ROSy+: 3D Object Pose Normalization based on PCA and Reflective Object Symmetry with Application in 3D Object Retrieval

Konstantinos Sfikas · **Theoharis Theoharis** · **Ioannis Pratikakis**

Abstract A novel pose normalization method based on 3D object reflective symmetry is presented. It is a general purpose global pose normalization method; in this paper it is used to enhance the performance of a 3D object retrieval pipeline. Initially, the axis-aligned minimum bounding box of a rigid 3D object is modified by requiring that the 3D object 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 object and its symmetric object, is calculated. By minimizing the combined distance, the 3D object fits inside its modified axis-aligned bounding box and alignment with the coordinate system is achieved. The proposed method is incorporated in a hybrid scheme, that serves as the alignment method in a 3D object retrieval system. The effectiveness of the 3D object retrieval system, using the hybrid pose normalization scheme, is evaluated in terms of retrieval accuracy and demonstrated using both quantitative and qualitative measures via an extensive

Konstantinos Sfikas
Computer Graphics Laboratory, Department of Informatics and Telecommunications, University of Athens, Athens, Greece
E-mail: ksfikas@di.uoa.gr

Theoharis Theoharis
Computer Graphics Laboratory, Department of Informatics and Telecommunications, University of Athens, Athens, Greece
E-mail: theotheo@di.uoa.gr

Ioannis Pratikakis
Department of Electrical and Computer Engineering,
Democritus University of Thrace,
GR-67100 Xanthi, Greece
E-mail: ipratika@ee.duth.gr

consistent evaluation on standard benchmarks. The results clearly show performance boost against current approaches.

1 Introduction

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. *Pose normalization* of 3D objects is a common preprocessing step in various computer graphics applications (Bustos et al 2004; Shilane et al 2004; Tangelder and Veltkamp 2008; Zaharia and Prêteux 2004). 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 normalization, for every 3D object, a corresponding set of normalization transformations (translation, scaling and rotation) in 3D space must be defined.

In most cases, translation and scale normalization can be achieved by standard techniques. The most frequently used method for performing translation normalization is to position the centroids of 3D objects at the origin. Scale normalization can be performed through the definition of a fixed surrounding object (a sphere or a rectangle) within which every 3D object is contained exactly. Rotation normalization (or 3D object alignment), however, is the most difficult part and still under investigation (Chaouch and Verroust-Blondet 2009; Chen and Ouhyoung 2002; Kazhdan 2007; Paquet 2000; Rustamov 2007; Vranić et al 2001). Although it is relatively easy to perform manual alignment of a 3D
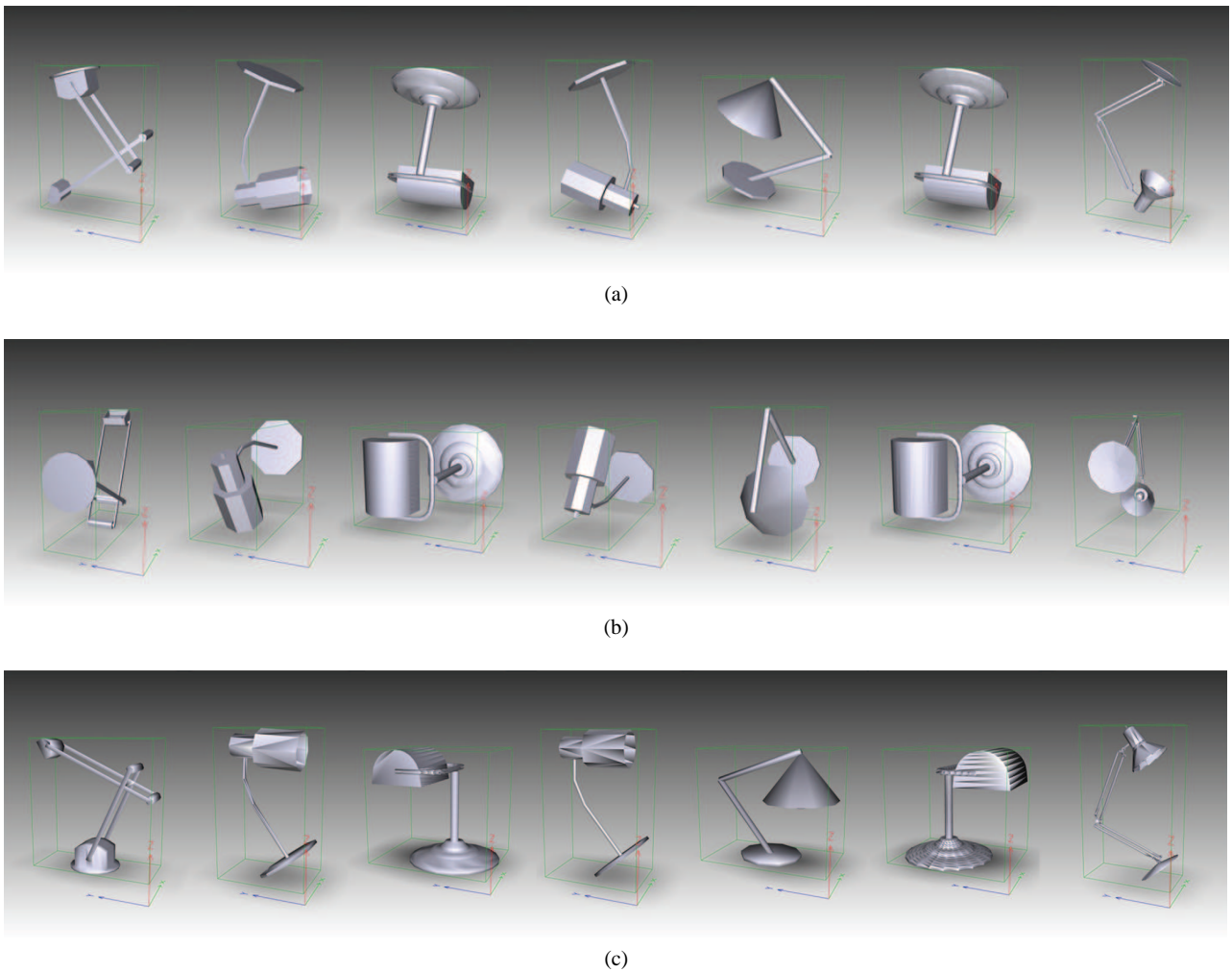
(a)

(b)

(c)

Fig. 1: Comparison between the alignment results of the CPCA (Vranić 2004), NPCA (Papadakis et al 2007) and ROSy methods on objects of the 'LAMP' class of the PSB dataset. (a) CPCA, (b) NPCA, (c) ROSy.

object with a fixed number of rotations and acceptable accuracy, the high complexity and the numerous variations of 3D objects render the automation of such a procedure difficult.

In this paper, rotation normalization is achieved by constraining the 3D objects' bounding boxes to be minimal based on PCA and reflective object symmetry. For this, an axis-aligned bounding box is defined and a distance measure, that estimates the degree of parallelization of a 3D object to the faces of that structure, is given. This distance measure is based on the spatial and angular relation of a 3D object and its symmetric object, with respect to the principal planes, which are set as the planes of symmetry. Experimental results of the proposed method show that the qualitative normalization outcome is improved, compared to current approaches (Figure 1). Additionally, when the proposed method is incorporated in a hybrid pose normalization scheme, it can significantly enhance the discriminative power of a 3D object retrieval system.

