

3D registration by textured spin-images

Nicola Brusco
brusco@dei.unipd.it

Marco Andreetto
cyb@dei.unipd.it

Andrea Giorgi
seleuda@dei.unipd.it

Guido M. Cortelazzo
corte@dei.unipd.it

*Dept. of Information Engineering
University of Padova*

Abstract

This work is motivated by the desire of exploiting for 3D registration purposes the photometric information current range cameras typically associate to range data. Automatic pairwise 3D registration procedures are two steps procedures with the first step performing an automatic crude estimate of the rigid motion parameters and the second step refining them by the ICP algorithm or some of its variations. Methods for efficiently implementing the first crude automatic estimate are still an open research area. Spin-images are a 3D matching technique very effective in this task. Since spin-images solely exploit geometry information it appears natural to extend their original definition to include texture information. Such an operation can clearly be made in many ways. This work introduces one particular extension of spin-images, called textured spin-images, and demonstrates its performance for 3D registration. It will be seen that textured spin-images enjoy remarkable properties since they can give rigid motion estimates more robust, more precise, more resilient to noise than standard spin-images at a lower computational cost.

1 Introduction

Establishing point correspondences between different 3D views of the same object is a well known problem, instrumental to various applications, among which 3D registration, object recognition and indexing of 3D models.

There exist a number of 3D point matching solutions which can be grouped in two fundamental approaches. One rests on the use of invariant statistics for each point of the mesh [4] [7] [10] [12] [26] [14] [16], the other on the use of special features based either on geometry [19] [24] [25] [8], or on texture [21] [23]. In principle the latter approaches can be more efficient, but less robust (for certain types of objects the wanted features may be missing or not easy to find). All methods, expectedly, have difficulties with highly symmetric objects or parts of objects.

Current range cameras typically deliver range data asso-

ciated with some kind of photometric information, either in terms of gray-level intensity or of (R,G,B) data. This makes it natural to devise 3D point matching techniques capable of exploiting texture and shape information.

Spin-images, introduced by Johnson and Hebert in [16] [14], are a 3D point matching instrument based on shape's characteristics only. This work proposes an extension capable to incorporate texture information in the concept of spin-images which we call "textured spin-images". For simplicity in the following we will also refer to texture spin-images as TSI and to standard spin-images as SI.

There can be many ways to include photometric information within spin-images. We experimented indeed various approaches. In spite of their a priori motivations, not all of them turned out effective. This work only reports about textured spin-images whose experimental performance turned out rather remarkable. We will see that textured spin-images are both more robust and computationally competitive with standard spin-images.

We will consider automatic pairwise registration of 3D views as the application in which to compare TSI versus standard SI. Standard spin-images are a tool of recognized effectiveness for a first automatic detection of the rigid motion between two partially overlapping 3D views (based on the recognition of the rigid motion affecting the surface's region that is common to the pair of 3D views [15] [3] [12]). Such a first rigid motion's estimate is necessary for making automatic 3D registration procedures based on ICP variations [5] [6] [27] [17] [18] [9] [22] or other methods [20]. It is also worth mentioning that the precision of the first automatic rigid motion estimate affects the result of refinement methods used in the second pass. ICP methods indeed need starting points rather close to the global rigid motion parameters in order not to get trapped by local minima. The starting point's position has also a direct impact on the convergence rate. It will be seen that TSI are not only more robust than SI, but also TSI lead to rigid motion's estimates more precise than those obtained by SI.

Therefore when texture is available together with range data, TSI can profitably use it.

This work has three other sections. Section 2 introduces

the concept of textured spin-images and their usage for 3D registration. Section 3 shows the experimental performance of textured spin-images in 3D registration analyzing the effects of various parameters choices. Section 4 draws the conclusions.

2 Textured spin-images and their application to 3D registration

2.1 3D registration by standard spin-images

The concept of spin-image is here quickly recalled for presentation's clarity. The reader is referred to [14] [16] for a detailed presentation. An oriented point at a surface mesh vertex is defined as the pair formed by the 3D vertex coordinates \mathbf{p} and the surface normal \mathbf{n} at the vertex. Consider a point \mathbf{x} on the 3D surface and define in the following way its coordinates pair (α, β) with respect to the reference system (\mathbf{p}, \mathbf{n}) associated to an oriented point: radial coordinate α is the perpendicular distance to the line through the surface normal at the vertex point, elevation coordinate β is the signed perpendicular distance to the tangent plane defined by vertex normal and position. The record of the (α, β) coordinates of all the points of the 3D mesh is called "spin-map". Formally the spin-map is defined as:

$$SI(\mathbf{x}) \rightarrow (\alpha, \beta) = \frac{1}{\sqrt{\|\mathbf{x} - \mathbf{p}\|^2 - (\mathbf{n} \cdot (\mathbf{x} - \mathbf{p}))^2}} \mathbf{n} \cdot (\mathbf{x} - \mathbf{p}), \quad (1)$$

A spin-image is a spatially discretized version of a spin-map, where the gray values are associated to the count of points of the spin-map falling in each discrete cell or bin.

