3D FREE FORM SURFACE MATCHING BASED ON ORIENTATION DIFFERENCE LENGTH DISTRIBUTION

Yonghuai Liu, Guoqiang Fei — Computer Science Dept., University of Wales, Aberystwyth, Ceredigion SY23 3DB, Wales

Baogang Wei — Computer Science School, Zhejiang University, Hangzhou 320027, China

Longzhuang Li — Computer Science Dept., Texas A & M University, Corpus Christi, TX-78412

ABSTRACT

In this paper, we propose using orientation difference length distribution (ODLD) to represent a 3D view of each object with free form surface. ODLD has a number of advantages over the shape distribution. A comparative study of ODLD and the shape distribution has shown that the former is significantly more accurate for 3D free form surface matching than the latter.

1. INTRODUCTION

Free form surface matching is a fundamental problem in the machine vision and image processing community. Recently, since a large number of 3D images have been put onto the World Wide Web with different purposes, it is realistic to search 3D images for the construction of 3D models, instead of capturing 3D images directly [1, 4]. Inherently, 3D free form surface matching is challenging due to the following two reasons: (1) 3D images on the web are beyond prediction in the sense of genre, quality, or even format; (2) full 3D model is often not available. Finding an accurate partially overlapping 3D image captured at different viewpoints with various quantities of overlapping and occlusion and appearance and disappearance of points remain an open problem. As a result, a powerful scheme for the representation of each view of an object is required.

The existing 3D free form surface matching methods can be broadly classified into two categories: local feature based methods and global feature based methods. While local features can be spin images [2] or surface signature [7], the global features can be shape distribution [4], harmonic shape representation [5] or skeleton [6], and the like. The former first extracts local features attached to each point, then matches these features and thus, establishes point correspondences between different images. The latter first extracts an overall object appearance representation from each image, then matches the appearance representation. While the former is computationally expensive and often more accurate, the latter is computationally efficient and often less accurate.

2. ORIENTATION DIFFERENCE LENGTH DISTRIBUTION

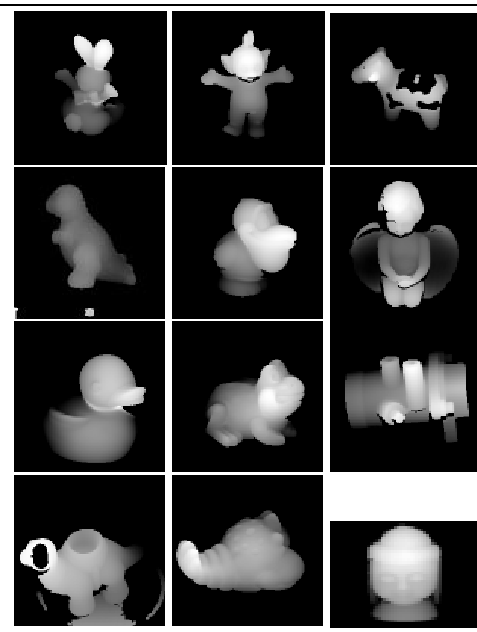

Fig. 1. The real range images used. From top to bottom, from left to right: bunny (8 views, 20°), tubby (7 views, 20°), cow (10 views, 10°, 20°, or 30°), red dinosaur (5 views, 36°), bird (6 views, 20°), angel (3 views, 20°), duck (2 views, 20°), frog (5 views, 20°), valve (2 views, 10°), dinosaur (2 views, 35°), lobster (8 views, 20°), and buddha (7 views, 20°). Here, x° denotes that the object undergoes a motion with a rotation angle of x° around an unknown rotation axis in 3D space.

Since global feature based methods are efficient and they can incorporate human-computer interaction, they are attractive not only for 3D object recognition, but also for 3D image search through the World Wide Web. While the shape distribution [4] represents a randomly selected interpoint distance distribution, in this paper, we propose using orientation difference length distribution to represent the overall appearance for each view of an object. The details for the construction of orientation difference length distribution are described as follows.

