

# STATISTICAL ANALYSIS OF THE GLOBAL GEODESIC FUNCTION FOR 3D OBJECT CLASSIFICATION

*Djamila Aouada, Shuo Feng, Hamid Krim*

North Carolina State University  
Electrical and Computer Engineering Department  
Raleigh, NC 27695, USA.  
{daouada, sfeng2, ahk}@ncsu.edu

## ABSTRACT

This paper presents a novel classification strategy for 3D objects. Our technique is based on using a *Global Geodesic Function* to intrinsically describe the surface of an object. The choice of the *Global Geodesic Function* ensures the invariance of the classification procedure to scaling and all isometric transformations. Using the *Jensen-Shannon Divergence*, feature parameters are extracted from the probability distribution functions of the *Global Geodesic Function* for each one of the classes. These parameters are used in the decision of a class membership of an object. This approach demonstrates low computational cost, efficiency, and robustness to resolution over many different data sets.

**Index Terms**— Object classification, Geodesic, Jensen-Shannon Divergence, Feature extraction.

## 1. INTRODUCTION

Research interest in 3D object analysis has witnessed an explosive growth over the last few years. While this may in part be explained by an equally impressive growth in computing power, the availability of 3D data acquisition systems and ready-access to it at relatively low cost, have been a key in addressing the numerous problems which arise in applications. Laser scanners, ranging cameras and others have indeed made solutions to multimedia applications, biometrics, and computer graphics more realistic and affordable. The wide distribution of 3D data over the internet at no-cost is also testimony to the high level of interest in the area [1], [2], [3]. Two “face” and “vertex” matrices are the usual digital representation of a 2D surface, and are often of prohibitive size. This in turn, unveils a computational challenge often encountered in practice and particularly in 3D shape classification and recognition applications. A number of approaches addressing this issue have recently appeared in the literature. Shape distribution methods based on the *probability density functions (pdf)* of features carrying intrinsic information about 2D surfaces

(area, curvature, ...) are techniques that proved their simplicity and effectiveness for objects' classification [4]. However, the performances of these shape descriptors are directly related to the choice of the shape function and its properties. In [5], a shape function based on the geodesic distance between points on the surface was proposed. The invariance properties of this function to object pose meet the requirements stated earlier for a unique surface characterization.

In this work we propose to use an approximated, yet an as efficient *Global Geodesic Function (GGF)* to describe 3D objects. The statistical variability of this new shape function is analyzed using an information theoretic measure known as the *Jensen-Shannon Divergence (JSD)* [6]. To efficiently compare objects, we characterize each class of objects by two parameters in the learning stage. This drastically simplifies the search in a testing evaluation. In addition, we introduce a more discriminative comparison of objects by defining an interval based *JSD*. The remainder of the paper is organized as follows. The next section gives a formulation of the problem along with a presentation of the key tools that are used namely the *GGF*, the *JSD* and the *zoom-in* operation. The detailed algorithm is explained in Section 3. In Section 4, experimental results demonstrate the robustness and efficiency of the method. Finally, Section 5 summarizes the paper and discusses future work to address outstanding issues.

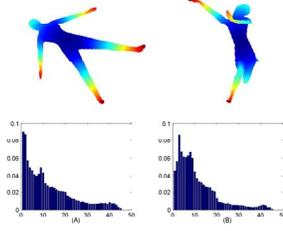
## 2. PROBLEM STATEMENT AND BACKGROUND

Given a set of  $N$  classes of objects  $\{C_1, C_2, \dots, C_N\}$ , our goal is to decide on a class membership of an object  $O$ . To efficiently carry out such a task, we describe an object with an appropriate shape function, *i.e.*, the *GGF*. Using the *JSD* we compare the statistical properties of two different *GGFs*.

### 2.1. Global Geodesic Function

A 3D object may be viewed as a 2D surface  $S$  embedded in  $\mathbb{R}^3$ . Our goal is to make our characterization of the surface invariant to non-elastic deformations, also referred to as iso-

Thanks to AFOSR F49620-98-1-0190 grant.



