates, with respect to the corresponding triangular face's centroid:

$$\phi^f = \sum_{i=1}^{3} b_i^f \phi_i^f, \quad f \in F \tag{20}$$

where b_i^f, denotes the barycentric coordinates of vertex i that belongs to triangle f, with regard to its centroid.

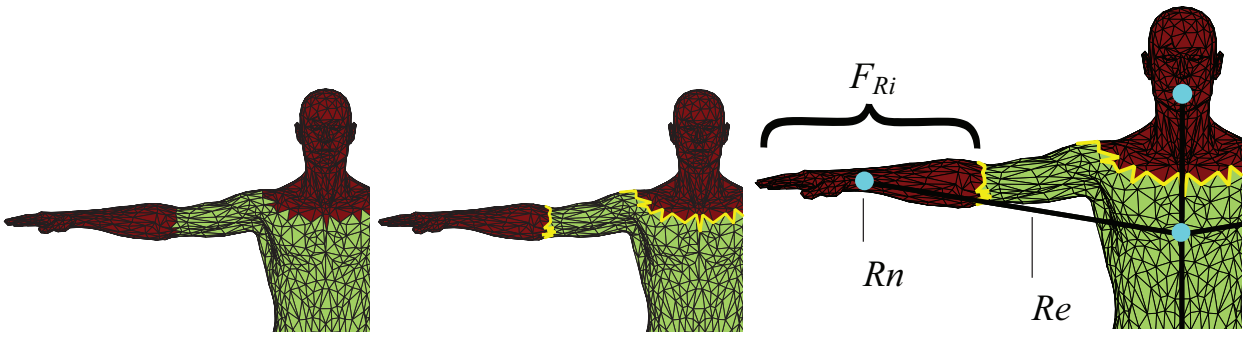


Figure 23: Detailed steps for the construction of the graph for a 3D mesh. Example Rn, Re, and F_{Ri} are illustrated.

The algorithm for the computation of the graph is composed of the following steps:

- Quantization of the 3D mesh surface into q discrete intervals, based on conformal factors as the multi-thresholding criterion. (Fig. 23a).
- Estimation of boundary edges between adjacent connected components (Fig. 23b).

- Creation of graph edges between neighboring nodes that represent adjacent connected components (Fig. 23c).

The outcome of the graph construction is:

Rn**:** The nodes of the graph.

Re**:** The edges are represented by an $N \times N$ adjacency matrix, where N is the cardinality of the sets of nodes.

F_{Ri}**:** The set of faces (triangles) corresponding to the graph node Ri (partition).

F_{Ri}^{area}**:** The area of the faces in F_{Ri}. F_{Ri}^{area} is normalized with respect to the total area of the 3D mesh.

B_{RiRj}**:** The set of common border vertices between the partitions corresponding to nodes Ri and Rj.

$\bar{\phi}_{Ri}$**:** The mean value of the conformal factor for the partition corresponding to node Ri.

In our implementation we have selected $q = 8$, which creates an eight-level graph. An example of a quantized mesh along with its corresponding graph is illustrated in Fig. 24. This choice has been experimentally determined as it yields good results while simultaneously preserving acceptable computational speed. During the graph construction procedure, if any *small* connected components are found, then these sets are removed as outliers. In our implementation we define a connected component as *small*, if it is composed of less than $(3/q)$% of the total 3D mesh surface.

Once the graph of the 3D mesh has been constructed, one final step is required before the matching can be performed. The connected partitions of the graph need to be ordered according to their distance from the *core* partition of the 3D mesh. To define the *core* partition and perform the ordering, an all-pairs shortest path algorithm, (i.e. the Floyd-Warshall [44] algorithm) is applied on the node (Rn) and edge (Re) sets of the graph. This algorithm calculates the graph distance of each node to all others (i.e. least number of nodes that need to be traversed to reach the target node). Therefore, the partition for which the corresponding graph node has the minimum graph distance from all other nodes of the graph is defined as the *core* partition; every other partition is ordered according to its

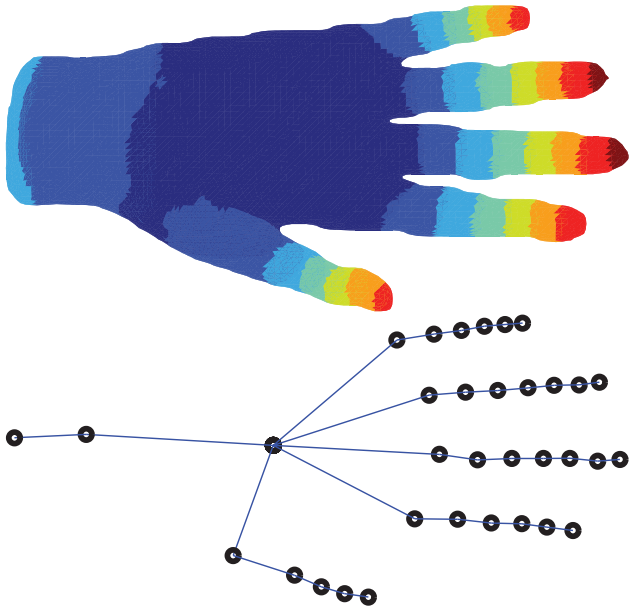


Figure 24: Illustration of an eight-level quantized mesh and the corresponding graph.

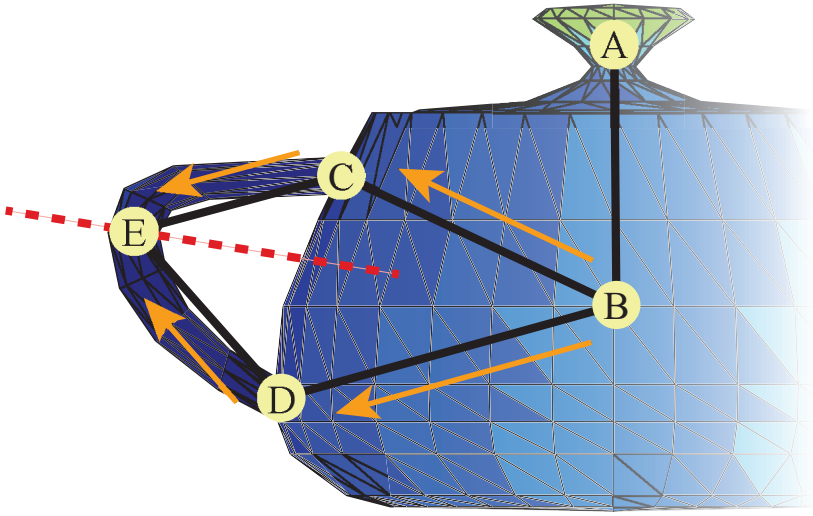


Figure 25: Illustration of the way a cycle is *broken* and considered as two different branches by the shortest path algorithm.

distance from the *core* partition. Since the all-pairs shortest path algorithm is applied on the graph of the 3D mesh and not on the full mesh, this procedure is very efficient.

The result of this procedure is an acyclic graph. Note that if any cycles exist these are detected and *broken* by the shortest path algorithm. This is achieved by considering the node of the circle which is most distant from the *core* partition as a leaf node and each path leading to it as a different branch (see Fig. 25).

Every path from the root node of the acyclic graph down to each of its leaves is considered as a *string*, used in the mesh matching algorithm described in the following section

(see Figs 5 and 7).

4.3 Mesh Matching

Mesh matching compares both geometrical and topological features as a measure of similarity between two 3D meshes in a unified manner. During matching, the topological equivalence between the graphs of two 3D meshes is examined and enhanced by node-to-node comparison of geometrical features.

The matching procedure described in this section is based on string matching. Each ordered path, of graph nodes, that extends from the *core* partition of the 3D mesh down to each of its articulations is considered a *string*. Furthermore, besides the ordered connectivity of the string (graph) nodes, a number of features are also attached to them, which will be used for the geometrical matching.

To address the matching problem between two 3D meshes that possibly have: (i) different numbers of strings and (ii) large string length variations, we consider the following algorithm:

Let M_1, $G_1 = (Rn_1, Re_1)$ and M_2, $G_2 = (Rn_2, Re_2)$ be two 3D meshes and their graphs respectively. Also, let $\mathbf{S_1}$ be the set of m strings derived from the graph of M_1 and $\mathbf{S_2}$ be the set of n strings derived from the graph of M_2, respectively. Each string is a sequence of nodes i, where $i > 0$. The set of features of each node si of the string is denoted as $si.ftrs$.

For each string in $\mathbf{S_1}$, the features of its nodes are compared to the features of the corresponding nodes of every string in $\mathbf{S_2}$. In the case of unequal lengths between two compared strings, the features of the exceeding nodes are added as is to the string difference, thus penalizing that the particular articulations are not of the same topological structure. This procedure is mathematically formulated below and graphically illustrated in Fig. 26.

$$D(p,q) = \frac{\left(\begin{array}{l} \sum_{k=1}^{\text{length}(p)} |p[k].ftrs - q[k].ftrs| + \\ \\ \sum_{l=\text{length}(p)+1}^{\text{length}(q)} q[l].ftrs \end{array} \right)}{\text{length}(q)} \tag{21}$$

where p, q denote the two strings compared, assuming (without loss of generality) length(q)$\geq$length(p),

where length is the number of nodes in the corresponding tree path.

The specific feature set that we used consists of:

1. Area of corresponding partition, normalized with respect to the area of the full mesh (F_{Ri}^{area}).

2. Mean conformal factor value of corresponding partition ($\bar{\phi}_{Ri}$).

3. The geodesic length between the common borders of the current partition and the adjacent ones, where adjacency is considered in string order.

The last feature can achieve isometric invariance (i.e. preservation of lengths before and after the transformations, or across shapes) [10]. For each two adjacent partitions of the 3D mesh, the set of vertices that are common to their faces define their common border vertex set (B_{RiRj}). The geodesic measure that we use is defined as the mean geodesic distance between each two consecutive common border vertex sets. The method used for the computation of the geodesic distances is the Fast Marching method by Sethian [6, 93]. To alleviate any inconsistencies due to variations in the resolution of the 3D mesh, the

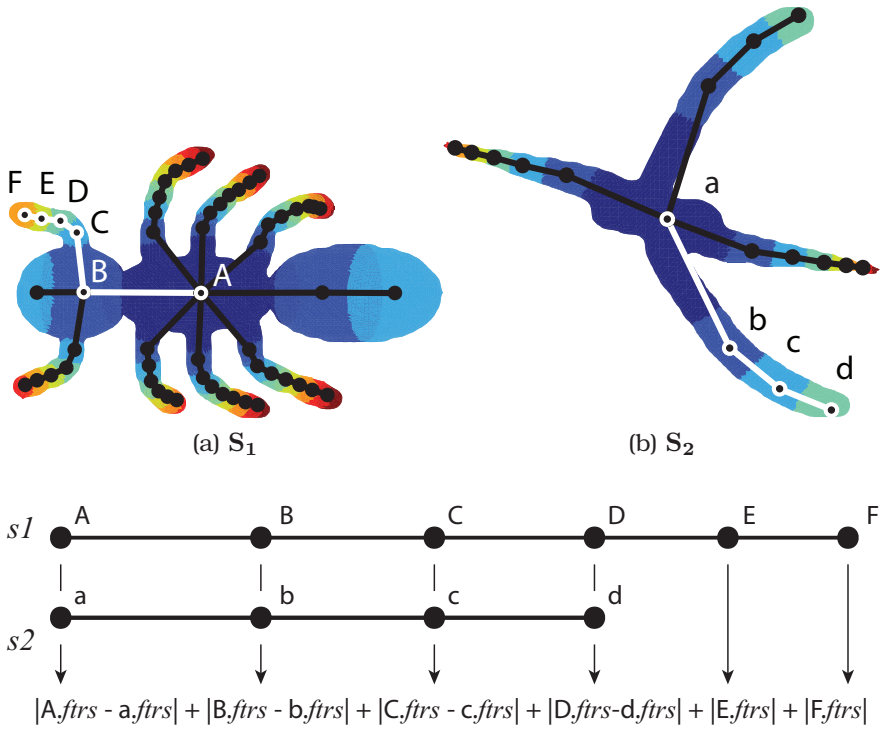


Figure 26: Feature comparison between two strings that are composed of different numbers of nodes. String $s1 \in \mathbf{S_1}$ is marked in 26a and string $s2 \in \mathbf{S_2}$ is marked in 26b

above measure is normalized with respect to the sum of geodesic distances between the common border vertex sets of the longest string.

Once all the strings in $\mathbf{S_1}$ have been compared to all strings in $\mathbf{S_2}$, an $m \times n$ matrix containing the difference between their features, is created (*String Difference Matrix* - SDM). If $m \neq n$ we further penalize that the two 3D meshes do not have the same number of articulations extruded from their *core* partition. Therefore, the smallest dimension of the SDM is padded with elements containing a penalizing factor (PF), set as the difference between the number of nodes in each graph divided by their sum:

$$PF = |m - n|/(m + n) \tag{22}$$

The final similarity measure between M_1 and M_2 is computed by applying the Hungarian algorithm [68] on the padded SDM. The Hungarian algorithm is a combinatorial optimization algorithm which solves the assignment problem and returns the minimum cost (ranging in $[0, \infty)$), which in this case is taken to be the difference between the two 3D meshes.

The mesh matching algorithm is summarized in Algorithm 4.

Algorithm 4 String Matching Algorithm

Input: $\mathbf{S_1}, \mathbf{S_2}$ ▷ String sets

$m = card(\mathbf{S_1}), n = card(\mathbf{S_2})$

$MaxStrings = max(m, n)$ ▷ maximum # of strings of the two 3D meshes.

$SDM[MaxStrings, MaxStrings] = PF$ ▷ see eq.22

for $i = 1 \rightarrow m$ **do**

  for $j = 1 \rightarrow n$ **do**

    $SDM[i, j] = D(\mathbf{S_1}[i], \mathbf{S_2}[j])$ ▷ see eq.21

    ▷ $\mathbf{S_1}[i], \mathbf{S_2}[j]$ denote the i^{th}, receptively j^{th} string of sets S_1, S_2

  end for

end for

Diff = Hungarian(SDM) ▷ combinatorial optimization algorithm [68] **return** Diff

5 3D object retrieval based on 2D Range Image Queries

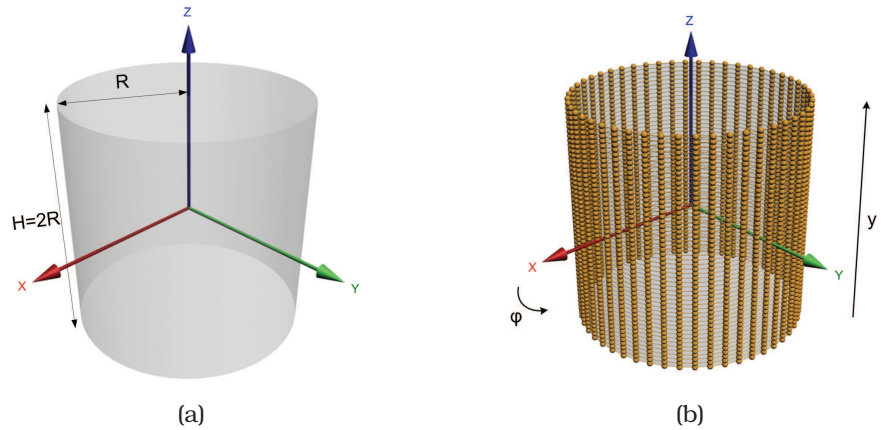
Since we can't change reality,
let's change the eyes which see reality

– N. Kazantzakis

In this chapter, the proposed methodology for 3D object retrieval via range image queries will be detailed. The motivation behind the proposed method, is to use a 2D image in order to query a database of 3D objects and bridge the representation gap between the two in an efficient manner. For the complete 3D objects of the database, shape description is achieved through a set of panoramic views and a Bag-of-Visual-Words model is built using SIFT features extracted from these views. To address the problem of partial matching, a histogram computation scheme, on the panoramic views, that represents local information by taking into account spatial context, is suggested. Furthermore, a number of optimization techniques are applied throughout the process, for enhancing the retrieval performance [107].

5.1 Panoramic Views Computation

For each 3D model of the database a number of panoramic views (or cylindrical projections) are extracted. These projections are computed on cylindrical axes that are perpendicular to and uniformly distributed over the surface of the 3D model's circumscribed sphere, in accordance with the PANORAMA [89] projection methodology. To obtain a panoramic view, we project the 3D model to the lateral surface of a cylinder of radius R and height $H = 2R$, centered at the origin with its axis parallel to one of the selected axes (in this



(a) (b)

Figure 27: 27a A projection cylinder for the acquisition of a 3D model's panoramic view and 27b the corresponding discretization of its lateral surface to the set of points $s(\phi_u, y_v)$

example the principal axis z, see Fig. 27a). We set the value of R to $2 * d_{max}$ where d_{max} is the maximum distance of the model's surface from its centroid. In the following, we parameterize the lateral surface of the cylinder using a set of points $s(\phi, y)$ where $\phi \in [0, 2\pi]$ is the angle in the XY plane, $y \in [0, H]$ and we sample the ϕ and y coordinates at rates $6B$ and B, respectively (we set $B = 360$). Thus we obtain the set of points $s(\phi_u, y_v)$ where $\phi_u = u * 2\pi/(2B)$, $y_v = v * H/B$, $u \in [0, 2B - 1]$ and $v \in [0, B - 1]$. These points are shown in Fig. 27b.

The next step is to determine the value at each point $s(\phi_u, y_v)$. The computation is carried out iteratively for $v = 0, 1, ..., B - 1$, each time considering the set of coplanar $s(\phi_u, y_v)$ points, i.e. a cross section of the cylinder at height y_v and for each cross section we cast rays from its center c_v in the ϕ_u directions. To capture the position of the model's surface, for each cross section at height y_v we compute the distances from c_v to the intersections of the model's surface with the rays at each direction ϕ_u.

Let $pos(\phi_u, y_v)$ denote the distance of the furthest from c_v point of intersection between the ray emanating from c_v in the ϕ_u direction and the model's surface; then $s(\phi_u, y_v) = pos(\phi_u, y_v)$. The value of a point $s(\phi_u, y_v)$ lies in the interval $[0, R]$, where R denotes the radius of the cylinder.

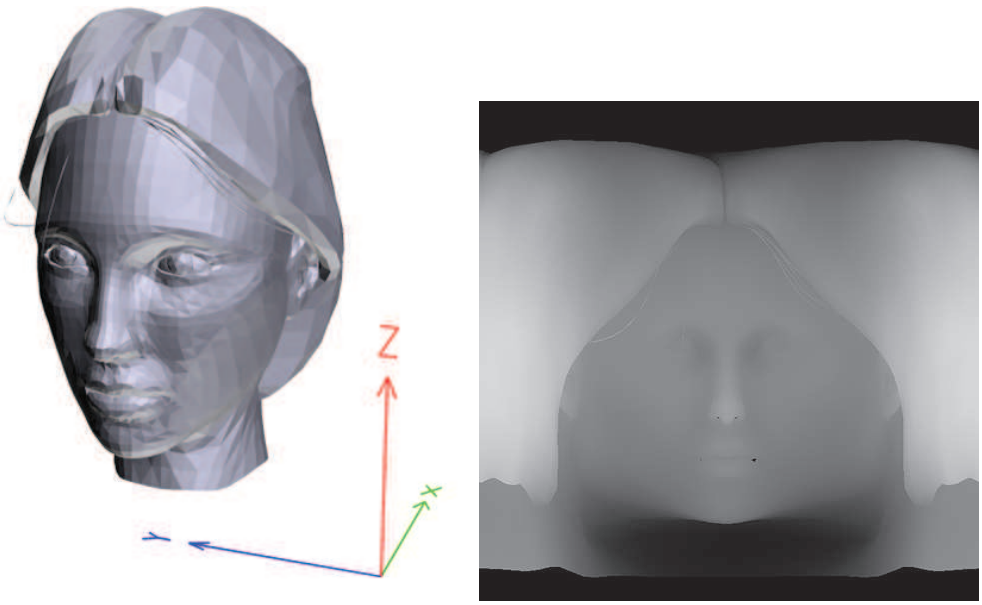
A cylindrical projection can be viewed as a 2D gray-scale image where pixels correspond to the $s(\phi_u, y_v)$ intersection points in a manner reminiscent of cylindrical texture mapping [124] and their values are mapped to the $[0, 1]$ interval. The number of extracted cylindrical projections for each complete 3D model is 60, which maintains acceptable

processing speed and coverage of the surface of the 3D model's circumscribed sphere. In Fig. 28a, we show an example 3D model and in Fig. 28b the unfolded visual representation of its corresponding cylindrical projection $s(\phi_u, y_v)$.