We shall define *pose normalization* as the complete normalization procedure which includes rotation, translation and scaling normalization; *alignment* will refer to rotation normalization only. Furthermore, throughout this paper the terms *3D model* and *3D object*, unless otherwise stated, will refer to both vertices and normals sets of a 3D mesh. The term *symmetric object* will refer to a 3D object that is the result of a reflective symmetry transformation on a given 3D object, against a specified principal plane of symmetry.

The remainder of the paper is structured as follows. In Section 2, previous work in pose normalization and 3D object retrieval is discussed. In Section 3, the problem of pose normalization is defined and preliminaries for the presented work are given. Section 4 details the proposed pose normalization method and Section 5 presents detailed experimental results achieved in the course of the method's evaluation. Finally, conclusions are drawn in Section 6.

2 Related Work

In this section an overview of the state-of-the-art in pose normalization methods, with a particular focus on the alignment phase, is presented. A discussion on the state-of-the-art in 3D object retrieval techniques is also included.

2.1 Pose normalization

Pose normalization methods can be divided into three major categories: (i) methods that are based on principal component analysis of the 3D objects; (ii) methods that exploit symmetry characteristics of the 3D objects and (iii) methods that achieve alignment by defining a shape descriptor of the 3D objects. Based on this categorization, an overview of pose normalization methods follows.

The best-known approach for computing the alignment of 3D objects is Principal Component Analysis (PCA) or Karhunen - Loeve transformation (Paquet 2000; Shilane et al 2004; Theodoridis and Koutroumbas 2006; Vranić et al 2001; Zaharia and Prêteux 2004). 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 (Chen et al 2003). To alleviate these problems, Vranic introduced an improvement to the original method, the Continuous PCA (CPCA) algorithm (Vranić 2004; Vranić et al 2001; Vranic 2005). 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 (Papadakis et al 2007, 2008), 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 (Theodoridis and Koutroumbas 2006). In (Elad et al 2001; Osada et al 2002), 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.

Another major category of normalization methods exploits symmetry characteristics found in a large number of 3D objects. Kazhdan et al. (Kazhdan et al 2002) 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. In (Podolak et al 2006), 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 (Rustamov 2007). Minovic et al. (Minovic et al 1993) 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. Martinet et al. (Martinet et al 2006) use generalized moments to detect perfect symmetries in 3D shapes and Mitra et al. (Mitra et al 2006) compute partial and approximate symmetries in 3D objects. Sun and Sherrah (Sun and Sherrah 1997) convert the symmetry detection problem to the correlation of the Gaussian image. Using both PCA-alignment and planar reflective symmetry, Chaouch and Verroust - Blondet (Chaouch and Verroust-Blondet 2009) 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. (Lian et al 2009) 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, (Chaouch and Verroust-Blondet 2009), 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.

A third category of methods achieves rotation invariance by the definition of the shape descriptor. These descriptors are invariant under rotation, but usually discard discriminative information regarding the 3D object. Descriptors based on spherical harmonics (Kazhdan et al 2003; Tangelder and Veltkamp 2008), Zernike moments (Novotni and Klein 2004; Tangelder and Veltkamp 2008) and shell histograms (Ankerst et al 1999; Xiang et al 2007; Yu et al 2007) are examples of representation methods able to achieve rotation invariance by definition. Kazhdan et al. in (Kazhdan et al 2003) 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 (Kazhdan et al 2004). Novotni and Klein in (Novotni and Klein 2004) use 3D Zernike invariants as de-

scriptors for 3D shape retrieval and Ankerst et al. (Ankerst et al 1999) 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 (Chen and Ouhyoung 2002) 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 handle 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 complementarity involved.

2.2 3D Object Retrieval

Content-based 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 3D shape descriptors for generic 3D object retrieval is presented.

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. (Chen et al 2003). 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. (Lian et al 2009) proposed an enhancement to the Light Field descriptor, by computing the same features on projections taken from the vertices of geodesic spheres generated by the regular unit octahedron. Vranic (Vranić 2004) 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. (Zarpalas et al 2007) introduced a 3D shape descriptor called the spherical trace transform, which is the generalization of the 2D trace transform. In this method,

a variety of 2D features are computed for a set of planes intersecting the volume of a 3D object. A newly proposed method is the depth line descriptor proposed by Chaouch and Verroust - Blondet (Chaouch and Verroust-Blondet 2009; Chaouch and Verroust-blondet 2007) 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 (Papadakis et al 2009) 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.

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. (Ankerst et al 1999), 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. (Osada et al 2002) 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 (Zaharia and Preteux 2001) 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 (Xiang et al 2007) propose a rigid transformation insensitive descriptor, called the Poisson shape histogram descriptor, extracted by a voxelized representation of the 3D objects. In (Yu et al 2007) 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. In Zhang et al. (Zhang et al 2005) consider the use of medial surfaces to compute an equivalent directed acyclic graph of an object. In the work of Sundar et al. (Sundar et al 2003), 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. (Cornea et al 2005) 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. The P3DS descriptor developed by Kim et al. (Kim et al 2004) 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 (Kazhdan et al 2004, 2002) 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. Extending this work to every possible plane Podolak et al. presented the planar reflective symmetry transformation (PRST) in (Podolak et al 2006).

Besides the previous categories, combinations of different methods have been considered in order to enhance the overall performance. Vranic (Vranić 2004) 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. (Kazhdan et al 2003) 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. (Papadakis et al 2007), 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 (Novotni and Klein 2004) presented the 3D Zernike descriptor, Daras et al. (Daras et al 2006) introduced the generalized radon transform and Ricard et al. (Ricard et al 2005) developed the 3D ART descriptor by generalizing the 2D angular radial transform. Vranic (Vranic 2005) 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. (Papadakis et al 2008) 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.

3 Preliminaries

In this section, 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 (van den Bergen 1997; Goldsmith and Salmon 1987). Furthermore, using the properties of the reflective symmetry transformation the problem will be mathematically formulated and the foundation for the proposed solution will be set.

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 (Kazhdan et al 2004; Tangelder and Veltkamp 2008; Vranić 2004). 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 2). Pose normalization is the procedure of finding a set of homogeneous transformations (translation, scaling, rotation) that fit a 3D object into its SoMBB.
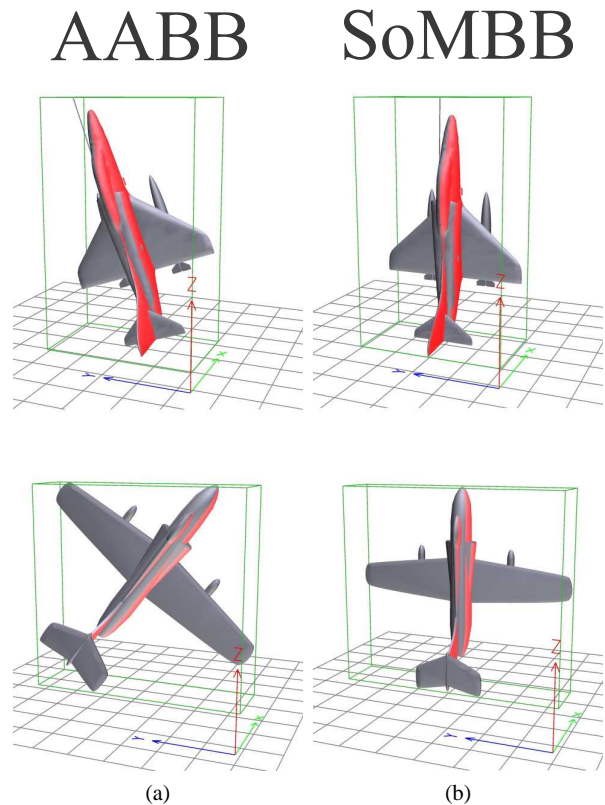