**Fig. 1.** Invariance to isometric transformations: (A) Original pose (B) Same subject after deformation and noise addition. At the bottom the *pdf* of *GGFs*.

metric transformations. Thus, we should uniquely represent two isometric surfaces  $S$  and  $\hat{S}$ .

**Definition:** (Isometric surfaces) We say that two surfaces  $S$  and  $\hat{S}$  are isometric if there is a bijective function  $f$  between  $S$  and  $\hat{S}$  such that:

$$\forall p, q \in S, \quad d(p, q) = d(f(p), f(q)), \quad (1)$$

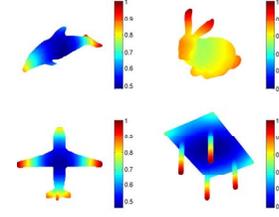
where  $d(p, q)$  is the shortest path between the two points  $p$  and  $q$  along the surface of interest, also called *geodesic distance*.

This definition includes transformations such as translations, rotations and all non-elastic deformations. This implies a natural choice of a geodesic distance to describe a given object. Fig.1 illustrates a practical example where isometries are encountered and similarly represented by a geodesic function. The independence of the selected shape function of any reference point is also an important property, which thus defines, as first introduced in [7], an intrinsic global function  $g(p) = \int_{q \in S} d(p, q) dS$  at each point  $p$  on the surface  $S$ . Assuming that the mesh representing a surface is uniform and sufficiently fine, we consider the approximated and discrete form of  $g(\cdot)$  as  $g(p_i) = \sum_{j=1}^m d(p_i, p_j) \delta S_j$ , where  $p_i$ ,  $i = 1, \dots, m$ , are all the vertices constituting the mesh, and  $\delta S_i$  is the area that a vertex  $p_i$  occupies.

Based on our assumption on the quality of the mesh,  $\delta S_i$  may be considered small, constant and equal to  $\delta S$ . We may hence achieve the property of scale invariance by normalizing the function  $g(\cdot)$  by its maximum value over all vertices.

$$\begin{aligned} g_n(p_i) &= \frac{g(p_i)}{\max_{j=1, \dots, m} g(p_j)} \\ &= \frac{\sum_{j=1}^m d(p_i, p_j) \delta S}{\max_{j=1, \dots, m} \sum_{k=1}^m d(p_j, p_k) \delta S} \\ &= \frac{\sum_{j=1}^m d(p_i, p_j)}{\max_{j=1, \dots, m} \sum_{k=1}^m d(p_j, p_k)}, \end{aligned} \quad (2)$$

where the subscript  $n$  denotes normalization. This operation ensures the convergence of  $g_n(\cdot)$ . The function  $g_n(\cdot)$  is what we refer to as the *GGF*. It ranges over  $]0, 1]$ , with the zero value being, in theory, unattainable. Some examples are shown in Fig.2. In practice, the *GGF* for each ver-



**Fig. 2.** Examples of the global geodesic function on different 3D objects.(Best visualized in color)

tex is obtained by computing the latter's geodesic distance to all other vertices. This normalization also obviates explicit computation of a surface element. This is efficiently realized with the well known *Dijkstra* algorithm whose complexity is  $\mathcal{O}(N^2 \log N)$  [8].

## 2.2. Statistical Analysis

### 2.2.1. Jensen-Shannon Divergence

The comparison of the statistical properties of two distinct distributions of the *GGF* may be carried out by the *JSD*. The *JSD*, indeed, enables us to quantify the difference between an arbitrary number of *pdfs* and is defined for two distributions  $P_1$  and  $P_2$  as follows:

$$JSD(P_1, P_2) = H \left( \sum_{l=1,2} \frac{1}{2} P_l \right) - \sum_{l=1,2} \frac{1}{2} H(P_l), \quad (3)$$

where  $H$  is the *Shannon entropy* defined by  $H(P_l) = -\sum_{i=1}^L P_l(i) \log P_l(i)$ , with  $l = 1, 2$ , and  $P_l(i)$ ,  $i = 1, \dots, L$ , are the elements of the discrete *pdf* vector  $P_l$ . It has been proven in [?] that in the framework of information theory the *JSD* may be interpreted as the mutual information between shape descriptors. We use this information to determine a characteristic resolution  $\mathcal{R}^1$  for each 3D object.

