

Scale and Resolution Invariant Spin Images for 3D Object Recognition

Jihad H'roua¹, Aissam Bekkari²,
Driss Mammass³, Ali Bouzit⁴
IRF-SIC Laboratory
Ibn Zohr University
Agadir, Morocco

Patrice Méniel⁵
ARTEHIS laboratory,
University of Bourgogne Franche-Comté
Dijon, France

Alamin Mansouri⁶, Michaël Roy⁷
Le2i Laboratory
University of Bourgogne Franche-Comté
Auxerre, France

Abstract—Until the last decades, researchers taught that teaching a computer how to recognize a bunny, for example, in a complex scene is almost impossible. Today, computer vision system do it with a high score of accuracy. To bring the real world to the computer vision system, real objects are represented as 3D models (point clouds, meshes), which adds extra constraints that should be processed to ensure a good recognition, for example the resolution of the mesh. In this work, based on the state of the art method called Spin Image, we introduce our contribution to recognize 3D objects. Our motivation is to ensure a good recognition under different conditions such as rotation, translation and mainly scaling, resolution changes, occlusions and clutters. To that end we have analyzed the spin image algorithm to propose an extended version robust to scale and resolution changes, knowing that spin images fails to recognize 3D objects in that case. The key idea is to approach the representation of spin images of the same object under different conditions by the mean of normalization, either these conditions result in linear or non-linear correlation between images. Our contribution, unlike spin image algorithm, allows to recognize objects with different resolutions and scale. Plus it shows a good robustness to occlusions up to 60% and clutters up to 50%, tested on two datasets: Stanford and ArcheoZoo3D.

Keywords—3D object; recognition; spin image; resolution; scaling

I. INTRODUCTION

New information and communication technologies have emerged in 1990s and have grown exponentially in power. The digital revolution which has been developing since its birth at the end of 20th century, has affected different sectors throughout the world. This revolution has led to the emergence of new type of data, resulting in new and broader databases, particularly 3D data. This requires technological advances in image processing or overall computer vision. Given the very wide spectrum of industrial, military and medical applications that can be considered, this field in its turn has developed very quickly. In the context of this digital revolution, notably in cognitive sciences, scientists in computer vision have redirected their efforts to put into place a variety of interactive applications with 3D real world, like 3D object recognition. To that aim, a better understanding of how the human visual system works is necessary. A first classical hypothesis assumes that, to recognize an object, the human brain starts by extracting features from objects captured by eyes. Then, depending on his previous knowledge, he elaborates a matching process. Nevertheless, with the development of neuro-sciences, scientists

assume that data in human brain travels in the neural networks where each node performs a separate task, to achieve the visual cortex where the recognition is performed based on its previous trained data. So, inspiring from this discovery, researchers in computer vision give another perspective called deep learning. Concerning classical hypothesis, different approaches have been proposed depending on the explored levels of the object and also extracted features. If the targeted level is global and tends to describe the overall shape of the object, we talk about global approaches. In the other hand, if the method focuses on extracting only local features, they are called local methods. Proposed approaches, either local or global aim to ensure the robustness to different condition 3D objects in real scenes can be through for example, rotation, translation, geometric deformations, occlusions, clutters, scaling, etc. In this respect, when it comes to occlusions and clutters, local approaches are known to be more efficient. Other strengths of this category is that they are popular to not requiring any segmentation and the pose estimation is simpler. However, the fact that local approaches are founded on local neighborhood, which is highly affected by the resolution changes, make them less discriminating. In addition, a verification, step is always needed to eliminate incorrect correspondences and the spatial information is missed. Concerning global methods, they are more discriminating since they provide a global description of the shape of the object. Besides, by only computing the nearest neighbor of the descriptor, we can perform matching, which makes it easier. In the opposite to local methods, they do not handle occlusions and clutters, the pose estimation is more complicated and they usually require a segmentation as a pre-processing. In this paper we introduce a novel local shape based approach approach for 3D object recognition, crafted to deal with resolution and scaling changes of the object in occluded and cluttered scenes. Our contribution, called Invariant to scale and resolution spin images (ISRSI), is based on a state of the art method called spin images. Spin images fails when the resolution and the scale of objects change. By performing a normalization step and defining efficiently the required parameters, we succeed to make this descriptor invariant to scale and resolution changes. Our contribution has shown good robustness to occlusions up to 60% and clutters up to 50%. The paper is laid out as follow. We briefly quote some related works in Section 2. Then, in Section 3 we describe the background method. The Section 4 is dedicated to introduce more details about our contribution. While experiments are conducted in Section 5. And finally, a conclusion is given in

Section 6.