Fig. 1a) shows a 3D view of an object, Fig. 1b) its spin-map with respect to the reference point marked on Fig. 1a) and Fig. 1c) shows its spin-image. The idea of accumulating the points of the spin-map into discrete bins is equivalent to a linear smoothing of the spin-map with an impulse response of value one over the bin and zero elsewhere. Indeed, since spin-images from different 3D views are used for mutual comparison, it is rather useful to remove fine detail and noise in general. For guidelines about mesh simplification and bin's dimensions see [14].

As Fig. 1 exemplifies, spin-images depend only from the intrinsic surface characteristics (or local statistical properties) and not from the surface's spatial position and orientation. In other words the spin-image associated with a vertex point is invariant with respect to rigid rototranslations. It is also very important to observe that the spin-images of neighbour points are very similar because the features of the local surface area near the point have a greater impact in the spin-image formation. Therefore, by way of spin-images, one may associate a collection of images to a 3D surface mesh, as every point of the surface can generate a

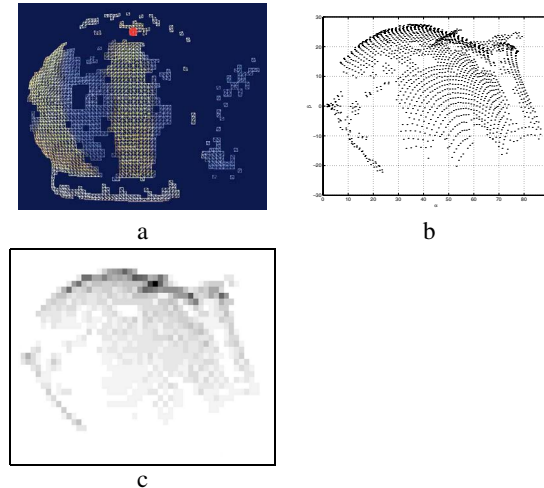


Figure 1. a) 3D View b) Spin-map c) Spin-image

spin-image. A pair of surfaces representing the same object from different view-points will be associated to a pair of sets of different spin-images: corresponding points in the common region between the two 3D views will have similar spin-images because, as already pointed out, spin-images strongly depend on local shape's characteristics. The spin-images of corresponding points of two partially overlapping 3D views will not be identical because of surface discretization effects and because the two patches only share a portion of their surface. However, if the overlap is substantial (no less than 30% of regions characterized by an adequate presence of geometrical features) corresponding points of the two 3D views located in the common region will have similar spin-images.

The detection of the common region between two partially overlapping 3D views, or, equivalently, the matching of a number of 3D points, in this way can be turned into the recognition of the most similar images of two sets of (spin) images, a problem for which a number of techniques is available.

For 3D registration purposes it is advisable to verify the correspondences against basic distance invariance constraints, i.e., if \mathbf{p}_1 and \mathbf{p}_2 are two points of the first mesh and \mathbf{q}_1 and \mathbf{q}_2 are their putative correspondences in the second mesh, they are retained only if $\|\mathbf{q}_1 - \mathbf{q}_2\| - \varepsilon < \|\mathbf{p}_1 - \mathbf{p}_2\| \leq \|\mathbf{q}_1 - \mathbf{q}_2\| + \varepsilon$ with a suitable ε . The correspondences can be organized within groups and for each group one can determine the rototranslations (\mathbf{R}, \mathbf{t}) moving the points of the first 3D view as close as possible to their corresponding points in the second 3D view, via Horn's algorithm [11]. The rototranslation (\mathbf{R}, \mathbf{t}) giving the widest overlap between the two views is selected as the correct registration. The final estimate of the rotation and translation

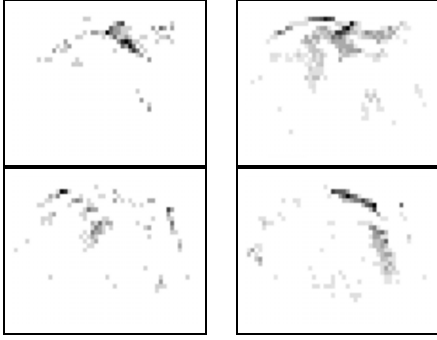


Figure 2. Textured Spin-Images with four levels, from the view of Figura 1a)

(R, t) between two 3D views is typically accomplished by the ICP algorithm [5] or some of its variants [6] [27] [17] [18] [9] [22] although other methods are available [20].