5.2 SIFT Descriptors Extraction

After the panoramic view extraction, the SIFT (Scale Invariant Feature Transform) [74] descriptor is calculated on the produced cylindrical depth images. The first step to the SIFT descriptor computation is the definition of an interest point set on the image, upon which the descriptors are calculated. The original implementation by Lowe, defines these interest points through the Difference of Gaussians (DoG) algorithm, which is geared towards enhancing the edges and other details present in the image. It has been experimentally found that the calculation of the SIFT descriptors over the complete image for a large number of randomly selected points [16, 17] (frequently defined as Dense SIFT/ DSIFT, in the literature), instead of selecting a limited number of interest points, yields better results in terms of retrieval accuracy.

At each interest point, an image descriptor is computed. The SIFT descriptor is defined as a position-dependent histogram of local gradient directions around the interest point.



(a) (b)

Figure 28: 28a An example 3D model and 28b its corresponding cylindrical projection on the z axis.

To achieve scale invariance of the descriptor, the size of this local neighbourhood is normalized in a scale-invariant manner. To achieve rotational invariance of the descriptor, the dominant orientation of the neighbourhood is determined from the orientations of the gradient vectors in this neighbourhood and is used for orienting the grid over which the position-dependent histogram is computed.

One recently proposed improvement technique for SIFT, by Arandjelovic and Zisserman [8], aims at enhancing the similarity measure used when comparing the descriptors (RootSIFT). The authors show that using a square root (Hellinger) kernel (also known as the Bhattacharyya's coefficient [12]) instead of the standard Euclidean distance measure, for the comparison of the SIFT descriptors (or SIFT histograms), increases performance. The intuition behind this proposal is based on the observation that Euclidean distance can be dominated by large bin values, whereas Hellinger distance is more sensitive to smaller bin values (See Fig. 29.

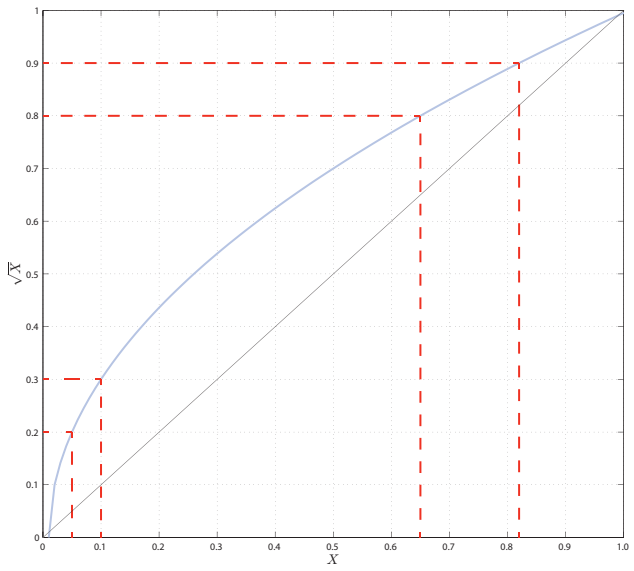


Figure 29: Comparison between the Euclidean distance and the Hellinger distance.

The Hellinger kernel for two L_1 normalized histograms, x and y is defined as:

$$H(x, y) = \sum_{i=1}^{n} \sqrt{x_i y_i} \tag{23}$$

To compare the SIFT vectors with a Hellinger kernel is a simple two-step algebraic manipulation (thus easy to implement in any existing SIFT implementation). First L_1

normalization of the SIFT vector, which originally has L_2 norm [74, 8], and then square-rooting each of its elements. Computing Euclidean distance in the feature map space is equivalent to Hellinger distance in the original space:

$$\sqrt{x}^T \sqrt{y} = H(x, y) \tag{24}$$

5.3 Bag-of-Visual-Words Modelling

Once the panoramic views extraction and the SIFT descriptor calculation steps are complete, the Bag-of-Visual-Words (BoVW) visual model for the database is built. In visual information retrieval, the BoVW model defines that each image contains a number of local *visual features*. Since every visual feature, or collection of similar visual features, can appear with different frequencies on each image, matching the visual feature frequencies of two images, achieves correspondence. In our case, the SIFT descriptors are defined as the BoVW model's visual features.

The basic step in the process of building the BoVW model for the 3D model database is the generation of a *codebook* (or a vocabulary), a collection of visual features that appear on each image. The codebook is generated by considering the visual features of a representative number of training database models (see Fig. 30). To achieve greater flexibility, rather than generating the codebook by selecting individual visual features of the training models, the corresponding panoramic views are clustered into several similar patches, the *codewords*. One simple method is performing *k-means* clustering [73] over all the visual features. *Codewords* are then defined as the centers of the learned clusters. The number of the clusters is the codebook size. Thus, each patch in an image is mapped to a certain codeword through the clustering process and the image can be represented by the histogram of the codewords.

After the codebook generation procedure, the next step is the description of the database 3D models using the corresponding codewords. In a similar manner, for each panoramic view of the database 3D model set, the visual features are computed and matched to their closest codewords, by comparing them to the corresponding clusters, generated in the previous step. Again, the *k-means* algorithm is used for matching. Note that the *k-means* clustering method makes use of the Euclidean distance for the com-

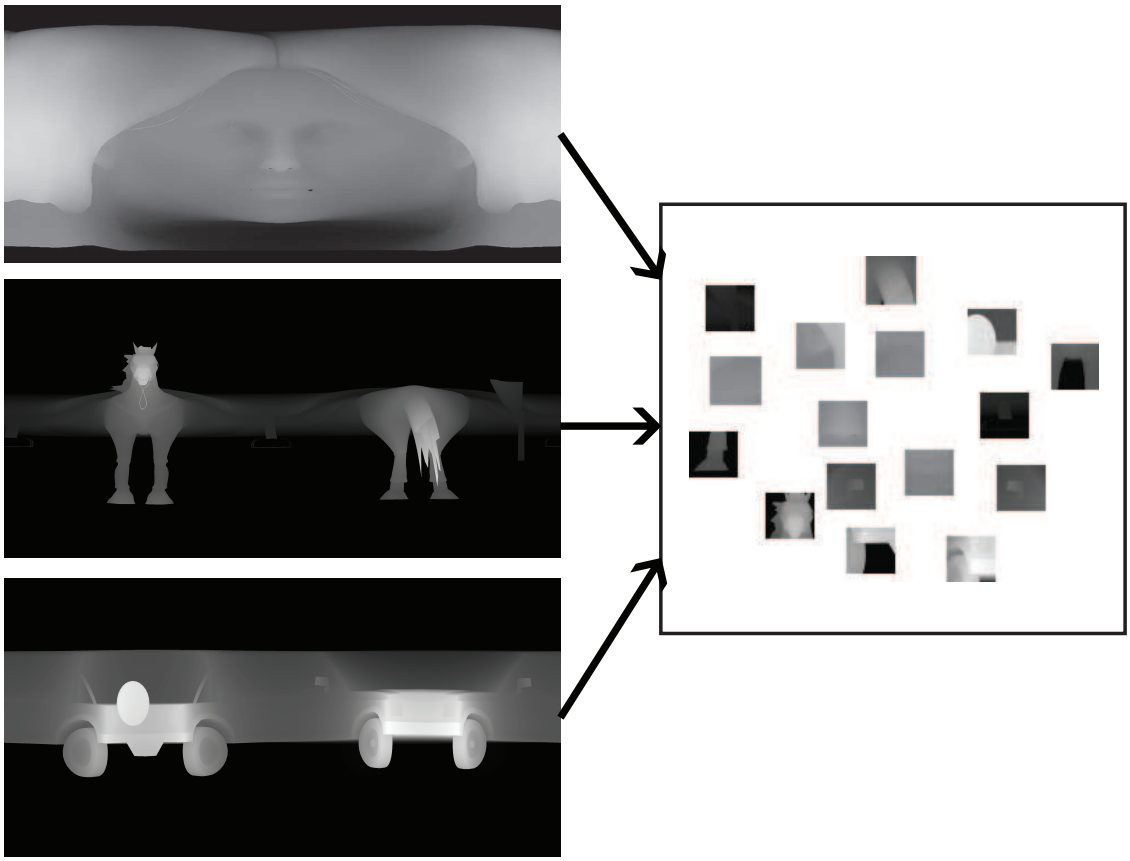


Figure 30: Sample BoVW codebook creation from a number of panoramic views extracted from 3 different corresponding 3D models

parison between the clusters and the visual features. According to the RootSIFT transformation, applied on the visual features, this results in using the Hellinger distance for the operation. The set of histograms describing the frequency of occurrence of the generated codewords, for each 3D model's panoramic views, is stored as the corresponding 3D model's signature.

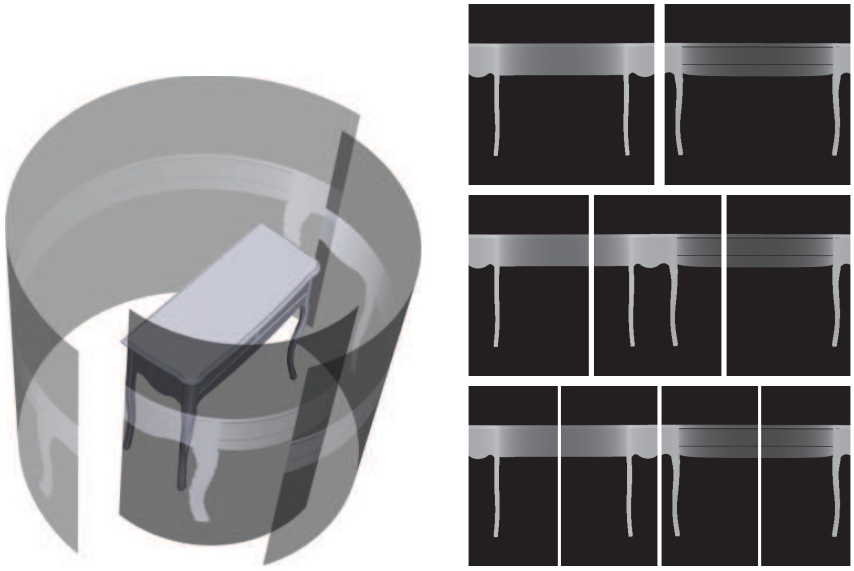
5.4 Spatial Histograms

In an extension of the standard Bag-of-Visual-Words model, described in the previous subsection, and targeting the matching between a complete 3D model and a range image representing a partial query, we have modified the histogram generation in the following manner. Since a panoramic view of a complete 3D model contains 360°of information, an attempt to match it to a query range image, which contains only a portion of that information, will produce poor results in the majority of cases. To alleviate this problem

we suggest a progressive partitioning scheme for the panoramic views of the database 3D models. Each panoramic view image is iteratively split along the horizontal dimension (width), which is perpendicular to the axis of the corresponding projection cylinder (see Fig. 31).

The spatial histograms are then computed for each resulting subpart of the image. For example, on a first level of progressive partitioning, the spatial histogram is computed on the complete panoramic view image of the 3D model. On the second level, the panoramic view image is partitioned once along the horizontal dimension and two spatial histograms are computed for the resulting left and right subimage, respectively. On the third level, the complete panoramic view image is partitioned twice along the horizontal dimension and three spatial histograms are computed for each of the three resulting subimages. The process continues until a certain level of progressive partitioning is reached, which in our case has been selected to be 6.

As the complete panoramic view images contain 360°of information, each level of progressive partitioning produces spatial histograms that reflect a fraction of that information (i.e. 180°, 120°, 90°, ..., etc). Therefore, the matching between each spatial histogram of a



(a) (b)

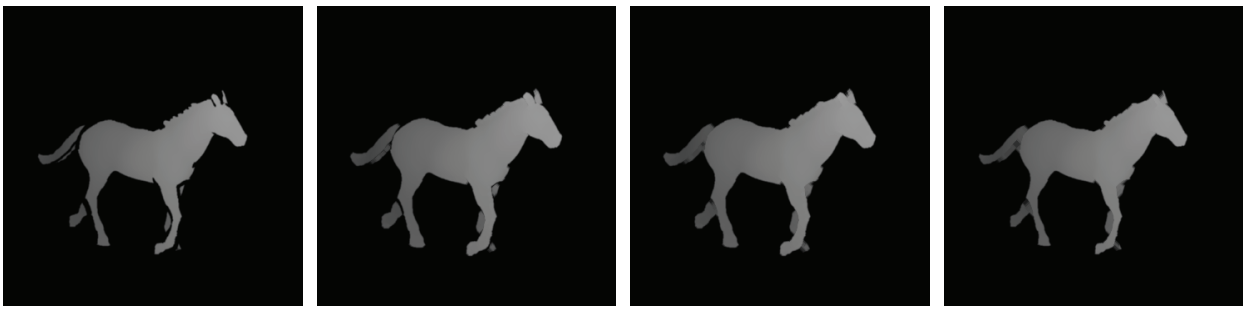
Figure 31: 31a simulated rendering of three progressive partitioning levels for a complete 3D model. At each level only one subimage is displayed. 31b unwrapped cylindrical projections of the aforementioned progressive partitioning levels with all subimages illustrated.

3D model and the corresponding histograms of a query range image needs to be weighted based on a ratio that measures the possibility of achieving a match between the contained information of the two (spatial) histograms, at each level of progressive partitioning:

$$PQ_{ratio}(i) = \left| \frac{Q_{angle}}{P_{angle}(i)} \right| \tag{25}$$

where $P_{angle}(i)$ is the Field-of-View [124] of the panoramic view (sub)image at the i^{th} level of progressive partitioning and Q_{angle} is the Field-of-View of the query range image, measured in degrees. The Q_{angle} is based on the properties of the camera used for capturing the range images and in our experiments, we have estimated this angle at 60°, which simulates the projection of the query range image to one of the faces of a hexagon.

5.5 Range Image Matching



(a) (b) (c) (d)

Figure 32: An example query range image 32a before preprocessing, 32b after dilation, 32c closing and 32d erosion steps, in consecutive order.

A *range image* is defined as a 2D image whose pixel values store the distance of each corresponding scene point from the projection plane [117]. Usually, the projection plane is associated with the range image camera sensor (see Fig 32a).

Based on the constructed BoVW model, for the database 3D models, the matching of the query range images is performed. The query range images are usually very noisy, due to the capturing process and an extra preprocessing step is often necessary before the actual matching. Here, we have followed a simple strategy that attempts to fill any holes, resulting from the object scanning and/or eliminate any outliers that do not belong to the actual objects (i.e. parts of the background).

Initially, morphological dilation [38] is performed on the original query range image. This step grows the range image regions, so that any moderate holes, that could have occurred due to errors in the capturing process, become small enough, in order to be considered as candidates for closing (see Fig. 32b). On the second step, morphological closing [38] is performed on the dilated image, in order to achieve closing of any small open areas that have been produced due to the digitization process (see Fig. 32c). The next step of the preprocessing strategy is the morphological erosion [38] of the image, so that it returns to its original form, with any small to moderate holes closed (see Fig. 32d). All of the aforementioned morphological operators are applied using disk-shaped structuring elements of size 3. The final step of the preprocessing strategy involves smoothing the range image by convolving it with an isotropic Gaussian kernel, which ensures that any rough edges are leveled and any small outlying regions, remote from the main object, are discarded.

Following the preprocessing, the range images are used as queries for the 3D model database. In a strategy similar to the histogram computation for the panoramic views of the database 3D models, the queries are compared to the models of the database. The SIFT descriptor is extracted from the range image, RootSIFT transformation is applied on the descriptor and finally, based on the generated Bag-of-Visual-Words, a histogram describing the codebook frequencies of occurrence is computed as the query's descriptor.

The similarity between the spatial histograms H of a 3D model and the histogram h of a query range image is calculated as follows: for each level l of progressive partitioning, the spatial histograms H_l are compared to the query range image histogram h using as a metric the Normalized Histogram Intersection distance ($D_{NHI}(H, h)$), defined by [30] as:

$$D_{NHI}(H, h) = \sum_{i=1}^{n} \frac{\min(H(i), h(i))}{H(i) + h(i)} \tag{26}$$

The best match is recorded and weighted by PQ_{ratio} for the corresponding level. Then, for all progressive partitioning levels, the best matches are summed to create the final distance between the query and the corresponding database model. We define max_l to be the maximum level of progressive partitioning set to 6.

$$Dist(H,h) = \sum_{l=1}^{max\_l} PQ_{ratio}(l)[\arg\min H_l(D_{NHI}(H_l,h))] \quad (27)$$

The complete descriptor extraction and matching algorithm is outlined in **Algorithm 5** and **Algorithm 6**, respectively.

Algorithm 5 Bag_of_Visual_Words Model Building Algorithm

Input: $3D\_models$ ▷ Database 3D model set

$n\_models = card(3D\_models)$ ▷ The number of 3D models of the database

$n\_axes = 60$ ▷ Define 60 random axis points

$max\_l = 6$
for $n = 1 \rightarrow n\_models$ **do**
  $axes = rand(n\_axes)$
  for $m = 1 \rightarrow n\_axes$ **do**
    $pan(n,m)$ = EXTRACT_PVIEW($3D\_model(n), axes(m)$)
    $sift(n,m)$ = DSIFT($pan(n,m)$)
    $rsift(n,m)$ = ROOTSIFT($sift(n,m)$)
  end for
end for
$train\_set$ = rand($rsift, n\_model/10$)
$codebook$ = TRAIN_BOVW($train\_set$)
for $n = 1 \rightarrow n\_models$ **do**
  for $m = 1 \rightarrow n\_axes$ **do**
    for $l = 1 \rightarrow max\_l$ **do**
      $hist_l(n,m)$ = EXTRACT_HIST($codebook, rsift(n,m)$)
    end for
  end for
end forreturn $codebook, hist$

Note that the EXTRACT_PVIEW, DSIFT, ROOTSIFT, TRAIN_BOVW and EXTRACT_HIST functions refer to the panoramic view extraction, Dense SIFT descriptor calculation, Root-SIFT transformation, Bag-of-Visual-Words codebook generation and (spatial) Histogram extraction operations of the pipeline, respectively.

Algorithm 6 Query Range Image Matching Algorithm

Input: $codebook, hist, query\_image$
Input: n_models, n_axes ▷ The number of 3D models of the database and the random axis points, as defined in **Algorithm 5**

1:
2: $siftq =$ DSIFT($query\_image$)
3: $rsiftq =$ ROOTSIFT($siftq$)
4: $histq =$ EXTRACT_HIST($codebook, rsiftq(n)$)
5: **for** $n = 1 \rightarrow n\_models$ **do**
6: **for** $m = 1 \rightarrow n\_axes$ **do**
7: $query\_dist(n, m) = Dist(hist(n), histq)$
8: **end for**
9: **end for**
10: $final\_dist = min(query\_dist)$ **return** $final\_dist$

6 Experimental Results

God ever geometrizes

– Plato

In this chapter, the experimental evaluation of the presented methods will be discussed in detail. Both quantitative and qualitative (applicable to the pose normalization methods) means of evaluation were used. First, we shall define the framework chosen for conducting the experiments. This framework is uniform across every presented method, allowing the reader to easily extract comparative conclusions.

The experimental evaluation is based on the Precision-Recall curves and five quantitative measures: Nearest Neighbor (NN), First Tier (FT), Second Tier (ST), E-measure (E) and Discounted Cumulative Gain (DCG) [110] for the classes of each corresponding dataset. For every query model that belongs to a class C, Recall denotes the percentage of models of class C that are retrieved and Precision denotes the proportion of retrieved models that belong to class C over the total number of retrieved models. The best score is 100% for both quantities. Nearest Neighbor (NN) indicates the percentage of queries where the closest match belongs to the query class. First Tier (FT) and Second Tier (ST) statistics, measure the recall value for the $(D-1)$ and $2(D-1)$ closest matches respectively, where D is the cardinality of the query's class. E-measure combines precision and recall metrics into a single number and the DCG statistic weighs correct results near the front of the list more than correct results later in the ranked list under the assumption that a user is less likely to consider elements near the end of the list [59, 110].