II. RELATED WORKS

3D free form objects recognition is a very challenging task due to the presence of different conditions revealed in the real world to take into account, like occlusions, clutters and other transformations such as scaling, rotation and translation. Besides, the 3D reconstruction of real objects adds more constraints mainly mesh resolution changes. To that end, researchers have proposed different range of methods. The state of the art introduces different survey on 3D object recognition approaches [1] and [2]. One can classify those methods to shape based approaches, local shape-based approaches, topological approaches and view-based approaches. Global shape-based approaches: As their name indicates it, they aim to describe the coarse shape of the 3D model. In this direction, Osada et al. [3] represent an object as a shape distribution by elaborating five functions based of the choice of a random set of points. Authors have shown that their approach is invariant to geometric transformations. Another approach have been proposed by Paquet et al. in [4] that can be used in the same time for 2D and 3D objects, have shown good robustness to resolution, translation and rotation. Local shape based approaches: are also known sometimes as key point based methods. In this branch we find a multiscale approach proposed by Nouri et al. [5]. They use patches with adaptive size to detect salient regions on the surface of a 3D model. Tang et al. [6] have proposed a local descriptor based on geometric centroids. Another method have been introduced by Maes et al. [7] as an extension of SIFT descriptor [8] to the 3D domain. Spin image descriptor [15] is another approach that aims to explore the local distribution of vertices on the surface area of the object to create a set of 2D images considered as the descriptor of the object. View based approaches: or in other words 3D/2D approaches describe object based on its projections in a 2D space. For example, Xiang et al. [9] have introduced a new descriptor called 3DVP for 3D voxel pattern encodes the object by a triplet (appearance, 3D shape, occlusions). In another contribution in [10] authors compute different features for object's views, such as 2D Fourier descriptor, 2D Zernike moments and 2D Krawtchouk moments. Topological approaches: we cite here for example the contribution of Pickup et al. [11]. It consists of constructing the skeleton based on Au et al's technique [12]. Then 3D pose normalization is performed using the canonical form of the skeleton of the object. And finally, utilizing Yan et al's approach [13] a deformation of the mesh is fulfilled in order to match the canonical transformation of its skeleton. Another approach aims to improve Reeb Graph of an object has been presented by Thierny et al. in [14] following three steps: 1) Extraction of salient vertices. 2) Emphasis of the overall shape of the object using an application function. 3) Refinement of Reeb graph into topological skeleton by the mean of constrictions.

III. BACKGROUND: SPIN IMAGE ALGORITHM

Spin image descriptor is an algorithm that has been first introduced by Johnson et al. [15]. The 3D mesh model is described by a set of its 2D projections on a well-defined 2D local coordinate systems. In order to define a local coordinate

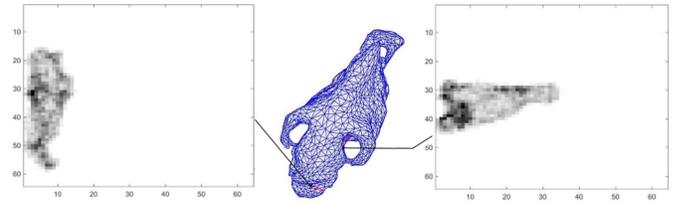


Fig. 1. Two spin images from two oriented points on the surface mesh of skull model.

system, authors first define an oriented point O as the center of this local basis. The oriented point in its turn is defined by a vertex $p(x, y, z)$ and a normal surface n . The normal surface is the plan tangent to the vertex p and perpendicular to its vector normal n . Then to define the two cylindrical coordinates α and β are computed for each other vertex x on the surface mesh such as:

$$\alpha = \sqrt{\|x - p\|^2 - (\vec{n} \cdot (x - p))^2} \quad (1)$$

$$\beta = \vec{n} \cdot (x - p) \quad (2)$$

So for each vertex a corresponding spin image is obtained using this projection function below:

$$S_O : R^3 \mapsto R^2$$

$$S_O(x) \mapsto (\alpha, \beta) = (\sqrt{\|x - p\|^2 - (\vec{n} \cdot (x - p))^2}, \vec{n} \cdot (x - p)) \quad (3)$$

During projection of vertices, authors have specified three parameters to take into account. First, we have bin size b which specifies the size of bins used to accumulate points projected. Then the angle support ϕ , it is the angle between the normal vector of each vertex to project and the normal vector of oriented point. Lastly, is the width W of the spin image. Equations (4) and (5) shows the relation between those three parameters.

$$i = \left\lfloor \frac{\frac{W}{2} - \beta}{b} \right\rfloor \quad j = \left\lfloor \frac{\alpha}{b} \right\rfloor \quad (4)$$

$$a = \alpha - ib \quad b = \beta - jb \quad (5)$$

Fig. 1 illustrates two spin images from two oriented point on the surface mesh of a horse's skull from ArcheoZoo3D database.

For the purpose of performing a matching between two objects, authors have put into place a surface matching algorithm following different stages. We summarize the different phases in the pipeline below. See Fig. 2. For detailed description readers can refer to [16].

