The Canonically Posed 3D Objects Dataset

Panagiotis Papadakis

Eurographics Workshop on 3D Object Retrieval (2014)
B. Bustos, H. Tabia, J.P. Vandeborre, and R. Veltkamp (Editors)

The Canonically Posed 3D Objects Dataset

P. Papadakis

Team LAGADIC, INRIA Méditerranée, Sophia-Antipolis, France

Abstract

Shape matching methodologies of generic 3D objects are conventionally preceded by a pose normalization stage, that transforms objects to a canonical coordinate frame wherein feature extraction and shape matching is performed. Arguably, the canonical pose of a 3D object depends not only on its geometry but also on its semantic meaning, a characteristic that generally complicates the extraction of ground truth data.

This paper introduces the first ground-truth dataset of 3D objects that allows an objective evaluation of methods which obtain the canonical pose of objects within extrinsic space. By virtue of the protocol that was followed to assemble the dataset, 3D objects of the same class share a fixed pose in terms of object center, scale and rotation while undergoing diverse shape deformations. The dataset is publicly disclosed and relevant use cases are discussed.

Categories and Subject Descriptors (according to ACM CCS): Collection [H.3.7]: Standards—

Keywords: normalization, canonical pose, ground truth, dataset, 3D models

1. Introduction

3D shape matching and retrieval has witnessed a significant progress in the past years while gradually expanding its scope into novel applications. SHREC [LGA*12] (SHape REtrieval Contest)[†] traces the growing range of applications for a number for years by evaluating the performance of different methodologies while a recent review [KPC13] reporting on the variety of benchmarks that are used as ground truth datasets, is representative of the elevated interest and instructive of future trends. Among them, shape matching and retrieval of generic (inter-class) polygon-soup like 3D objects has shown to remain highly prioritized through the years, followed by the increase in the number of repositories.

A decomposition of the problem of inter-class 3D object retrieval prescribes that a retrieval methodology is generally composed of three main stages, namely, *pose normalization*, *feature extraction* and *matching*. 3D pose normalization [VSR01, Vra04] regards the computation of a so-called *canonical* 3D coordinate frame that is parameterized by its center, scale and rotation that are characteristic of the object's category and the consecutive transformation of the object into that coordinate frame. Analogously, if this transformation matrix is equal to the identity matrix then we say that the object is in its canonical pose. Despite the fact that feature extraction and matching may bear invariance properties to the respective similarity transformations, experience has shown that this is often accompanied by a non-trivial loss of discriminative information in the resulting shape signature and in turn in retrieval performance.

While the determining role of pose normalization in generic shape retrieval methodologies is generally acknowledged, the currently available datasets are inadequate to further serve as ground-truth for giving the canonical pose of 3D objects. The main issue that has prevented the construction of such a ground-truth is that the canonical pose of a 3D object may vary depending on the shape interpretation, i.e. the semantics of its class. At the same time, a manual configuration of an object to its canonical coordinate frame is a tedious task that requires setting multiple degrees of freedom of the coordinate frame, namely, the position, orientation and scale, which is prone to discretization errors.

This paper discloses the first dataset serving as ground truth for the canonical pose of 3D objects and provides a protocol that expedites the construction of similar future datasets. It detail, it introduces the notion of the *semantic* canonical pose and distinguishes it from the geometrically