Fig. 2: Difference between the AABB and SoMBB on objects of the same class. On column (a) two 3D objects are enclosed inside their AABB, while on column (b), the same 3D objects are enclosed inside their SoMBB. Marked faces indicate faces whose normals are parallel to the SoMBB's face normals.
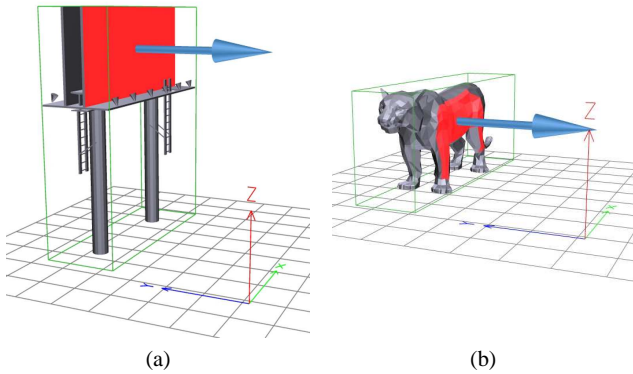
(a)   (b)

Fig. 3: 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). (a) illustrates a 3D object that has a planar surface defined by structure and (b) illustrates a 3D object that defines a planar surface through many small triangles with similar orientation.

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 (Gottschalk et al 1996). Example 3D object classes that explicitly define large planar areas are buildings, airplanes, ships, tables, billboards, etc (Figure 3a). 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 3b).

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

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) (Theoharis et al 2008) is a linear transformation which computes the symmetric object M_{Π}^{-1} about a candidate plane of symmetry Π (Π :

$ax + by + cz + d = 0$). The planes of symmetry used, are the three principal planes of the Cartesian coordinate system (i.e. XY, XZ, YZ). Figure 4 illustrates example reflective symmetry transformations against the principal planes of the Cartesian coordinate system.
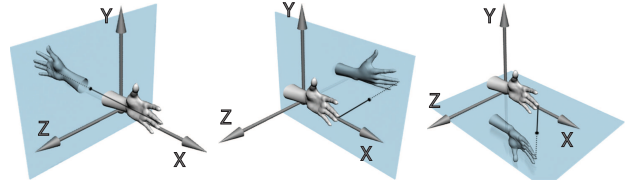

Fig. 4: Examples of the reflective symmetry transformation on planes YZ, XY and XZ, respectively.

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 (Mitchell 1965; Goldstein and Poole 2001). 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.

Note that in the remainder of the paper, when referring to axes or planes that belong to the *target* coordinate system, they will be denoted with uppercase lettering (eg. axis *X* and plane *YZ*), while when referring to axes or planes that belong to the *source* coordinate system of the 3D object, they will be denoted with lowercase lettering (eg. axis *x* and plane *yz*).

4 The proposed 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 (Chan and Tan 2001) 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 (Ahn et al 2005,

2008) 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 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 2a. As has been described before, our motivation is to achieve an alignment which is intuitively shown in Figure 2b. 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 5 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.
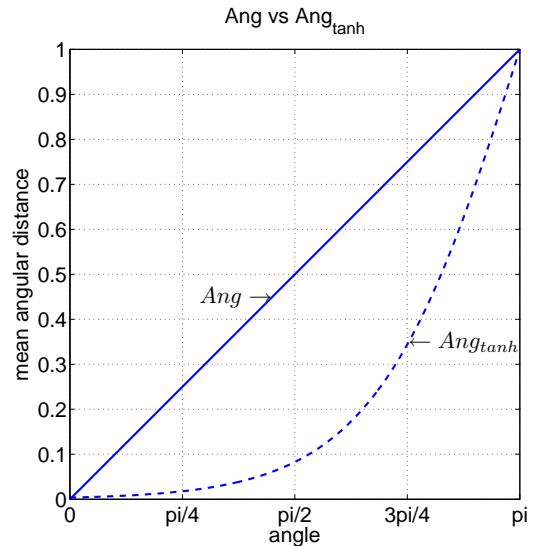
Fig. 5: Graphical representation of $Ang(M,M_\Pi^{-1})$ and $Ang_{tanh}(M,M_\Pi^{-1})$, respectively.

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

---

[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.

$Ang_{tanh}(M, M_\Pi^{-1})$ 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.

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

$Dist(M, M_\Pi^{-1})$ 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 (Papadakis et al 2007; Vranić 2004; Vranić et al 2001). 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 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 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 6). 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.

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

---

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 $(XY < YZ < ZX)$;
28: **return** M_{best};

---

with the remaining axes X and Y of the coordinate system (Figure 7).

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.
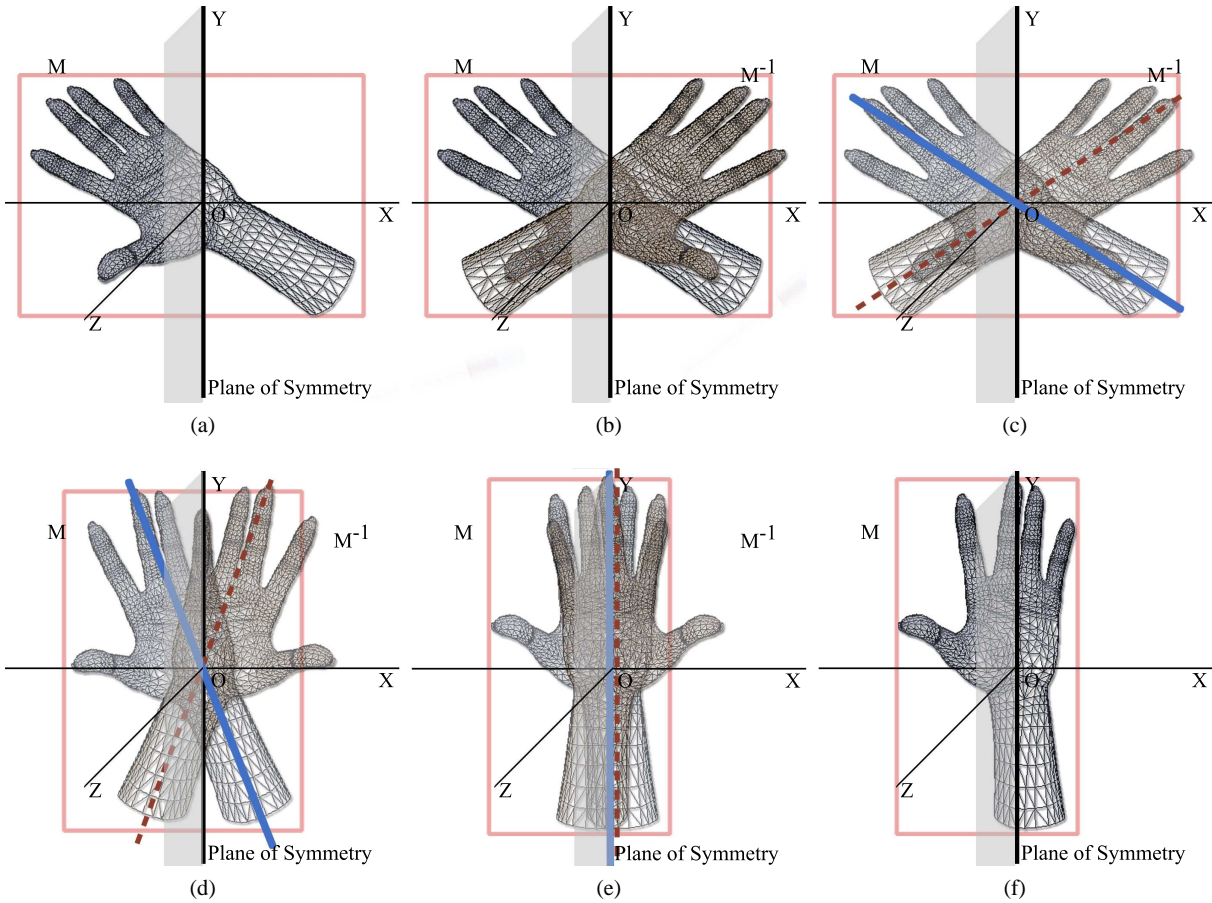
Fig. 6: Illustration of the alignment procedure for a single rotation step on the YZ plane of symmetry. The bounding rectangle illustrates the SoMBB of the 3D object's projection on plane XY. (a) Original 3D object orientation. (b) Original (M) and its symmetric (M_{Π}^{-1}) 3D objects. (c) Selected principal axes projection of the original (dashed axis) and the symmetric (solid axis) 3D object, on principal plane YZ. (d) - (f) Stepwise minimization of the distance between the original and its symmetric 3D object.