A 3D image is a rendering of a set of 3D structured data points and represents a free form surface of an object. Such an image is called a range image in the machine vision community. In each raster range image file, for any valid point $\mathbf{p}$, within its eight nearest neighbours, if there are more than two valid points $\mathbf{p}_{n1}$, $\mathbf{p}_{n2}$, $\cdots$, $\mathbf{p}_{nr}$ ($r \leq 8$), then we call this point $\mathbf{p}$ a non-boundary point. Otherwise, this point $\mathbf{p}$ is called a boundary point. For a non-boundary point $\mathbf{p}$, a plane can be used to fit its valid neighbouring points and the surface normal $\mathbf{N}$ at this point $\mathbf{p}$ can then be estimated as the normal of that plane. The surface normal $\mathbf{N}$ at point $\mathbf{p}$ is the eigenvector of matrix $\mathbf{A}$=$\sum_i (\mathbf{p}_{ni} - \bar{\mathbf{p}_n})(\mathbf{p}_{ni} - \bar{\mathbf{p}_n})^T$ that correspond to the smallest eigenvalue of matrix $\mathbf{A}$ where $\bar{\mathbf{p}_n}$ is the centroid of $\mathbf{p}$, $\mathbf{p}_{n1}$, $\mathbf{p}_{n2}$, $\cdots$, $\mathbf{p}_{nr}$, superscript T denotes the transpose of a point vector. Clearly, the surface normal is independent of translation vector of the viewpoint from which the image was captured.

Then we compute the orientation difference of neighbouring points. For point $\mathbf{p}$, its orientation difference is computed as: $OD_p = \mathbf{N} - \bar{\mathbf{N}}$, where $\bar{\mathbf{N}}$ is the normalized mean of the surface normals at $\mathbf{p}$, $\mathbf{p}_{n1}$, $\mathbf{p}_{n2}$, $\cdots$, $\mathbf{p}_{nr}$, respectively. Then we compute the Euclidean norm of orientation difference vector: $d_p = ||OD_p||$. Clearly, this orientation difference length is independent of the rotation of the viewpoint. Since both $\mathbf{N}$ and $\bar{\mathbf{N}}$ are of unit length, $d_p = ||\mathbf{N} - \bar{\mathbf{N}}|| \leq ||\mathbf{N}|| + ||\bar{\mathbf{N}}|| = 2$. This is an advantage of orientation difference length, since it provides a good opportunity to determine a fixed bin size for the histogram of orientation difference length.

Finally, we compute the histogram of orientation difference length. As long as the scale m of the histogram has been given, then the fixed bin size s can be computed as: $s = 2/m$. Once the bin size has been determined, the construction of histogram with regard to orientation difference length is straightforward. For the sake of facilitating the matching of different histograms, we normalise the histogram $\mathbf{H}_i = \{h_{ik} | k = 1, 2, \cdots m\}$ through dividing each frequency by the number of points used for the construction of the histogram $h_{ik} = \frac{h_{ik}}{\sum_k h_{ik}}$. Examples of the histograms are presented in Figure 2.

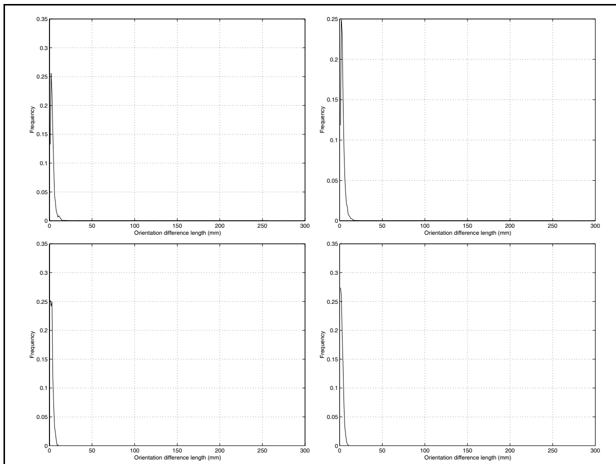

Fig. 2. The histogram constructed. Top two for two views of bunny. Bottom two for two views of cow. m=256.

3. ADVANTAGES OF ORIENTATION DIFFERENCE LENGTH DISTRIBUTION

The orientation difference magnitude describes the change in the turning angle as we move along the curve that is the intersection between the free form surface and the plane containing $\bar{\mathbf{N}}$ and $\mathbf{N}$ and thus is powerful in representing the overall appearance of objects as demonstrated in the next section. Moreover, it has the following three advantages:

- The ODLD histogram is independent of rigid body transformation (rotation and translation), scale and robust to noise. These properties are critical for any global feature extraction and matching [4].

- It facilitates the determination of the fixed bin size for the construction of histogram. Since the length of orientation difference is at most 2, thus, as long as the scale of the histogram has been given, then the fixed bin size can be easily decided. This determination of the fixed bin size has nothing to do with any specific images of objects. This is in contrast with the shape distribution [4] that applies mean normalisation to eliminate the effect of object scale on the final classification accuracy. While the mean of the interpoint distances is unpredictable, depending on the actual size of image and the specific random point selection scheme at each trial of computation, it either is difficult to determine the fixed bin size or leads to variable scales of histogram. Such a kind of characteristics makes the algorithm less reliable.