We start by considering  $K$  different resolutions,  $R_0 > R_1 > \dots > R_{K-1}$ , for the same object  $O$  and the  $K$  corresponding *pdfs*  $P_{R_i}$ ,  $i = 0, 1, \dots, K-1$  of the *GGF*. That is we represent the surface of  $O$  with  $R_i$  points and compute the *GGF* on each one of them to finally get the resulting distribution  $P_{R_i}$ . We define

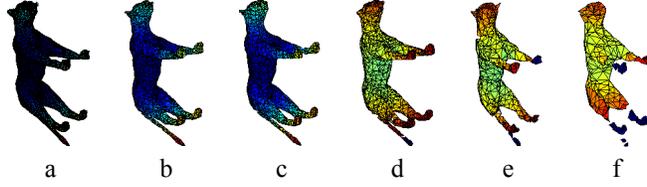
$$\zeta(R_i) = (JSD(P_{R_0}, P_{R_i}) - JSD(P_{R_0}, P_{R_{i+1}}))^2. \quad (4)$$

The characteristic resolution  $\mathcal{R}$  of the object  $O$  is then:

$$\mathcal{R} = \arg \max_{i=1, \dots, K-1} (\zeta(R_i)). \quad (5)$$

The parameter  $\mathcal{R}$  translates the trend observed for all 3D objects, as illustrated, for instance, for the object *tiger* in Fig.

<sup>1</sup>**Resolution:** is the number of vertices used to represent the surface of an object



**Fig. 3.** Illustration of the visual effect of resolution reduction on the *GGF* of a 3D object. Resolution is decreasing from (a) to (f). An abrupt change in the *GGF* occurs at (d). (Best visualized in color)

3.  $\mathcal{R}$  is basically the resolution at which an abrupt and sharp change in the overall distribution of the *GGF* occurs. Any lower resolution will then fail to exhibit the desired invariance of the distribution of the *GGF* for an object  $O$ .

### 2.2.2. Zoom-in Operation

To detect dissimilarities between two objects from the same class  $O_j$  and  $O_k$ , with  $j \neq k$ , we compare the corresponding *pdfs* over smaller regions of their support, which we focus on by a zoom-in operation. To that end, denote by  $L$  the maximal support of the two *pdfs* (histograms)  $P^{(j)}$  and  $P^{(k)}$  of the *GGFs* for  $O_j$  and  $O_k$ , respectively. We proceed by looking at  $N$  fixed intervals  $L_i$ , with  $\cup_{i=1}^N L_i = L$  and  $L_i \cap L_j = \emptyset$  if  $i \neq j$ . We compute over each interval the normalized *pdfs*  $P^{(j)}(L_i)$  and  $P^{(k)}(L_i)$ . The *JSD* between the two objects is now a function of the interval  $L_i$  and is defined as

$$\alpha(L_i) = JSD \left( P^{(j)}(L_i), P^{(k)}(L_i) \right). \quad (6)$$

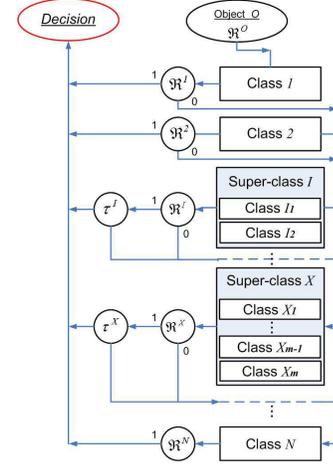
Establishing a critical region over which two *pdfs* are most similar, namely  $L_0 = \arg \min_{i=1, \dots, N} \alpha(L_i)$ , we redefine our interval of interest as  $\lambda = L - L_0$ . The interval  $\lambda$ , hence, includes the largest dispersion between  $O_j$  and  $O_k$ .