For each presented method, the evaluations were conducted on widely used datasets,

allowing the comparison of the results with both state-of-the-art and future methods. In the following sections, the experimental evaluation for each of the methods introduced in this thesis will be presented.

6.1 Rigid 3D Object Retrieval

This section provides detailed performance results of the ROSy and SymPan, SymPan+ pose normalization methods. The first step of the evaluation process is to set up the testing framework. The result of pose normalization is an aligned placement of the input 3D object in space. As it is not trivial to directly quantify the quality of the alignment, an indirect way of testing and comparison will be used. Since pose normalization procedures are primarily used as a preprocessing step in graphics applications like visualization, reconstruction from broken fragments and 3D object retrieval, it is possible to evaluate the performance of the proposed method through the final results of such a system. Our evaluation is based on rigid 3D object retrieval.

6.1.1 ROSy for 3D Object Retrieval

For the evaluation of the ROSy pose normalization method, we have chosen a state-of-the-art 3D object retrieval methodology, by Papadakis et al. [88] as the evaluation vehicle. The datasets, on which the experiments were conducted, are the following: the training and test sets of the Princeton Shape Benchmark (PSB) [110], the classified objects of the National Taiwan University database (NTU) [28], the MPEG-7 dataset [129], the Engineering Shape Benchmark dataset (ESB) [60] the National Institute of Standards and

Table 2: Categories and cardinalities of evaluation datasets

3D object dataset	# of Categories	# of objects
PSB training	90	907
PSB test	92	907
NTU	41	549
MPEG-7	135	1300
ESB	48	866
NIST	40	800
McGill Articulated	10	254
McGill Non Articulated	9	202

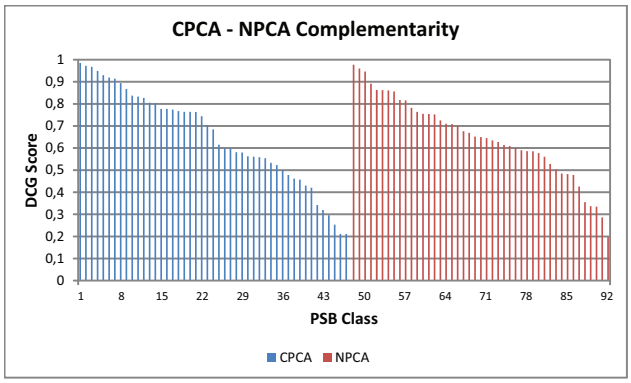
Technology dataset (NIST), which contains shape normalized and visually categorized 3D objects found in the SHREC 2009 competition [43] and both the Articulated and Non Articulated objects of the McGill dataset [144]. From the NTU dataset, only the classified objects were used, as unclassified objects would not give accurate retrieval results. Table 2 shows the number of categories and the total number of objects in each dataset used for the experiments.

In previous section, $\frac{1}{2}Dist(\cdot)$ and $Ang_{tanh}(\cdot)$ were defined as equally weighted in the equation of the minimization criterion k_{Π}. This choice is justified through a series of retrieval tests with differently normalized weight factors. In these tests ROSy was used as the pose normalization procedure of the 3D object retrieval system and the resulting Discounted Cumulative Gain (DCG) [59] was measured on the PSB test, the NIST, the ESB and the MPEG7 datasets. The DCG statistic gives a sense of how well the overall retrieval would be viewed by a human. Correct shapes near the front of the list are more likely to be seen than correct shapes near the end of the list. Table 3 confirms that the most suitable choice is the use of equal weight factors in the minimization criterion.

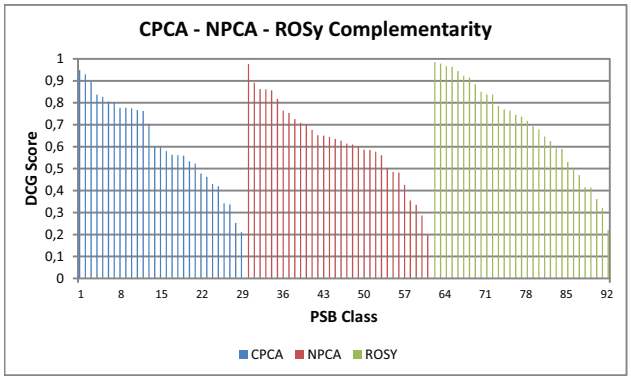
Table 3: Impact of the weight factor (in the minimization criterion k_{Π}) on DCG score for four datasets. Higher DCG score is better.

Weight Factors		DCG Score			
$\frac{1}{2}Dist(\cdot)$	$Ang_{tanh}(\cdot)$	PSB test	NIST	ESB	MPEG7
0	1	0.665	0.719	0.729	0.805
0.25	0.75	0.666	0.756	**0.732**	0.812
0.5	**0.5**	**0.678**	**0.764**	**0.732**	**0.821**
0.75	0.25	0.667	0.763	0.700	0.809
1	0	0.645	0.724	0.698	0.790

Papadakis' 3D object retrieval system, in its original form, uses a combination of the CPCA and NPCA algorithms to achieve pose normalization of a 3D object set. This approach defines a successful hybrid scheme that could be further improved by the proposed method. However, to test if the three pose normalization methods can benefit the retrieval process, without adding any redundant complexity to it, a complementarity test needs to be performed. This test assesses the number of classes that are best aligned by each method in terms of retrieval accuracy, by performing the DCG test on the retrieval results of the test PSB dataset. If the percentages of success of the three methods are similar, then



(a)



(b)

1 (62)	sedan	47 (92)	church
2 (63)	hourglass	48 (29)	covered_wagon
3 (64)	monster_truck	49 (65)	electrical_guitar
4 (1)	hammer	50 (66)	fighter_jet
5 (2)	human	51 (30)	gazebo
6 (67)	dining_chair	52 (70)	enterprise_like
7 (68)	biplane	53 (31)	head
8 (3)	race_car	54 (32)	stealth_bomber
9 (69)	handgun	55 (33)	desk_chair
10 (4)	glider	56 (34)	chess_set
11 (71)	glass_with_stem	57 (72)	motorcycle
12 (5)	face	58 (73)	commercial
13 (6)	horse	59 (35)	bench
14 (7)	flying_saucer	60 (84)	rectangular
15 (8)	human_arms_out	61 (36)	computer_monitor
16 (9)	potted_plant	62 (78)	umbrella
17 (10)	tie_fighter	63 (37)	submarine
18 (11)	jeep	64 (79)	barren
19 (75)	door	65 (38)	ship
20 (74)	streetlight	66 (39)	hot_air_balloon
21 (77)	mailbox	67 (40)	dog
22 (76)	sword	68 (80)	sea_turtle
23 (12)	helicopter	69 (41)	desktop
24 (13)	ant	70 (42)	hand
25 (82)	geographic_map	71 (43)	fish
26 (14)	bush	72 (44)	skull
27 (15)	large_sail_boat	73 (45)	fireplace
28 (83)	knife	74 (46)	shovel
29 (16)	gear	75 (47)	shelves
30 (17)	hat	76 (48)	flying_bird
31 (18)	ladder	77 (88)	train_car
32 (19)	eyeglasses	78 (49)	semi
33 (81)	rabbit	79 (50)	single_leg
34 (20)	billboard	80 (51)	conical
35 (21)	snake	81 (52)	flowers
36 (86)	vase	82 (85)	walking
37 (22)	city	83 (53)	axe
38 (23)	butterfly	84 (54)	book
39 (87)	wheel	85 (55)	newtonian_toy
40 (24)	slot_machine	86 (89)	cabinet
41 (25)	standing_bird	87 (56)	school_desk
42 (26)	two_story_home	88 (57)	pail
43 (91)	barn	89 (58)	one_peak_tent
44 (90)	sink	90 (59)	one_story_home
45 (27)	staircase	91 (60)	skyscraper
46 (28)	satellite	92 (61)	satellite_dish

(c)

Figure 33: DCG retrieval scores for the CPCA, NPCA and ROSy pose normalization methods, color coded by the method that achieves the best results per class. 33a Per class CPCA and NPCA complementarity results. 33b Per class CPCA, NPCA and ROSy complementarity results. 33c The correspondence between class IDs and class names. Parenthesized IDs refer to plot
33b.

the methods can be considered complementary. Firstly, a test of the complementarity between the CPCA and the NPCA methods was performed and then the same evaluation was conducted for all three pose normalization methods. The results, illustrated in Figure 33, confirm that both CPCA and NPCA methods are between them complementary in terms of per class retrieval accuracy and that the proposed method is also complementary to them. Therefore, the addition of the ROSy method to the pose normalization procedure, could potentially improve the overall performance of the retrieval system, by achieving better alignment (in terms of retrieval accuracy) on a subset of objects where the original two component approach fares badly.

The next step is to test whether the retrieval system can actually benefit by using ROSy

in addition to the original pose normalization approach.

In Figure 34, the P-R plot of the retrieval process on the PSB test dataset is illustrated. The proposed triple (CPCA, NPCA and ROSy) approach, identified as ROSy+ for the remainder of the text, is compared against the CPCA, NPCA and ROSy standalone methods and the dual (CPCA, NPCA) approach.

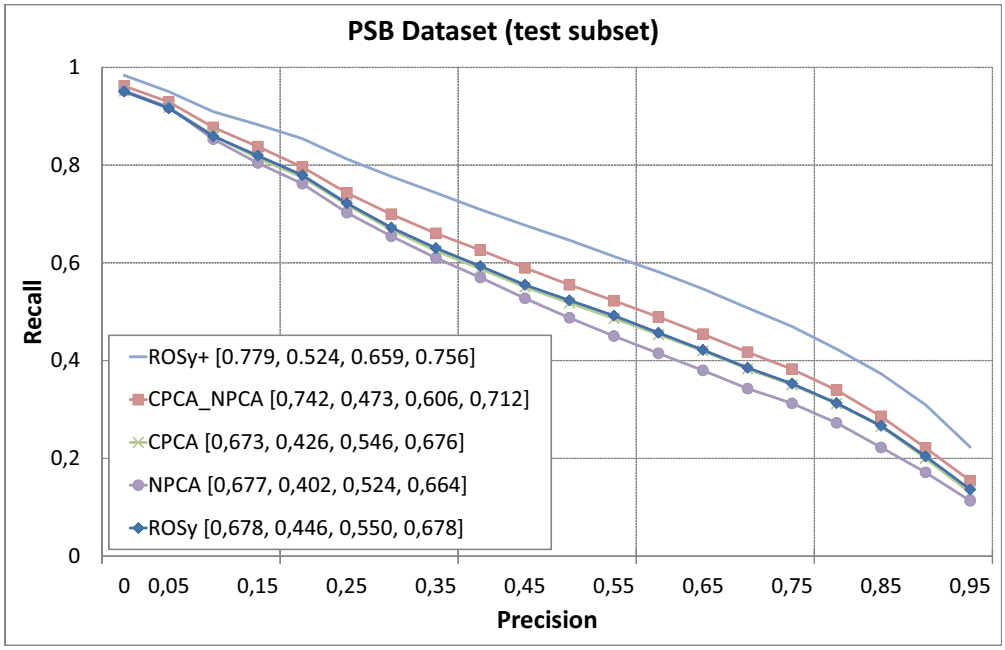


Figure 34: Precision-Recall plot for the Princeton Shape Benchmark test dataset. The use of ROSy alongside the original approach significantly boosts the performance of the retrieval process.

ROSy itself has similar performance to CPCA and NPCA. However, the combination of the three pose normalization methods (namely ROSy+) gives a significant boost to the discriminative power of the retrieval process, outperforming the original hybrid (CPCA, NPCA) approach. Similar to the original methodology, the descriptor consists of three sets of coefficients corresponding to the three aligned versions of the object (using CPCA, NPCA and ROSy). The comparison between two objects is done between the corresponding aligned sets, consequently, the CPCA aligned query object is compared with the CPCA aligned version of the gallery object, the NPCA aligned query object is compared with the NPCA aligned version of the gallery object and similarly for the ROSy version. The 2D and 3D features are computed for three alternative rotation normalized versions of a 3D object. Thus, the final hybrid 3D shape descriptor si of an object i is the concatenation of the 2D and 3D features for each aligned version of the 3D object, giving:

$$si = (2Df_i^{CPCA}, 2Df_i^{NPCA}, 2Df_i^{ROSy}, \\ 3Df_i^{CPCA}, 3Df_i^{NPCA}, 3Df_i^{ROSy}) \quad (28)$$

where $2Df_i^j$ and $3Df_i^j$ are the 2D and 3D feature vectors of model i, respectively. Each feature vector is computed by an alignment of model i, using CPCA, NPCA or ROSy, denoted by $j \in \{CPCA, NPCA, ROSy\}$.

To compare the descriptors $s1$ and $s2$ of two 3D objects the following schema is adopted, to compute their distance:

$$Distance(s1, s2) = dist_{2Df} + dist_{3Df} \quad (29)$$

where $dist_{2Df}$ and $dist_{3Df}$ is the distance between the 2D and 3D features, respectively, computed as:

$$dist_{2Df} = min_j(L_1(2Df_1^j, 2Df_2^j)) \\ dist_{3Df} = min_j(L_1(3Df_1^j, 3Df_2^j)) \quad (30)$$

where $j \in \{CPCA, NPCA, ROSy\}$ and L_1 is the Manhattan distance between the corresponding features.

The comparison giving the minimum distance sets the distance score between the query and gallery objects. The notion of taking the minimum distance is based on the expectation that the best establishment of correspondences between two objects is achieved when the difference between the shape descriptors is minimum.

In Figure 35 it is further illustrated that the 3D object retrieval system using ROSy+ outperforms two recent pose normalization methods: DLA [27] and GSMD+SHD+R [72], the PANORAMA descriptor [89] and also three classic 3D object retrieval methods: Light-field [28], SH-GEDT [65] and DESIRE [132] approaches. Again, the P-R plot of the retrieval process on the PSB test dataset and the four quantitative measures (NN, FT, ST, DCG) are displayed.

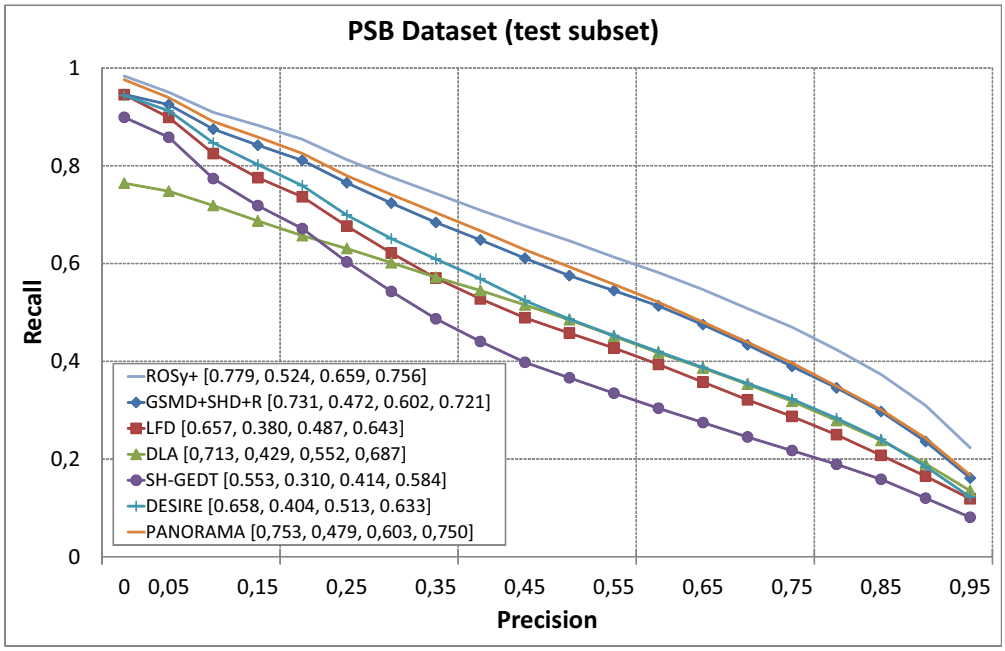
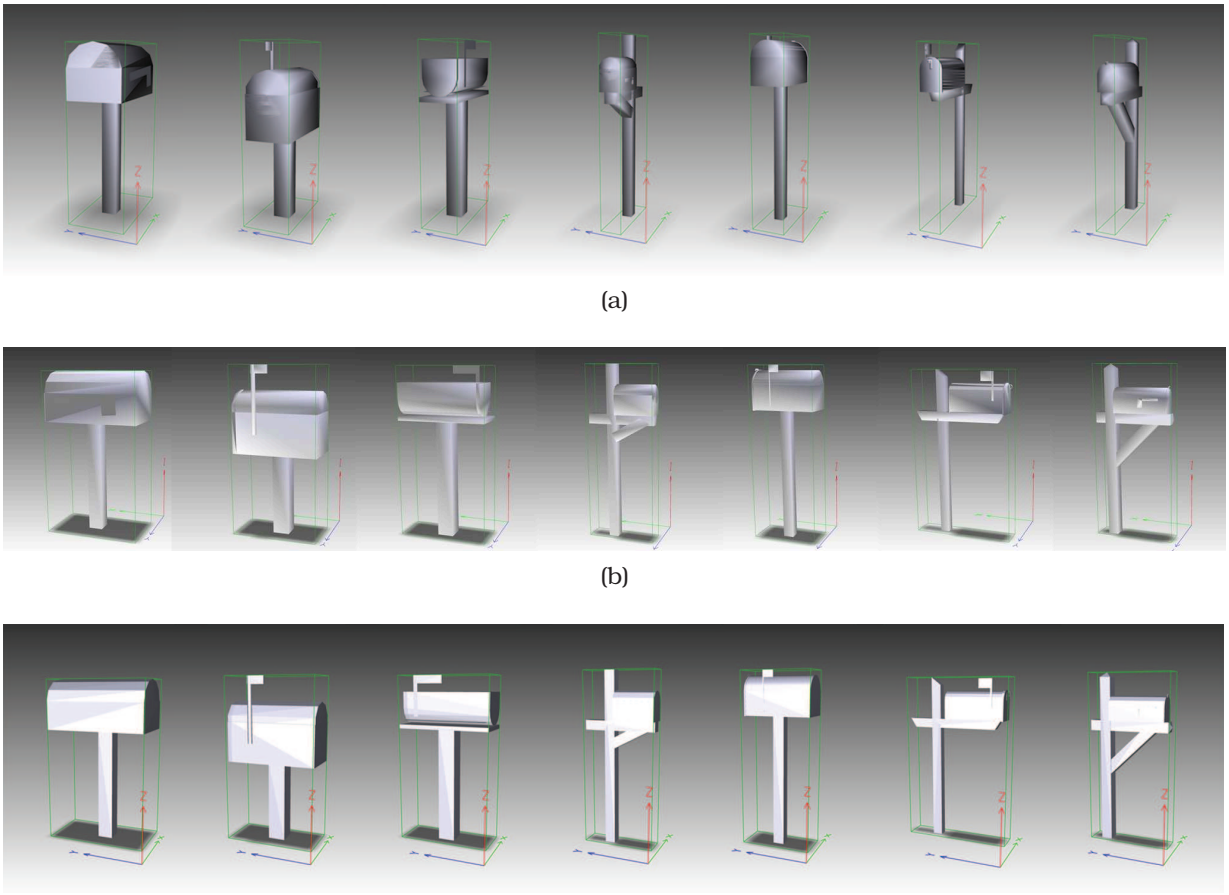


Figure 35: Precision-Recall plot for the Princeton Shape Benchmark test dataset. ROSy+ retrieval results are compared against state-of-the-art 3D object retrieval techniques.

To establish that increase in the discriminative power is not dependent on the PSB dataset, the dual (CPCA, NPCA) and the ROSy+ approaches were tested on the rest of the available datasets. The quantitative measure scores of the results are shown in Table 4. These scores show that the results are consistent throughout the datasets, revealing the stability of the proposed approach and the gain with respect to the original hybrid (CPCA, NPCA) system. Furthermore, it is clear that ROSy+ performs better than previous methods and the recently proposed methods by Chaouch and Verroust -Blondet, using the Depth Line Approach descriptor [27], by Lian et al. using the combined GSMD - SHD descriptors with Rectilinearity [72] and by Papadakis et al. using the PANORAMA descriptor [89] by about 2% - 5%.