5 Experimental Evaluation

This section provides detailed performance results of the ROSy pose normalization method. 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 addressed.

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. We have chosen a state-of-the-art 3D object retrieval methodology, by Papadakis et al. (Papadakis et al 2008) 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) (Shilane et al 2004), the classified objects of the National Taiwan University database (NTU) (Chen et al 2003), the MPEG-7 dataset (Vranić 2005), the Engineering Shape Benchmark dataset (ESB) (Jayanti et al 2006) the National Institute of Standards and Technology dataset (NIST), which contains shape normalized and visually categorized 3D objects found in the SHREC 2009 competition (Fang et al 2008) and both the Articulated and Non Articulated objects of the McGill dataset (Zhang et al 2005). From the NTU dataset, only the classified objects were used, as unclassified objects would not give accurate retrieval results. Table 1 shows the number of categories and the total number of objects in each dataset used for the experiments.

In the 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
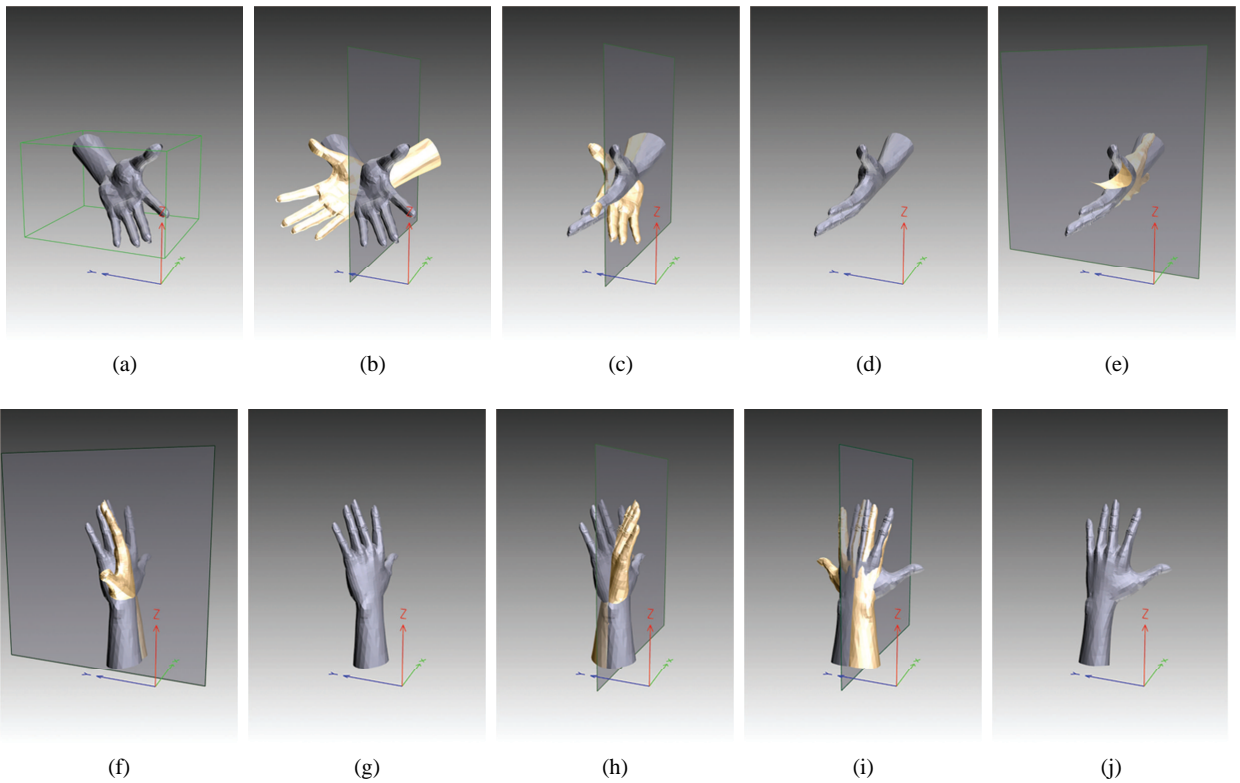
Fig. 7: Overview of all the steps of the alignment procedure. (a) initial 3D object orientation. (b) symmetric 3D object against plane XZ. (c) - (d) minimum k_{XZ} point after rotation of the 3D object and its symmetric object about axis Z. (e) symmetric object against plane YZ. (f) - (g) minimum k_{YZ} point after rotation of the 3D object and its symmetric object about axis Y. (h) symmetric 3D object against plane XZ. (i) minimum k_{XZ} point after rotation of the 3D object and its symmetric object about axis Z. (j) final aligned 3D object.

Table 1: 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

Table 2: 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

these tests ROSy was used as the pose normalization procedure of the 3D object retrieval system and the resulting Discounted Cumulative Gain (DCG) (Jarvelin and Kekalainen 2002) 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 2 confirms that the most suitable choice is the use of equal weight factors in the minimization criterion.