## 3. PROPOSED CLASSIFICATION APPROACH

With the tools described in Section 2 in hand, we propose a classification strategy that heavily relies on a training procedure. During the training, a class parametrization is achieved. To further refine discrimination among objects, we introduce a post processing *zoom-in* procedure on *pdfs*' comparison to focus on more detailed dissimilarities among objects from the same class.

### 3.1. Class Characterization

As defined in Eq. (5), we determine a characteristic resolution  $\mathcal{R}_i$  for each training object  $O_i$ ,  $i \in \Omega = \{1, \dots, M\}$ , from a class  $C$ . We obtain a *class characteristic resolution* as being the maximum of all the characteristic resolutions within that class, *i.e.*,  $\mathcal{R}^C = \max_{i=1, \dots, M} \{\mathcal{R}_i\}$ . Upon establishing



**Fig. 4.** Algorithm of the discrete classification decision for 3D objects.  $N$  classes and  $X$  super-classes are represented. They are arranged according to increasing values of their characteristic resolutions  $\mathcal{R}^i$ . Parametric comparison between object  $O$  and classes is represented by circular shapes.

$\mathcal{R}_C$ , any comparison within a class is carried out at  $\mathcal{R}_C$ . We compute a pairwise *JSD* among all training objects in  $C$ . The largest value obtained from the last *JSDs* is the highest deviation that no object within a class  $C$  would never exceed. A threshold  $\tau = \max_{i,j \in \Omega} \{JSD(P^{(i)}, P^{(j)})\}$  is hence defined as a second *class characteristic parameter*.

### 3.2. Algorithm

The distinct steps of the algorithm are sketched in Fig. 4, and summarized next.

1. Define  $N$  object classes  $\{C_1, C_2, \dots, C_N\}$ .
2. In the learning phase, associate to each class  $C_i$  a corresponding  $\mathcal{R}_{C_i}$  and the threshold  $\tau_{(C_i)}$ .
3. Sort all classes in an increasing resolution order  $\mathcal{R}_{C_i}$ .
4. Construct super-classes by merging classes sharing the same parameter  $\mathcal{R}_{C_i}$ .
5. Start from the lowest resolution  $\mathcal{R}_{C_1}$  and set  $l = 1$ .
6. Compute the *GGF* of  $O$  at  $\mathcal{R}_{(C_l)}$  and get its resolution parameters.
7. Compare the resolution parameters of  $O$  with those of  $C_l$ . If similarity is established, *i.e.*, decision is 1;
  - Termination of search if  $C_l$  is a class.
  - Apply  $\tau$ -thresholding if  $C_l$  is a super-class.

Otherwise the decision is 0 and  $l = l + 1$ .

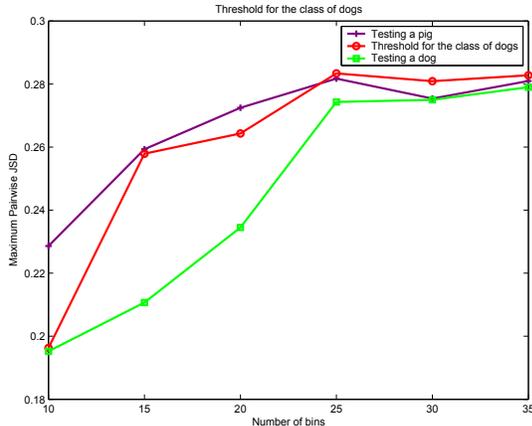


Fig. 5. Illustration of the  $\tau$ -thresholding for super-classes

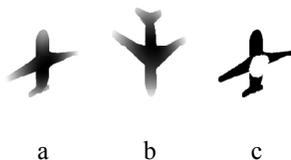


Fig. 6. Detection of the area of dissimilarity between two airplanes using the GGF (White blob in c).

