

# Loosely Distinctive Features for Robust Surface Alignment

Andrea Albarelli, Emanuele Rodolà, and Andrea Torsello

Dipartimento di Informatica - Università Ca' Foscari di Venezia

**Abstract.** Many successful feature detectors and descriptors exist for 2D intensity images. However, obtaining the same effectiveness in the domain of 3D objects has proven to be a more elusive goal. In fact, the smoothness often found in surfaces and the lack of texture information on the range images produced by conventional 3D scanners hinder both the localization of interesting points and the distinctiveness of their characterization in terms of descriptors. To overcome these limitations several approaches have been suggested, ranging from the simple enlargement of the area over which the descriptors are computed to the reliance on external texture information. In this paper we offer a change in perspective, where a game-theoretic matching technique that exploits global geometric consistency allows to obtain an extremely robust surface registration even when coupled with simple surface features exhibiting very low distinctiveness. In order to assess the performance of the whole approach we compare it with state-of-the-art alignment pipelines. Furthermore, we show that using the novel feature points with well-known alternative non-global matching techniques leads to poorer results.

## 1 Introduction

Feature detection and characterization is a key step in many tasks involving the recognition, registration or database search of 2D and 3D data. Specifically, when suitable interest points are available, all these problems can be tackled by working with the set of extracted features, rather than dealing with the less stable and noisier information carried by the whole data. Of course, for an interest point to be reliable it must exhibit two properties: repeatability and distinctiveness. A feature is highly repeatable if it can be detected with good positional accuracy over a wide range of noise levels and sampling conditions as well as different scales and transformations of the data itself. Further, description vectors calculated over interesting points are said to be distinctive if they are well apart when related to different features, yet coherent when associated to multiple instances of the same point. These properties are somewhat difficult to attain since they are subject to antithetical goals. In fact, to achieve good repeatability despite of noise, larger patches of data must be considered. Unfortunately this leads to a lower positional precision and a less sharp culling of uninteresting points. Moreover, for descriptor vectors to be distinctive among different features, they

need to adopt a large enough basis, which, owing to the well known “dimensionality curse”, also affects their coherence over perturbed versions of the same feature. In the last two decades these quandaries have been addressed with great success in the domain of 2D images where salient points are localized with sub-pixel accuracy by detectors exploiting strong local variation in intensity, such as Harris Operator [1] and Difference of Gaussians [2], or by using techniques that are able to locate affine invariant regions, such as Maximally stable extremal regions (MSER) [3] and Hessian-Affine [4]. Among the most used descriptors are the Scale-invariant feature transform (SIFT) [5], the Speeded Up Robust Features (SURF) [6] and Gradient Location and Orientation Histogram (GLOH) [7]. While these approaches work well with 2D intensity images, they cannot be easily extended to handle 3D surfaces since no intensity information is directly available. Of course several efforts have been made to use other local measures, such as curvature or normals. One of the first descriptor to capture the structural neighborhood of a surface point was described by Chua and Jarvis that with their Point Signatures [8] suggest both a rotation and translation invariant descriptor and a matching technique. Later, Johnson and Hebert introduced Spin Images [9], a rich characterization obtained by a binning of the radial and planar distances of the surface samples respectively from the feature point and from the plane fitting its neighborhood. Given their ability to perform well with both surface registration and object recognition, Spin Images have become one of the most used 3D descriptors. More recently, Pottmann et al. proposed the use of Integral Invariants [10], stable multi-scale geometric measures related to the curvature of the surface and the properties of its intersection with spheres centered on the feature point. Finally, Zaharescu et al. [11] presented a comprehensive approach for interest point detection (MeshDOG) and description (MeshHOG), based on the value of any scalar function defined over the surface (i.e. curvature or texture, if available). MeshDOG localizes feature points by searching for scale-space extrema over progressive Gaussian convolutions of the scalar function and thus by applying proper thresholding and corner detection. MeshHOG calculates a histogram descriptor by binning gradient vectors with respect to a rotational invariant local coordinate system.