Comparing the plots and the four quantitative measures, it can be concluded that the combined use of the three complementary pose normalization methods significantly elevates the discriminative power of the 3D object retrieval system. On all 8 datasets, ROSy+ is able to achieve an average performance gain of about 3% over the previous dual approach (mean value over the quantitative measures used). This gain is significant, because it is accomplished exclusively by enhancing the pose normalization procedure and not the core retrieval algorithm. Note that Papadakis' object retrieval system [88] has achieved state-of-the-art performance.



(a)

(b)

(c)

Figure 36: Alignments of the 'MAILBOX' class: 36a ROSy pose normalization results 36b a consistently rotated version of ROSy pose normalization results by 90 degrees around the z axis 36c results of the method proposed in [27]

A visual qualitative evaluation is next provided. In Figure 36, comparative alignments between ROSy and the proposed method by Chaouch and Verroust - Blondet [27], on the complete 'MAILBOX' class of the PSB dataset, are illustrated. These alignments show that ROSy is able to produce accurate alignments, similar to those of the method proposed by Chaouch and Verroust - Blondet, while simultaneously achieving better quantitative scores. As is illustrated in Figure 37, ROSy is also capable of producing accurate alignment results that, regardless of the originating class or the morphology of the input objects, are consistent and stable. In Figure 37, 3D objects 37a - 37d show perfect global symmetry against one principal plane. 3D Objects 37e - 37h show global symmetry against one principal plane, that is not perfect however, because of minor parts of the objects that don't fully match. 3D Objects 37i and 37j show global symmetry against two principal planes simultaneously, while 3D objects 37k - 37n have local symmetries in

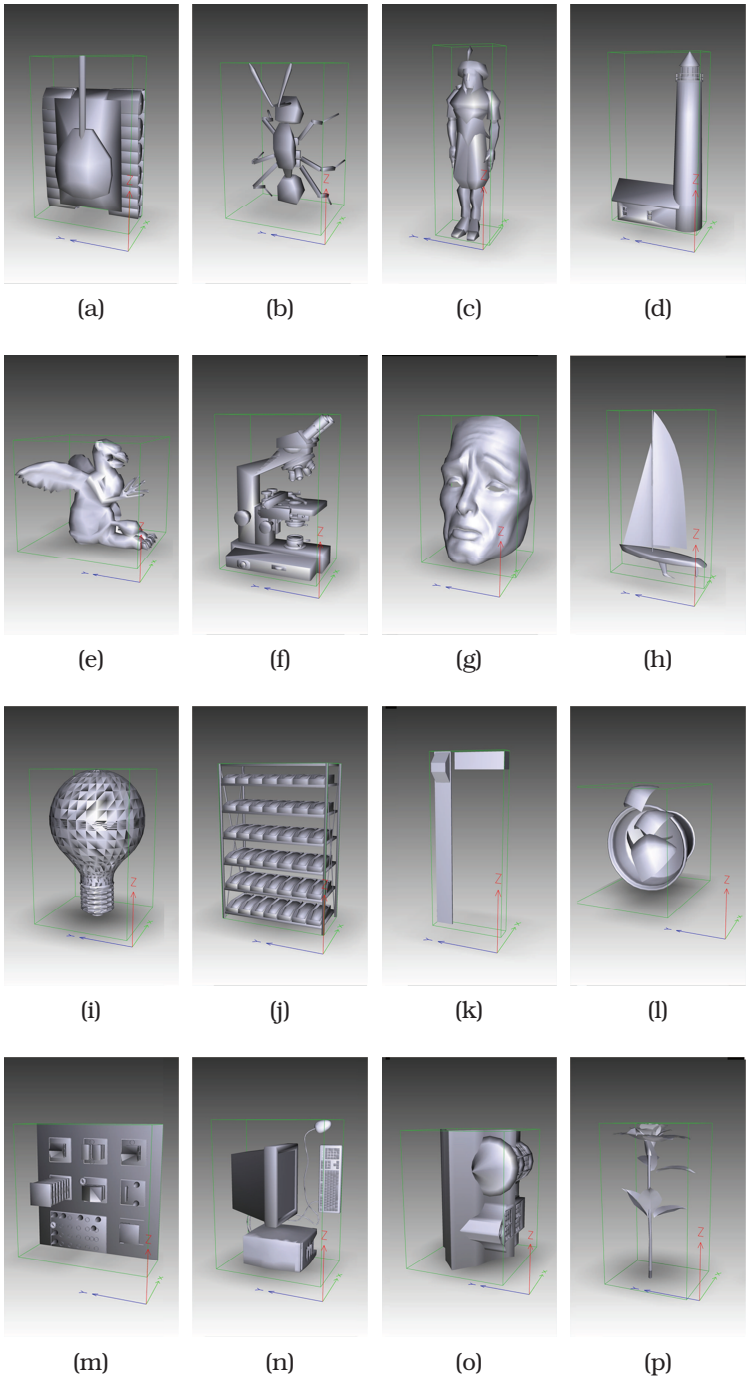
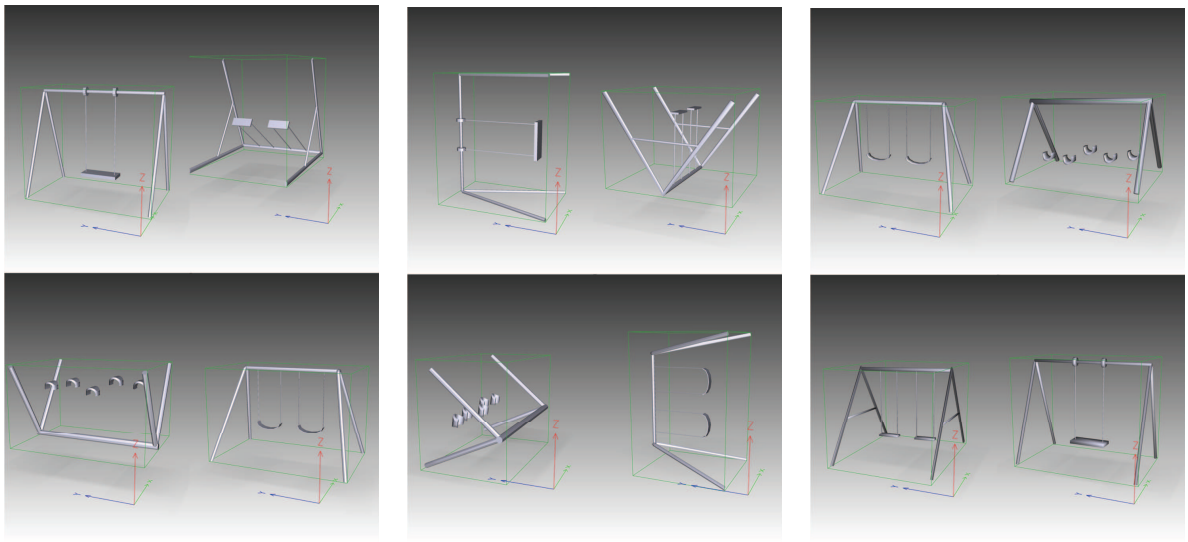


Figure 37: Sample alignments of 3D objects originating from different PSB classes, using ROSy pose normalization method. The illustrated 3D objects exhibit various types of global symmetries 37a - 37j, local symmetries 37k - 37n or no symmetries, at all 37o - 37p.



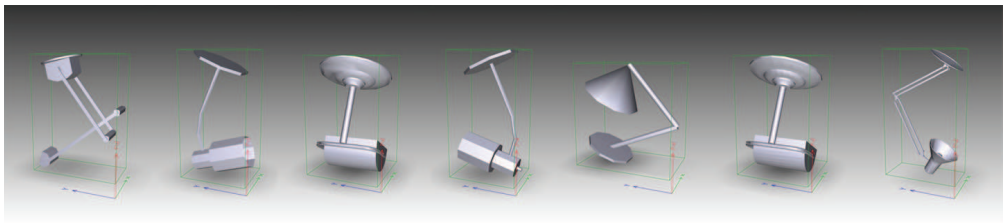
(a) (b) (c)

Figure 38: Alignments of the PSB class 'SWINGSET'. 38a CPCA, 38b NPCA, 38c ROSy.

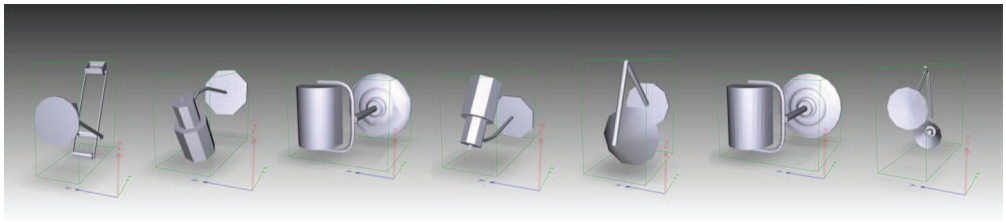
their structures. 3D Objects 37o and 37p exhibit no symmetry at all.

With respect to the CPCA and NPCA approaches, ROSy uses a combination of spatial (vertices) and angular (normals) features, to achieve 3D object alignment. The NPCA method performs better with objects that have dominant flat surfaces, while CPCA best aligns objects that are composed of bumpy surfaces. The ROSy method, while able to handle well 3D objects composed of either flat or bumpy surfaces, also exhibits no degradation of performance in the alignment of 3D objects that have both types of surfaces. Comparative examples of alignments, against the CPCA and NPCA methods, on the 'SWINGSET' class which contains round-edged 3D objects are illustrated in Figure 38. Also, Figure 39 shows comparative alignment results for 3D objects that belong to the 'LAMP' class and are composed of both flat and bumpy surfaces.

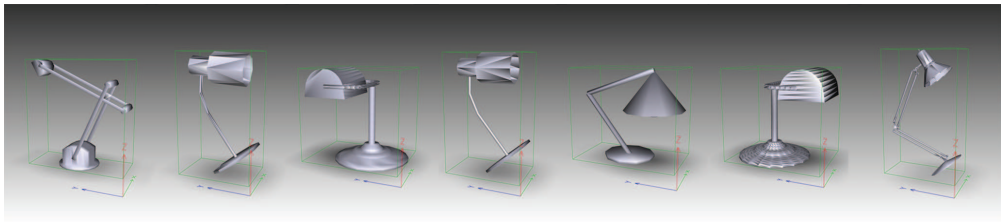
At this point a paradox arises. Although ROSy alone clearly produces visually better alignments than CPCA or NPCA, its standalone *retrieval* results are not spectacularly different from those of CPCA and NPCA (Figure 34). Since a 3D object retrieval system is a complex procedure, only speculations can be made about the cause. However the experimental process reveals an interesting fact: 3D objects that belong to different classes, but have related structure (e.g. trucks and cars) are also aligned similarly by ROSy. Although this is correct, in terms of alignment, it possibly interferes with the retrieval process because it enhances the similarities between the 3D objects and hides their differences.



(a)



(b)



(c)

Figure 39: Comparison between the alignment results of the CPCA [131], NPCA [87] and ROSy methods on objects of the 'LAMP' class of the PSB dataset. 39a CPCA, 39b NPCA, 39c ROSy.

6.1.1.1 Remark: For the corresponding numerical results, please refer to the appropriate tables in Appendix A.

6.1.2 SymPan, SymPan+ on PANORAMA

Similar to the way that the ROSy+ system has been used for the quantitative evaluation of the ROSy pose normalization method, for SymPan and SymPan+ we have chosen the PANORAMA state-of-the-art 3D object retrieval system, by Papadakis et al. [89] as the evaluation vehicle. The proposed method replaces the NPCA pose normalization method in the existing hybrid scheme.

The main dataset, on which the experiments were conducted, is the test subset of the Princeton Shape Benchmark (PSB) [110]. This dataset is composed of 907 3D models classified into 92 classes. The direct effect of the proposed alignment methods can be evaluated by comparing against the original PANORAMA performance. In terms of object

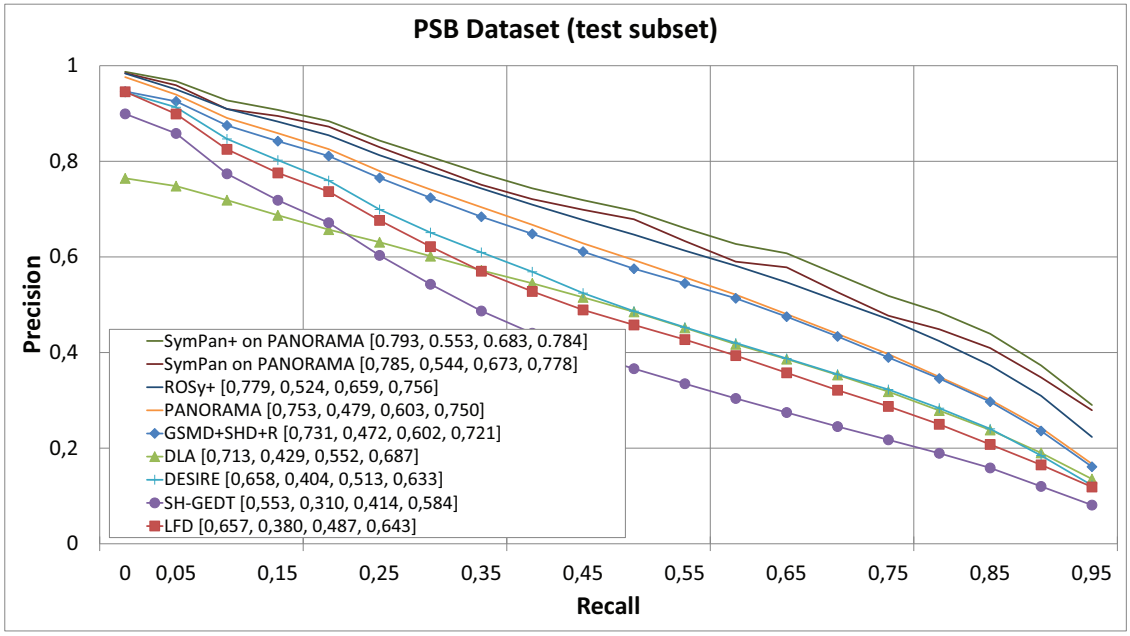


Figure 40: Precision-Recall plot for the Princeton Shape Benchmark test dataset. SymPan, SymPan+ 3D model pose normalization methods on PANORAMA retrieval results are compared against state-of-the-art 3D object retrieval techniques.

retrieval performance, we compare against DLA [27], GSMD+SHD+R [72], ROSy+ [105], Lightfield [28], SH-GEDT [65] and DESIRE [132].

Regarding SymPan+, to further establish that the discriminative power of the method is not dependent on the PSB dataset we also conducted a number of experiments on the following datasets: the MPEG-7 dataset [140, 67], the Engineering Shape Benchmark dataset (ESB) [60] and the National Institute of Standards and Technology dataset (NIST), which contains shape normalized and visually categorized 3D objects found in the SHREC 2009 competition [43]. The proposed method was tested against the performance of the original PANORAMA and ROSy+ approaches.

In Fig. 40, using the experimental results given in [105, 104], we illustrate the P-R scores for the test subset of the PSB dataset, for the PANORAMA 3D object retrieval system enhanced by the SymPan and SymPan+ pose normalization methods. Table 5 shows quantitative measures for the same methods. The quantitative measure scores for the remaining test datasets of the SymPan+ pose normalization method are shown in Table 6.

The results demonstrate that the proposed scheme outperforms state-of-the-art methods and significantly increases the performance of the PANORAMA 3D object retrieval

system compared to its original pose normalization scheme. Furthermore, these scores show that the results are consistent throughout the datasets, revealing the stability of the proposed approach and the gain with respect to the original system.

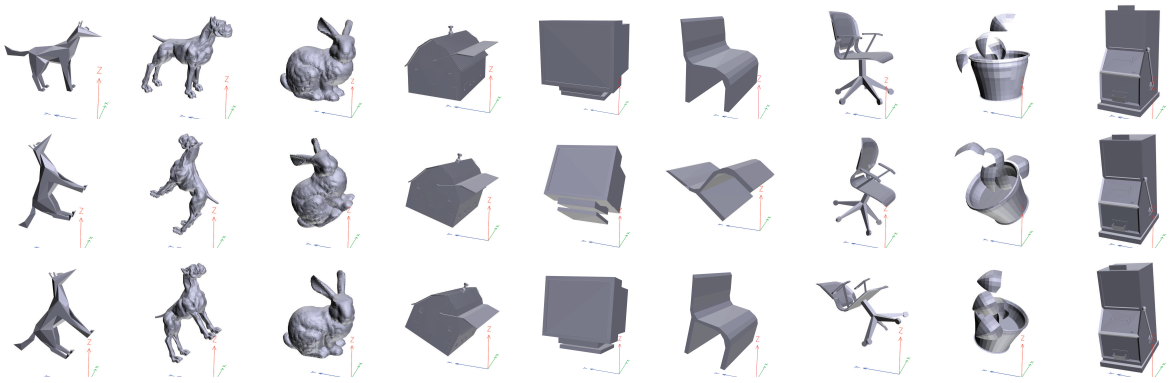


Figure 41: Sample alignments of 3D objects originating from different PSB classes. Top Row: SymPan+ alignment method, middle row: CPCA alignment method, bottom row: NPCA alignment method.

In Fig. 41 comparative alignments between SymPan+ and CPCA, NPCA pose normalization methods on various 3D models from the PSB dataset, are illustrated. These alignments show that SymPan+ is able to produce accurate alignment results that, regardless of the originating class or the morphology of the input objects, are consistent and stable.

An interesting observation that has motivated our research, is that most of the objects found in the benchmark datasets, exhibit some degree of planar symmetry. This is especially true for real life objects, e.g. humans, animals, insects, faces etc. and also for many man-made items, like tools, cars, bikes, aeroplanes etc. We have also found that the majority of these objects only have one plane of symmetry, a moderate portion of the dataset has two or more (e.g. mechanical objects found in the ESB dataset) and only a small fraction exhibit no symmetry at all. In a preliminary research on the PSB dataset, using Eq. 9 on the panoramic views extracted from 3 the orthogonal principal axes after alignment of the 3D models, we have found that: approx. 31.25% of the 3D model have 2 or more symmetry planes, approx. 64.58% of the 3D models have one symmetry plane and approx. 4.16% of the 3D models exhibit no symmetry plane.

6.1.2.1 Remark: For the corresponding numerical results, please refer to the appropriate tables in Appendix B.

6.2 Non-Rigid 3D Object Retrieval

In this section we show the performance results of the proposed non-rigid 3D object retrieval methodology (denoted as *ConTopo++*) on a number of standard datasets and compare the updated implementation with the one proposed in [102]. The datasets that we used for the experimental evaluation are the following: (i)TOSCA [19, 18], (ii) SHREC'07 *Watertight Models* [50], (iii) SHREC'10 *Non-rigid 3D Models* [71] and (iv) SHREC'11 *Non-rigid 3D Watertight Meshes* [70].

We first conducted a number of experiments comparing to the segmentation method proposed by Agathos et al. [2] (denoted as *EMD-PPPT*), in order to justify if the selected feature set as well as the proposed quantization and graph matching strategy are suitable for the non-rigid 3D mesh matching problem. These experiments were conducted on the SHREC'10 *Non-rigid 3D Models* dataset and are composed of the following evaluations, in terms of retrieval performance:

1. Compare against EMD-PPPT implementation.
2. Use ConTopo++ feature set, based on EMD-PPPT segmentation and graph matching strategy.
3. Use EMD-PPPT feature set, based on ConTopo++ quantization and graph matching strategy.

The results on Table 7 show that the proposed scheme (i.e. both the feature set and the partitioning/ matching strategy used in ConTopo++) performs better than the complete EMD-PPPT method and/or any combination between the feature sets and the partitioning/ matching strategies involved. Moreover, it is evident that the use of the proposed quantization and graph matching strategy yields better results than the corresponding segmentation strategy of EMD-PPPT. We can assume that this is due to the multi-level hierarchical partitioning of the proposed scheme over the single-level partitioning of the EMD-PPPT method.

We next compare ConTopo++ against other state-of-the-art methods on standard datasets.