8. Go to 6.
9. Repeat operations until a decision 1 is reached.

#### 4. EXPERIMENTAL RESULTS

An important characteristic of a classification algorithm is its consistent performance independently of what data-base or measurement system the data originates from. So as long as the format is in readable form, we have shown that our proposed approach enjoys such a property. We made extensive use of Princeton’s benchmark [?] for all our learning and testing. We have subsequently tested data from INRIA’s benchmark [?] to validate our claim. We have set the tolerance decision for parameters’ similarity to 10% as a maximum absolute error  $\frac{|\mathcal{R}_{learning} - \mathcal{R}_{testing}|}{\mathcal{R}_{learning}}$ . The corresponding overall recognition performance, *i.e.*, number of correct decisions over all the testing experiments, is 93.88%. This result reflects the combined usage of the parameters  $\mathcal{R}$  and  $\tau$ . Indeed, the  $\tau$  thresholds contributed in improving the final performance as they permitted to separate the classes having close values for the parameter  $\mathcal{R}$ . In Fig. 5, we illustrate an example of confusion among the two super-classes of *pigs* and *dogs* and show how it is solved using the second class parameter  $\tau$  which, in this case, is a vector whose elements are thresholds computed at different sampling rates, *i.e.*, numbers of bins used to compute the histogram of the *GGF*.

The *GGF* is sufficiently powerful to detect and measure dissimilarity for even relatively close 3D objects. In Fig. 6., applying the *zoom-in* at  $\lambda = 1/10$ , we look at the area of highest difference between two airplanes. In (a), a commercial airplane is represented. In (b), a more recent model of commercial airplane is shown. The zoom-in technique highlights in (c) (in white) the difference in  $\lambda$ -surface subsection. This difference restricted to the fuselage, is intuitively pleasing as it confirms our visual interpretation of a fatter first airplane.

#### 5. CONCLUSION

In this paper, we presented a new 3D object classification strategy based on two feature parameters. We have shown that it generally provides a quick discrete decision on object comparison. The technique is based on the statistical properties of the *GGF* which, in turn, provides very interesting advantages such as robustness to noise, complexity of representation, and invariance to isometric transformations and scaling. Moreover, a new and precise dissimilarity evaluation has been introduced. A first estimation of the accuracy of this technique is of 93.88%. Additional work is required for a more thorough evaluation.

#### 6. ACKNOWLEDGEMENT

The authors would like to thank Sajjad Baloch for valuable discussions, and INRIA, Princeton and McGill Universities for making their 3D data sets available for research purposes, and AFOSR F49620-98-1-0190 grant for supporting this work.

#### 7. REFERENCES

- [1] <http://www.cim.mcgill.ca/shape/benchMark/>
- [2] <http://shape.cs.princeton.edu/benchmark/>
- [3] <http://www-pc.inria.fr/Eric.Saltel/download/download.php>
- [4] R. Osada, T. Funkhouser, B. Chazelle, and D. Dobkin: “Shape Distributions”, *ACM Transactions on Graphics*, Vol. 21, NO. 4, pp. 807-832, October 2002.
- [5] A. B.Hamza, and H. Krim: “Geodesic Matching of Triangulated Surfaces”, *IEEE transactions on image processing*, Vol 15, NO. 8, pp 2249-2258, August 2006.
- [6] I. Grosse, P. Benaola-Galvan, P. Carpena, R. Roman-Roldan, J. Oliver, H. Eugene Stanley: “Analysis of symbolic sequences using the Jensen-Sannon divergence”, *Physical Review E*, Vol 65, 041905, pp 1-16, 2002.
- [7] M. Hilaga, Y. Shinagawa, T. Kohmura, and T. L.Kunii: “Topology Matching for Fully Automatic Similarity Estimation of 3D Shapes”, *Proc. SIGGRAPH*, pp. 203-212, August 2001.
- [8] Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, Clifford Stein: “Introduction to Algorithms”, Second Edition, September 2001. MIT Press, Cambridge, Massachusetts.