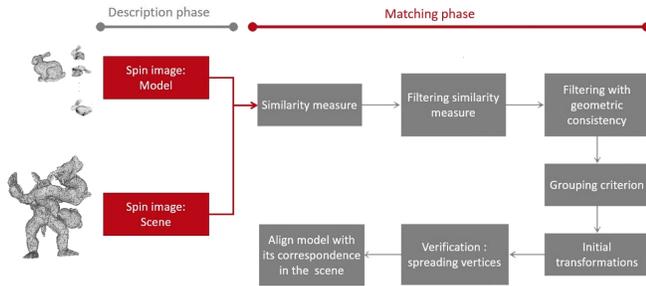


Fig. 2. Pipeline of spin images matching.

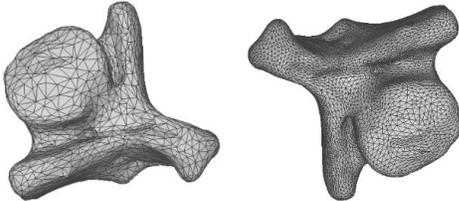


Fig. 3. An example of 3D mesh of caudal with two different resolutions.

IV. SCALE AND RESOLUTION INVARIANT SPIN IMAGES: SRISI

A. Invariance to Resolution

At the end of the eighties, efforts to reproduce three dimensional world have borne fruit and the first 3D scanning systems, based on imaging triangulation, were installed for industrial applications. After decades, high definition 3D scanners are at the forefront of archeology field, for a wide range of items, small-sized artifacts such as coins, teeth and bones, fragments and scripts up to significant figures, statues and small buildings. Thus, bringing history back to the life. According to a 3D scanning pipeline [17], a complete 3D model of the object is provided, in general in the form of 3D meshes. Here comes other challenges related to the representation of the object to take into consideration, in order to insure a good recognition. One of these parameters is the resolution of the mesh, which is defined here as the lengths of edges of the mesh, or precisely, the median of the lengths of all edges of the mesh. So the same object can be represented with different resolutions, implicitly different number of vertices, see Fig. 3.

As we have cited above, the spin image algorithm is known to be robust to occlusions and clutters, but when it comes to spot the same object with different resolution in the scene, or when we change the scale, the process of recognition fails. For the purpose of making spin image algorithm robust to resolution changes, we need first to understand what the impact of resolution changes is on the description phase that makes this algorithm fails and at which level of the matching phase the process crashes?

To that aim, let us back up a moment and talk about the creation of spin images, mainly, how we generate a spin map for each oriented point. The generation of spin images is controlled by three parameters. The first parameter is the bin size, which is defined as a multiple of the resolution of the surface mesh. Then the support angle, that controls the

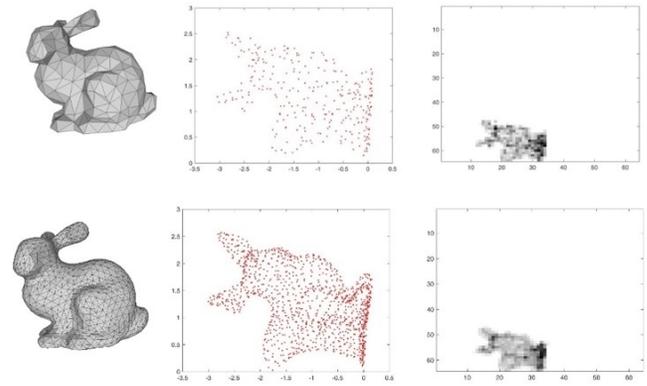


Fig. 4. Two spin maps and their two spin images of bunny with different resolution.

vertices to be projected based on the angle between their normal vectors and the normal vector of the oriented point. The third parameter is W the width of the spin image. So, when the resolution changes, the number of vertices is not the same and their space partitioning is different, which leads to a difference in the set of normal vectors to be managed. All these changes can be clearly seen on the results of equations (3), (4) and (5). In Fig. 4, we show the difference between two spin maps and their two corresponding spin images of the same object bunny with different resolution.

After visualizing results obtained of resolution changes during the extraction of the descriptor, we need to understand now how it impacts the matching phase. Let us first analyze the first step of matching algorithm: the computation of the similarity measure. In order to find for each model image the one that is most similar to it in the scene, authors have defined a similarity measure eq. (7) based on the correlation coefficient eq. (6).

$$R(P, Q) = \frac{N \sum p_i q_i - \sum p_i \sum q_i}{\sqrt{(N \sum p_i^2 - (\sum p_i)^2)(N \sum q_i^2 - (\sum q_i)^2)}} \quad (6)$$

$$C(P, Q) = (\arctan(R(P, Q)))^2 - \lambda \left(\frac{1}{N-3} \right) \quad (7)$$

The correlation coefficient provides a measure of the intensity and direction of the linear relationship between two variables. Further, this metric is useful in measuring linear relationships. But when the relationship between two images is nonlinear, this measure may give somewhat misleading information. Since resolution changes cause weak linearity or even nonlinearity between images, the algorithm using the correlation coefficient doesn't provide good matches. We establish a correlation diagram to illustrate the impact of resolution changes on the relation between intensities of two spin images, Fig. 5 shows the results.