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

(a)

(b)

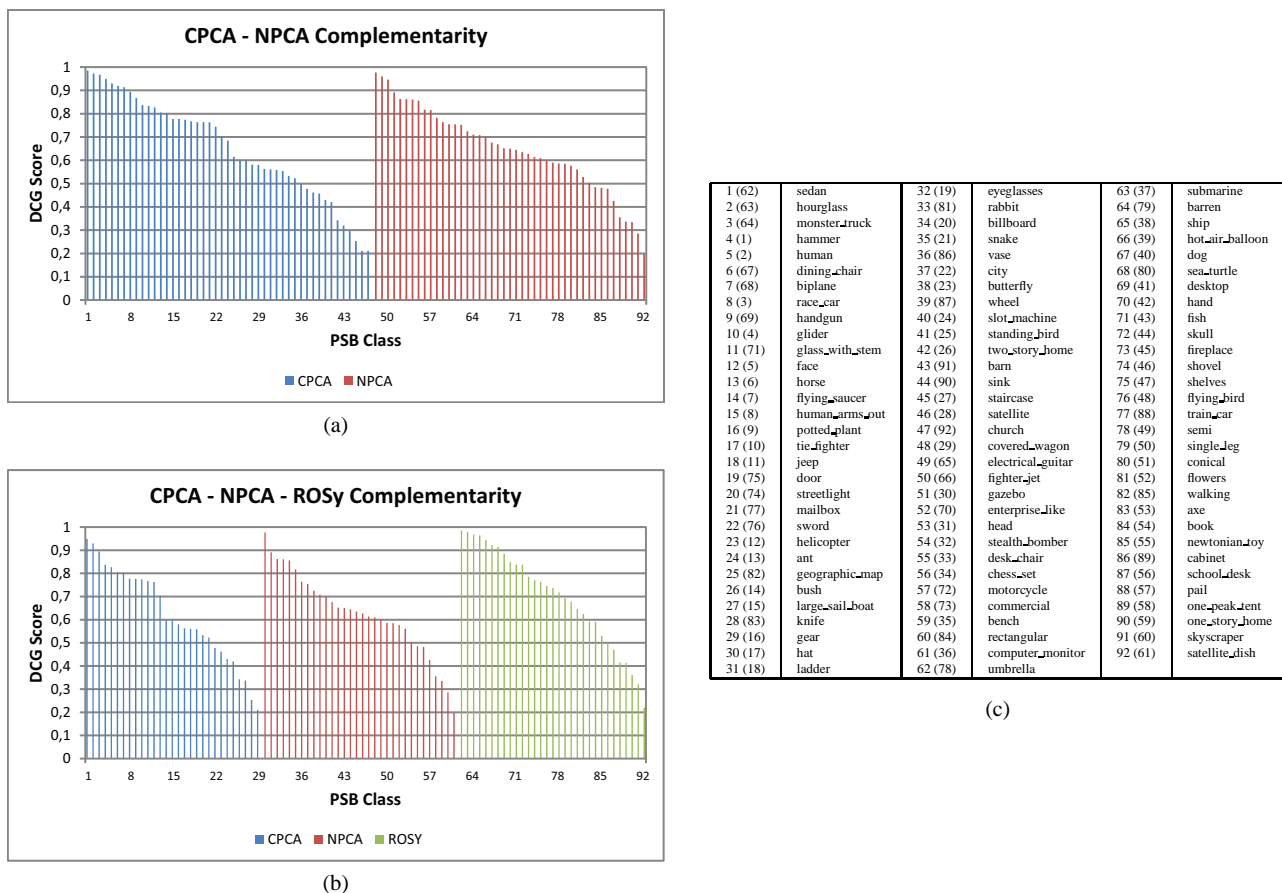| 1 (62)	sedan	32 (19)	eyeglasses	63 (37)	submarine
2 (63)	hourglass	33 (81)	rabbit	64 (79)	barren
3 (64)	monster_truck	34 (20)	billboard	65 (38)	ship
4 (1)	hammer	35 (21)	snake	66 (39)	hot_air_balloon
5 (2)	human	36 (86)	vase	67 (40)	dog
6 (67)	dining_chair	37 (22)	city	68 (80)	sea_turtle
7 (68)	biplane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

(c)

Fig. 8: DCG retrieval scores for the CPCA, NPCA and ROSy pose normalization methods, color coded by the method that achieves the best results per class. (a) Per class CPCA and NPCA complementarity results. (b) Per class CPCA, NPCA and ROSy complementarity results. (c) The correspondence between class ids and class names. Parenthesized ids refer to plot (b).

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 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 8, 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. This quantitative evaluation is based on Precision-Recall (P-R) plots. In this test, for every

query object that belongs to a class C, recall is the percentage of objects of class C that are retrieved and precision is the proportion of retrieved objects that belong to class C over the total number of retrieved objects. The best score is 100% for all plots.

In Figure 9, 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 paper, is compared against the CPCA, NPCA and ROSy standalone methods and the dual (CPCA, NPCA) approach. To make the illustrated results more concrete, four quantitative measures are also displayed: Nearest Neighbor (NN), First Tier (FT), Second Tier (ST) and Discounted Cumulative Gain (DCG) (Shilane et al 2004). 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 for the $(D-1)$ and $2(D-1)$ closest matches respectively, where D is the cardinality of the query's class.
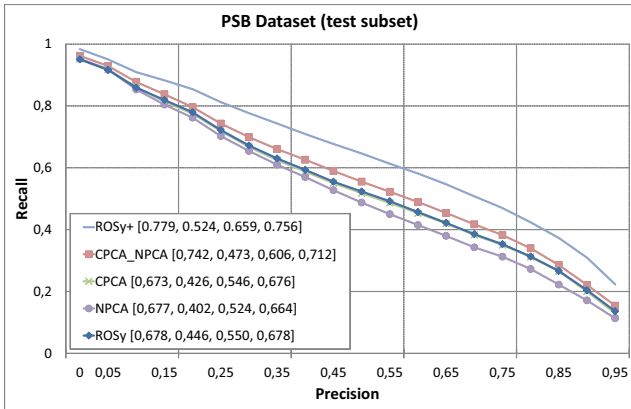
Fig. 9: Precision-Recall plot for the Princeton Shape Benchmark test dataset. The use of ROSy alongside the original approach significantly boosts the performance of the retrieval process.
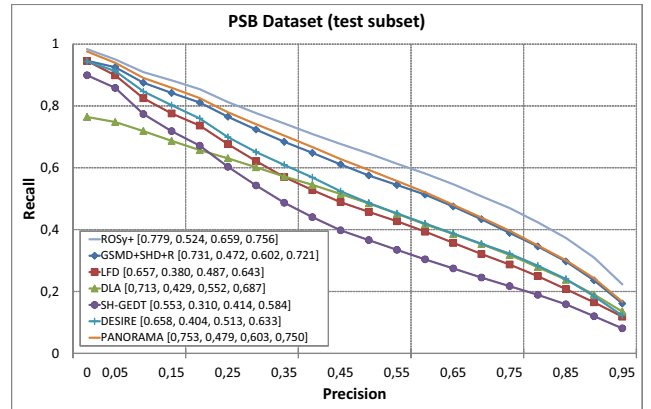

Fig. 10: Precision-Recall plot for the Princeton Shape Benchmark test dataset. ROSy+ retrieval results are compared against state-of-the-art 3D object retrieval techniques.

ROSy itself has similar performance to CPCA and NPCA. However, the combination of the three pose normalization methods (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:

$$si = (2Df_i^{CPCA}, 2Df_i^{NPCA}, 2Df_i^{ROSy},$$
$$3Df_i^{CPCA}, 3Df_i^{NPCA}, 3Df_i^{ROSy}) \qquad (7)$$

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:

$$Distance(s1, s2) = dist_{2Df} + dist_{3Df} \qquad (8)$$

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

$$dist_{2Df} = min_j(L_1(2Df_1^j, 2Df_2^j))$$
$$dist_{3Df} = min_j(L_1(3Df_1^j, 3Df_2^j)) \qquad (9)$$

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 10 it is further illustrated that the 3D object retrieval system using ROSy+ outperforms two recent pose normalization methods: DLA (Chaouch and Verrout-Blondet 2009) and GSMD+SHD+R (Lian et al 2009), the PANORAMA descriptor (Papadakis et al 2009) and also three classic 3D object retrieval methods: Lightfield (Chen et al 2003), SH-GEDT (Kazhdan et al 2003) and DESIRE (Vranic 2005) 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.