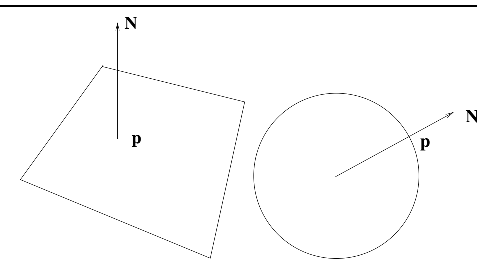

Fig. 3. Planar and spherical objects

- It is easy to examine that special objects have special ODLD (Fig. 3). For example, since all the points on the same planar patch have the same orientation, then the length of orientation difference will always be zero and thus, occupies the same bin on the histogram. For a valid point on a sphere, if its eight nearest neighbouring points are all valid, then the length of orientation difference is also zero and thus, occupies the same bin on the histogram. Thus, the objects with special surfaces have special ODLD. However, their shape distribution may be unpredictable, as discussed above. As a result, the shape distribution does not facilitate the matching of regular surfaces, while the ODLD does.

4. MATCHING OF HISTOGRAMS

Once the histogram $\mathbf{H}_i = \{h_{ik}|k = 1, 2, \cdots m\}$ has been constructed, it is critical to devise a good measurement for the matching of different histograms. For a comparative study, we implemented the following commonly-used measurements:

- Bhattacharyya divergence:

$$div_B(\mathbf{H}_i, \mathbf{H}_j) = 1 - \sum_k \sqrt{h_{ik}h_{jk}}$$

- Kullback-Leibler divergence:

$$div_{KL}(\mathbf{H}_i, \mathbf{H}_j) = \sum_k (h_{ik} - h_{jk}) \log(h_{ik}/h_{jk})$$

- χ^2 divergence: $div_{\chi^2}(\mathbf{H}_i, \mathbf{H}_j) = \sum_k \frac{(h_{ik}-h_{jk})^2}{h_{ik}+h_{jk}}$.

For all these measurements, if none of any orientation difference length falls into a bin, then the occurrence frequency of that bin is zero (Fig. 2). This describes such a scenario that a certain orientation difference length does not occur. Obviously, such a situation may dominate the final free form surface matching. Thus, we propose introducing virtual training examples for learning [3] so that a more accurate histogram can be constructed. Since we have no knowledge to believe that one frequency is higher than another, we shift all frequencies up about a small positive real number ϵ: $\mathbf{H}_i = \{h_{ik} \leftarrow h_{ik} + \epsilon | \epsilon > 0\}$. The rationale behind this operation is that we model explicitly the occlusion and appearance and disappearance of points with a small probability. Points in the image then reinforce this small probability. Thus, our new method explicitly models occluded and appearing and disappearing points with a small probability, instead of zero. Finally, we normalise the histogram again: $h_{ik} = \frac{h_{ik}}{\sum_k h_{ik}}$. In the experiments described below, we let $\epsilon = 0.0015$.

5. EXPERIMENTAL RESULTS

In order to provide a better understanding of the performance of the proposed orientation difference length distribution (ODLD) algorithm, the shape distribution (SD) algorithm [4] was also implemented and applied to the same data set without and with shifting the histograms (SODLD and SSD). For the performance measurement, we ran a series of "leave one out" tests. In each test, we compared the ODLD with and without shifting the histogram of each view of each object in the database against all others. As long as it can best match one of the views for that object, we regard the algorithm correctly classifying this image.

All the images (Figure 1) used in this paper were downloaded from a publicly available database maintained by the Signal Analysis and Machine Perception Laboratory at Ohio State University. The images were captured using a Minolta vivid 700 range camera and are of size 200 by 200. Different objects are represented by different views. The views of objects undergo motions with a rotation angle around an unknown rotation axis in 3D space respectively. The rotation angles are generally 10°, 20°, 30°, 35°, or 36°. This is the case that on the world wide web, no one has regulated what motion the object should undergo. Thus, our experiments represent the actual imaging condition for free form surface matching and the actual situation for the images on the World Wide Web. The experimental results are presented in Tables 1, 2, and 3.

Table 1. The average time in seconds on Pentium III computer for the construction of ODLD for different classes of objects with different sizes m of histogram.