2.2 Textured Spin-Images

In order to account for texture information we extended Definition (1) in the following way. Let $I(\mathbf{x})$ indicate the luminance value associated to the (R, G, B) chromatic values of the surface texture at \mathbf{x} and let's quantize $I(\mathbf{x})$ into L uniform gray level intervals $l_i, i = 1, 2, \dots, L$. A textured spin-map with respect to an oriented point is the set of triplets determined by all the points \mathbf{x} of the surface and their intensity $I(\mathbf{x})$ in the following way:

$$TSI(\mathbf{x}) \rightarrow (\alpha, \beta, l_i) = (\sqrt{\|\mathbf{x} - \mathbf{p}\|^2 - (\mathbf{n} \cdot (\mathbf{x} - \mathbf{p}))^2}, \mathbf{n} \cdot (\mathbf{x} - \mathbf{p}), l_i), \quad (2)$$

A textured spin image is obtained from the textured spin map by accumulating in each bin not just the number of (α, β) coordinates falling in it as in the case of standard spin-images, but all the discrete gray level values l_i of the (α, β) coordinates falling in the bin. Hence a textured spin image is essentially the set of spin-images recording the (α, β) values associated to each l_i gray level interval, with an important difference with respect to the computation of spin-image: the bin value for each $i = 1, 2, \dots, L$ is not the count of points of the spin-map falling in the bin as for the SI, but the sum of all the luminance values l_i falling in the bin. In that way, even a single-level TSI ($L = 1$) takes into account texture's information.

Fig. 2 shows a TSI with $L=4$, obtained from the 3D view of Fig. 1 with respect to the reference point marked on Fig. 1a). In principle, the larger the number of intensity intervals L is, the greater the texture information recorded by the textured spin-image will be. In practice it will be shown that a sort of saturation takes place beyond $L=4$. With respect to Definition (2), standard spin-images can be consid-

ered as a special case of TSI with $L = 0$ (no luminance information).

It may be worth reporting that, among other extensions, we also experimented a variation of Definition (2) which we called (R, G, B) -spin-image, given by the set of 3 spin-images, the first one recording the (α, β) coordinates associated with the R values, the second the (α, β) coordinates associated with the G values, and the third one the (α, β) coordinates associated with the B values. Such an extension of the concept of spin-images proved to be more effective than SI, but not as effective as TSI.

The use of TSI for 3D registration purposes is completely similar to that of SI as indicated in the previous section.

3 Experimental Performance

In order to get a large number of partially superposed textured 3D views of objects we decided to use a library of digital 3D models [1] and to synthetically mimic actual object acquisition by way of range cameras. Since what one typically does when acquiring an object is to go around it and to cover it with views having at least 30% overlap with their neighbours, we decided to take an equatorial belt of equally spaced 3D views around the main axis of each 3D model. As the extent of mutual superposition between pairs of 3D views is a rather important factor for 3D registration, we chose two types of angular spacing, namely 36 and 72 degrees respectively. Spacing neighbour 3D views by 36 degrees obviously gives 10 3D views per model. The actual superposition between pairs of views depends on the shape of the 3D model, therefore some views turned out to have too little surface's superposition with their neighbours and were discarded as not suited to be registered even by supervised methods. Other views were discarded for the opposite reason, i.e., because their superposition with their neighbours was excessive. In this case 3D registration would not be a problem by any method, however excessive superposition would cause too many point correspondences which would just increase computation without adding any real information about the registration performance of TSI.

The first 400 3D models, forming the first four groups, of the Princeton data-base [1] (all the 3D models of such a library come with geometry information only) were textured by projecting on them the images of Fig. 3 painted on the interior of a cylinder centered around the principal axis of the 3D model. Such a set of 1200 textured 3D models was synthetically acquired in order to get the 3D views used in our tests.

The 3D views relative to the 36 degrees spacing should have given 4000 3D views per texture which become 3110 per texture after the above mentioned 3D view elimination. A first set of 2000 3D views per texture relative to 72 degrees spacing was obtained by taking the first 3D view, the third, the fifth and so on from the 3D views spaced 36



Figure 3. Images used in order to texture the 3D models of [1]

degrees apart; a second set of 2000 3D views per texture spaced 72 degrees apart were obtained by taking the second, the fourth, and so on. After the above mentioned 3D views elimination from the two sets we obtained 3164 3D views per texture spaced 72 degrees apart.