† http://www.aimatshape.net/event/SHREC/

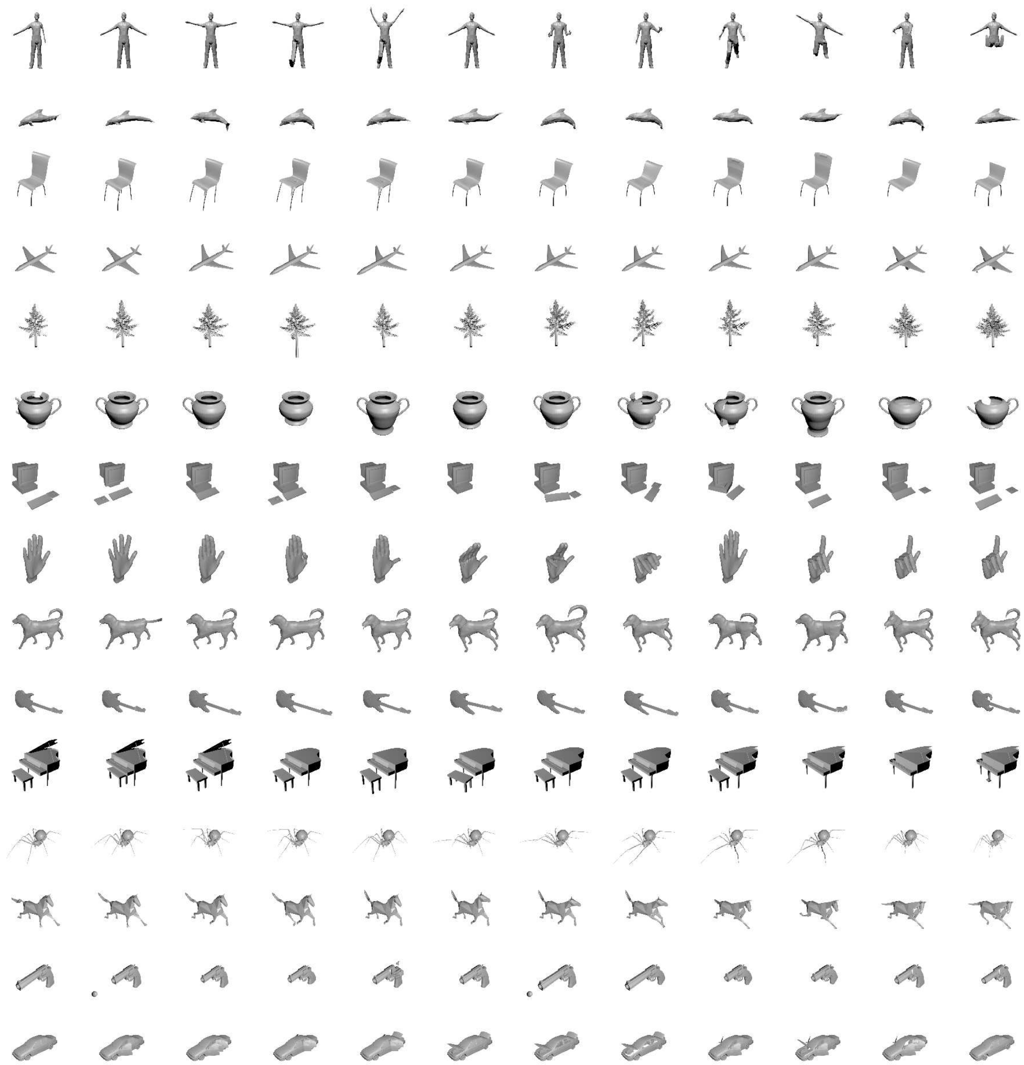

Figure 1: *Contents of the CAPOD dataset.*

inspired approach which has governed the development of earlier approaches, by prioritizing the class where the object belongs to rather than its geometric characteristics. Methodologically, the construction of the dataset is inspired by findings recently reported in [Pap14] and provides an augmented corpus of 3D objects and object categories together with instructive details on new use cases. By constituting the first publicly available dataset of its kind, it can give new insights in the fields of shape cognition, matching and retrieval.

2. Dataset

In order to construct a ground-truth dataset of canonically posed 3D objects, the semantics subjectivity and error proneness issues were alleviated by following a reversed design perspective. In particular, instead of the conventional approach where 3D objects are first collected, then classified and finally pose normalized, the reverse direction was followed. The corresponding steps are listed as follows:

1. Selection of categories of 3D objects that span a desired semantics range which depends on the desired categorization resolution.
2. Acquisition of a reference 3D object for each category which exhibits the semantics of the particular class.
3. Production of 3D object instances for each distinct class by applying permissible modifications of the corresponding reference object.

Implementing the third and final step further implies that modifications can be of any type, excluding those altering the predefined object class semantics or the semantically determined canonical pose. Essentially, this means that similarity transformations are not applicable while on the contrary, non-rigid transformations such as limbs articulation and removal or addition of minor object parts are plausible transformations. In this context, while the purely geometry-based canonical pose of a produced object may be altered as a result of the redistribution of its mass, its semantically-based canonical pose remains unaltered. Therefore, all the 3D object instances of a class share the same, objectively maintained canonical pose.