Object	Time (s)		
	m=128	m=256	m=512
Bunny	20.5	24.1	30.7
Tubby	15.8	17.4	22.5
Cow	4.9	5.7	7.2
Red Dinosaur	2.6	3.2	4.0
Bird	24.0	27.3	33.6
Angel	4.7	5.0	6.3
Duck	27.5	33.5	43.0
Frog	13.0	15.8	20.6
Valve	19.5	23.5	29.5
Dinosaur	3.0	3.0	4.5
Lobster	8.7	9.5	13.0
Buddha	3.0	3.3	4.14

From Table 1, it can be seen that the larger the histogram, the longer time its construction takes. In general, the more points a view has, the more time it needs to construct the ODLD histogram. This is expected. The shape distribution for any view of an object generally takes less than one second on a Pentium III computer. Note that ODLD can be constructed offline and needs to be constructed only once.

Table 2. The average classification accuracy of different global features using different measurements over different sizes m of histogram for 6 classes of objects with 39 images.

m	Measure	ODLD	SODLD	SD	SSD
128	Bhattacharyya	87.2	87.2	11.1	68.2
	KL	86.8	87.2	68.2	68.2
	χ^2	87.2	92.7	68.2	68.2
256	Bhattacharyya	96.6	96.6	11.1	65.1
	KL	94.4	96.6	65.1	65.1
	χ^2	94.4	100.0	65.1	65.1
512	Bhattacharyya	96.6	96.6	13.8	63.4
	KL	94.4	96.6	63.4	63.4
	χ^2	94.4	96.6	63.4	63.4

Table 3. The average classification accuracy of different global features using different measurements over different sizes m of histogram for 12 classes of objects with 65 images.

m	Measure	ODLD	SODLD	SD	SSD
128	Bhattacharyya	63.3	63.3	5.5	42.7
	KL	60.9	63.3	41.5	42.7
	χ^2	63.3	67.3	40.3	42.7
256	Bhattacharyya	69.7	69.7	7.2	37.2
	KL	65.8	69.7	38.3	41.4
	χ^2	65.8	69.7	38.3	36.2
512	Bhattacharyya	70.9	69.7	4.4	39.5
	KL	66.9	69.7	35.7	39.5
	χ^2	69.7	69.7	38.5	39.5

Thus, it can still satisfy the requirements of real time applications, such as 3D image internet search.

6. EXPERIMENTAL RESULT ANALYSIS AND CONCLUSION

In this paper, a powerful scheme for the representation of each view of an object has been proposed and validated. In addition, we have made the following observations from Tables 2 and 3:

- The ODLD produces a significantly better result than the shape distribution. In addition, while the Bhattacharyya divergence fails to match the shape distribution for different images, it produces very good results for the matching of ODLD without shifting histogram. This shows that ODLD is more expressive for the representation of objects than the shape distribution. This is because ODLD extracts information from all points, while the shape distribution extracts interpoint distance information of a limited number of randomly selected point pairs (10,000 in this paper).
- The shifted ODLD (SODLD) generally improves the classification accuracy relative to that without shifting. The shifting of histograms clearly improves the classification accuracy for the shape distribution based matching, especially when $m = 512$. Thus, the explicit model of occlusion and appearance and disappearance of points is successful.
- The fewer candidate images used for matching, the higher the classification accuracy. This is expected, since a large number of candidate images render it difficult to distinguish correct matches from false ones.
- The size, m=256, of histogram provides a good compromise between the computation of histogram and its resulting classification accuracy.

In practice, it is useful to combine different algorithms where some algorithms are used to roughly match free form surfaces and the others are used to refine the matching results. Also, here we just employ ODLD for free form surface matching. In practice, we can use other information like text, 2D sketching, 3D sketching [1], or 2D projective images for matching. As a result, the free form surface matching accuracy is expected to be improved.

7. REFERENCES

[1] T. Funkhouser, P. Min, et al. A search engine for 3D models. *ACM Trans. Graphics*, 22(2003) 83-105.

[2] A. Johnson and M. Herbert. Using spin images for efficient multiple model recognition in cluttered 3-D scenes. *IEEE Trans. PAMI*, 21(1999) 433-449.

[3] T. Mitchell. Matching Learning. The McGraw-Hill Companies, Inc., 1997.

[4] R Osada, T. Funkhouser, B. Chazelle, and D. Dobkin. Shape distributions. *ACM Trans. Graphics*, 21(2002) 807-832.

[5] M. Kazhdan and T. Funkhouser. Harmonic 3D Shape Matching. SIGGRAPH 2002 Technical Sketches, p. 191, 2002.

[6] H. Sundar, D. Silver, N. Gagbani, and S. Dichison. Skeleton based shape matching and retrieval. *Proc. SMI*, 2003, pp. 130-139.

[7] S.M. Yamany and A.A. Farag. Surface signatures: an orientation independent free-form surface representation scheme for the purpose of objects registration and matching. *IEEE Trans. PAMI*, 24(2002) 1105-1120.