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 3. 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 Verrout-Blondet, using the Depth Line Approach descriptor (Chaouch and Verroust-Blondet 2009), by Lian et al. using the combined GSMD - SHD descriptors with Rectilin-
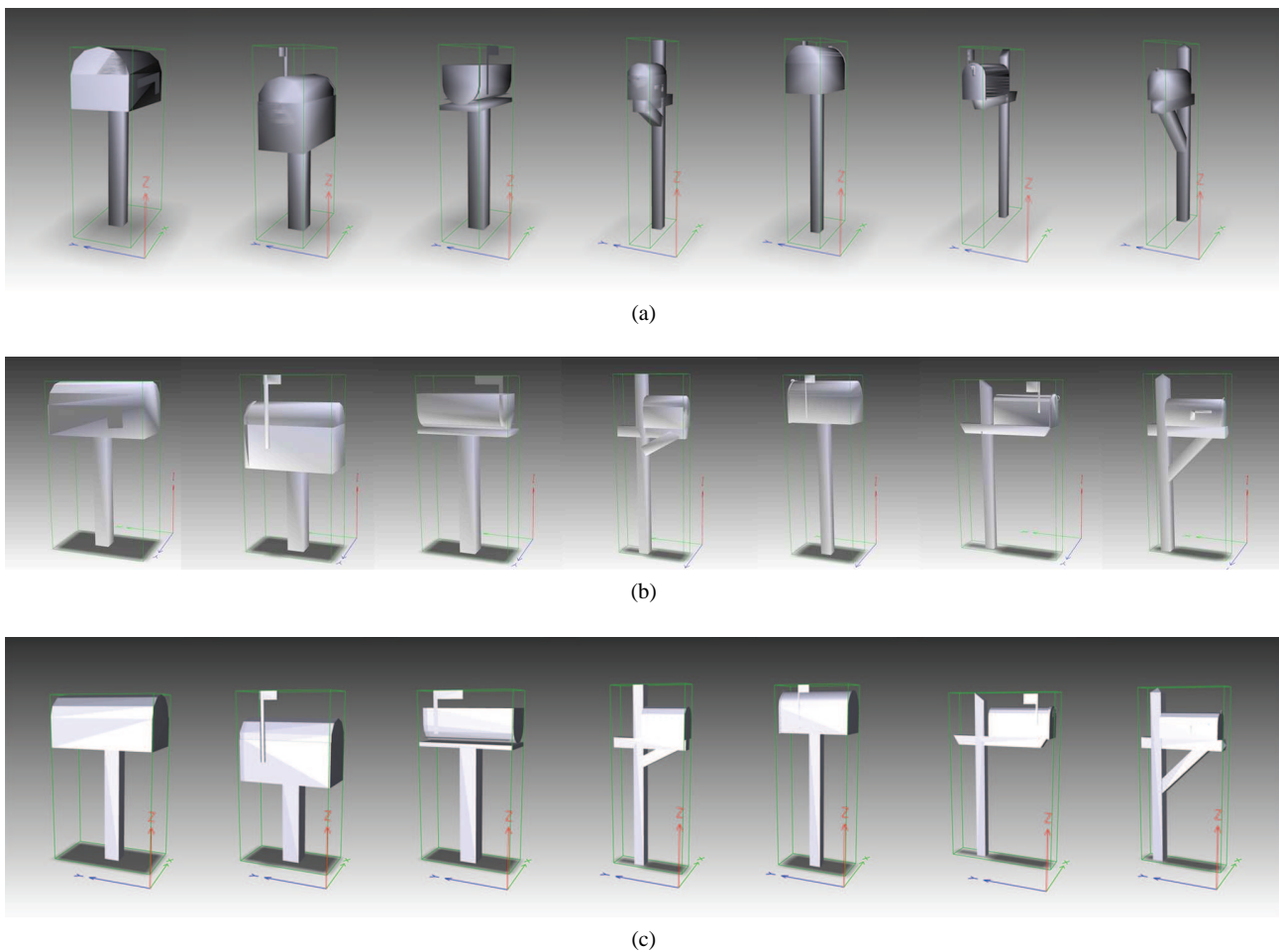
Fig. 11: Alignments of the 'MAILBOX' class: (a) ROSy pose normalization results (b) a consistently rotated version of ROSy pose normalization results by 90 degrees around the Z axis (c) results of the method proposed in (Chaouch and Verroust-Blondet 2009)

earity (Lian et al 2009) and by Papadakis et al. using the PANORAMA descriptor (Papadakis et al 2009) 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. 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 (Papadakis et al 2008) has achieved state-of-the-art performance.

A visual qualitative evaluation is next provided. In Figure 11, comparative alignments between ROSy and the proposed method by Chaouch and Verroust - Blondet (Chaouch and Verroust-Blondet 2009), 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 12, 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 12, 3D objects (a) - (d) show perfect global symmetry against one principal plane. 3D Objects (e) - (h) 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 (i) and (j) show global symmetry against two principal planes simultaneously, while 3D objects (k) - (n) have local symmetries in their structures. 3D Objects (o) and (p) 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

Table 3: 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 9 and 10 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

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 13. Also, Figure 1 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 9). 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)   (d)
(e)   (f)   (g)   (h)
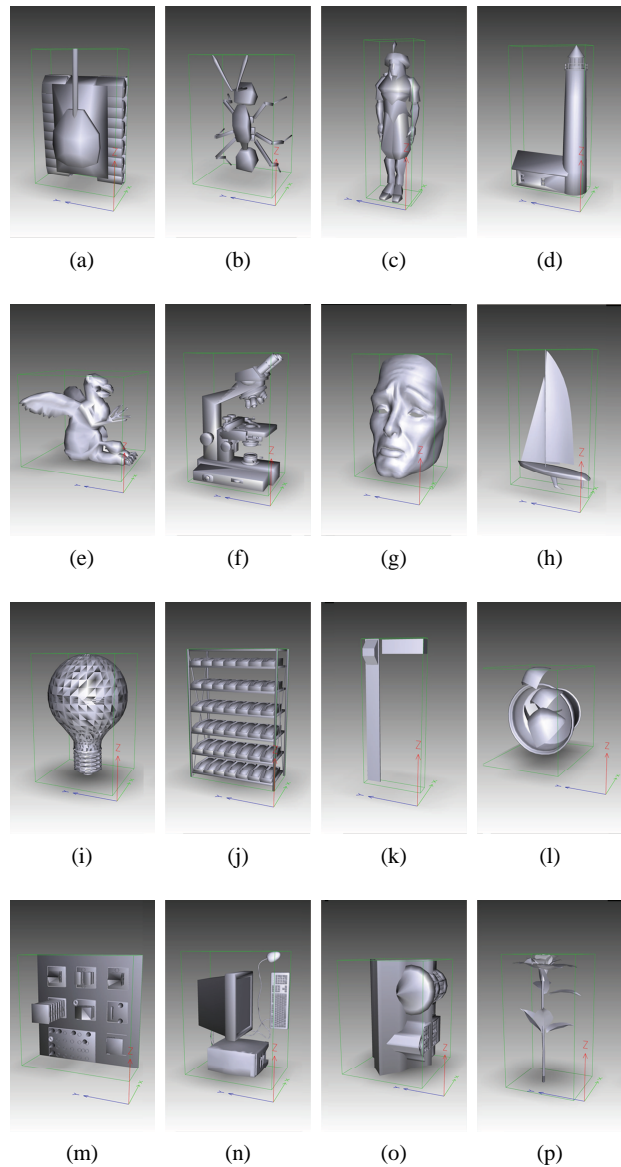(i)   (j)   (k)   (l)
(m)   (n)   (o)   (p)

Fig. 12: 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 (a) - (j), local symmetries (k) - (n) or no symmetries, at all (o) - (p).