In this paper we introduce a novel technique to detect and describe 3D interest points and to use them for robust surface registration. Unlike previous approaches we do not aim to obtain a very distinctive characterization. Instead, we settle for very simple descriptors, named *Surface Hashes*, that span only 3 to 5 dimensions. As their name suggests, we expect Surface Hashes to be repeatable through the same feature point, yet to suffer a high level of clashing due to their limited distinctiveness. In order to overcome this liability we avoid the use of classical RANSAC-based matchers; rather we adopt a robust game-theoretic inlier selector which exploits rigidity constraints among surfaces to guarantee a global geometric consistency. The combination of these loosely distinctive features and our robust matcher leads to an effective surface alignment approach. In the experimental section we point out this symbiosis by showing that standard matching techniques are not able to make the most of our descriptors.

## 2 Game-Theoretic Matching

Before describing in detail the Surface Hashes features we need to introduce some basic concepts about Evolutionary Game Theory and to present the idea of a Matching Game, originally presented in [12] and exploited by our technique both as an inlier selector and a robust matcher.

Evolutionary Game Theory [13] considers an idealized scenario where pairs of individuals are repeatedly drawn at random from a large population to play a two-player game. Each player obtains a payoff that depends only on the strategies played by him and its opponent. Players are not supposed to behave rationally, but rather they act according to a pre-programmed behavior, or mixed strategy. It is supposed that some selection process operates over time on the distribution of behaviors favoring players that receive larger payoffs. More formally, let  $S = \{1, \dots, n\}$  be the set of available strategies (*pure strategies* in the language of game theory) and  $C = (c_{ij})$  be a matrix specifying the payoff that an individual playing strategy  $i$  receives against someone playing strategy  $j$ . A *mixed strategy* is a probability distribution  $\mathbf{x} = (x_1, \dots, x_n)^T$  over the available strategies  $S$ ; being probability distributions, mixed strategies lie in the  $n$ -dimensional standard simplex  $\Delta^n = \{\mathbf{x} \in \mathbb{R}^n : \forall i \in 1 \dots n \ x_i \geq 0, \sum_{i=1}^n x_i = 1\}$ . The *support* of a mixed strategy  $\mathbf{x} \in \Delta$ , denoted by  $\sigma(\mathbf{x})$ , is defined as the set of elements chosen with non-zero probability:  $\sigma(\mathbf{x}) = \{i \in S \mid x_i > 0\}$ . The expected payoff received by a player choosing element  $i$  when playing against a player adopting a mixed strategy  $\mathbf{x}$  is  $(C\mathbf{x})_i = \sum_j c_{ij}x_j$ , hence the expected payoff received by adopting the mixed strategy  $\mathbf{y}$  against  $\mathbf{x}$  is  $\mathbf{y}^T C\mathbf{x}$ . The *best replies* against mixed strategy  $\mathbf{x}$  is the set of mixed strategies

$$\beta(\mathbf{x}) = \{\mathbf{y} \in \Delta \mid \mathbf{y}^T C\mathbf{x} = \max_{\mathbf{z}} (\mathbf{z}^T C\mathbf{x})\}.$$

A strategy  $\mathbf{x}$  is said to be a *Nash equilibrium* if it is the best reply to itself, i.e.,  $\forall \mathbf{y} \in \Delta, \mathbf{x}^T C\mathbf{x} \geq \mathbf{y}^T C\mathbf{x}$ . This implies that  $\forall i \in \sigma(\mathbf{x})$  we have  $(C\mathbf{x})_i = \mathbf{x}^T C\mathbf{x}$  that is, the payoff of every strategy in the support of  $\mathbf{x}$  is constant. A strategy  $\mathbf{x}$  is said to be an *evolutionary stable strategy* (ESS) if it is a Nash equilibrium and  $\forall \mathbf{y} \in \Delta \quad \mathbf{x}^T C\mathbf{x} = \mathbf{y}^T C\mathbf{x} \Rightarrow \mathbf{x}^T C\mathbf{y} > \mathbf{y}^T C\mathbf{y}$ . This condition guarantees that any deviation from the stable strategies does not pay. The search for a stable state is performed by simulating the evolution of a natural selection process. Under very loose conditions, any dynamics that respect the payoffs is guaranteed to converge to Nash equilibria [13] and (hopefully) to ESS's; for this reason, the choice of an actual selection process is not crucial and can be driven mostly by considerations of efficiency and simplicity. We chose to use the replicator dynamics, a well-known formalization of the selection process governed by the following equation

$$\mathbf{x}_i(t+1) = \mathbf{x}_i(t) \frac{(C\mathbf{x}(t))_i}{\mathbf{x}(t)^T C\mathbf{x}(t)}$$