On the TOSCA dataset we compared against the non-rigid discrete Conformal Factor (Cfact) [11] and ConTopo descriptors [102] as well as the rigid LightField (LF) [28] and

Spherical Harmonics (SH) [65] descriptors. On the SHREC'07 *Watertight Models* dataset we compared against the non-rigid discrete Conformal Factor (Cfact) descriptor, the ConTopo descriptor and the augmented Multiresolution Reeb Graph (aMRG) method [127]. On the SHREC'10 *Non-rigid 3D Models* and SHREC'11 *Non-rigid 3D Watertight Meshes* datasets we compared against the published results of the participating contestants (for the presented contestant methodologies refer to the corresponding papers [71] and [70], respectively). The LF and SH benchmark code is publicly available, the discrete Conformal Factor code was obtained from the author of [11], whom we gratefully acknowledge, while the aMRG code is not publicly available and therefore, the original graphs from the SHREC'07 *Watertight Models* [50] competition were used instead, since it was necessary for making a comparison.

On the first two datasets, in accordance with the published results in [102], the ex-

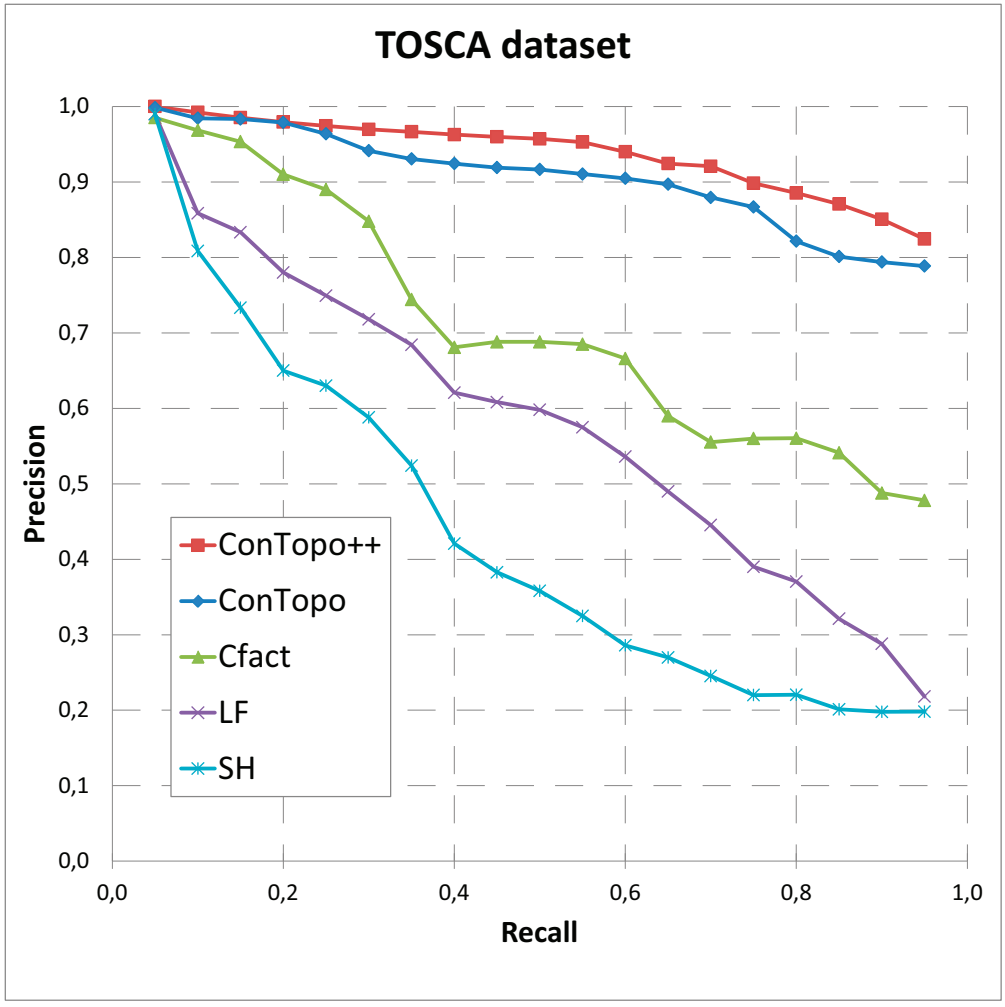


Figure 42: Comparative results based on the average P-R scores for the TOSCA dataset.

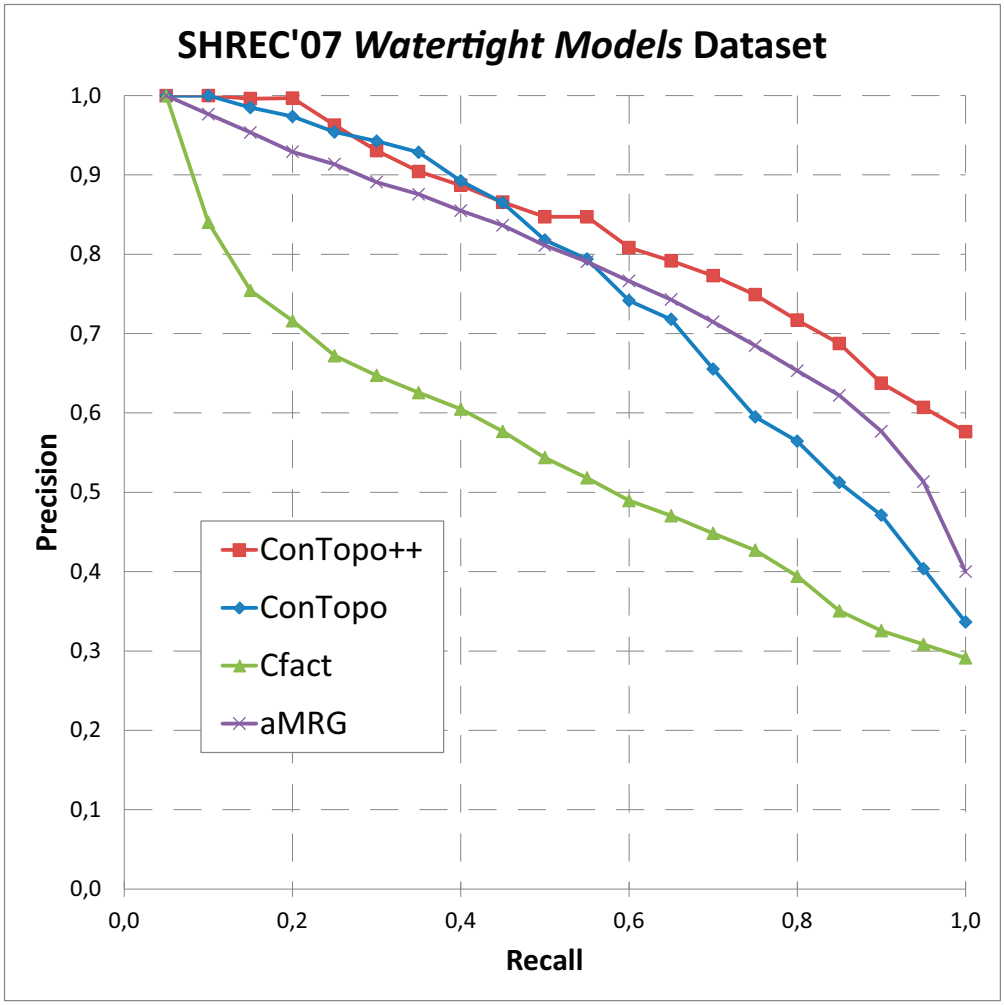


Figure 43: Comparative results based on the average P-R scores for the SHREC'07 *Watertight Models* dataset.

perimental evaluation is based on Precision-Recall plots for the classes of the corresponding datasets. On the remaining two datasets, five more quantitative measures are also computed: Nearest Neighbor (NN), First Tier (FT), Second Tier (ST), E-measure (E) and Discounted Cumulative Gain (DCG).

In Fig. 42, in accordance to the experimental results shown in [11] and [102], we illustrate the P-R plots for the complete TOSCA dataset for the proposed non-rigid 3D object retrieval methodology (ConTopo++), the ConTopo descriptor, the discrete Conformal Factor (Cfact) descriptor and two state-of-the-art shape descriptors for 3D objects: the LightField (LF) and the Spherical Harmonics (SH) descriptor, which are used to give a lower bound on the performance of any shape descriptor for non-rigid objects. The P-R scores of the methods clearly illustrate the increase in accuracy of the proposed method

ConTopo++, which outperforms the original ConTopo non-rigid 3D object descriptor.

According to the SHREC'07 *Watertight Models* classification scheme, the dataset, composed of 400 3D objects, is classified into 20 classes, each of which contains 20 objects. Fig. 43 illustrates the P-R plots for the complete dataset. The proposed method performs better than the original ConTopo descriptor and both the discrete Conformal Factor and the aMRG approaches.

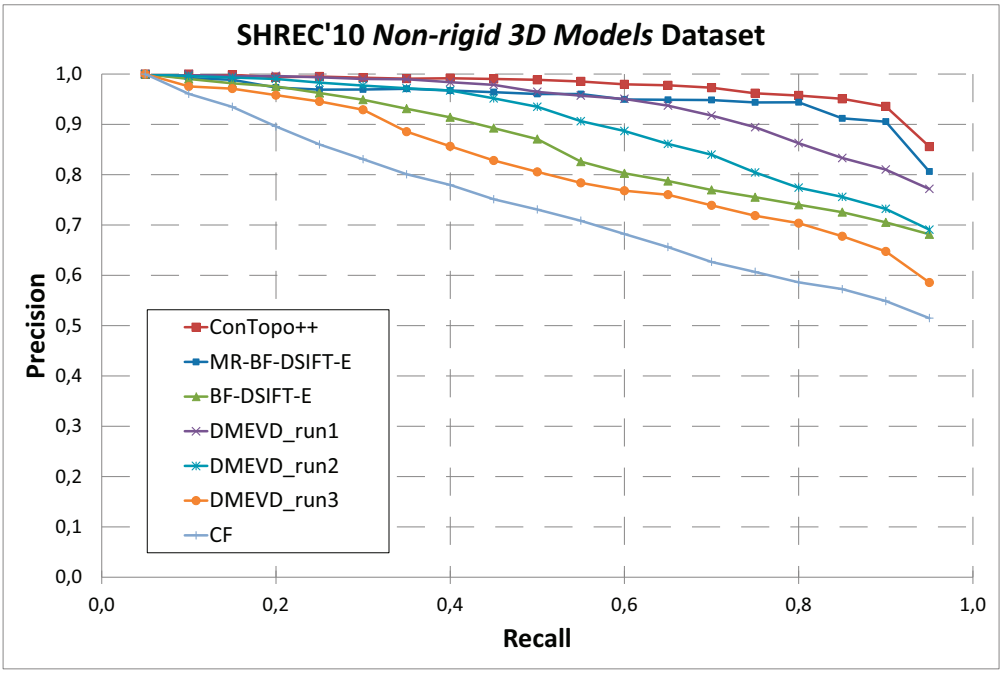


Figure 44: Comparative results based on the average P-R scores for the SHREC'10 *Non-rigid 3D Models* dataset.

In Figures 44 and 45 we illustrate the P-R scores of the proposed method (ConTopo++) against the published results of the SHREC'10 *Non-rigid 3D Models* and SHREC'11 *Non-rigid 3D Watertight Meshes* datasets, respectively. The corresponding quantitative measures are displayed in Tables 8 and 9, respectively. In Fig. 44 it is clear that the proposed method outperforms the track contestants, even though the published results were already of high performance. This can also be seen by the quantitative measures of Table 8 where the proposed method has the highest scores on three out of five measures (and is also second on the remaining two measures, by a small margin). In the SHREC'11 *Non-rigid 3D Watertight Meshes* dataset, the proposed method performs better or exhibits mixed results with respect to approximately 65% of the contestant results. Fig. 45 and Table 9 illustrate this.

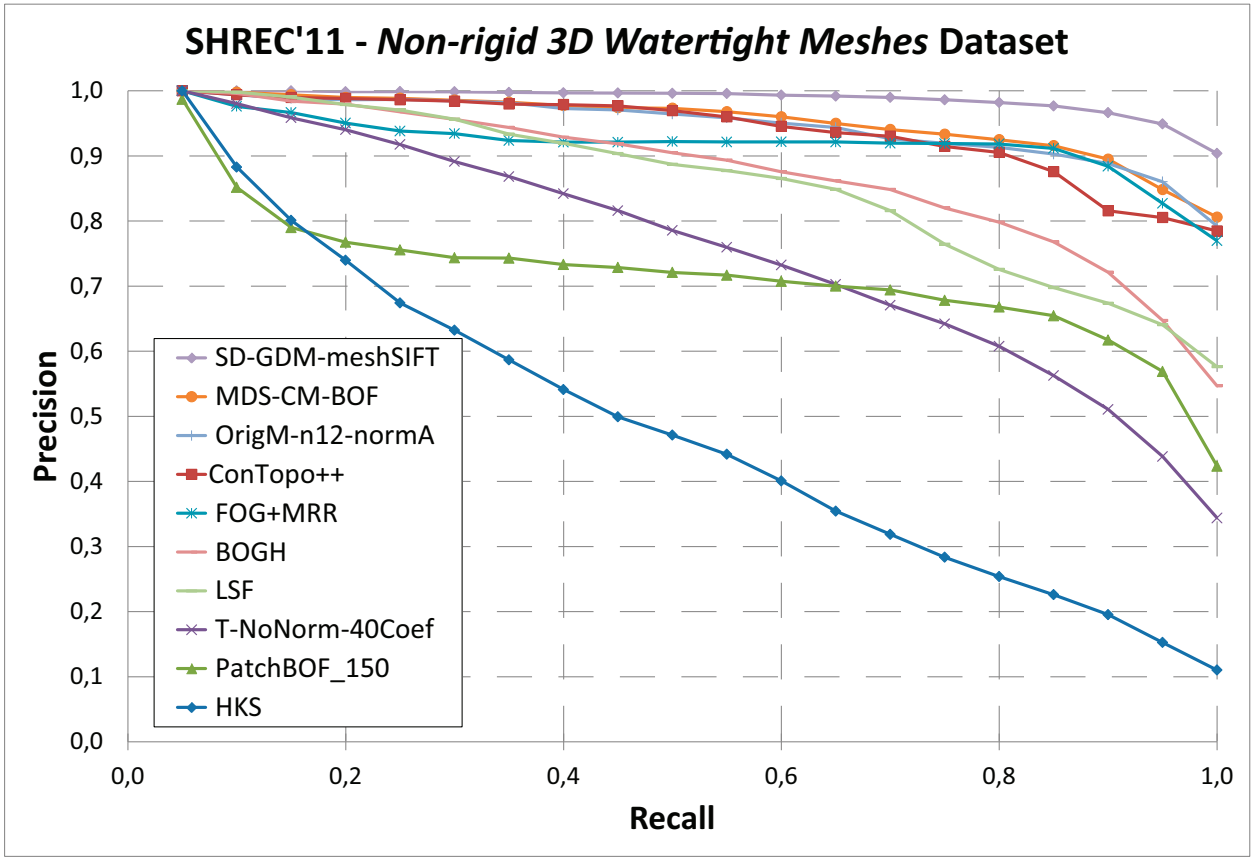


Figure 45: Comparative results based on the average P-R scores for the SHREC'11 *Non-rigid Watertight 3D Meshes* dataset.

6.2.0.2 Remark: For the corresponding numerical results, please refer to the appropriate tables in Appendix C.

6.3 3D object Retrieval based on 2D Range Image Queries

The datasets that we used for the experimental evaluation of our method are the following: (i) SHREC'09 *Querying with Partial Models* [39], (ii) SHREC'10 *Range Scan Retrieval* [40] and (iii) SHREC'11 *Shape Retrieval Contest of Range Scans* [113]. The target subset of the datasets used is based on the generic shape benchmark constructed at NIST [43].

On the first two datasets, SHREC'09 *Querying with Partial Models* and SHREC'10 *Range Scan Retrieval*, we compared against existing results of the participating contestants. More specifically, on the SHREC'09 *Querying with Partial Models* we compared against the variations of CMVD (Compact MultiView Descriptor) by Daras and Axenopoulos [34] and the BF-SIFT and BF-GridSIFT methods by Furuya and Ohbuchi. On the

SHREC'10 *Range Scan Retrieval* dataset we compared against the variations of the BF-DSIFT-E method proposed by Ohbuchi and Furuya [83] and the variations of the SURFLET method proposed by Hillebrand et al. [133]. Furthermore, on this dataset we compared against the initial version of our method (without Bag-of-Visual-Words) presented in [103]. In the SHREC'11 *Shape Retrieval Contest of Range Scans* competition, due to lack of participants no results were published. We publish our results on this dataset for future reference.

According to the SHREC'09 classification scheme, the target subset is composed of 720 complete 3D models, classified into 40 classes, each of which contains 18 models. The query set is composed of 20 range images taken from 20 objects from arbitrary view directions.

Figure 46, using the experimental results given in [39], illustrates the P-R scores for the complete SHREC'09 *Querying with Partial Models* dataset for the proposed 3D model retrieval method (PV - BoVW) and the methods by Daras and Furuya. Table 10 shows the corresponding five quantitative measures for the same methods.

Both the P-R scores of Fig. 46, as well as the quantitative measures of Table 10 illustrate that the proposed method achieves superior performance compared to the variations of the CMVD, as well as both the BF-SIFT and the BF-GridSIFT retrieval methods.

The next dataset, SHREC'10 *Range Scan Retrieval* is composed of the following two subsets: the target subset which contains 800 complete 3D models, classified into 40 classes, each of which has 20 models and the query subset which contains 120 range images that have been acquired by capturing 3 range scans of 40 objects from arbitrary view directions.

In Figure 47, using the experimental results given in [39], we show the P-R scores for the complete SHREC'10 *Range Scan Retrieval* dataset for the proposed 3D object retrieval method and the methods by Ohbuchi and Hillebrand. Table 11 shows the corresponding five quantitative measures for the same methods.

Table 11 shows that the proposed method has the highest scores on three out of five measures (and is also close on the remaining two measures, by a small margin). The P-R scores of Fig. 47, illustrate that the proposed method outperforms the track contestants, as well as our previously proposed method (PanoramicViews - SIFT), presented in [103].

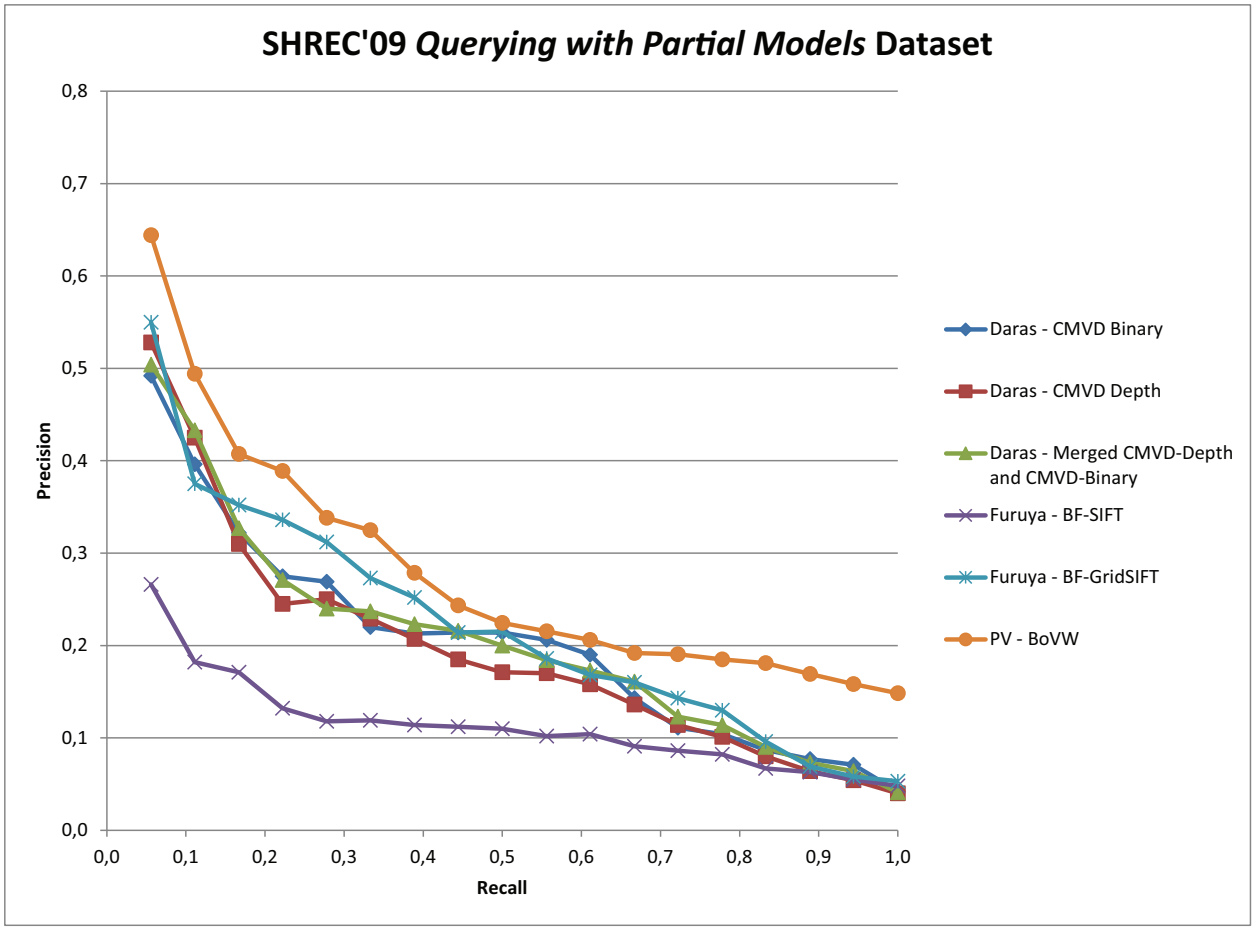


Figure 46: Comparative results based on the average P-R scores for the SHREC'09 *Querying with Partial Models* dataset.

