ABSTRACT

Despite the popularity of spin images in surface matching and registration, disadvantages such as noise sensitivity and low discriminative ability still hindered their usefulness in real applications. In this paper, a novel approach was proposed for improving the spin images. The proposed method modified the standard spin images by using angle information between the normals of reference point and neighboring points. This information largely increased the robustness to noise without losing the intrinsic advantages of spin images. Moreover, signs were defined to incorporate the directions of angles which were shown to be able to further improve the descriptive power. Experiments were also conducted to show the outperformance of improved spin images under different levels of noise, and good agreements were obtained by comparing with the standard spin images and a recent popular 3D descriptor.

Index Terms—Surface matching, spin images, 3D descriptor

1. INTRODUCTION

3D shape matching is a fundamental and challenging task in computer vision. With the rapid increasing number of 3D models, its role becomes more attracted in various research areas such as shape registration, retrieval, object detection, biometrics, and 3D recognition. Over the past decades, a large number of surface matching techniques have been proposed [1-3], among which matching with local shape descriptors has become a popular research trend due to its promising ability to handle with missing data, occlusion, clutter, rotation, translation and resolution variance.

One of the most excellent local shape descriptors is Spin Images (Fig.1) which was first proposed by Johnson and Hebert in 1997 [4] for surface registration, and has also been used for recognition problem [5-7]. After its first invention, numerous improvements and extensions were made. For example, authors in [8] and [9] tackled the variant resolution problem by mapping the surface area rather than surface points. And the sampling problem was also solved through using distinct landmarks. A multi-resolution representation for spin images was presented by Dinh and Kropac [10] which was shown to be able to increase the efficiency by comparing spin images in a low-to-high resolution manner, but this technique can also adds the risk of aliasing. In [11], textured spin image was proposed for 3D registration. This method did not modify the original structure of spin images. Instead, the enhancement was achieved by adding texture information. Another extension of spin image was 3D shape context [12] which accumulated 3D histograms of points within a sphere centered at the basis point. This method was shown to be more robust than spin images but incurred significantly higher computational cost. More recently, Tombari [13] proposed a new descriptor by combining a unique and repeatable local reference frame with a 3D descriptor of hybrid signature and histogram. But one drawback of this method is that the supporting angle is not considered while it is well defined in spin images. This may affect its performance under clutter. And the cosine function values may be also not discriminative enough.

This paper aims to further improve the descriptiveness and robustness of spin images and in the mean time preserve the good properties of spin images such as efficient to build and resistant to partial views. In section 2, the proposed Improved Spin Image is illustrated in detail. A kind of new descriptive information named ‘signed angle’ is proposed to modify and enhance the standard spin image. Moreover, a repeatable local Reference Frame technique is adopted for creating the normals. To validate the advantages of the proposed method, experiments are conducted with results shown in section 3 to compare the proposed improved spin images with standard spin images and a recent popular 3D descriptor. Finally conclusions are drawn in section 4. For simplicity, in the following sections ISI and SI are short for Improved Spin Image and standard Spin Image respectively.

Fig.1. A spin image built around reference point \( p \)
2. METHODOLOGY

For a vertex \( p \) on a surface (Fig.1 left), a Spin Image (SI) is computed in a cylindrical coordinate system defined by vertex \( p \) and its corresponding normal. All points within local support region are encoded into a 2D representation which is a histogram with one axis \( \alpha \) indicates the perpendicular distance from the reference normal ray while the other axis \( \beta \) stands for signed perpendicular distance to the tangent plane (Fig.1 right). Nice properties such as rotation, translation and pose invariant allow SI to work for many problems. However, less descriptive power and noise sensitivity are the intrinsic deficiencies of SI. The authors of SI suggested that grouping point matches with geometric consistency should be used to enhance the robustness. But this can decrease the efficiency. In fact, these disadvantages are due to the fact that SI solely encodes the geometry information which is very sensitive to noise and less discriminative. Based on this observation, an Improved Spin Image (ISI) is proposed in this section, which encodes a new feature named Signed Angles.

Before establishing ISI, the normals need to be built up first. A popular way to generate the normal of each feature point on the surface is to calculate the normals of its surrounding faces first, and then the sum of these normals which are weighted by their corresponding areas is used as the normal of this feature point. Despite the simple calculation, non-unique and ambiguity can be introduced when these normals are used as reference axes to generate SI or ISI, the number of mismatches will increase during matching stage. In this work, a repeatable local Reference Frame technique [13] is utilized to build the normals such that both uniqueness and unambiguity can be achieved.

Having built the normals, the ISI can be established by replacing one of the axes of SI with angle information. Here, the angle indicates the included angle between the normals of reference point and the neighboring points. It seems that either \( \alpha \) or \( \beta \) can be replaced. To facilitate intuitive understanding, \( \beta \) is chosen for replacement. And the ISI can be explained as angles’ distribution among different rings. Here each ring corresponds to a specific \( \alpha \) value (Fig.4 left). There has been similar modification which utilizes cosine function of these included angles within local grid [13]. This results in a coarser binning for direction close to the reference normal direction and finer one for orthogonal directions. Although this representation can limit the noise to some extent, it is not suitable for a large support region like that in SI, because in a large support region neighboring points far away from reference point with small included angles should be treated equally as those with large angles. So in this paper the included angles are used directly.