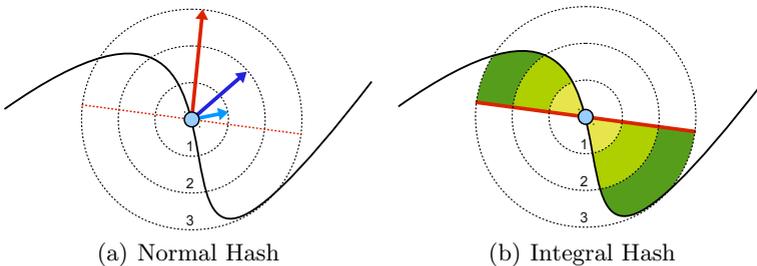
where  $\mathbf{x}_i$  is the  $i$ -th element of the population and  $C$  the payoff matrix.

Once the population has reached a local maximum, all the non-extinct pure strategies (i.e.  $\sigma(\mathbf{x})$ ) can be considered selected by the process.

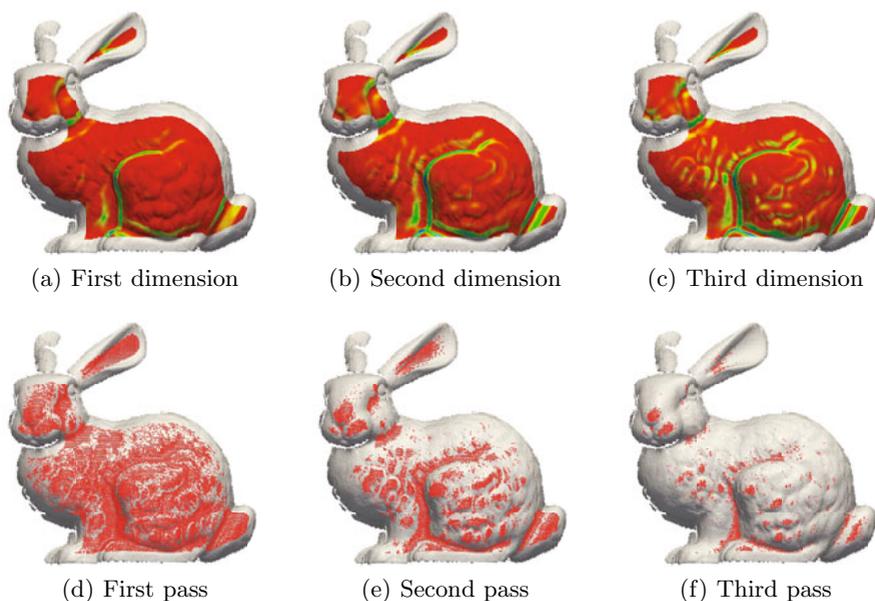
Following [12] and [14], we define a *Matching Game* as a non-cooperative game where the set of strategies  $S$  is a subset of all the possible correspondences, and the payoff  $c_{ij}$  between two strategies is proportional to some notion of compatibility between correspondences. By using different sets to be matched and alternative payoff functions, we are able to define games specially crafted to solve specific problems. In the following section we will define more formally two Matching Games. Respectively the first game will be dedicated to the localization of interest points over a surface described by Surface Hashes, while the second one will address the search for reliable correspondences between feature points extracted from two different meshes.

### 3 Surface Hashes

Intuitively, a Surface Hash is a concise point feature descriptor which exhibit the property of being highly repeatable at the cost of a relatively high probability of clashing. In practice this happens with any low-dimensional descriptor, such as the Gaussian or Mean Curvature (1 dimension), the first two Principal Components of a patch (2 dimensions), or the normal vector associated to a point (2 dimensions). While those descriptors could be used with our registration pipeline, we prefer to introduce some multiscale Surface Hashes based respectively on the dot product between normals and a local surface integral. Each of our descriptors corresponds to a vector of scalar measures evaluated at different scales. By increasing or reducing the number of scales, we are able to obtain vectors of different length, thus being more or less distinctive. The *Normal Hash* (see Fig. 1(a)) is obtained by setting as a reference the average surface normal over a patch that extends to the largest scale (red arrow in figure) and then, for each smaller scale, calculate the dot product between the reference and the average normal over the reduced patches (blue arrows in figure). This measure finds its rationale in the observation that at the largest scale the average normal is more stable with respect to noise and that the dot product offers a concise representation of the relation between the vectors obtained at various scales. The *Integral Hash* (see Fig. 1(b)) is similar in spirit to the Normal Hash. In this case we search for the best fitting plane (in the least squares sense) with respect