All the 3D views were randomly rotated. Such rotation matrixes, denoted by \mathbf{R} , served as ground truth.

The number of points on each 3D view, a parameter very important for its performance and computational implications in 3D registration, was determined by suitably setting the sensor's spatial resolution in our synthetic 3D acquisitions. The synthetic sensor's spatial resolution was chosen in order to obtain a version of approximately 4000 points for each 3D view. The actual number of points may vary from 3500 to 4500 from one 3D view to another according to the specific object's shape. We also obtained for each 3D view one version of about 2000 points and one of about 500 points, by setting different spatial resolutions.

The TSI were used in order to determine matching points from which to obtain an estimate $\hat{\mathbf{R}}$ of the ground truth rotation \mathbf{R} by the procedure described in Section 2. The estimate's quality can be measured in terms of matrix norms by defining the error $\epsilon' \doteq \|\mathbf{I} - \mathbf{R}\hat{\mathbf{R}}'^{-1}\| = \|\mathbf{I} - \mathbf{R}\hat{\mathbf{R}}'^T\|$, where \mathbf{I} is the 3×3 identity matrix and the matrix norm $\|\mathbf{M}\|$ is chosen as the largest singular value of \mathbf{M} [13].

Fig. 4a) shows an example of a typical histogram of ϵ' obtained by registering a set of 500 3D views taken 72 degrees apart by standard spin-images and Fig. 4b) an histogram of ϵ' obtained by registering them by TSI. It can be seen that both histograms are bimodal with a large peak around 0 (since in the large majority of the cases the estimated $\hat{\mathbf{R}}$ was close to the ground truth \mathbf{R}) and a small peak somehow above $\epsilon' = 1$. Similar histogram behaviours are also found in the registration by standard spin-images and TSI of 3D views taken 36 degrees apart (Fig. 4c) and Fig. 4d), with the only difference that in this case the peaks around 0 are higher. This is because larger overlaps produce better matches which in turn give a higher number of correct registrations. As a counterpart of such an increase of the peaks around zero, the peaks above 1 decrease down to really small values and these histograms essentially exhibit a single peak around 0.

In front of the behaviour of the registration quality measure ϵ' , we qualify as correct registrations the ones asso-

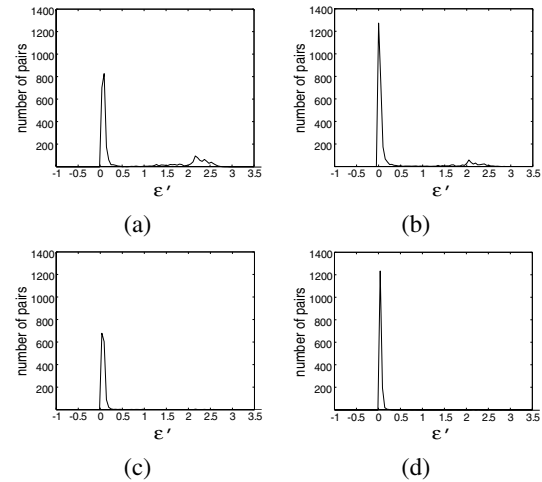


Figure 4. Histograms of ϵ' a) SI taken 72 degrees apart b) TSI with 4 intensity levels taken 72 degrees apart c) SI taken 36 degrees apart d) TSI with 4 intensity levels taken 36 degrees apart

ciated with $\epsilon' \leq 0.5$ and as wrong registrations the ones associated with $\epsilon' > 0.5$. In order to give a feeling for this choice, let's observe that a registration error of 15 degrees on the x axis, 15 degrees on the y axis, 15 degrees on the z axis, is associated with $\epsilon' = 0.45$.

It is also worth pointing out that the main peaks around 0 of the histograms relative to the use of TSI are always much higher than those relative to the use of standard spin-images. This first general indication that TSI can effectively exploit texture information will be further analyzed in the following sections.

3.1 5x5 textured spin-images

Spin-images and hence their extension TSI are characterized by several parameters among which the spin-image size plays a major role both for its feature matching and computational implications. Recommended sizes for spin-images typically range from 15×15 to 30×30 bins. Small size spin-images are powerful for recovering local shape description [15] as they focus on rather small regions. However as size decreases, spin-images become understandably prone to false correspondences, therefore one should avoid too small sizes. It will be seen instead that size 5×5 , truly not recommendable with standard spin-images, turns out rather effective with TSI, as Fig. 5 and Fig. 6 show.