The proposed method was tested on a Core2Quad 2.5 GHz system, with 6 GB of RAM, running Matlab R2009a. The system'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 5,000 vertex object, the time required for the pose normalization process is about 0.8 seconds.
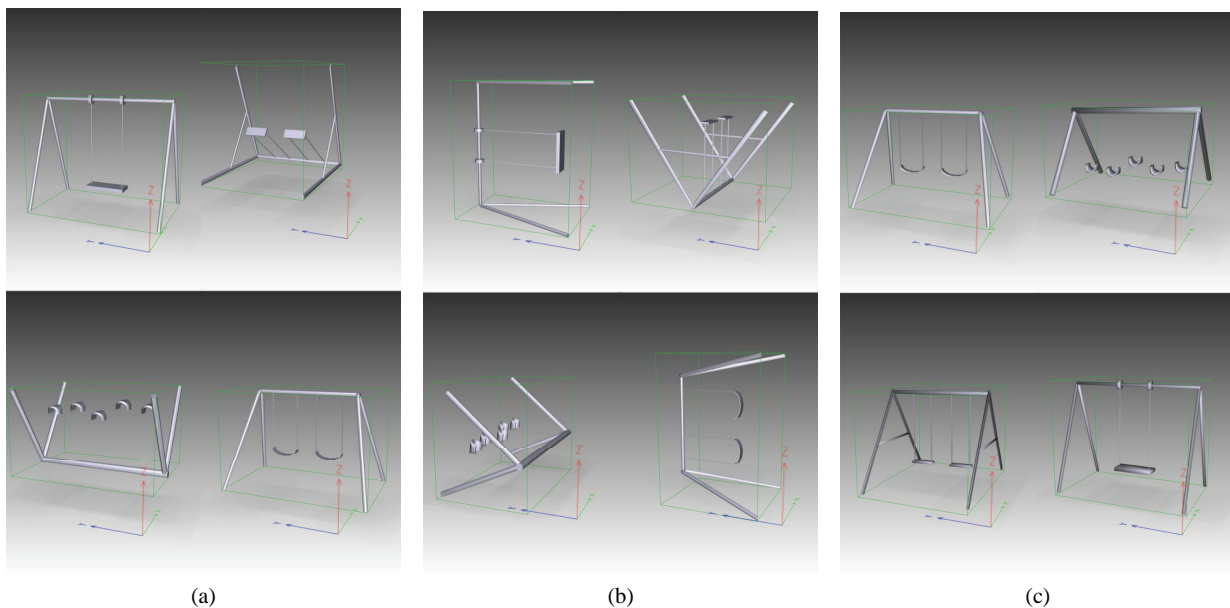
Fig. 13: Alignments of the PSB class 'SWINGSET'. (a) CPCA, (b) NPCA, (c) ROSy.

6 Conclusion

In this paper a novel method for 3D object pose normalization, based on the reflective symmetry properties of 3D objects, is presented. 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 about 3%, over the previous best approach. Furthermore, the proposed method is able to produce high quality alignments of 3D objects, regardless of their originating class or morphology. These alignments are both stable and consistent.

References

Ahn HK, Cheong O, Park CD, Shin CS, Vigneron A (2005) Maximizing the overlap of two planar convex sets under rigid motions. In: SCG '05: Proceedings of the twenty-first annual symposium on Computational geometry, ACM, New York, NY, USA, pp 356–363

Ahn HK, Brass P, Shin CS (2008) Maximum overlap and minimum convex hull of two convex polyhedra under translations. Comput Geom Theory Appl 40(2):171–177

Ankerst M, Kastenmüller G, Kriegel HP, Seidl T (1999) 3d shape histograms for similarity search and classification in spatial databases. In: SSD '99: Proceedings of the 6th International Symposium on Advances in Spatial Databases, Springer-Verlag, London, UK, pp 207–226

van den Bergen G (1997) Efficient collision detection of complex deformable models using aabb trees. J Graph Tools 2(4):1–13

Bustos B, Keim D, Saupe D, Schreck T, Vranic D (2004) An experimental comparison of feature-based 3d retrieval methods. 3D Data Processing Visualization and Transmission, International Symposium on 0:215–222

Chan CK, Tan ST (2001) Determination of the minimum bounding box of an arbitrary solid: an iterative approach. Computers & Structures 79(15):1433 – 1449, URL http://www.sciencedirect.com/science/article/B6V28-43CT96W-5/2/52fce6656f09dd9535a6ef8861d9d9f2

Chaouch M, Verroust-blondet A (2007) 3d model retrieval based on depth line descriptor. In: in IEEE International Conference on Multimedia & Expo (ICME07

Chaouch M, Verroust-Blondet A (2009) Alignment of 3d models. Graphical Models 71(2):63–76, URL http://dx.doi.org/10.1016/j.gmod.2008.12.006

Chen DY, Ouhyoung M (2002) A 3d model alignment and retrieval system. In: Proceedings of International Computer Symposium, Workshop on Multimedia Technologies, pp 1436–1443

Chen DY, Shen YT, Tian XP, Ouhyoung M (2003) On visual similarity based 3d model retrieval. In: Eurographics computer graphics forum, pp 223–232

Cornea ND, Demirci MF, Silver D, Shokoufandeh A, Dickinson SJ, Kantor PB (2005) 3d object retrieval using many-to-many matching of curve skeletons. Shape Modeling and Applications, International Conference on 0:368–373, DOI http://doi.ieeecomputersociety.org/10.1109/SMI.2005.1

Daras P, Zarpalas D, Tzovaras D, Strintzis MG (2006) Efficient 3-d model search and retrieval using generalized 3-d radon transforms. IEEE Transactions on Multimedia 8(1):101–114

Elad M, Tal A, Ar S (2001) Content based retrieval of vrml objects - an iterative and interactive approach. In: Proceedings of the 6th Eurographics workshop on multimedia, pp 97–108

Fang R, Godil A, Li X, Wagan A (2008) A new shape benchmark for 3d object retrieval. In: ISVC '08: Proceedings of the 4th International Symposium on Advances in Visual Computing, Springer-Verlag, Berlin, Heidelberg, pp 381–392

Goldsmith J, Salmon J (1987) Automatic creation of object hierarchies for ray tracing. IEEE Comput Graph Appl 7(5):14–20

Goldstein H, Poole CP (2001) Classical Mechanics. Addison Wesley, URL http://www.worldcat.org/isbn/0321188977

Gottschalk S, Lin MC, Manocha D (1996) Obbtree: a hierarchical structure for rapid interference detection. In: SIGGRAPH '96: Proceedings of the 23rd annual conference on Computer graphics and interactive techniques, ACM, New York, NY, USA, pp 171–180, DOI http://doi.acm.org/10.1145/237170.237244

Jarvelin K, Kekalainen J (2002) Cumulated gain-based evaluation of ir techniques. ACM Transactions on Information Systems 20:422–446