**Fig. 1.** Example of the two basic Surface Hashes proposed in this paper



**Fig. 2.** Example of a 3-dimensional Normal Hash and the related detection process

to the surface patch associated to the largest scale. We calculate the volume enclosed between the surface and such a plane. In practice, it is not necessary to evaluate this volume accurately: even naive approximations, such as the sum of the distances of the surface points from the plane, have shown to provide a reasonable approximation in all the empirical tests. Note that Normal Hashes evaluated over  $n$  scales yield descriptor vectors of length  $n - 1$  (since the larger scale is used only to calculate the reference normal), while Integral Hashes provide  $n$ -dimensional vectors. In Fig. 2 a Normal Hash of dimension 3 (respectively from (a) to (c)) evaluated over 4 scales is shown. Note that the descriptor is not defined on the points for which the larger support is not fully contained in the surface, i.e., points close to the surface boundary.

### 3.1 Interest Points Detection

Given the large number of points contained in typical 3D objects, it is not practical for any matching algorithm to deal with all of them. In addition, the isolation of a relatively small number of interest points can enhance dramatically the ability of the matcher to avoid false correspondences, usually due to a large number of features with very common characterizations. This is particularly true when using Surface Hashes, which are loosely distinctive by design. Paradoxically, we use exactly this property to screen out features exhibiting descriptors that are too common over the surface. This happens by defining a Matching Game

where the strategy set  $S$  corresponds to the set of all the surface points and the payoff matrix is defined by:

$$C(ij) = e^{-\alpha|d_i-d_j|}, \quad (1)$$

where  $d_i$  and  $d_j$  are the descriptor vectors associated to surface point  $i$  and  $j$ , and  $\alpha$  is a parameter that controls the level of selectivity. Clearly, features that are similar in terms of Surface Hashes will get a large mutual payoff and thus are more likely to be selected by the evolutive process. In this sense, our goal is to let the population evolve to an ESS and then remove from the set of interest points the features that survived the evolutive process. At the beginning we can initialize the set of retained features to the whole surface and run a sequence of Matching Games until the desired number of points are left. At this point, the remaining features are those characterized by less-common descriptors which are more likely to represent good cues for the matching. It should be noted that by choosing high values for  $\alpha$  the payoff function decreases more rapidly with the growth of the distance between the Surface Hashes, thus the Matching Game becomes more selective and less points survive after reaching an ESS. In the end this results in a blander decimation and thus in a larger ratio of retained interest points. By converse, a low value for  $\alpha$  leads to a more greedy filtering and thus to a more selective interest point detector. In Fig. 2 (from (d) to (f)) we show three steps of the evolutive interest point selection with respect to the 3-dimensional Normal Hash shown from (a) to (c). In Fig. 2(d) we see that after a single pass of the Matching Game most of the surface points are still considered interesting, while after respectively two and three passes only very distinctive points (belonging to areas with less common curvatures) are left.

### 3.2 Matching Surface Hashes

After obtaining a reduced set of interest points from the two surfaces, we could proceed to align them using some robust algorithm such as a basic RANSAC [15], that would use just the point locations and some initial match hypotheses, or PROSAC [16], that could better exploit the prior expressed by the descriptors. Unfortunately, Surface Hashes, despite the proposed filtering technique, are still not distinctive enough to be used directly by such methods. For this reason we define another Matching Game that ignores the information given by the descriptors and takes advantage of the rigidity constraint to be enforced in the surface registration problem. While this can sound counterintuitive, the main idea of this approach is to limit the use of the weak features to the selection of interest points and to use a more reliable global approach (that does not depend on descriptors) for the registration itself.