Finally, in the SHREC'11 classification scheme, the target subset is composed of 1000 complete 3D models, categorized into 50 classes, each of which contains 20 models. The query set is composed of 150 range images acquired by capturing 3 range scans, of each of 50 objects that correspond to the above classes, from arbitrary view directions.

In Table 12 we show the five quantitative measures for the complete SHREC'11 *Shape Retrieval Contest of Range Scans* dataset for the proposed 3D model retrieval method.

6.3.0.3 Remark: For the corresponding numerical results, please refer to the appropriate tables in Appendix D.

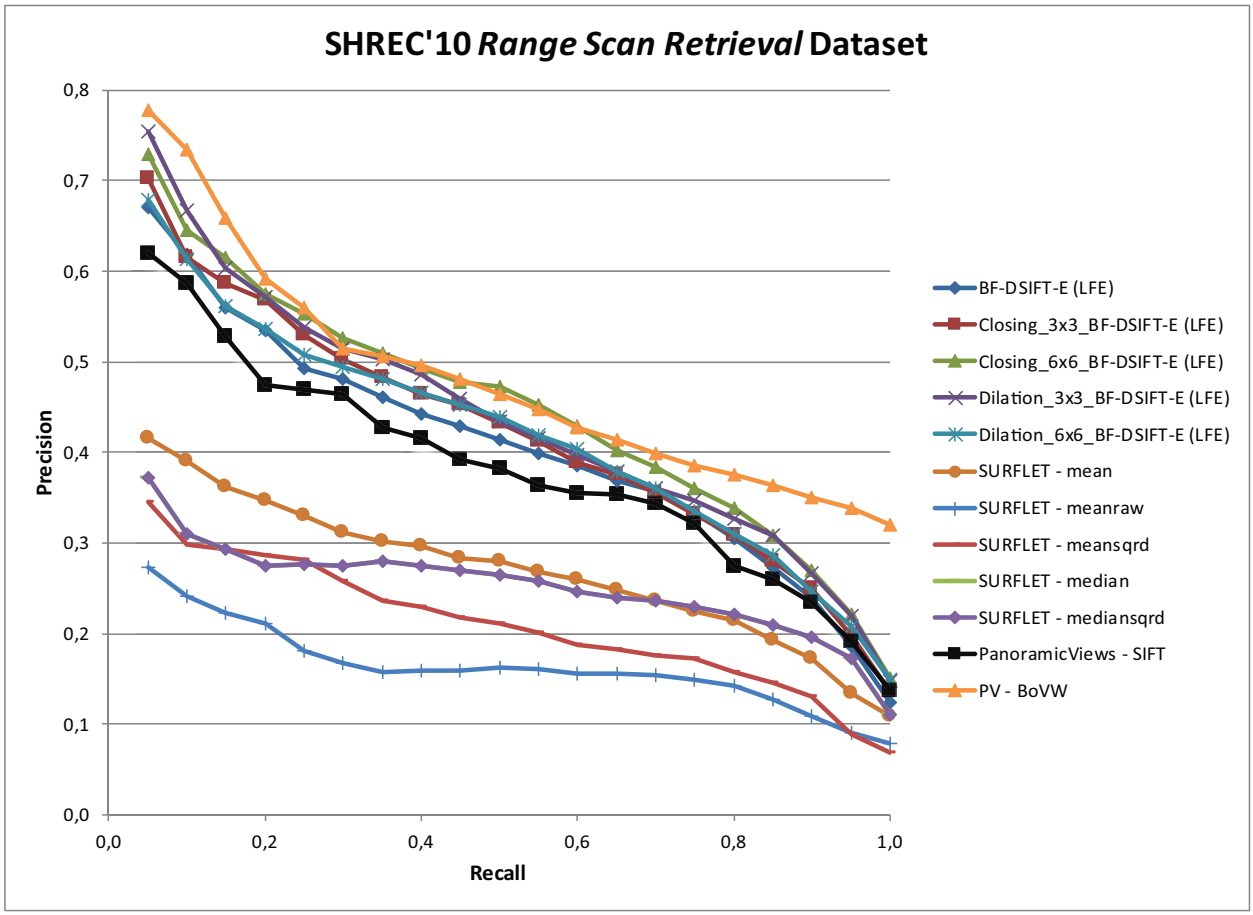


Figure 47: Comparative results based on the average P-R scores for the SHREC'10 *Range Scan Retrieval* dataset.

6.4 Computational Time

The methods presented in this dissertation were developed and tested on an Intel®-Core2Quad 2.5 GHz CPU system, with 6 GB of RAM, running a recent version of Matlab (R2009 - R2012). Based on this system, the computational times of the proposed algorithms are presented in the following.

6.4.1 ROSy+

The methods's speed is dependent on the number of 3D object vertices. The iterations are exhaustive in the current implementation but an optimization method could be developed to improve its speed. For a typical 100,000 vertex object, the time required for the pose normalization process is approximately 60 seconds.

6.4.2 SymPan, SymPan+ on PANORAMA

The methods were developed in a hybrid Matlab/C++/OpenGL architecture, which resulted in low computational times. The average pose normalization time of SymPan for a typical 100,000 vertex 3D model is approximately 15 seconds. The SymPan+ method requires approximately 25 seconds to perform pose normalization for a 100,000 vertex 3D model.

6.4.3 Non-Rigid 3D object Retrieval

The time complexity of the proposed method can be broken down into its two main constituents: (i) graph construction/ feature extraction and (ii) string matching. The first, requires linear time to complete, since both graph construction (including the computation of the conformal factor) and feature extraction are performed in linear time with respect to the faces of the 3D mesh ($O(f)$). The second part, is dependent on the number of string matches between two 3D shapes, dominated by the Hungarian Algorithm, which requires $O(MaxStrings^3)$ time to complete. The total complexity sums up to $O(f) + O(MaxStrings^3)$.

The average descriptor extraction time for a typical 100,000 vertex 3D mesh is approximately 4 seconds.

6.4.4 3D object retrieval based on 2D Range Image Queries

The system was developed in a hybrid Matlab/C++/OpenGL architecture, which resulted in low computational times. The average computation time for *the Bag_of_Visual_Words Model Building* algorithm (see **Algorithm 5**) on a 1,000 3D model database is approximately 4,5 hours (an offline process). The average computation time required for querying a range image object on the aforementioned database (see **Algorithm 6**) is approximately 5 seconds.

7 Conclusions - Further Work

Any element of our past
is compared with the end result

– Demosthenes

In this final chapter, conclusions regarding each of the proposed methods will be drawn. A general review summarising the overall contribution of this thesis, as well as a brief discussion on directions for future work, will be made. To address the problems of 3D model pose normalization, 3D object retrieval with applications to rigid and non-rigid models, as well as image based 3D object retrieval, four novel methodologies have been developed.

In the field of 3D model pose normalization three novel methods, based on the reflective symmetry properties of 3D objects, have been proposed. ROSy successfully complements the CPCA and NPCA methods as a pose normalization preprocessing step for a 3D object retrieval system. The addition of the proposed method increases the discriminative power of the system by approximately 3%, over the previously best approach. The SymPan methodology for 3D model pose normalization is based on a Normals' Deviation Map of the 3D model's surface as well as reflective object symmetry properties. The proposed pose normalization method uses information extracted from the 3D models by projecting them on their circumscribed cylinder. Based on the same principles as the SymPan pose normalization method and driven by the motivation that most 3D objects posses at least one plane of symmetry, in the SymPan+ method we attempt to determine this plane and consecutively align it with one of the principal planes of space. In the sequel, the principal axes of the 3D model are estimated via the computation of the variance on the pixel values of the panoramic views. The proposed pose normalization method is

based on information extracted from the 3D models by projecting them on their circumscribed cylinder. The quality of the proposed alignment methods is proven both visually and through the performance of a state-of-the-art 3D object retrieval system. SymPan improves the discriminative power of the PANORAMA 3D object retrieval system by an average of 6% over the original approach. SymPan+ further improves the discriminative power of the SymPan on PANORAMA 3D object retrieval system by an average of 1%, thus giving a 7% increase over the original approach. Both SymPan and SymPan+ methods exhibit improved performance over ROSy+ by an average of 2 - 3%. All three proposed methods are able to produce high quality alignments of 3D objects, regardless of their originating class or morphology. These alignments are both stable and consistent. The experimentations performed, indicate that the use of reflective symmetry is a promising feature for pose normalization of 3D models. The three proposed methods exploited this feature separately in both 3D and 2D representations. Further work should integrate the estimation of 3D and 2D reflective symmetry as well as the use of PCA analysis for the iterative steps of the procedures.

To address the problem of non-rigid 3D object retrieval the ConTopo++ descriptor has been proposed. This descriptor integrates both geometrical and topological features in a unified descriptor extraction procedure. Furthermore, a graph matching technique based on string matching is introduced. This improved 3D object retrieval methodology, was evaluated not only against the primary form of the algorithm (ConTopo) and the corresponding datasets shown in [102] but also on standard datasets from the SHREC'10 *Non-rigid 3D Models* and SHREC'11 *Non-rigid 3D Watertight Meshes* tracks and the corresponding state-of-the-art 3D shape descriptors. In every case, the proposed non-rigid 3D object retrieval methodology is able to achieve high levels of retrieval accuracy and outperform many of the competing descriptors at a low computational cost. Fig. 48 illustrates some retrieval samples from the SHREC'10 *Non-rigid 3D Models* dataset. As a future step of improvement on the presented non-rigid 3D object retrieval methodology, a segmentation procedure could be used in place of the quantization step during the graph node creation. This could take into account the special features of each 3D object class and thus result into more consistent graph representations. Moreover, the use of dynamic programming could benefit the graph matching strategy by reducing penalties induced

due to poorly correlated segmentations.

Figure 48: Sample queries from the SHREC'10 *Non-rigid 3D Models* dataset. First column indicates the query model and results are illustrated in ranking order. The thumbnails have been taken from the SHREC'10 *Non-rigid 3D Models* dataset.

In the field of image-based 3D object retrieval, we proposed a spatial histograms strategy in a Bag-of-Visual-Words context that fits the information present in panoramic views of 3D objects to the task of partial matching. Special attention has been given to the query image preprocessing stage, where a number of consecutive filters are applied on the images in order to alleviate problems introduced by the digitization process. This improved 3D object retrieval methodology, was evaluated not only against our previous approach [103] and the corresponding SHREC'10 *Range Scan Retrieval* track dataset but also on standard datasets from the SHREC'09 *Querying with Partial Models* track and the corresponding state-of-the-art 3D object retrieval methodologies and the SHREC'11 *Shape Retrieval Contest of Range Scans* track. In every case, the proposed 3D object retrieval method outperforms competing descriptors. Future steps in the field of image-based 3D object retrieval should include alignment of the 3D models and the partial queries, so that the extraction of the 2D image representations could directly match the captured features. This step could greatly increase the speed of the algorithm, as well as its accuracy, by

disqualifying early non-corresponding representations from the matching procedure.

This dissertation successfully addressed a number of problems in the field of 3D Object Retrieval. 3D Object Retrieval can be considered as a relatively new field that falls under Computer Vision, Computer Graphics and Content Based Information Retrieval. The described methodologies have proven to be robust in terms of retrieval accuracy and outperformed previous state-of-the-art methods in the corresponding evaluation tests. These tests were conducted on publicly available datasets.

Appendices

A ROSy+ 3D Object Retrieval Results

Table 4: Quantitative measures of ROSy+ and the CPCA-NPCA pose normalization methods for the PSB train, NTU, MPEG-7, ESB, NIST, McGill datasets. The quantitative measures of Figures 34 and 35 are also presented. All measures are normalized.

Dataset	Method	NN	FT	ST	DCG
PSB Test	ROSy+	**0.779**	**0.524**	**0.659**	**0.756**
	CPCA-NPCA	0.742	0.473	0.606	0.712
	CPCA	0.673	0.426	0.546	0.676
	NPCA	0.677	0.402	0.524	0.664
	ROSy	0.678	0.446	0.550	0.678
	PANORAMA	0.753	0.479	0.603	0.750
	DLA	0.713	0.429	0.552	0.687
	GSMD+SHD+R	0.731	0.472	0.602	0.721
	LFD	0.642	0.375	0.484	0.642
	SH-GEDT	0.553	0.310	0.414	0.584
	DESIRE	0.658	0.404	0.513	0.663
PSB Train	ROSy+	**0.799**	**0.521**	**0.655**	**0.765**
	CPCA-NPCA	0.730	0.460	0.598	0.718
NTU	ROSy+	**0.434**	**0.237**	**0.326**	**0.521**
	CPCA-NPCA	0.413	0.222	0.300	0.503
MPEG-7	ROSy+	**0.879**	**0.619**	**0.731**	**0.837**
	CPCA-NPCA	0.861	0.596	0.707	0.819
ESB	ROSy+	**0.874**	**0.508**	**0.657**	**0.796**
	CPCA-NPCA	0.829	0.465	0.605	0.747
NIST	ROSy+	**0.918**	**0.634**	**0.776**	**0.867**
	CPCA-NPCA	0.881	0.556	0.721	0.841
McGill Articulated	ROSy+	**0.965**	**0.599**	**0.753**	**0.871**
	CPCA-NPCA	0.941	0.568	0.721	0.857
McGill Non Articulated	ROSy+	0.881	**0.517**	**0.696**	**0.822**
	CPCA-NPCA	**0.891**	0.513	0.689	0.817

For a typical 100,000 vertex object, the time required for the pose normalization process is approximately 60 seconds.

B SymPan, SymPan+ on PANORAMA 3D Object Retrieval Results

Table 5: Quantitative measures of SymPan and SymPan+ 3D model pose normalization methods on PANORAMA, and the corresponding 3D object retrieval methods illustrated in Fig 40. All measures are normalized.

Method	NN	FT	ST	DCG
SymPan+, PANORAMA	**0.793**	**0.553**	**0.683**	**0.784**
SymPan, PANORAMA	0.785	0.544	0.673	0.778
PANORAMA	0.753	0.479	0.603	0.750
ROSy+	0.779	0.524	0.659	0.756
GSMD+SHD+R	0.731	0.472	0.602	0.721
DLA	0.713	0.429	0.552	0.687
DESIRE	0.658	0.404	0.513	0.663
LFD	0.642	0.375	0.484	0.642
SH-GEDT	0.553	0.310	0.414	0.584

Table 6: Quantitative measures of SymPan+ 3D model pose normalization method on PANORAMA, the original PANORAMA and the ROSy+ 3D object retrieval methods. Datasets presented are the MPEG-7, ESB and NIST. All measures are normalized.

Method	NN	FT	ST	DCG
MPEG-7				
SymPan+, PANORAMA	**0.883**	**0.641**	**0.751**	**0.858**
PANORAMA	0.872	0.618	0.731	0.831
ROSy+	0.879	0.619	0.731	0.837
ESB				
SymPan+, PANORAMA	**0.880**	**0.526**	**0.678**	**0.815**
PANORAMA	0.865	0.494	0.641	0.795
ROSy+	0.874	0.508	0.657	0.796
NIST				
SymPan+, PANORAMA	**0.925**	**0.656**	**0.800**	**0.901**
PANORAMA	0.906	0.634	0.775	0.852
ROSy+	0.918	0.634	0.776	0.867

The average pose normalization time of SymPan for a typical 100,000 vertex 3D model is approximately 15 seconds. The SymPan+ method requires approximately 25 seconds to perform pose normalization for a 100,000 vertex 3D model.

C ConTopo++ 3D Object Retrieval Results

Table 7: Comparison between ConTopo++ and EMD-PPPT methodologies, as well as possible combinations between their feature sets and partitioning/graph matching strategies using five quantitative measures, on the SHREC'10 *Non-rigid 3D Models* dataset. All measures are normalized.

Feature Set	Partitioning/ Matching Strategy	NN	FT	ST	E	DCG
ConTopo++	ConTopo++	**0.995**	**0.907**	**0.978**	**0.714**	**0.978**
EMD-PPPT	EMD-PPPT	0.965	0.862	0.938	0.675	0.967
ConTopo++	EMD-PPPT	0.870	0.558	0.801	0.550	0.859
EMD-PPPT	ConTopo++	0.923	0.875	0.906	0.750	0.965

Table 8: Five quantitative measures of ConTopo++ and the participants of the SHREC'10 *Non-rigid 3D Models* dataset. All measures are normalized.

Method	NN	FT	ST	E	DCG
ConTopo++	0.9950	0.9070	**0.9780**	**0.7140**	**0.9780**
MR-BF-DSIFT-E	0.9850	**0.9092**	0.9632	0.7055	0.9763
BF-DSIFT-E	0.9800	0.7658	0.8924	0.6447	0.9409
DMEVD-run1	**1**	0.8611	0.9571	0.7012	0.9773
DMEVD-run2	0.9950	0.7884	0.9442	0.6796	0.9612
DMEVD-run3	0.9600	0.7189	0.8505	0.6157	0.9203
CF	0.9200	0.6347	0.7800	0.5527	0.8781

Table 9: Five quantitative measures of ConTopo++ and the participants of the SHREC'11 *Non-rigid Watertight 3D Meshes* dataset. All measures are normalized.

Method	NN	FT	ST	E	DCG
SD-GDM-meshSlFT	1.0000	0.9720	0.9901	0.7358	0.9955
MDS-CM-BOF	0.9950	0.9127	0.9691	0.7166	0.9822
ConTopo++	0.9930	0.8850	0.9520	0.6950	0.9810
OrigM-n12-normA	0.9917	0.9153	0.9569	0.7047	0.9783
FOG+MRR	0.9600	0.8810	0.9461	0.6958	0.9586
BOGH	0.9933	0.8111	0.8839	0.6469	0.9493
LSF	0.9950	0.7988	0.8631	0.6327	0.9432
T-NoNorm-40Coef	0.9550	0.6717	0.8026	0.5791	0.8972
PatchBOF_150	0.7483	0.6419	0.8333	0.5881	0.8367
HKS	0.8367	0.4061	0.4973	0.3525	0.7304

The average descriptor extraction time for a typical 100,000 vertex 3D mesh is approximately 4 seconds.

D 3D object Retrieval based on 2D Range Image Queries Results

Table 10: Comparison between the proposed method and the methods presented on the SHREC'09 *Querying with Partial Models* track using five quantitative measures. All measures are in the interval $[0, 1]$.

Method	NN	FT	ST	E	DCG
CMVD-BINARY	0.350	0.217	0.283	0.200	0.521
CMVD-DEPTH and CMVD-BINARY	0.450	0.197	0.267	0.174	0.511
BF-SIFT	0.150	0.114	0.186	0.116	0.423
BF-GridSIFT	0.450	0.225	**0.297**	0.204	0.532
PV - BoVW	**0.600**	**0.251**	0.292	**0.206**	**0.553**

Table 11: Comparison between the proposed method and the methods presented on the SHREC'10 *Range Scan Retrieval* track using five quantitative measures. All measures are in the interval [0,1].