Besides, when we take a look at the values of intensities of model spin images and scene spin images, we can see clearly how different the ranges are, in Fig. 6 we provide an example

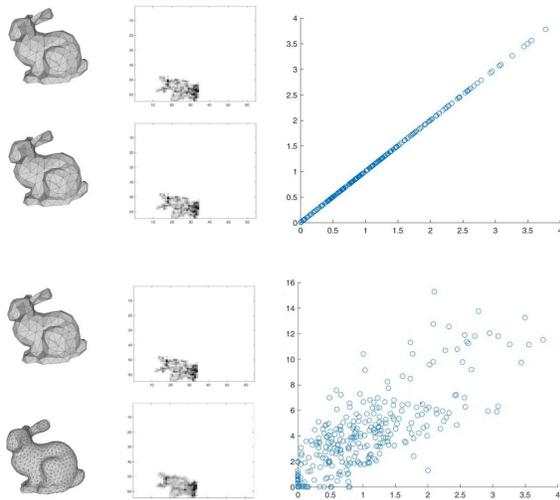


Fig. 5. An illustration of the impact of resolution changes on relation between images using the correlation diagram. Up the correlation between spin images of the same object with the same resolution is linear. Down, the difference in resolution results in a non-linearity of correlation.

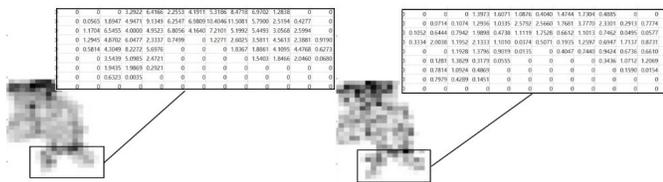


Fig. 6. An example of the difference between two spin images resulted from bunny under different resolution. Left: the correspondent spin image of vertex 100 from bunny with resolution 0.3 and number of vertices equal to 302. Right: the correspondent spin image of vertex 100 from bunny with resolution 0.15 and number of vertices equal to 1202.

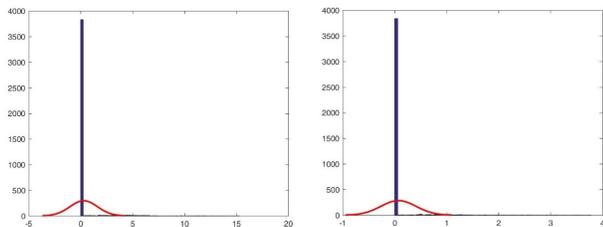


Fig. 7. The corresponding histograms of the two spin images from bunny under different resolutions. In the histogram left, range of intensity values varies between -5 and almost 20. Right, values are between -1 and almost 4.

to show the difference between intensities of two spin images that are meant to be similar.

In Fig. 7, we illustrate their corresponding histograms to show clearly the difference between ranges.

As the histogram is of essential importance in terms of characterizing the global appearance of a given image one needs to represent the values of compared histograms in the same range in order perform an effective comparison. As known, the min-max normalization approach is the simplest normalization technique in which we fit the data, in a pre-

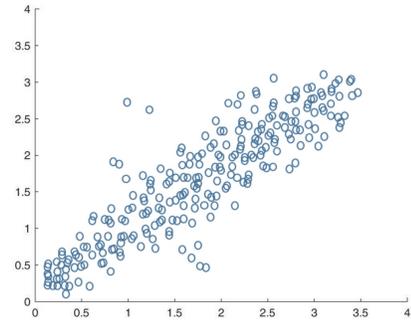


Fig. 8. correlation diagram of two spin images after normalization.

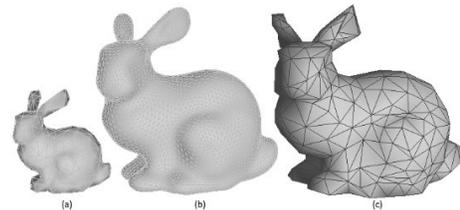


Fig. 9. Two different scenarios of scaling of bunny: (a) The original object. (b) Scaling and resolution scaling changes. (c) Scaling changes only.

defined range, as it is very common and usually more efficient. To normalize the data in the boundary of $[A,B]$, the min-max normalization is defined as:

$$x_i - normalized = \frac{(x_i - \min(x)) * (B - A)}{\max(x) - \min(x)} + A \quad (8)$$

So, the idea here is to bring the two spin images to the same range $[0,1]$, in order to normalize bin values for all spin images to be able to compute the correlation coefficient efficiently.