Given a set of model interest points  $M$  and a set of data interest points  $D$  we define the set of strategies for our Matching Game as all the possible correspondences between them:  $S = \{(a_1, a_2) | a_1 \in M \text{ and } a_2 \in D\}$ . Of course for practical reasons it is perfectly reasonable to limit the size of  $S$  by including only pairs that show similar descriptors.

Once  $S$  has been selected, our goal becomes to extract from it the largest subset that includes only correctly matched points: that is, strategies that associate a point in the model surface with the same point in the data surface. To enforce this we assign to each pair of strategies a payoff that is inversely proportional to a measure of violation of the rigidity constraint. This violation can be expressed in several ways, but since all the rigid transformations preserve Euclidean distances, we choose this property to express the coherence between strategies.

**Definition 1.** *Given a function  $\pi : S \times S \rightarrow \mathbb{R}^+$ , we call it a rigidity-enforcing payoff function if for any  $((a_1, a_2), (b_1, b_2))$  and  $((c_1, c_2), (d_1, d_2)) \in S \times S$  we have that  $\| |a_1 - b_1| - |a_2 - b_2| \| > \| |c_1 - d_1| - |c_2 - d_2| \|$  implies  $\pi((a_1, a_2), (b_1, b_2)) < \pi((c_1, c_2), (d_1, d_2))$ . In addition, if  $\pi((a_1, a_2), (b_1, b_2)) = \pi((b_1, b_2), (a_1, a_2))$ ,  $\pi$  is said to be symmetric.*

A rigidity-enforcing payoff function is a function that is monotonically decreasing with the absolute difference of the Euclidean distances between respectively the model and data points of the strategies compared. In other words, given two strategies, their payoff should be high if the distance between the model points is equal to the distance between the data points and it should decrease as the difference between such distances increases.

Further, if we want matching to be one-to-one, we must put an additional constraint on the payoffs, namely that mates sharing a point are incompatible.

**Definition 2.** *A rigidity-enforcing payoff function  $\pi$  is said to be one-to-one if  $a_1 = b_1$  or  $a_2 = b_2$  implies  $\pi((a_1, a_2), (b_1, b_2)) = 0$ .*

Given a set of strategies  $S$  and an enumeration  $O = \{1, \dots, |S|\}$  over it, a *mating game* is a non-cooperative game where the population is defined as a vector  $\mathbf{x} \in \Delta^{|S|}$  and the payoff matrix  $C = (c_{ij})$  is defined as  $c_{ij} = \pi(s_i, s_j)$ , where  $s_i, s_j \in S$  are enumerated by  $O$  and  $\pi$  is a symmetric one-to-one rigidity-enforcing payoff function. Intuitively,  $\mathbf{x}_i$  accounts for the percentage of the population that plays the  $i$ -th strategy. By using a symmetric one-to-one payoff function in a mating game we are guaranteed that ESS's will not include mates sharing either model or data nodes (see [12]). Moreover, a mating game exhibits some additional interesting properties.

**Theorem 1.** *Given a set of model points  $M$ , a set of data points  $D = TM$  that are exact rigid transformations of the points in  $M$ , and a set of strategies  $S \subseteq M \times D$  with  $(m, Tm) \in S$  for all  $m \in M$ , and a mating game over them with a payoff function  $\pi$ , the vector  $\hat{\mathbf{x}} \in \Delta^{|S|}$  defined as*

$$\hat{\mathbf{x}}_i = \begin{cases} 1/|M| & \text{if } s_i = (m, Tm) \text{ for some } m \in M; \\ 0 & \text{otherwise,} \end{cases}$$

*is an ESS and obtains the global maximum average payoff.*

This theorem states that when matching a surface with a rigidly transformed copy of itself the optimal solution (i.e., the population configuration that selects

all the strategies assigning each point to its copy) is the stable state of maximum payoff. Clearly, aligning a surface to an identical copy is not very useful in practical scenarios, where occlusion and measurement noise come into play. While the quality of the solution in presence of noise will be assessed experimentally, we can give some theoretical results regarding occlusions.