Fig. 5 and Fig. 6 on the x axis report the number of quantization levels, with the convention that zero level corresponds to standard spin-images, i.e., to spin-images without texture. The y axis of these figures reports the percentage of correctly registered views. Fig. 5 refers to 3D views 72

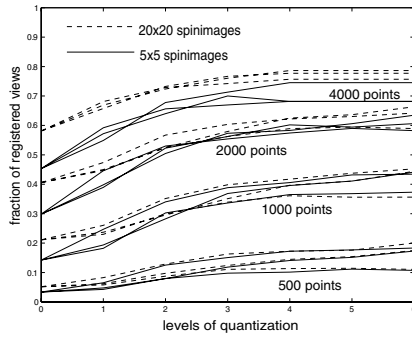


Figure 5. Results of TSI of size 5x5 and 20x20, spaced 72 degrees

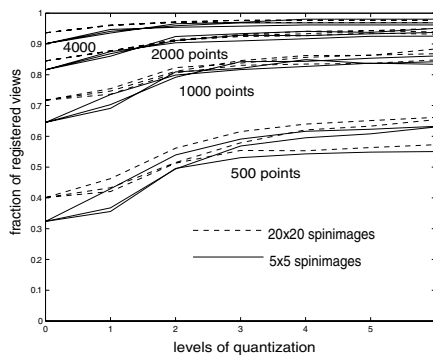


Figure 6. Results of TSI of size 5x5 and 20x20, spaced 36 degrees apart

degrees apart; Fig. 6 to 3D views 36 degrees apart, which are obviously easier to register. The impact of different parameters (kind of texture, distance between views, number of levels, number of points per 3D view, noise) with respect to registration quality will be next assessed for 5x5 TSI.

3.1.1 Texture's impact

Fig. 5 and Fig. 6 first of all show that the curves relative to different textures (baboon, lena and clown) are rather close. The specific texture's characteristics only lead to slight variations of the registration's quality. The overall shape of the plots is not altered by the specific texture. It is fair to say that the availability of texture's information makes a big difference which is definitely worth exploiting for registration purposes. In the following we will typically refer to the baboon's texture in the lack of different explicit indications.

3.1.2 Number of levels

The registration's performance increases from SI to TSI. With the baboon texture, 4000 points per view and 72 de-

grees distance, SI register 45% of the views, TSI 74% with four quantization levels. In the case of 36 degrees, SI register 90% and TSI 98% of the 3D views. In the case of 36 degrees, the improvement of TSI with respect to SI is less significant, since SI already work rather well. However, the ϵ' histogram associated with TSI have a larger peak near zero. This means that the registration is more precise with TSI than with SI. Such a precision will help subsequent fine registration steps. For instance the ICP algorithm will not be trapped by local minima only if started sufficiently close to the global minimum and the closer the starting point is to the global minimum, the fastest the convergence will be. Fig. 5 and Fig. 6 indicate that with TSI there is very little to gain by using a number of intensity levels L greater than 4. Sometimes the performance might even decrease as for the baboon's case. The plots similarly indicate that with $L = 1$ (which means that the intensity levels $I(\mathbf{x})$ are directly added into each bin) there is little improvement over standard spin-images.

3.1.3 Common region's impact

As expected, 3D scans with larger common regions (36 degrees apart) are easier to register. The 4000 points 3D views textured with the baboon's image, when texture is ignored, i.e., when 5x5 SI are used are correctly registered in the 90% of the cases, with 36 degrees distance, and in the 45% of the cases with 72 degrees distance. With four quantization levels, these 3D views are correctly registered in the 98% of the cases with 36 degrees distance and in the 74% of the cases with 72 degrees distance. This is an indication that texture is exploited by TSI when most needed, i.e., in order to match points in far apart views.

3.1.4 Number of points

The registration's performance improves as the average number of points per 3D view increases. However, beyond 4000 points the improvement rapidly tends to saturate (Fig. 8). Furthermore, since beyond 4000 points computation times become very large, registration becomes impractical. It is interesting to note from Fig. 5 and Fig. 6 that TSI with 2000 points per 3D view work as well as SI with 4000 points per 3D view. This is relevant in front of the computing times shown in Fig. 9 where it can be seen that TSI can achieve performance comparable with that of SI at lower computational cost. The computing times of Fig. 9 refer to an INTEL Xeon 3.2 GHz, 512 kb Cache L1, 2 Mb Cache L2, 1 Gb Memory.