In Fig. 8, we show the impact of normalization in the correlation diagram of two spin images with different resolution after normalization.

Here we can see clearly that the two images are more correlated.

B. Invariance to Scale

One other drawback of spin image algorithm is the scaling. We have two scenarios about scaling. The first one concerns the object with the same number of vertices, but the scale is different, see Fig. 9(c) and the second one is when the same object is represented with different resolution and scale in the scene, see Fig. 9(b).

The first case is simpler. As the scaling here does not change the normal vectors of vertices, the number of vertices to project controlled by the parameter A (Angle support) is the same on each spin map. Since the image width is fixed for both spin image model and scene, to deal with changes which influences the accumulation of points in each bin of the spin image, the bin size of the scene spin image should be set to

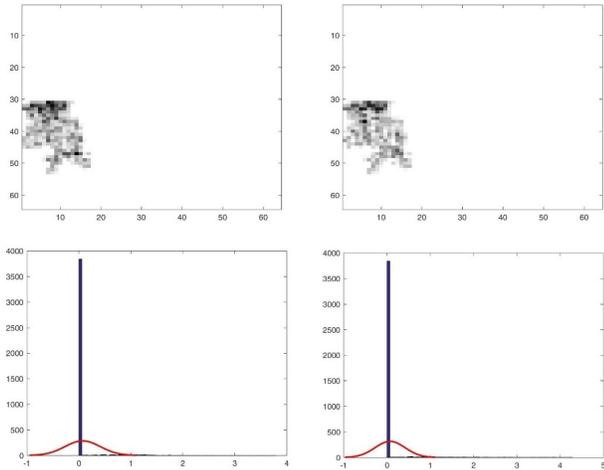


Fig. 10. Two spin images and their corresponding histograms of bunny and its scaled version with no resolution changes. On left the spin image of bunny and its histogram. On the right side the spin image of the scaled version and its histogram.

a multiple of scaling factor λ and bin size of the model spin image:

$$b_s = \lambda b_m \quad (9)$$

Which helps to get spin images with the same intensity values of spin images when we have no scaling.

In Fig. 10, we illustrate an example of a spin image of the object and its scaled version. As in practice we don't usually know the scaling factor, the bin size of the scene is determined empirically as a multiple of a multiple of the bin size of the model. This will reduce the effect of discrete location and individualization effect of points on the surface scene.

In order to show the importance of bin size for spin image matching, we experiment the effect of bin size on match distance, which is defined as the median of all distances computed between each computed correspondences during the phase of the similarity measure. A good match is established, which means correct correspondences are computed when the match distance is low. Results are shown in Fig. 11.

The second case which is more complicated is when both resolution and the scale are different. Combining the resolution changes and scaling has the same effect as changing only resolution. In that case also both the spatial distribution and the number of vertices are different. To explain with more details during the description phase different vertices on the surface mesh of the scene are falling into different bins. This is due to the difference in the number of vertices which leads to the difference of their spatial location from the ones of the model. Consequently, the normal vectors are also different, which has an impact on the choice of vertices based on the angle support α . Therefore, spin images that are meant to be similar will be dissimilar. In Fig. 12, we provide an example of this case showing the spin image of the same vertex for a bunny model with three different resolutions.

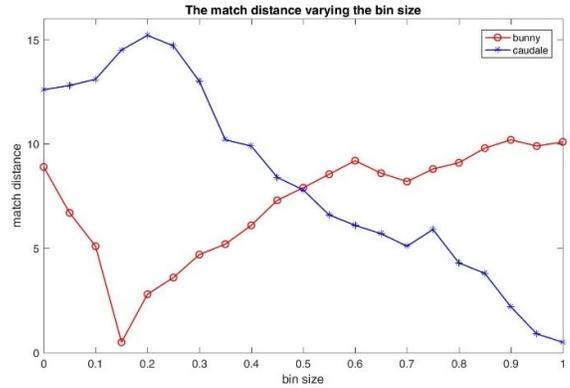


Fig. 11. The impact of varying bin size on the distance match.

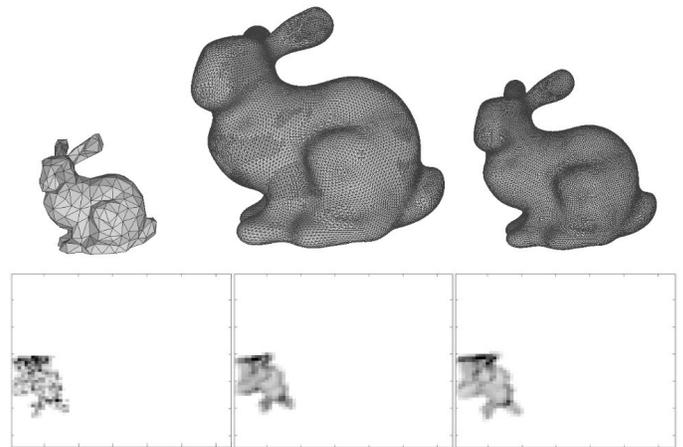


Fig. 12. Scaling and resolution changes and its impact on the generation of spin images.