Method	NN	FT	ST	E	DCG
BF-DSIFT-E (LFE)	0.573	0.380	0.524	0.367	0.683
Closing_3x3_BF-DSIFT-E (LFE)	0.598	0.393	0.535	0.382	0.696
Closing_6x6_BF-DSIFT-E (LFE)	0.650	**0.424**	0.569	**0.398**	0.713
Dilation_3x3_BF-DSIFT-E (LFE)	0.675	0.405	0.557	0.392	0.713
Dilation_6x6_BF-DSIFT-E (LFE)	0.547	0.395	0.550	0.386	0.696
SURFLET - mean	0.325	0.244	0.363	0.252	0.556
SURFLET - meanraw	0.171	0.153	0.242	0.163	0.462
SURFLET - meansqrd	0.231	0.197	0.322	0.213	0.513
SURFLET - median	0.282	0.226	0.325	0.224	0.528
SURFLET - mediansqrd	0.282	0.226	0.325	0.224	0.528
PanoramicViews - SIFT	0.512	0.374	0.466	0.256	0.598
PV - BoVW	**0.691**	0.413	**0.570**	0.386	**0.720**

Table 12: Five quantitative measures for the proposed 3D object retrieval method on the SHREC'11 *Shape Retrieval Contest of Range Scans* dataset. All measures are normalized.

Method	NN	FT	ST	E	DCG
PV - BoVW	0.512	0.374	0.466	0.256	0.598

The average computation time for *the Bag_of_Visual_Words Model Building* algorithm (see **Algorithm 5**) on a 1,000 3D model database is approximately 4,5 hours (an offline pro-

cess). The average computation time required for querying a range image object on the aforementioned database (see **Algorithm 6**) is approximately 5 seconds.

Notation

$\mathbb{R}$	Real numbers
$\mathbb{R}^m$	m-dimensional vector space
$\mathbb{R}^{m\times n}$	$m \times n$ vector space
$\mathbb{Z}$	Integers
$\mathbb{Z}^{m\times n}$	$m \times n$ integer grid
$\mathcal{C}^m$	Class of m-times continuously differentiable functions
$\mathcal{C}^\infty$	Class of smooth functions
$\lvert\cdot\rvert$	Absolute value
$\lVert\cdot\rVert$	Norm
$\lVert\cdot\rVert_p$	p-norm
∇f	Gradient of f
$\nabla^2\phi = f$	Poisson equation
$\mathbf{a} = \begin{bmatrix} a_1 & \cdots & a_m \end{bmatrix}^T$	Vector $\mathbf{a} \in \mathbb{R}^m$
$a_i = [a]_i = a(i)$	i^{th} component of a vector
$\mathbf{A} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}$	Matrix $\mathbf{A} \in \mathbb{R}^{m\times n}$
$a_{ij} = [A]_{ij} = A(i,j)$	(i,j) element of a matrix
$\mathbf{A}_j = \mathbf{A}(:,j)$	j^{th} column-vector of a matrix
$\mathbf{A}^T$, $\mathbf{a}^T$	Transpose of a matrix or vector
$\mathbf{A}^{-1}$	Inverse of a matrix
$det(\mathbf{A})$	Determinant of a matrix
$trace(\mathbf{A})$	Trace of a matrix
$rank(\mathbf{A})$	Rank of a matrix
$diag(\mathbf{A})$	Diagonal elements of a matrix in vector form
$\mathbf{1}$	Identity matrix
$\odot$	Element-wise (Hadamard) product
f , $\mathbf{f}$	Scalar/vector function
df	Differential of f

Δf	Laplace-Beltrami operator on f
I	Intensity (B/W) image
I_G	Geometry map/image
I_N	Normal map/image
I_Z	Depth map/image
$\mathrm{Pr}[\cdot]$	Probability
$\mathrm{Var}[\cdot]$	Variance
$\mathrm{Covar}[\cdot,\cdot]$	Covariance
D	Distance measure
S	Similarity measure
g	Low-pass Haar filter
h	High-pass Haar filter
$\leftarrow$, $\rightarrow$	Mapping
$:=$	Assignment
$\ldots$	Up to
$O(\cdot)$	Algorithm complexity

Acronyms

2D	Two-Dimensional
3D	Three-Dimensional
4D	Four-Dimensional / 3D Video
AABB	Axis Aligned Bounding Box
ARG	Attributed Reeb Graph
ART	Angular Radial Transform
BoW	Bag-of-Words
BoF	Bag-of-Features
BoVW	Bag-of-Visual-Words
CMVD	Compact Multi-View Descriptor
ConTopo	Conformal geometry and graph-based Topological information, non-rigid 3D object retrieval method
CPCA	Continuous PCA
CPU	Central Processing Unit
CRSP	Concrete Radialized Spherical Projection
DB	Data Base
DCG	Discounted Cumulative Gain
DGI	Depth Gradient Image
DLA	Depth Line-based Approach
DoG	Difference of Gaussians
DPD	Dynamic Programming Distance
DSIFT	Dense SIFT
E	E-Measure
EMD	Earth Mover's Distance
ESB	Engineering Shape Benchmark
FT	First Tier
GEDT	Gaussian Euclidean Distance Transform
GPU	Graphics Processing Unit
GSMD	Geodesic Sphere based Multi-view Descriptor
LF	Light Field

LTIC	Local Translational Invariance Cost
MODFT	Multiple Orientation Depth Fourier Transform
MPEG	Motion Pictures Experts Group
MRG	Multiresolution Reeb Graph
NDM	Normals' Deviation Map
NHI	Normalized Histogram Intersection
NIST	National Institute of Standards and Technology, USA
NN	Nearest Neighbor
NPCA	Normal PCA
NTU	National Taiwan University, Taiwan
P-R	Precision - Recall
P3DS	Perceptual 3D Shape Descriptor
PCA	Principal Component Analysis
PF	Penalizing Factor
PRSD	Planar Reflective Symmetry Descriptor
PRST	Planar Reflective Symmetry Transform
PSB	Princeton Shape Benchmark
PV	Panoramic View
RAM	Random Access Memory
ROSy	Reflective Object Symmetry pose normalization method
SDM	String Difference Matrix
SEGI	Sorted Extended Gaussian Image
SH	Spherical Harmonics
SHD	Spherical Harmonics Descriptor
SHD+R	Spherical Harmonics Descriptor + Rectilinearity
SHREC	SHape Retrieval Contest
SIFT	Scale Invariant Feature Transform
SoMBB	Surface-oriented Minimum Bounding Box
ST	Second Tier
SVD	Singular Value Decomposition
SymPan	Panoramic views and reflective Symmetry pose normalization method
TOSCA	TOols for non-rigid Shape Comparison and Analysis

Index

References

[1] A. Adán, P. Merchán, and S. Salamanca, "3D scene retrieval and recognition with depth gradient images," *Pattern Recognition Letters*, vol. 32, no. 9, pp. 1337–1353, 2011.

[2] A. Agathos, I. Pratikakis, P. Papadakis, S. J. Perantonis, P. N. Azariadis, and N. S. Sapidis, "3D articulated object retrieval using a graph-based representation," *The Visual Computer*, vol. 26, no. 10, pp. 1301–1319, 2010.

[3] H.-K. Ahn, P. Brass, and C.-S. Shin, "Maximum overlap and minimum convex hull of two convex polyhedra under translations," *Comput. Geom.*, vol. 40, no. 2, pp. 171–177, 2008.

[4] H.-K. Ahn, O. Cheong, C.-D. Park, C.-S. Shin, and A. Vigneron, "Maximizing the overlap of two planar convex sets under rigid motions," in *Symposium on Computational Geometry*, J. S. B. Mitchell and G. Rote, Eds. ACM, 2005, pp. 356–363.

[5] F. Alizadeh and A. Sutherland, "3D model retrieval using the 2D poisson equation," in *CBMI*, P. Lambert, Ed. IEEE, 2012, pp. 1–6.

[6] A. M. Andrew, "Level set methods and fast marching methods: Evolving interfaces in computational geometry, fluid mechanics, computer vision, and materials science, by j.a. sethian, cambridge university press, cambridge, uk, 2nd edn 1999 (first published 1996 as *level set methods*) xviii + 420 pp., isbn (paperback) 0-521-64557-3, (hardback) 0-521-64204-3 (pbk, £18.95)," *Robotica*, vol. 18, no. 1, pp. 89–92, 2000.

[7] M. Ankerst, G. Kastenmüller, H.-P. Kriegel, and T. Seidl, "3D shape histograms for similarity search and classification in spatial databases," in *SSD*, ser. Lecture Notes in Computer Science, R. H. Güting, D. Papadias, and F. H. Lochovsky, Eds., vol. 1651. Springer, 1999, pp. 207–226.

[8] R. Arandjelovic and A. Zisserman, "Three things everyone should know to improve object retrieval," in *CVPR*. IEEE, 2012, pp. 2911–2918.

[9] A. Axenopoulos, G. Litos, and P. Daras, "3D model retrieval using accurate pose estimation and view-based similarity," in *ICMR*, F. G. B. D. Natale, A. D. Bimbo, A. Hanjalic, B. S. Manjunath, and S. Satoh, Eds. ACM, 2011, p. 41.

[10] F. S. Beckman and Jr, "On isometries of euclidean spaces," *Proceedings of the American Mathematical Society*, vol. 4, no. 5, pp. 810–815, 1953.

[11] M. Ben-Chen and C. Gotsman, "Characterizing shape using conformal factors," in *3DOR*, S. J. Perantonis, N. S. Sapidis, M. Spagnuolo, and D. Thalmann, Eds. Eurographics Association, 2008, pp. 1–8.

[12] A. Bhattacharyya, "On a measure of divergence between two statistical populations defined by their probability distributions," *Bulletin of the Calcutta Mathematical Society*, vol. 35, pp. 99–109, 1943.

[13] S. Biasotti, D. Giorgi, M. Spagnuolo, and B. Falcidieno, "Reeb graphs for shape analysis and applications," *Theor. Comput. Sci.*, vol. 392, no. 1-3, pp. 5–22, 2008.

[14] S. Biasotti, G. Patanè, M. Spagnuolo, B. Falcidieno, and G. Barequet, "Shape approximation by differential properties of scalar functions," *Computers & Graphics*, vol. 34, no. 3, pp. 252–262, 2010.

[15] J.-D. Boissonnat and P. Alliez, Eds., *Second Eurographics Symposium on Geometry Processing, Nice, France, July 8-10, 2004*, ser. ACM International Conference Proceeding Series, vol. 71. Eurographics Association, 2004.

[16] A. Bosch, A. Zisserman, and X. Muñoz, "Scene classification via pLSA," in *ECCV (4)*, ser. Lecture Notes in Computer Science, A. Leonardis, H. Bischof, and A. Pinz, Eds., vol. 3954. Springer, 2006, pp. 517–530.

[17] ——, "Image classification using random forests and ferns," in *ICCV*. IEEE, 2007, pp. 1–8.

[18] A. Bronstein, M. Bronstein, and R. Kimmel, *Numerical Geometry of Non-Rigid Shapes*, 1st ed. Springer Publishing Company, Incorporated, 2008.

[19] A. M. Bronstein, M. M. Bronstein, and R. Kimmel, "Efficient computation of isometry-invariant distances between surfaces," *SIAM J. Scientific Computing*, vol. 28, no. 5, pp. 1812–1836, 2006.

[20] A. M. Bronstein, M. M. Bronstein, R. Kimmel, M. Mahmoudi, and G. Sapiro, "A gromov-hausdorff framework with diffusion geometry for topologically-robust non-rigid shape matching," *International Journal of Computer Vision*, vol. 89, no. 2-3, pp. 266–286, 2010.

[21] B. Bustos, D. A. Keim, D. Saupe, T. Schreck, and D. V. Vranic, "An experimental comparison of feature-based 3D retrieval methods," in *3DPVT*. IEEE Computer Society, 2004, pp. 215–222.

[22] G. Carlsson, A. Zomorodian, A. D. Collins, and L. J. Guibas, "Persistence barcodes for shapes," in *Symposium on Geometry Processing*, ser. ACM International Conference Proceeding Series, J.-D. Boissonnat and P. Alliez, Eds., vol. 71. Eurographics Association, 2004, pp. 124–135.

[23] C. Chan and S. Tan, "Determination of the minimum bounding box of an arbitrary solid: an iterative approach," *Computers & Structures*, vol. 79, no. 15, pp. 1433 – 1449, 2001.

[24] M. Chaouch and A. Verroust-Blondet, "Enhanced 2D/3D approaches based on relevance index for 3D-shape retrieval," in *SMI*. IEEE Computer Society, 2006, p. 36.

[25] ——, "3D model retrieval based on depth line descriptor," in *ICME*. IEEE, 2007, pp. 599–602.

[26] ——, "A new descriptor for 2D depth image indexing and 3D model retrieval," in *ICIP (6)*. IEEE, 2007, pp. 373–376.

[27] ——, "Alignment of 3D models," *Graphical Models*, vol. 71, no. 2, pp. 63–76, 2009.

[28] D.-Y. Chen, X.-P. Tian, Y.-T. Shen, and M. Ouhyoung, "On visual similarity based 3D model retrieval," *Comput. Graph. Forum*, vol. 22, no. 3, pp. 223–232, 2003.

[29] H. Chen and B. Bhanu, "3D free-form object recognition in range images using local surface patches," *Pattern Recognition Letters*, vol. 28, no. 10, pp. 1252–1262, 2007.

[30] E. Cheng, N. Xie, H. Ling, P. R. Bakic, A. D. A. Maidment, and V. Megalooikonomou, "Mammographic image classification using histogram intersection," in *ISBI*. IEEE, 2010, pp. 197–200.

[31] K. Cole-McLaughlin, H. Edelsbrunner, J. Harer, V. Natarajan, and V. Pascucci, "Loops in reeb graphs of 2-manifolds," *Discrete & Computational Geometry*, vol. 32, no. 2, pp. 231–244, 2004.

[32] N. D. Cornea, M. F. Demirci, D. Silver, A. Shokoufandeh, S. J. Dickinson, and P. B. Kantor, "3D object retrieval using many-to-many matching of curve skeletons," in *SMI*. IEEE Computer Society, 2005, pp. 368–373.

[33] M. Daoudi, T. Schreck, M. Spagnuolo, I. Pratikakis, R. C. Veltkamp, and T. Theoharis, Eds., *Eurographics Workshop on 3D Object Retrieval, Norrköping, Sweden, May 2, 2010, Proceedings*. Eurographics Association, 2010.

[34] P. Daras and A. Axenopoulos, "A compact multi-view descriptor for 3D object retrieval," in *CBMI*, S. D. Kollias and Y. S. Avrithis, Eds. IEEE Computer Society, 2009, pp. 115–119.

[35] P. Daras, D. Zarpalas, D. Tzovaras, and M. G. Strintzis, "Efficient 3D model search and retrieval using generalized 3D radon transforms," *IEEE Transactions on Multimedia*, vol. 8, no. 1, pp. 101–114, 2006.

[36] M. de Berg and M. J. van Kreveld, "Trekking in the alps without freezing or getting tired," in *ESA*, ser. Lecture Notes in Computer Science, T. Lengauer, Ed., vol. 726. Springer, 1993, pp. 121–132.

[37] T. K. Dey, J. Giesen, and S. Goswami, "Shape segmentation and matching with flow discretization," in *WADS*, ser. Lecture Notes in Computer Science, F. K. H. A. Dehne, J.-R. Sack, and M. H. M. Smid, Eds., vol. 2748. Springer, 2003, pp. 25–36.

[38] E. R. Dougherty, *An Introduction to Morphological Image Processing*, 3rd ed., ser. Tutorial Texts in Optical Engineering. SPIE, 1992, no. 9.

[39] H. Dutagaci, A. Godil, A. Axenopoulos, P. Daras, T. Furuya, and R. Ohbuchi, "SHREC'09 track: Querying with partial models," in *3DOR*, M. Spagnuolo, I. Pratikakis, R. C. Veltkamp, and T. Theoharis, Eds. Eurographics Association, 2009, pp. 69–76.

[40] H. Dutagaci, A. Godil, C. P. Cheung, T. Furuya, U. Hillenbrand, and R. Ohbuchi, "SHREC'10 track: Range scan retrieval," in *3DOR*, M. Daoudi, T. Schreck, M. Spagnuolo, I. Pratikakis, R. C. Veltkamp, and T. Theoharis, Eds. Eurographics Association, 2010, pp. 109–115.

[41] M. Elad, A. Tal, and S. Ar, "Content based retrieval of VRML objects: an iterative and interactive approach," in *Proceedings of the sixth Eurographics workshop on Multimedia 2001*. Springer-Verlag New York, Inc., 2002, pp. 107–118.

[42] A. E. (Elbaz) and R. Kimmel, "On bending invariant signatures for surfaces," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 10, pp. 1285–1295, 2003.

[43] R. Fang, A. Godil, X. Li, and A. Wagan, "A new shape benchmark for 3D object retrieval," in *ISVC (1)*, ser. Lecture Notes in Computer Science, G. Bebis, R. D. Boyle, B. Parvin, D. Koracin, P. Remagnino, F. M. Porikli, J. Peters, J. T. Klosowski, L. L. Arns, Y. K. Chun, T.-M. Rhyne, and L. Monroe, Eds., vol. 5358. Springer, 2008, pp. 381–392.

[44] R. W. Floyd, "Algorithm 97: Shortest path," *Commun. ACM*, vol. 5, no. 6, p. 345, 1962.

[45] A. Frome, D. Huber, R. Kolluri, T. Bülow, and J. Malik, "Recognizing objects in range data using regional point descriptors," in *Proceedings of the European Conference on Computer Vision (ECCV)*, ser. Lecture Notes in Computer Science, T. Pajdla and J. Matas, Eds., vol. 3023. Springer Berlin Heidelberg, 2004, pp. 224–237.

[46] T. Funkhouser and P. Shilane, "Partial matching of 3D shapes with priority-driven search," in *Fourth Eurographics symposium on Geometry processing*, 2006, pp. 131–142.

[47] T. Furuya and R. Ohbuchi, "Dense sampling and fast encoding for 3D model retrieval using bag-of-visual features," in *CIVR*, S. Marchand-Maillet and Y. Kompatsiaris, Eds. ACM, 2009.

[48] R. Gal, A. Shamir, and D. Cohen-Or, "Pose-oblivious shape signature," *IEEE Trans. Vis. Comput. Graph.*, vol. 13, no. 2, pp. 261–271, 2007.

[49] P. Geurts, D. Ernst, and L. Wehenkel, "Extremely randomized trees," *Machine Learning*, vol. 63, no. 1, pp. 3–42, 2006.

[50] D. Giorgi, S. Biasotti, and L. Paraboschi, "SHape REtrieval contest 2007: Watertight models track," 2007.

[51] J. Goldsmith and J. Salmon, "Automatic creation of object hierarchies for ray tracing," *IEEE Comput. Graph. Appl.*, vol. 7, no. 5, pp. 14–20, 1987.

[52] H. Goldstein and C. P. Poole, *Classical Mechanics*. Addison Wesley, 2001.

[53] L. Gorelick, M. Galun, E. Sharon, R. Basri, and A. Brandt, "Shape representation and classification using the poisson equation," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 28, no. 12, pp. 1991–2005, 2006.

[54] S. Gottschalk, M. C. Lin, and D. Manocha, "OBBTree: A hierarchical structure for rapid interference detection," in *SIGGRAPH*, 1996, pp. 171–180.

[55] A. B. Hamza and H. Krim, "Geodesic object representation and recognition," in *DGCI*, ser. Lecture Notes in Computer Science, I. Nyström, G. S. di Baja, and S. Svensson, Eds., vol. 2886. Springer, 2003, pp. 378–387.

[56] G. Hetzel, B. Leibe, P. Levi, and B. Schiele, "3D object recognition from range images using local feature histograms," in *CVPR (2)*. IEEE Computer Society, 2001, pp. 394–399.

[57] M. Hilaga, Y. Shinagawa, T. Komura, and T. L. Kunii, "Topology matching for fully automatic similarity estimation of 3D shapes," in *SIGGRAPH*, 2001, pp. 203–212.