By employing the aforementioned methodology the CAnonically Posed 3D Objects Dataset (CAPOD) was built, incorporating in total 15 generic object categories and 180 objects uniformly distributed across categories, giving a fixed 12 object members per class. The reference 3D objects were collected from publicly available internet repositories. The chosen object categories are in order: (i) Human, (ii) dolphin, (iii) chair, (iv) airplane, (v) tree, (vi) vase, (vii) desktop computer, (viii) hand, (ix) dog, (x) electric guitar, (xi) piano, (xii) spider, (xiii) horse, (xiv) handgun and (xv) car. The contents of the dataset are depicted in Figure 1.

The initial translation, scale and rotation for each reference 3D object was arbitrary and in turn of no practical interest. Since CAPOD is designed in order to serve as a ground-truth dataset that provides the canonical pose strictly for 3D objects that belong to the same class, any relation (geometric or semantic) among the reference 3D objects and in turn among object categories is beyond scope.

The modifications that were applied to the reference objects for the derivation of their class members, were chosen in order to reflect scenarios common to the respective class. Characteristically, subpart articulations were applied to the 3D objects that contained moving parts, extrusions or suppression of parts were mainly applied to deformable objects while the addition or deletion of sub-parts was performed within the limits of the respective class semantics. Figure 2 provides some representative examples that demonstrate the nature of modifications that have been considered in order to produce the class members of objects' categories.

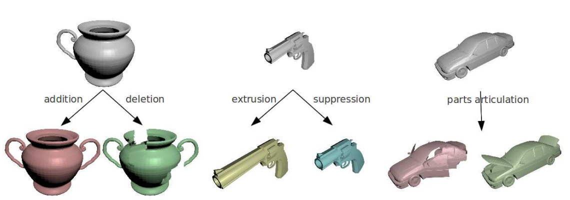

Figure 2: *Repertoire of shape deformations used to produce object class members.*

No constraints have being imposed on the properties of 3D meshes themselves such as watertight restrictions or surfaces of certain genus, therefore, objects are generally viewed as polygon-soup like representations. In pursuing compatibility with 3D mesh processing software and other ground-truth datasets, 3D objects are given in a standard 3D polygonal representation format, namely, the Wavefront Object file format (.obj). The dataset is made available to the community by the web-link `sites.google.com/site/pgpapadakis/home/CAPOD` accompanied with a Creative Commons 3.0 Attribution licence[‡].

3. Use cases

Using a smaller corpus of data, the experiments presented in [PP11, Pap14] reported on the comparative performance of translation and scale normalization approaches. This was accomplished by measuring the standard deviation σ in the computation of the canonical object centroid and scale, for each class individually and collectively by averaging among classes. The idea behind using the σ statistic relies on the hypothesis that the performance in estimating the canonical pose is proportional to certainty, or stability of computation. Furthermore, it generalizes earlier approaches in the domain of 2D shape normalization [CAdlT04, JT08] that employed the euclidean distance between the estimated canonical center in pairs of objects as a measure of performance.

In a similar context, CAPOD offers the possibility for the evaluation of rotation normalization techniques that have so far only been evaluated through the performance of content-based 3D object retrieval. Using CAPOD, the performance of rotation normalization techniques can be evaluated independently of a shape description methodology and assist in a better comprehension of the comparative performances. This is made feasible by establishing a distance measure for rotation transformations that give the canonical rotation and quantifying the difference within each class and collectively by averaging among all classes. Inspired by earlier works in the 2D [CAdlT04,JT08] and 3D domain [WLL*12] that have employed the difference between the angles of the computed principal axes among pairs of objects, it appears straightforward to employ the standard deviation statistic for evaluations that involve multiple objects.

By isolating the dependence of 3D shape descriptors from the preceding pose normalization stage, CAPOD also allows a focused evaluation of the comparative discriminative power of features, within a shape retrieval evaluation framework. Shape description methodologies that use features which are not invariant to similarity transformations can now be compared on an equal basis by extracting them from the already canonically posed 3D objects. Exclusively towards making meaningful such an evaluation and allow

‡ http://creativecommons.org/licenses/by/3.0/

inter-class shape matching, all reference 3D objects were primarily transformed in order to fit within the unit cube centred at the coordinates origin.