To overcome this issue, we proceed in the same way as we did for resolution changes. Since the difference in image width is not handled by correlation coefficient, we start by fixing the image width for both models and scenes. Then to reduce the effect of discretization and in order to represent the shapes in spin images in the same scale level, we set the bin size of the scene to a multiple of the bin size of the model. Then we perform a normalization to bring intensity values to the same range to compute the correlation coefficient efficiently. To validate what we have explained above and the choice of bin size empirically we evaluate a plot of match correlation. We mention here that the match correlation measure is the median of the histogram of the correlation coefficient between spin-images computed for all point matches. When correlation is high, the correspondences are correctly computed. See Fig. 13.

V. EXPERIMENTAL RESULTS

The current section provides an evaluation of our suggested approach SRISI in comparison with the spin image algorithm. For this purpose, we perform a wide range of tests utilizing models from two datasets: ArcheoZoo3D and Stanford's 3D scanning repository. First, Section A briefly presents our

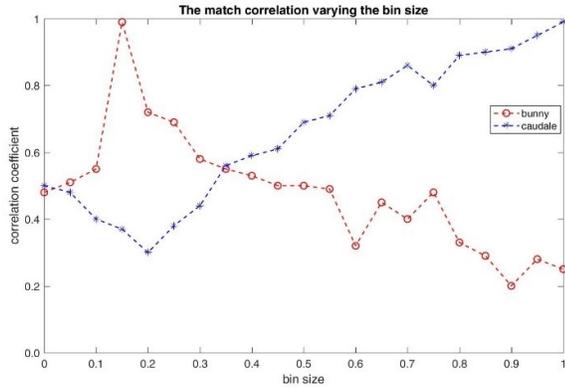


Fig. 13. The influence of varying the bin size on the match correlation for bunny and caudale.

database. Afterwards, in Section B, we provide detailed technical information on the implementation environment. Next, the experiment carried out is revealed in Section C. In the same section we measure the precision and recall to evaluate the performance of our contribution. And then we compare it to the standard algorithm with a discussion of strengths and shortcomings of our contribution.

A. Datasets

In this work we have validated our approach on two datasets. The first one is Stanford 3D scanning repository. A well known repository that provides some dense polygonal models publically. The second database is Archeozoo3D. It gathers 3D scans of horse's bones. Before recognition we have processed objects to remove all unreferenced vertices. Then we construct proper triangulated surfaces with screened Poisson surface method to remove holes. We sampled all objects to have the same resolution.

B. Implementation

In order to put the algorithm of spin images into action, we have based our implementation on the information provided in the thesis work [16]. We have implemented the whole phases of the algorithm from descriptor extraction to verification passing by the matching in Matlab. Concerning models in the two databases, they have been processed, whether for creating scenes, normalizing vectors or applying transformations, etc. with the aide of Meshlab, blender and using the "Toolbox Graph" of Peyre¹ in Matlab. About environmental information, our experiments were carried out on a computer with 2.50 GHz Intel i7 processor and 16GB of memory.

C. Results and Discussion

The purpose of this current section is to provide an evaluation of our proposed method SRISI in comparison with the original one SI. In order to provide a robust evaluation, the state of the art presents different metrics. We have chosen two of the most important ones utilized in the information retrieval

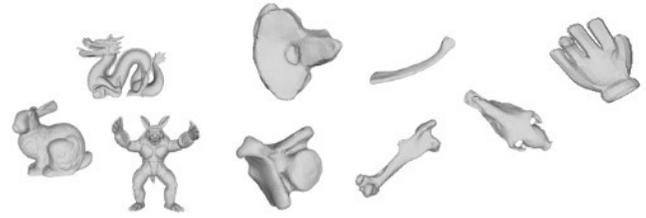


Fig. 14. Objects used to run tests. three first objects from 3D Stanford repository. Five second objects from ArcheoZoo3D database. And lastly glove model.

domain, Precision and Recall. The mathematical formula for each one is given in equations (10) and (11).

$$Precision = \frac{tp}{tp + fp} \quad (10)$$

$$Recall = \frac{tp}{tp + fn} \quad (11)$$

With tp : True positives is the number of times an existing object in different scenes with different conditions is correctly recognized. fp : False positives indicates the number of times a non-existing object in the scenes is mentioned to be recognized. It is to say that the algorithm finds correspondences on the scene, so the model is aligned with another object. Finally, fn : false negatives, when a model exists in the scene, but the algorithm fails to recognize them, in our case, it fails to find any correspondences. To test the validity of our approach, we used three objects from the Stanford repository, five objects from the ArcheoZoo3D database and one other object called glove modeled by Alexander Masliukivaky. The objects are listed in Fig. 14.