3.1.5 Noise

Gaussian noise was also added to 3D data in order to simulate the effects of real noise in 3D acquisition. The noise of commercial high quality range cameras may have σ of approximately 0,05 mm with volumes of $10 \times 10 \times 10 \text{ cm}^3$.

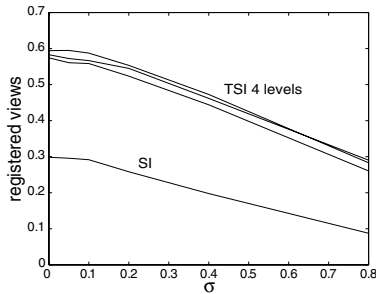


Figure 7. Registration performance of SI and 4 levels TSI in presence of noise (5x5, 2000 points, 72 degrees distance)

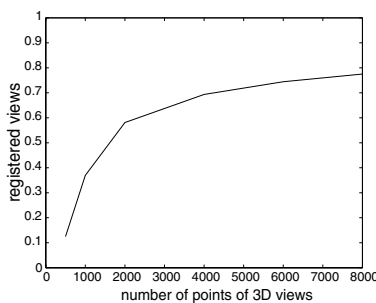


Figure 8. Registration performance of TSI versus average number of points per 3D view

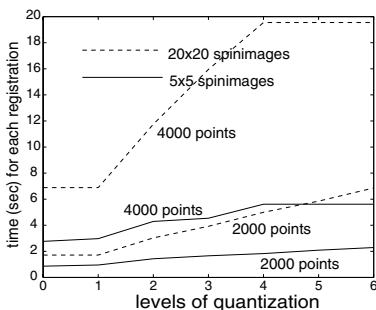


Figure 9. Computation time versus number of quantization levels for 5x5 and 20x20 SI and TSI

Hence in our noise experiments we first applied gaussian noise with σ equal to the 0,05% of the object's size. We then applied noise with progressively higher σ typical of low-cost range cameras and of inexpensive 3D acquisition procedures. Fig. 7 shows the percentage of correctly registered 3D views versus σ . Such a plot indicates that the registration performance is minimally affected by gaussian noise with σ below 0.1% the object's size and then it rapidly

decreases as σ increases beyond this value. It is interesting that in presence of noise (as it could be the case with a silhouette based 3D reconstruction's method) TSI's results tend to be equal to those of SI without noise. This indicates that TSI are rather robust feature matching tools, even in presence of significant noise. Noise was not added to texture, since in our real experiments we usually correct images to get colorimetric coherence [2].

3.2 Comparison with 20x20 textured spin-images

It is also worth examining the performance of 20x20 TSI, reported as dashed lines in the previous plots together with the performance of 5x5 TSI. Size 20x20 is rather common with SI [15]. With respect to 3D views made of 4000 points, with the baboon texture, direct comparison between 20x20 SI and TSI shows that the percentage of correct registration for SI is 65% and for 4 intensity levels TSI 75%, when the views are taken 72 degrees apart. With views taken 36 degrees apart the percentage of correct registration for SI is 96% and 98% for TSI. The percentage of successful registrations concerning 3D views made by smaller numbers of points is more favorable to TSI (in most cases the dashed and solid line plots of Fig. 5 and Fig. 6 overlap). However, the most interesting element exhibited by Fig. 5 and Fig. 6 is that the performance of 20x20 TSI although, as expected, superior to that of 5x5 TSI is not that superior. A size increase from 5x5 to 20x20 is rather relevant for SI, whose registration rate goes from 45% to 58% in the case of 3D views taken 72 degrees apart, but relatively relevant for TSI whose registration rate just goes from 74% to 76%.

This fact is most important in front of the computation requirements implied by size 5x5 and 20x20. Indeed, it can be seen from Fig. 9 that the computation required by 5x5 TSI with 4 intensity levels is inferior to that of 20x20 SI, but the registration rate of 5x5 TSI with 4000 points 3D views taken 72 degrees apart is 74% while that of 20x20 SI is 58%. It is therefore fair to draw the general conclusions that with textured 3D models TSI can achieve better results than SI with considerably less computation, and that 5x5 is a good size choice for TSI. Texture information is able to characterize rather well small surface patches just of size 5x5. Larger surface patches can of course add information, but not that much. At any rate the characteristics of a small 5x5 textured surface's patch which TSI can exploit appear superior to the pure geometric characteristics of larger patches, e.g., 20x20, which SI takes into account.