To conclude, it should be noted however that while the contribution of the dataset is to assist in the evaluation of pose normalization techniques and additionally of shape description methodologies, a joint evaluation of both applications at the same time requires a more careful consideration. This is due to the fact that improving the performance of pose normalization may not under all conditions assist in improving the performance of shape retrieval. Experimental evidence of this behaviour were first reported in the work of Sfikas et al. [STP11], suggesting that when objects of the same class are more consistently normalized for rotation, this occasionally results in reducing the differences not only among objects that belong to the same class but further between objects belonging to different classes which gives rise to undesired matches.

4. Conclusions

The Canonically Posed 3D Objects Dataset was presented, constituting the first ground-truth dataset of its kind for the evaluation of 3D pose normalization techniques. Following its construction methodology, it reduces the inherent ambiguity and subjectivity in determining the canonical pose of generic 3D objects and can serve as a basis for the construction of larger ground-truth datasets. In turn, it allows revisiting the problem of content-based retrieval by allowing the isolation of the discriminative power of 3D shape signatures from the pose normalization stage.

This work could further open the discussion for determining the semantics within the pose of objects and their significance in shape matching. For example, while shape matching is traditionally treated as a process that should be invariant to scale, in practice the scale of objects plays an important role in object discrimination as it conveys important information on the object' s semantic properties. Normalizing the rotation may occasionally impede retrieval performance as it can suppress inter-class object differences to a greater degree compared to intra-class differences. Clearly, there is a need for a better understanding of the semantics that are expressed through an object' s inherent pose and to explore the potential in using these semantics for discrimination purposes in future endeavours, possibly in combination with other non-purely geometric information.

References

[CAdlT04] CORTADELLAS J., AMAT J., DE LA TORRE F.: Robust normalization of silhouettes for recognition applications. *Pattern Recognition Letters 25*, 5 (2004), 591–601. 3

[JT08] JIANG T., TOMASI C.: Robust shape normalization based on implicit representations. In *International Conference on Pattern Recognition* (2008), pp. 1–4. 3

[KPC13] KOUTSOUDIS A., PRATIKAKIS I., CHAMZAS C.: On 3d object retrieval benchmarking. *3D Research 4*, 4 (2013), 1–12. 1

[LGA*12] LI B., GODIL A., AONO M., BAI X., FURUYA T., LI L., LÃŞPEZ-SASTRE R., JOHAN H., OHBUCHI R., REDONDO-CABRERA C., TATSUMA A., YANAGIMACHI T., ZHANG S.: SHREC'12 Track: Generic 3D Shape Retrieval. In *Eurographics Workshop on 3D Object Retrieval* (2012), pp. 119–126. 1

[Pap14] PAPADAKIS P.: Enhanced pose normalization and matching of non-rigid objects based on support vector machine modelling. *Pattern Recognition 47*, 1 (2014), 216 – 227. 2, 3

[PP11] PAPADAKIS P., PIRRI F.: Consistent pose normalization of non-rigid shapes using one-class support vector machines. In *IEEE International Conference of Computer Vision and Pattern Recognition, NORDIA Workshop* (2011), pp. 23–30. 3

[STP11] SFIKAS K., THEOHARIS T., PRATIKAKIS I.: Rosy+: 3d object pose normalization based on pca and reflective object symmetry with application in 3d object retrieval. *International Journal of Computer Vision 91*, 3 (2011), 262–279. 4

[Vra04] VRANIC D. V.: *3D Model Retrieval*. PhD thesis, University of Leipzig, 2004. 1

[VSR01] VRANIC D., SAUPE D., RICHTER J.: Tools for 3d-object retrieval: Karhunen-loeve transform and spherical harmonics. In *IEEE Fourth Workshop on Multimedia Signal Processing* (2001), pp. 293–298. 1

[WLL*12] WANG C., LIU Y.-S., LIU M., YONG J.-H., PAUL J.-C.: Robust shape normalization of 3d articulated volumetric models. *Computer-Aided Design 44*, 12 (2012), 1253–1268. 3