At first all objects have the same resolution. Resolution here refers to the median of the lengths of the edges between the vertices. The tests were done first for each isolated object. We initially change only the resolution and keep the scale fixed, then apply the transformations (translation, rotations) as well as truncating parts of the objects. We next carry out tests in reverse. We fixed the resolution and changed the scale. Lastly, we change both resolution and scale. In the second time we test the robustness of our method to oclusions and clutters. To do that, we have created 30 scenes from 4 objects of Stanford datasets, then we have changed the resolution of scenes two times, which results in 90 trials for each model. Then we have repeated the same process for Archeozoo3D dataset. So roughly, concerning SRISI we get 360 trials for Stanford and 360 for Archeozoo3D. For SI, as mentioned earlier, the algorithm does not find any correspondence. For the results presented in this work, image width is set to 64, the resolutions of models is set to 0.3, the bin size is 0.15 and the angle support equal 180. To show the effect of oclusions and clutters on our method, we will compute the recognition rate in terms of oclusions and clutters. To do this, for each scene of the 30 scenes created from the Stanford database, we run the recognition test. This will allow us to deduce the true positives, false positives and false negatives. Then we calculate

¹<http://www.mathworks.com/matlabcentral/fileexchange/5355-toolbox-graph>

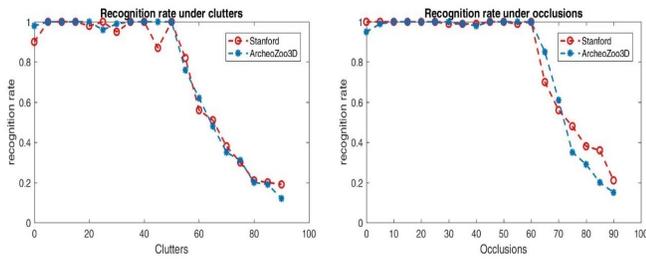


Fig. 15. Recognition rate under occlusions (left) and clutters (right) for both Stanford and Archeozoo3D.

for each test the occlusions and clutters given by the equations below.

$$\text{Occlusion} = 1 - \frac{\text{model surface match area}}{\text{total model surface area}} \quad (12)$$

$$\text{Clutter} = \frac{\text{clutter point in relevant volume}}{\text{total points in relevant volume}} \quad (13)$$

The surface area of a mesh is defined as the sum of the areas of its all faces. Clutter points are vertices in the scene surface mesh that does not belong to the model surface patch. Then, we repeat the same procedure for the thirteen scenes created from objects of Archeozoo3D database. Results for both databases are plotted bellow. See Fig. 15.

Examining the scatterplots in Fig. 15 we observe that the recognition rate is highly affected by occlusions. For both databases, from an amount of occlusions equal to 60%, the true positives rate starts to drop and in counterpart, at almost the same amount of occlusions true negatives and false positives increase. This is expressed by the failure of the algorithm to recognize object correctly in the scene for occlusions beyond 60%. In Fig. 15 (left), scatterplots show that clutters also influence the recognition rate. Up to 50% the algorithm still succeed at recognizing objects, but higher than this threshold, the recognition failures dominate. We assess the performance of our contribution in both Stanford and ArcheoZoo3D datasets by computing the precision and recall and comparing it to Spin Image algorithm and SHOT (Signature of Histograms of Orientations) [18]. In the table below we illustrate results obtained.

TABLE I. PERFORMANCE OF OUR CONTRIBUTION SRISI IN COMPARISON WITH SOME STATE-OF-THE-ART METHODS.

Dataset	Method	Precision	Recall	Inv. to Scale	Inv. to Resolution
ArcheoZoo3D	SI	0.00	0.00	NO	NO
	SRISI	0.61	0.51	YES	YES
	SHOT	0.50	0.49	NO	NO
Stanford 3D	SI	0.00	0.00	NO	NO
	SRISI	0.67	0.59	YES	YES
	SHOT	0.58	0.50	NO	NO

From results in Table I, we see that our contribution achieves good results in term of precision and recall for both datasets, while the original algorithm fails to recognize objects when we change resolution and scale. The changes of scale and resolution of objects results in changes of spatial

location of vertices and changes in the number of vertices, consequently, changes of normal vectors. As the creation of spin images is based on the projection of vertices on the surface mesh and also, this projection is controlled by angles between normal vector of the oriented point and normal vectors of other points, the accumulation of points in spin images becomes different between two same objects with different resolution and scale, resulting thus in a non-linearity between images so in difference intensity values. Knowing also that the correlation coefficient can only perform a good similarity between two spin images if only they have the same width and the transformation between them is linear. Defining efficiently good parameters for spin image generation, by setting the bin size of the scene to a multiple of the bin size of the model, then choosing to fix also the spin images width in order to represent shapes at the same scale level and finally bringing the intensity values to the same range by normalizing spin images before the matching algorithm, we have made the recognition of objects successful when resolution and scale change.