As we verified that TSI of size smaller than 5x5 do not give interesting results, we also checked TSI of size larger than 20x20. It was found that TSI of size 30x30 perform slightly worse than TSI of size 20x20 with considerable increase in the computation. One possible explanation is that the kind of "spinning" texture and geometry information gathered by TSI is rather powerful and precise with small patches where it can characterize very basic features, but

ϵ'				
Brain's model			Globe's model	
3D views	ϵ' of SI	ϵ' of TSI	ϵ' of SI	ϵ' of TSI
0-2	0.0236	0.0190	–	0.1191
1-3	0.0109	0.0249	–	–
2-4	0.1453	0.0324	–	0.0675
3-5	2.6443	0.0596	–	–
4-6	0.0764	0.0393	–	0.0454
5-7	1.4655	0.0193	–	–
6-8	2.0121	0.1282	–	0.1322
7-9	2.2910	0.0115	–	–
8-0	1.4972	0.0230	–	0.2101
9-1	2.4769	0.0302	–	–

Table 1. Registration of 3D brain's and globe's model: results

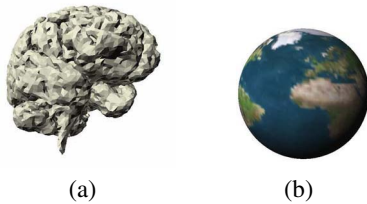


Figure 10. a) the brain model b) the globe

as the patch's size increases it loses recognition power because it ends up with mixing more and more basic features together.

3.3 Symmetrical objects

It is also worth considering the performance of TSI with symmetrical objects not manageable by SI, such as the globe's [23] and brain's 3D models shown in Fig. 10a) and of Fig. 10b). Indeed, since SI use geometric information only, they are not able to register portions of objects such as the globe (obviously all the points of a sphere have the same spin-image) or of the brain (the geometric pseudoperiodicity of which clearly creates matching problems). Instead TSI on the basis of the globe's texture were able to correctly register all the five views of the textured globe taken 72 degrees apart. It was also verified that the 10 3D views of a brain textured with image Lena taken 72 degrees apart were correctly registered by TSI. The relative ϵ' are shown in Table 1 and indicate a good match between the ground truth \mathbf{R} and its estimate $\hat{\mathbf{R}}$.

4 Conclusions

Spin-images are known to be an excellent 3D point matching and 3D registration tool. They were proposed a

few years ago when the availability of range cameras delivering photometric surface information was rather limited, differently than today. This work presents textured spin-images which are a way of including texture information within the concept of spin-images of proven practical effectiveness. It may be worth recalling that we tried and tested other ways of incorporating texture within spin-images but their performance was not as good as that of TSI.

The analysis of TSI's performance in 3D registration gives a number of interesting indications which is worth summarizing.

The texture's characteristics have a limited impact on TSI's performance. Of course, since uniform texture regions don't carry any texture information, high activity textures are expected to perform better. It appears that with TSI there is little to gain in using more than four gray levels for the texture. Computation is strongly impacted by the gray level number but four gray levels are still manageable for practical purposes. Other schemes, such as schemes directly using RGB values which seemed rather reasonable, were tried but their performance turned out consistently worse than that of four levels TSI.

Texture information makes remarkably effective 5x5 TSI. The possibility of resorting to TSI of such a limited size is very important as such TSI require less computation than SI of typical size in front of a performance both more robust and precise than them.

The TSI characteristics make them well suited to operate with reduced meshes, which is very important for practical applications.

TSI are remarkably robust even when range data are affected by large noise. Texture information makes TSI effectively cope with symmetrical and pseudo-periodic features.

How to best use geometry and texture information for 3D point matching is an open research question. One possible approach is to extend 3D point matching methods which work well with geometry information. This work follows this approach and shows what is at reach of one of such extensions based on spin-images.

References

- [1] <http://shape.cs.princeton.edu/benchmark/index.cgi>.
- [2] M. Andreetto, N. Brusco, and G. M. Cortelazzo. Color equalization of 3d textured surfaces. In *Proc. Eurographic Italian Chapter*, Milan, Italy, September 2002.
- [3] M. Andreetto, N. Brusco, and G. M Cortelazzo. Automatic 3D modeling of textured cultural heritage objects. *IEEE Transactions on Image Processing*, 13(3):354–369, March 2004.
- [4] A. Ashbrook, R. Fisher, C. Robertson, and N. Werghi. Finding surface correspondence for object recognition