Jayanti S, Kalyanaraman Y, Iyer N, Ramani K (2006) Developing an engineering shape benchmark for cad models. Computer-Aided Design 38(9):939 – 953, URL http://www.sciencedirect.com/science/article/B6TYR-4KV4MRW-2/2/f871141051c71e73b0f373c574b7783b, shape Similarity Detection and Search for CAD/CAE Applications

Kazhdan M (2007) An approximate and efficient method for optimal rotation alignment of 3d models. IEEE Transactions on Pattern Analysis and Machine Intelligence 29(7):1221–1229

Kazhdan M, Chazelle B, Dobkin D, Finkelstein A, Funkhouser T (2002) A reflective symmetry descriptor. In: European Conference on Computer Vision (ECCV), pp 642–656

Kazhdan M, Funkhouser T, Rusinkiewicz S (2003) Rotation invariant spherical harmonic representation of 3d shape descriptors. In: SGP '03: Proceedings of the 2003 Eurographics/ACM SIGGRAPH symposium on Geometry processing, Eurographics Association, Aire-la-Ville, Switzerland, Switzerland, pp 156–164

Kazhdan M, Funkhouser T, Rusinkiewicz S (2004) Symmetry descriptors and 3d shape matching. In: SGP '04: Proceedings of the 2004 Eurographics/ACM SIGGRAPH symposium on Geometry processing, ACM, New York,

NY, USA, pp 115–123

Kim DH, Park IK, Yun ID, Lee SU (2004) A new mpeg-7 standard: Perceptual 3-d shape descriptor. In: PCM (2), pp 238–245

Lian Z, Rosin P, Sun X (2009) Rectilinearity of 3d meshes. International Journal of Computer Vision DOI 10.1007/s11263-009-0295-0

Martinet A, Soler C, Holzschuch N, Sillion FX (2006) Accurate detection of symmetries in 3d shapes. ACM Trans Graph 25(2):439–464

Minovic P, Ishikawa S, Kato K (1993) Symmetry identification of a 3-d object represented by octree. IEEE Trans Pattern Anal Mach Intell 15(5):507–514

Mitchell E (1965) Quaternion parameters in the simulation of a spinning rigid body', simulation. Body, Simulation (18):390–396

Mitra NJ, Guibas LJ, Pauly M (2006) Partial and approximate symmetry detection for 3d geometry. In: SIGGRAPH '06: ACM SIGGRAPH 2006 Papers, ACM, New York, NY, USA, pp 560–568

Novotni M, Klein R (2004) Shape retrieval using 3d zernike descriptors. Computer Aided Design 36:1047–1062

Osada R, Funkhouser T, Chazelle B, Dobkin D (2002) Shape distributions. ACM Transactions on Graphics 21(4):807–832

Papadakis P, Pratikakis I, Perantonis S, Theoharis T (2007) Efficient 3d shape matching and retrieval using a concrete radialized spherical projection representation. Pattern Recognition 40(9):2437–2452

Papadakis P, Pratikakis I, Perantonis S, Theoharis T (2008) 3d object retrieval using an efficient and compact hybrid shape descriptor. In: Eurographics Workshop on 3D object retrieval, pp 9–16

Papadakis P, Pratikakis I, Theoharis T, Perantonis S (2009) Panorama: A 3d shape descriptor based on panoramic views for unsupervised 3d object retrieval. International Journal of Computer Vision DOI 10.1007/s11263-009-0281-6

Paquet E (2000) Description of shape information for 2-d and 3-d objects. Signal Processing: Image Communication 16(1-2):103–122

Podolak J, Shilane P, Golovinskiy A, Rusinkiewicz S, Funkhouser T (2006) A planar-reflective symmetry transform for 3d shapes. In: SIGGRAPH '06: ACM SIGGRAPH 2006 Papers, ACM, New York, NY, USA, pp 549–559

Ricard J, Coeurjolly D, Baskurt A (2005) Generalizations of angular radial transform for 2d and 3d shape retrieval. Pattern Recogn Lett 26(14):2174–2186, DOI http://dx.doi.org/10.1016/j.patrec.2005.03.030

Rustamov RM (2007) Augmented symmetry transforms. In: SMI '07: Proceedings of the IEEE International Conference on Shape Modeling and Applications 2007, IEEE

Computer Society, Washington, DC, USA, pp 13–20

Shilane P, Min P, Kazhdan M, Funkhouser T (2004) The princeton shape benchmark. Shape Modeling and Applications, International Conference on pp 167–178

Sun C, Sherrah J (1997) 3d symmetry detection using the extended gaussian image. IEEE Transactions on Pattern Analysis and Machine Intelligence 19(2):164–168

Sundar H, Silver D, Gagvani N, Dickinson S (2003) Skeleton based shape matching and retrieval. In: SMI '03: Proceedings of the Shape Modeling International 2003, IEEE Computer Society, Washington, DC, USA, p 130

Tangelder JW, Veltkamp RC (2008) A survey of content based 3d shape retrieval methods. Multimedia Tools and Applications 39(3):441–471

Theodoridis S, Koutroumbas K (2006) Pattern Recognition, Third Edition. Academic Press

Theoharis T, Papaioannou G, Platis N, Patrikalakis NM (2008) Graphics and Visualization: Principles & Algorithms. A. K. Peters, Ltd., Natick, MA, USA

Vranić DV (2004) 3D model retrieval. Phd dissertation, University of Leipzig

Vranić DV (2005) Content-based classification of 3d-models by capturing spatial characteristics, URL http://merkur01.inf.uni-konstanz.de/CCCC. Accessed on 28 July 2007

Vranic DV (2005) Desire: a composite 3d-shape descriptor. In: ICME, pp 962–965

Vranić DV, Saupe D, Richter J (2001) Tools for 3d-object retrieval: Karhunen-loeve transform and spherical harmonics

Xiang P, Hua CQ, Gang FX, Chuan ZB (2007) Pose insensitive 3d retrieval by poisson shape histogram. In: ICCS '07: Proceedings of the 7th international conference on Computational Science, Part II, Springer-Verlag, Berlin, Heidelberg, pp 25–32

Yu Z, Zhang S, Wong HS, Zhang J (2007) A filter-refinement scheme for 3d model retrieval based on sorted extended gaussian image histogram. In: MLDM '07: Proceedings of the 5th international conference on Machine Learning and Data Mining in Pattern Recognition, Springer-Verlag, Berlin, Heidelberg, pp 643–652

Zaharia T, Prêteux F (2004) 3d versus 2d/3d shape descriptors: A comparative study. In: SPIE Conference on Image Processing: Algorithms and Systems, vol 2004

Zaharia T, Preteux FJ (2001) 3d-shape-based retrieval within the mpeg-7 framework. SPIE, vol 4304, pp 133–145, DOI 10.1117/12.424969, URL http://link.aip.org/link/?PSI/4304/133/1

Zarpalas D, Daras P, Axenopoulos A, Tzovaras D, Strintzis MG (2007) 3d model search and retrieval using the spherical trace transform. EURASIP J Appl Signal Process 2007(1):207–207, DOI http://dx.doi.org/10.1155/2007/23912

Zhang J, Siddiqi K, Macrini D, Shokoufandeh A, Dickinson S (2005) Retrieving articulated 3-d models using medial surfaces and their graph spectra. In: Energy Minimization Methods in Computer Vision and Pattern Recognition, pp 285–300