VI. CONCLUSION

In this paper we have introduced an approach robust to resolution and scale changes based on spin image algorithm. By understanding on the one hand that the issue of spin image algorithm was mainly related to the accumulation of points projected on the image, which leads to a non-linear transformation on the spin images to be compared and on the other hand the correlation coefficient will not properly calculate the similarity in this case, as well as the in-depth study of the influence of the parameters on the creation of spin images, we have succeeded to make the spin image descriptor robust to scale and resolution changes, with an occlusion rate up to 60% and 50% for clutters. We have integrated normalization into the matching pipeline and chose the right parameters by fixing the size of the spin images and setting the bin size of the scene into a multiple of the bin size of the model. Our future work aims to pursue the field of artificial intelligence. To this end we are interested to integrate descriptor in a neural network framework to automate and improve the recognition results under different conditions.

REFERENCES

- [1] J. H'roua et al., "3D objects descriptors methods: Overview and trends," in *Advanced Technologies for Signal and Image Processing (ATSIP)*, 2017 International Conference on, 2017, pp. 1–9.
- [2] P. Loncomilla, J. Ruiz-del-Solar and L. Martinez, "Object recognition using local invariant features for robotic applications: A survey," *Pattern Recognit.*, vol. 60, pp. 499–514, 2016.
- [3] R. Osada, T. Funkhouser, B. Chazelle and D. Dobkin, "Shape Distributions," vol. 21, no. 4, pp. 807–832, 2002.
- [4] E. Paquet, M. Rioux, O. Ontario and K. I. A. Or "Nefertiti : a Query by Content Software for Three-Dimensional Models Databases Management," pp. 345–352, 1997.
- [5] A. Nouri, C. Charrier and O. Lézoray, "Multi-scale saliency of 3D colored meshes," in *Image Processing (ICIP)*, 2015 IEEE International Conference on, 2015, pp. 2820–2824.
- [6] K. Tang, P. Song and X. Chen, "3D object recognition in cluttered scenes with robust shape description and correspondence selection," *IEEE Access*, vol. 5, pp. 1833–1845, 2017.
- [7] C. Maes, T. Fabry, J. Keustermans, D. Smeets, P. Suetens and D. Vandermeulen, "Feature detection on 3D face surfaces for pose normalization and recognition," in *Biometrics: Theory Applications and Systems (BTAS)*, 2010 Fourth IEEE International Conference on, 2010, pp. 1–6.

- [8] D. G. Lowe, "Object recognition from local scale-invariant features," in *Computer vision, 1999. The proceedings of the seventh IEEE international conference on*, 1999, vol. 2, pp. 1150–1157.
- [9] Y. Xiang, W. Choi, Y. Lin and S. Savarese, "Data-driven 3d voxel patterns for object category recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 1903–1911.
- [10] X. Li, I. Guskov and A. Arbor, "3D object recognition from range images using pyramid matching," pp. 1–6, 2007.
- [11] D. Pickup, X. Sun, P. L. Rosin and R. R. Martin, "Skeleton-based canonical forms for non-rigid 3D shape retrieval," *Comput. Vis. media*, vol. 2, no. 3, pp. 231–243, 2016.
- [12] O. K.-C. Au, C.-L. Tai, H.-K. Chu, D. Cohen-Or and T.-Y. Lee, "Skeleton extraction by mesh contraction," *ACM Trans. Graph.*, vol. 27, no. 3, p. 44, 2008.
- [13] W. Jin, C. Yan, L. Ma, H. Ye and H. Wang, "Joint extended fractional Fourier transform correlator," *Opt. Commun.*, vol. 268, no. 1, pp. 34–37, 2006.
- [14] J. Tierny, J. Vandeborre and M. Daoudi, "Analyse topologique et géométrique de maillages 3D pour l' extraction de squelette," pp. 1–8, 2006.
- [15] A. E. Johnson and M. Hebert, "Using spin images for efficient object recognition in cluttered 3D scenes," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 21, no. 5, pp. 433–449, 1999.
- [16] A. E. Johnson, "Spin-Images: A Representation for 3-D Surface Matching," *Carnegie Mellon University*, 1997.
- [17] F. Bernardini and H. Rushmeier, "for um The 3D Model Acquisition Pipeline," vol. 21, no. 2, pp. 149–172, 2002.
- [18] F. Tombari, S. Salti, L. Di Stefano, Unique signatures of histograms for local surface description, *Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*. 6313 LNCS (2010) 356 – 369 *doi* : 10.1007/978 – 3 – 642 – 15558 – 1₂₆.