[58] V. Jain and H. Zhang, "A spectral approach to shape-based retrieval of articulated 3D models," *Computer-Aided Design*, vol. 39, no. 5, pp. 398–407, 2007.

[59] K. Järvelin and J. Kekäläinen, "Cumulated gain-based evaluation of IR techniques," *ACM Trans. Inf. Syst.*, vol. 20, no. 4, pp. 422–446, 2002.

[60] S. Jayanti, Y. Kalyanaraman, N. Iyer, and K. Ramani, "Developing an engineering shape benchmark for CAD models," *Computer-Aided Design*, vol. 38, no. 9, pp. 939–953, 2006.

[61] A. E. Johnson and M. Hebert, "Using spin images for efficient object recognition in cluttered 3D scenes," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 21, no. 5, pp. 433–449, 1999.

[62] A. Kadyrov and M. Petrou, "The trace transform and its applications," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 23, no. 8, pp. 811–828, 2001.

[63] M. M. Kazhdan, "An approximate and efficient method for optimal rotation alignment of 3D models," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 7, pp. 1221–1229, 2007.

[64] M. M. Kazhdan, B. Chazelle, D. P. Dobkin, A. Finkelstein, and T. A. Funkhouser, "A reflective symmetry descriptor," in *ECCV (2)*, ser. Lecture Notes in Computer Science, A. Heyden, G. Sparr, M. Nielsen, and P. Johansen, Eds., vol. 2351. Springer, 2002, pp. 642–656.

[65] M. M. Kazhdan, T. A. Funkhouser, and S. Rusinkiewicz, "Rotation invariant spherical harmonic representation of 3D shape descriptors," in *Symposium on Geometry Processing*, ser. ACM International Conference Proceeding Series, L. Kobbelt, P. Schröder, and H. Hoppe, Eds., vol. 43. Eurographics Association, 2003, pp. 156–164.

[66] ——, "Symmetry descriptors and 3D shape matching," in *Symposium on Geometry Processing*, ser. ACM International Conference Proceeding Series, J.-D. Boissonnat and P. Alliez, Eds., vol. 71. Eurographics Association, 2004, pp. 115–123.

[67] D. H. Kim, I. K. Park, I. D. Yun, and S. U. Lee, "A new MPEG-7 standard: Perceptual 3D shape descriptor," in *PCM (2)*, ser. Lecture Notes in Computer Science, K. Aizawa, Y. Nakamura, and S. Satoh, Eds., vol. 3332. Springer, 2004, pp. 238–245.

[68] H. W. Kuhn, "The hungarian method for the assignment problem," *Naval Research Logistics Quarterly*, vol. 2, no. 1-2, pp. 83–97, 1955.

[69] H. Laga, T. Schreck, A. Ferreira, A. Godil, I. Pratikakis, and R. C. Veltkamp, Eds., *Eurographics Workshop on 3D Object Retrieval 2011, Llandudno, UK, April 10, 2011. Proceedings*. Eurographics Association, 2011.

[70] Z. Lian, A. Godil, B. Bustos, M. Daoudi, J. Hermans, S. Kawamura, Y. Kurita, G. Lavoué, H. V. Nguyen, R. Ohbuchi, Y. Ohkita, Y. Ohishi, F. Porikli, M. Reuter, I. Sipiran, D. Smeets, P. Suetens, H. Tabia, and D. Vandermeulen, "SHREC '11 track: Shape retrieval on non-rigid 3D watertight meshes," in *3DOR*, H. Laga, T. Schreck, A. Ferreira, A. Godil, I. Pratikakis, and R. C. Veltkamp, Eds. Eurographics Association, 2011, pp. 79–88.

[71] Z. Lian, A. Godil, T. Fabry, T. Furuya, J. Hermans, R. Ohbuchi, C. Shu, D. Smeets, P. Suetens, D. Vandermeulen, and S. Wuhrer, "SHREC'10 track: Non-rigid 3D shape retrieval," in *3DOR*, M. Daoudi, T. Schreck, M. Spagnuolo, I. Pratikakis, R. C. Veltkamp, and T. Theoharis, Eds. Eurographics Association, 2010, pp. 101–108.

[72] Z. Lian, P. L. Rosin, and X. Sun, "Rectilinearity of 3D meshes," *International Journal of Computer Vision*, vol. 89, no. 2-3, pp. 130–151, 2010.

[73] S. P. Lloyd, "Least squares quantization in PCM," *IEEE Transactions on Information Theory*, vol. 28, no. 2, pp. 129–136, 1982.

[74] D. G. Lowe, "Object recognition from local scale-invariant features," in *ICCV*, 1999, pp. 1150–1157.

[75] A. Mademlis, P. Daras, D. Tzovaras, and M. G. Strintzis, "3D object retrieval using the 3D shape impact descriptor," *Pattern Recognition*, vol. 42, no. 11, pp. 2447–2459, 2009.

[76] A. Martinet, C. Soler, N. Holzschuch, and F. X. Sillion, "Accurate detection of symmetries in 3D shapes," *ACM Trans. Graph.*, vol. 25, no. 2, pp. 439–464, 2006.

[77] M. Meyer, M. Desbrun, P. Schröder, and A. H. Barr, "Discrete differential-geometry operators for triangulated 2-manifolds," in *Proc. VisMath*, 2002, pp. 35–57.

[78] P. Minovic, S. Ishikawa, and K. Kato, "Symmetry identification of a 3D object represented by octree," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 15, no. 5, pp. 507–514, 1993.

[79] E. Mitchell and A. Rogers, "Quaternion parameters in the simulation of a spinning rigid body', simulation," *Simulation*, vol. 4, no. 6, pp. 390–396, 1965.

[80] N. J. Mitra, L. J. Guibas, and M. Pauly, "Partial and approximate symmetry detection for 3D geometry," *ACM Trans. Graph.*, vol. 25, no. 3, pp. 560–568, 2006.

[81] M. Novotni and R. Klein, "Shape retrieval using 3D zernike descriptors," *Computer-Aided Design*, vol. 36, no. 11, pp. 1047–1062, 2004.

[82] R. Ohbuchi, T. Minamitani, and T. Takei, "Shape-similarity search of 3D models by using enhanced shape functions," *Computer Applications in Technology*, vol. 23, no. 2-4, pp. 70–85, 2005.

[83] R. Ohbuchi and T. Furuya, "Scale-weighted dense bag of visual features for 3D model retrieval from a partial view 3D model," in *Computer Vision Workshops (ICCV Workshops), 2009 IEEE 12th International Conference on*, 2009, pp. 63 – 70.

[84] R. Ohbuchi, M. Nakazawa, and T. Takei, "Retrieving 3D shapes based on their appearance," in *Multimedia Information Retrieval*, N. Sebe, M. S. Lew, and C. Djeraba, Eds. ACM, 2003, pp. 39–45.

[85] R. Ohbuchi, K. Osada, T. Furuya, and T. Banno, "Salient local visual features for shape-based 3D model retrieval," in *Shape Modeling International*. IEEE, 2008, pp. 93–102.

[86] R. Osada, T. A. Funkhouser, B. Chazelle, and D. P. Dobkin, "Shape distributions," *ACM Trans. Graph.*, vol. 21, no. 4, pp. 807–832, 2002.

[87] P. Papadakis, I. Pratikakis, S. J. Perantonis, and T. Theoharis, "Efficient 3D shape matching and retrieval using a concrete radialized spherical projection representation," *Pattern Recognition*, vol. 40, no. 9, pp. 2437–2452, 2007.

[88] P. Papadakis, I. Pratikakis, T. Theoharis, G. Passalis, and S. J. Perantonis, "3D object retrieval using an efficient and compact hybrid shape descriptor," in *3DOR*, S. J. Perantonis, N. S. Sapidis, M. Spagnuolo, and D. Thalmann, Eds. Eurographics Association, 2008, pp. 9–16.

[89] P. Papadakis, I. Pratikakis, T. Theoharis, and S. J. Perantonis, "Panorama: A 3d shape descriptor based on panoramic views for unsupervised 3d object retrieval," *International Journal of Computer Vision*, vol. 89, no. 2-3, pp. 177–192, 2010.

[90] E. Paquet, M. Rioux, A. M. Murching, T. Naveen, and A. J. Tabatabai, "Description of shape information for 2D and 3D objects," *Sig. Proc.: Image Comm.*, vol. 16, no. 1-2, pp. 103–122, 2000.

[91] G. Passalis, T. Theoharis, and I. A. Kakadiaris, "PTK: A novel depth buffer-based shape descriptor for three-dimensional object retrieval," *Vis. Comput.*, vol. 23, no. 1, pp. 5–14, 2006.

[92] S. J. Perantonis, N. S. Sapidis, M. Spagnuolo, and D. Thalmann, Eds., *Eurographics Workshop on 3D Object Retrieval, 3DOR 2008, Crete, Greece, 2008. Proceedings*. Eurographics Association, 2008.

[93] G. Peyré and L. D. Cohen, "Geodesic remeshing using front propagation," *International Journal of Computer Vision*, vol. 69, no. 1, pp. 145–156, 2006.

[94] J. Podolak, P. Shilane, A. Golovinskiy, S. Rusinkiewicz, and T. A. Funkhouser, "A planar-reflective symmetry transform for 3D shapes," *ACM Trans. Graph.*, vol. 25, no. 3, pp. 549–559, 2006.

[95] M. Reuter, F.-E. Wolter, and N. Peinecke, "Laplace-beltrami spectra as 'Shape-DNA' of surfaces and solids," *Computer-Aided Design*, vol. 38, no. 4, pp. 342–366, 2006.

[96] J. Ricard, D. Coeurjolly, and A. Baskurt, "Generalizations of angular radial transform for 2D and 3D shape retrieval," *Pattern Recognition Letters*, vol. 26, no. 14, pp. 2174–2186, 2005.

[97] Y. Rubner, C. Tomasi, and L. Guibas, "A metric for distributions with applications to image databases," in *Computer Vision, 1998. Sixth International Conference on*, 1998, pp. 59–66.

[98] S. Ruiz-Correa, L. G. Shapiro, and M. Meila, "A new paradigm for recognizing 3-D object shapes from range data," in *Proceedings of the Ninth IEEE International Conference on Computer Vision*, ser. ICCV '03, vol. 2. Washington, DC, USA: IEEE Computer Society, 2003, pp. 1126–.

[99] R. M. Rustamov, "Augmented symmetry transforms," in *Shape Modeling International*. IEEE Computer Society, 2007, pp. 13–20.

[100] ——, "Laplace-beltrami eigenfunctions for deformation invariant shape representation," in *Symposium on Geometry Processing*, ser. ACM International Conference Proceeding Series, A. G. Belyaev and M. Garland, Eds., vol. 257. Eurographics Association, 2007, pp. 225–233.

[101] T. B. Sebastian, P. N. Klein, and B. B. Kimia, "Recognition of shapes by editing shock graphs," in *ICCV*, 2001, pp. 755–762.

[102] K. Sfikas, I. Pratikakis, and T. Theoharis, "ConTopo: Non-rigid 3D object retrieval using topological information guided by conformal factors," in *3DOR*, H. Laga, T. Schreck, A. Ferreira, A. Godil, I. Pratikakis, and R. C. Veltkamp, Eds. Eurographics Association, 2011, pp. 25–32.

[103] ——, "3D object retrieval via range image queries based on SIFT descriptors on panoramic views," in *3DOR*, M. Spagnuolo, M. M. Bronstein, A. M. Bronstein, and A. Ferreira, Eds. Eurographics Association, 2012, pp. 9–15.

[104] ——, "SymPan: 3D Model Pose Normalization via Panoramic Views and Reflective Symmetry," in *3DOR*, U. Castellani, T. Schreck, S. Biasotti, I. Pratikakis, A. Godil, and R. C. Veltkamp, Eds. Eurographics Association, 2013, pp. 41–48.

[105] K. Sfikas, T. Theoharis, and I. Pratikakis, "ROSy+: 3D object pose normalization based on PCA and reflective object symmetry with application in 3D object retrieval," *International Journal of Computer Vision*, vol. 91, no. 3, pp. 262–279, 2011.

[106] ——, "Non-rigid 3D object retrieval using topological information guided by conformal factors," *The Visual Computer*, vol. 28, no. 9, pp. 943–955, 2012.

[107] ——, "3d object retrieval via range image queries in a bag-of-visual-words context," *The Visual Computer*, vol. 29, no. 12, pp. 1351–1361, 2013.

[108] ——, "Pose normalization of 3D models via reflective symmetry on panoramic views," 2014, submitted to.

[109] J.-L. Shih, C.-H. Lee, and J. T. Wang, "A new 3D model retrieval approach based on the elevation descriptor," *Pattern Recognition*, vol. 40, no. 1, pp. 283–295, 2007.

[110] P. Shilane, P. Min, M. M. Kazhdan, and T. A. Funkhouser, "The princeton shape benchmark," in *SMI*. IEEE Computer Society, 2004, pp. 167–178.

[111] Y. Shinagawa, T. L. Kunii, and Y. L. Kergosien, "Surface coding based on morse theory," *IEEE Comput. Graph. Appl.*, vol. 11, no. 5, pp. 66–78, 1991.

[112] J. Shotton, A. W. Fitzgibbon, M. Cook, T. Sharp, M. Finocchio, R. Moore, A. Kipman, and A. Blake, "Real-time human pose recognition in parts from single depth images," in *CVPR*. IEEE, 2011, pp. 1297–1304.

[113] SHREC2011, "Shape retrieval contest of range scans," http://www.itl.nist.gov/iad/vug/sharp/contest/2011/RangeScans/, accessed on 08/2012.

[114] K. Siddiqi, A. Shokoufandeh, S. J. Dickinson, and S. W. Zucker, "Shock graphs and shape matching," *International Journal of Computer Vision*, vol. 35, no. 1, pp. 13–32, 1999.

[115] J. Song and F. Golshani, "Shape-based 3D model retrieval," in *Proceedings of the IEEE International Conference on Tools with Artificial Intelligence*, 2003, pp. 636–640.

[116] T. G. Stavropoulos, P. Moschonas, K. Moustakas, D. Tzovaras, and M. G. Strintzis, "3D model search and retrieval from range images using salient features," *IEEE Transactions on Multimedia*, vol. 12, no. 7, pp. 692–704, 2010.

[117] G. Stockman and L. G. Shapiro, *Computer Vision*, 1st ed. Upper Saddle River, NJ, USA: Prentice Hall PTR, 2001.

[118] C. Sun and J. Sherrah, "3D symmetry detection using the extended gaussian image," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 2, pp. 164–168, 1997.

[119] H. Sundar, D. Silver, N. Gagvani, and S. J. Dickinson, "Skeleton based shape matching and retrieval," in *Shape Modeling International*. IEEE Computer Society, 2003, pp. 130–142, 290.

[120] S. Takahashi, Y. Shinagawa, and T. L. Kunii, "A feature-based approach for smooth surfaces," in *Symposium on Solid Modeling and Applications*, 1997, pp. 97–110.

[121] J. W. H. Tangelder and R. C. Veltkamp, "A survey of content based 3D shape retrieval methods," *Multimedia Tools Appl.*, vol. 39, no. 3, pp. 441–471, 2008.

[122] S. P. Tarasov and M. N. Vyalyi, "Construction of contour trees in 3D in O(n log n) steps," in *Symposium on Computational Geometry*, 1998, pp. 68–75.

[123] S. Theodoridis and K. Koutroumbas, *Pattern recognition*. Academic Press, 1999.

[124] T. Theoharis, G. Papaioannou, N. Platis, and N. M. Patrikalakis, *Graphics and Visualization: Principles & Algorithms*. Natick, MA, USA: A. K. Peters, Ltd., 2007.

[125] J. Tierny, J.-P. Vandeborre, and M. Daoudi, "Partial 3D shape retrieval by reeb pattern unfolding," *Comput. Graph. Forum*, vol. 28, no. 1, pp. 41–55, 2009.

[126] R. Trudeau, *Introduction to Graph Theory*, ser. Dover books on advanced mathematics. Dover Pub., 1993.

[127] T. Tung and F. Schmitt, "The augmented multiresolution reeb graph approach for content-based retrieval of 3D shapes," *International Journal of Shape Modeling*, vol. 11, no. 1, pp. 91–120, 2005.

[128] G. van den Bergen, "Efficient collision detection of complex deformable models using AABB trees," *J. Graph. Tools*, vol. 2, no. 4, pp. 1–13, 1998.

[129] D. V. Vranić, "Content-based classification of 3D-models by capturing spatial characteristics," 2005, accessed on 28 July 2007. [Online]. Available: http://merkur01.inf.uni-konstanz.de/CCCC

[130] D. V. Vranić, D. Saupe, and J. Richter, "Tools for 3D-object retrieval: Karhunen-loeve transform and spherical harmonics," in *IEEE MMSP 2001*, 2001, pp. 293–298.

[131] D. V. Vranic, "3D model retrieval," Ph.D. dissertation, University of Leipzig, 2004.

[132] ——, "DESIRE: a composite 3D-shape descriptor," in *ICME*. IEEE, 2005, pp. 962–965.

[133] E. Wahl, U. Hillenbrand, and G. Hirzinger, "Surflet-pair-relation histograms: A statistical 3D-shape representation for rapid classification," in *3DIM*. IEEE Computer Society, 2003, pp. 474–482.

[134] S. Wang, Y. Wang, M. Jin, X. Gu, and D. Samaras, "3D surface matching and recognition using conformal geometry," in *CVPR (2)*. IEEE Computer Society, 2006, pp. 2453–2460.

[135] P. Xiang, C. Q. Hua, F. X. Gang, and Z. B. Chuan, "Pose insensitive 3D retrieval by poisson shape histogram," in *International Conference on Computational Science (2)*, ser. Lecture Notes in Computer Science, Y. Shi, G. D. van Albada, J. Dongarra, and P. M. A. Sloot, Eds., vol. 4488. Springer, 2007, pp. 25–32.

[136] M. Yu, I. Atmosukarto, W. K. Leow, Z. Huang, and R. Xu, "3D model retrieval with morphing-based geometric and topological feature maps," in *5th ACM SIGMM International Workshop on Multimedia Information Retrieval*, 2003, pp. 656–661.

[137] Z. Yu, S. Zhang, H.-S. Wong, and J. Zhang, "A filter-refinement scheme for 3D model retrieval based on sorted extended gaussian image histogram," in *MLDM*, ser. Lecture Notes in Computer Science, P. Perner, Ed., vol. 4571. Springer, 2007, pp. 643–652.

[138] D. yun Chen and M. Ouhyoung, "A 3D model alignment and retrieval system," in *In Proceedings of International Computer Symposium, Workshop on Multimedia Technologies (dec 2002*, 2002, pp. 1436–1443.

[139] T. Zaharia and F. Preteux., "Shape-based retrieval of 3D mesh models," 2002, pp. 437–440.

[140] T. Zaharia and F. J. Preteux, "3D-shape-based retrieval within the MPEG-7 framework," in *Nonlinear Image Processing and Pattern Analysis*, E. R. Dougherty and J. T. Astola, Eds., vol. 4304. SPIE, 2001, pp. 133–145.

[141] T. B. Zaharia and F. J. Prêteux, "3D versus 2D/3D shape descriptors: a comparative study," in *Image Processing: Algorithms and Systems*, ser. SPIE Proceedings, E. R. Dougherty, J. Astola, and K. O. Egiazarian, Eds., vol. 5289. SPIE, 2004, pp. 47–58.

[142] D. Zarpalas, P. Daras, A. Axenopoulos, D. Tzovaras, and M. G. Strintzis, "3D model search and retrieval using the spherical trace transform," *EURASIP J. Adv. Sig. Proc.*, vol. 2007, 2007.

[143] J. Zhang and K. Huebner, "Using symmetry as a feature in panoramic images for mobile robot applications," in *Robotik*, ser. VDI-Berichte, vol. 1679, 2002, pp. 263–268.

[144] J. Zhang, K. Siddiqi, D. Macrini, A. Shokoufandeh, and S. J. Dickinson, "Retrieving articulated 3D models using medial surfaces and their graph spectra," in *EMMCVPR*, ser. Lecture Notes in Computer Science, A. Rangarajan, B. C. Vemuri, and A. L. Yuille, Eds., vol. 3757. Springer, 2005, pp. 285–300.