In addition to the included angle, the directions of normals should also be encoded. For example, in Fig.2, \( n_0 \) and \( n_1 \) have the same angle relative to \( n \), but apparently they point to different directions: \( n_0 \) points towards \( n \) while \( n_1 \) points outward \( n \). Therefore, \( \theta_0 \) and \( \theta_1 \) should be discerned although they hold the same absolute value. To differentiate their directions, an angle is assigned positive when it points towards \( n \), otherwise it should be assigned negative. After defining the signed angles, the local 3D surface can be mapped to 2D domain as an improved spin image using equation 1.

\[
S_p : R^3 \rightarrow R^2
\]

\[
S_{p,x}(\alpha, \beta) = (x - p)^2 - (n \cdot (x - p))^2, D(\arccos(n \cdot n_{\alpha}))
\]

(1)

where \( p \) is the reference point and \( x \) is a surrounding point. \( n \) and \( n_\alpha \) are the normals of \( p \) and \( x \) respectively. \( D \) is the sign indicator determined by equation 2.

\[
D = \begin{cases} 
+1, & \text{dot}(n, x - p) \leq \text{dot}(n_\alpha, x - p) \\
-1, & \text{otherwise}
\end{cases}
\]

(2)

When \( n_\alpha \) points towards \( n \), the included angle between \( n \) and \( x - p \) is larger than the included angle between \( n_\alpha \) and \( x - p \), and the dot product of \( n \) and \( x - p \) is thus less than the dot product of \( n_\alpha \) and \( x - p \). So the sign indicator is assigned +1. Otherwise, it is assigned -1 (Fig. 3).
Finally, the ISI is generated as is shown in Fig.4, in which each row corresponds to a specific α determined by the perpendicular distance from the reference point while each column corresponds to a certain signed angle. From the figure, it is clear that ISI can also be seen as a distribution of the signed angles among each ring around the reference point.

3. EXPERIMENTAL RESULTS

3.1. Dataset and Evaluation Methodology

The dataset for the following experiments includes 6 models and 45 scenes which can be downloaded from [14]. The 6 models are originally from Stanford 3D scanning Repository [15]. The 45 scenes are built up by randomly rotating and translating different subsets of the 6 model set to create clutter. Similar to [13], three levels of Gaussian noise are added. Three levels of noise σ1, σ2 and σ3 correspond to 10%, 20% and 30% of the average mesh resolution (Fig.5).

For each model 1000 feature points are randomly selected, and n * 1000 feature points were extracted from each scene (n indicates the number of models contained in each scene). Feature vectors were built for every feature point using the shape descriptor. During the matching procedure, all the feature points of scenes are matched against the feature points of the 6 models by iterating the models one by one. If the Euclidean distance between feature vectors for a particular pair of feature points is below a given threshold, this pair is called a match. A correct positive indicates a match where the two feature points correspond to the same physical location, while a false positive is a match where two feature points are from different physical locations. And the total number of positives is prior known (number of scenes multiply by number of feature points of each scene). Having these values recall (3) vs. 1 - precision (4) curve can be generated by varying the threshold.

\[
\text{recall} = \frac{\text{number of correct positives}}{\text{total number of positives}} \quad (3)
\]

\[
1 - \text{precision} = \frac{\text{number of false positives}}{\text{total number of matches (correct or false)}} \quad (4)
\]

3.2. Comparisons of ISI with Other Descriptors

In this section, ISI is compared with standard SI and a recent 3D shape descriptor (SHOT). For these three descriptors, support radius is their common parameter which determines how many neighboring points will be involved in calculation of the local descriptor. To ensure a fair comparison, the same support radius is set for these three descriptors. In this experiment, it is set as 10 times of the mesh resolution. The rest parameters are set as follows. For ISI and SI, image size is set as 15. This results in a descriptor with length of 225. And the support angle is set as 90 degree, which can be used to limit the effects of partial views and clutter. For SHOT, the performance is also influenced by number of spatial bins which is set as 32 as is suggested in [13]. And bin number in each shape histogram is set as 10. These settings give rise to a SHOT descriptor with length of 320.

The comparison results are presented in Fig.6, Fig.7 and Fig.8 respectively corresponding to three levels of noise. The proposed ISI outperforms SHOT and SI in all experiments. The reason for the success of ISI lies in mainly three aspects. First, the signed angles include more discriminative information of local surface. Second, support angle help limit effects of clutter. In the above experiments the support angle is set as 90 degree for ISI and SI. This is not considered in SHOT which encodes all the neighboring point within the support radius. Third, the normals generated by the repeatable local Reference Frame technique help further enhance the performance since the non-uniqueness and ambiguity have been eliminated.

It is also worthy to notice that the length of ISI (225) is shorter than SHOT (320), but high accuracy is still achieved. And a descriptor with shorter length can allow faster matching. This is a big advantage when a large number of models need to be matched. Additionally, in the above experiments both ISI and SI use 90 degree support angle which means that at each reference point ISI and SI only encode the information contained in a hemisphere, whereas SHOT encodes the information of the whole sphere. This shows that ISI has an extremely powerful descriptive ability.

4. CONCLUSIONS

This paper presented a new 3D surface matching method for improving the Spin Images (SI). The Improved Spin Images (ISI) used signed angles as a replacement for the β value (perpendicular distance to the tangent plane) of the standard SI. The signed angles not only encoded the basic first-order differential of the normals between reference points and its neighboring points within the local support but also the directions of the normals. The α value (perpendicular distance from the reference normal ray) of the SI is inherited unchanged in ISI. Each α corresponds to a specific ring. In this way, the distribution of the signed angles in every ring was encompassed into the ISI. To
eliminate the non-uniqueness and ambiguity, a local Reference Frame technique was adopted to generate the normals. Finally, experiments showed that the proposed ISI outperformed standard SI and a latest 3D descriptor (SHOT) under different levels of noise.

5. ACKNOWLEDGEMENT

We acknowledge Dr. Federico Tombari’s help for providing us their testing codes and evaluation suggestions.

6. REFERENCES