- and registration using pairwise geometric histograms. In *Proc. European Conference on Computer Vision*, volume 1407, pages 674–786, 1998.
- [5] P. J. Besl and N. D. McKay. A method for registration of 3D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2):239–259, November 1992.
- [6] Y. Chen and G.G. Medioni. Object modeling by registration of multiple range images. *Image and Vision Computing*, 10(3):145–155, 1992.
- [7] C. Dorai, G. Wang, A. Jain, and C. Mercer. Registration and integration of multiple object views for 3D model construction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(1):83–89, November 1998.
- [8] L. Van Gool, D. Vandermeulen, G. Kalberer, T. Tuytelaars, and A. Zalesny. Modeling shapes and textures from images: new frontiers. In *Proc. of 1st International Symposium on 3D Data Processing Visualization and Transmission (3DPVT2002)*, pages 286–294, Padova, Italy, June 2002. IEEE Press.
- [9] M. Greenspan and G. Godin. A nearest neighbor method for efficient ICP. In *Proceedings of the Third International Conference on 3D Digital Imaging and Modeling*, pages 161–168, Quebec City, Canada, May 2001.
- [10] K. Higuchi, M. Hebert, and K. Ikeuchi. Building 3D models from unregistered range images. *Graphical Models and Image Proc.*, 57(4):315–333, 1995.
- [11] B. K. P. Horn. *Computer Vision*. McGraw-Hill, New York, N.Y., 1987.
- [12] D. Huber and M. Hebert. Fully automatic registration of multiple 3D data sets. *IEEE Workshop on Computer Vision Beyond the visible Spectrum*, pages 433–449, December 2001.
- [13] H. Hügli and C. Schultz. Geometric matching of 3d objects: Assessing the range of successful initial configurations. In *Proc. Int'l Conf. Recent Advances in 3D Digital Imaging and Modeling*, pages 101–106, May 1997.
- [14] A. E. Johnson. *Spin-Images: A Representation for 3-D Surface Matching*. PhD thesis, Carnegie Mellon University, Pittsburgh, August 1997.
- [15] A. E. Johnson. *Spin-Images: A Representation for 3-D Surface Matching*. PhD thesis, Carnegie Mellon University, Pittsburgh, August 1997.
- [16] A. E. Johnson and M. Hebert. Using Spin-Images for Efficient Multiple Model Recognition in Cluttered 3-D Scenes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(5):433–449, 1999.
- [17] T. Jost and H. Hügli. A Multi-Resolution Scheme ICP Algorithm for Fast Shape Registration. In *Proc. of 1st International Symposium on 3D Data Processing Visualization and Transmission (3DPVT2002)*, pages 540–543, Padova, Italy, 2002. IEEE Press.
- [18] C. Kapoutsis, C. Vavoulidis, and I. Pitas. Morphological iterative closest point algorithm. *IEEE Trans. on Image Processing*, 8(11):1644–1646, November 99.
- [19] P. Krsek, T. Pajdla, V. Hlavac, and R. Martin. Range image registration driven by hierarchy of surface differential features. In *Proc. 22nd Workshop of the Austrian Association for Pattern Recognition*, pages 175–183, May 1998.
- [20] L. Lucchese, G. Doretto, and G. M. Cortelazzo. A frequency domain technique for 3-d view registration. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 24(11):1468–1484, November 2002.
- [21] G. Roth. Registering two overlapping range images. In *Proceeding of the Second Intl. Conf. on 3D Digital Imaging and Modeling*, pages 191–200, Ottawa, Canada, October 1999.
- [22] S. Rusinkiewicz and M. Levoy. Efficient variants of the ICP algorithm. In *Proceedings of the Third International Conference on 3D Digital Imaging and Modeling*, pages 145–152, Quebec City, Canada, May 2001.
- [23] J. V. Wyngaerd and L. Van Gool. Combining texture and shape for automatic crude patch registration. In *Fourth International Conference on 3-D Digital Imaging and Modeling*, pages 179–186, October 2003.
- [24] J. Vanden Wyngaerd and L. Van Gool. Coarse registration of surface patches with local symmetries. In *Proc. European Conference on Computer Vision (ECCV'02)*, volume 2351, pages 572–586, May 2002.
- [25] J. Vanden Wyngaerd, L. Van Gool, R. Koch, and M. Proesmans. Invariant-based registration of surface patches. In *Proc. International Conference on Computer Vision (ICCV'99)*, pages 301–306, 1999.
- [26] D. Zhang and M. Hebert. Harmonic maps and their applications in surface matching. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR 99)*, pages 524–530, November 1999.
- [27] Z. Zhang. Iterative point matching for registration of free-form curves and surfaces. *International Journal of Computer Vision*, 13(2):119–152, 1994.