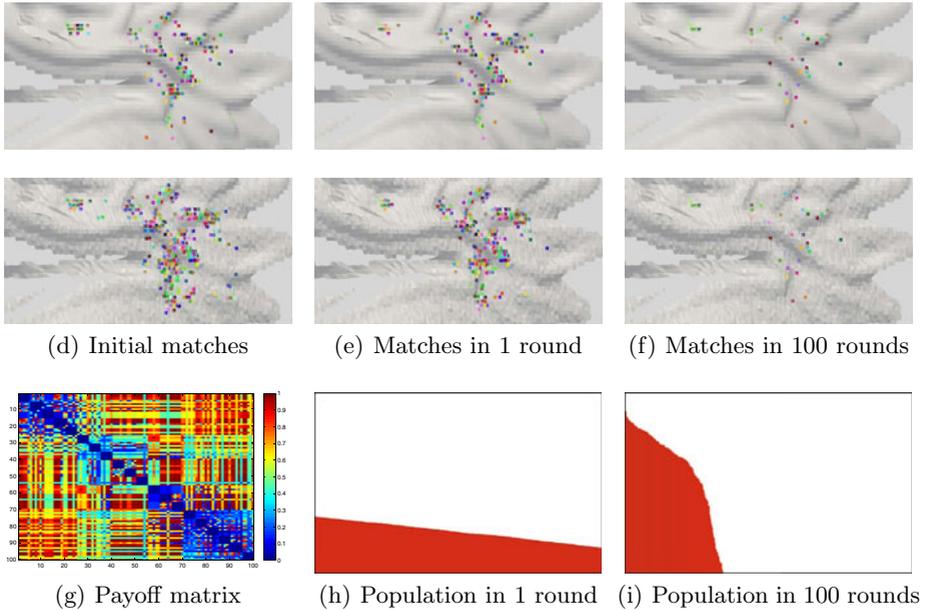
**Theorem 2.** *Let  $M$  be a set of points with  $M_a \subseteq M$  and  $D = TM_b$  a rigid transformation of  $M_b \subseteq M$  such that  $|M_a \cap M_b| \geq 3$ , and  $S \subseteq M_a \times D$  be a set of strategies over  $M_a$  and  $D$  with  $(m, Tm) \in S$  for all  $m \in M_a \cap M_b$ . Further, assume that the points that are not in the overlap, that is the points in  $E_a = M_a \setminus (M_a \cap M_b)$  and  $E_b = M_b \setminus (M_a \cap M_b)$ , are sufficiently far away such that for every  $s \in S, s = (m, Tm)$  with  $m \in M_a \cap M_b$  and every  $q \in S, q = (m_a, Tm_b)$  with  $m_a \in E_a$  and  $m_b \in E_b$ , we have  $\pi(q, s) < \frac{|M_a \cap M_b| - 1}{|M_a \cap M_b|}$ , then, the vector  $\hat{\mathbf{x}} \in \Delta^{|S|}$  defined as*

$$\hat{\mathbf{x}}_i = \begin{cases} 1/|M| & \text{if } s_i = (m, Tm) \text{ for some } m \in M_a \cap M_b; \\ 0 & \text{otherwise,} \end{cases}$$

is an ESS.

The result of theorem 2 is slightly weaker than theorem 1, as the face of the simplex corresponding to the “correct” overlap, while being an evolutionary stable state, is not guaranteed to obtain the overall highest average payoff. This is not a limitation of the framework as this weakening is actually due to the very nature of the alignment problem itself. The inability to guarantee the maximality of the average payoff is due to the fact that the original object ( $M$ ) could contain large areas outside the overlapping subset that are perfectly identical. Further, objects that are able to slide (for instance a plane or a sphere) could be allowed to move between different mixed strategies without penalty. These situations cannot be addressed by any algorithm without relying on supplementary information. However, in practice, they are quite unlikely extreme cases. In the experimental section we will show that our approach can effectively register a wide range of surface types.

In Fig. 3 we show a complete example of the evolutionary matching process. In order to make the example easy to understand we restricted our focus to a detail of a range scan of the Stanford “dragon”. In this example (and throughout all the experimental section)  $S$  is built by including all the strategy pairs composed by a feature point in the model and the 5 nearest feature points in the data in terms of Surface Hash (in this example we used an Integral Hash with 3 scales). In Fig. 3(g) we show, on a colored scale from 0 to 1, the payoff matrix of the rigid enforcing function used (which is discussed in the experimental section). Note that in the diagonal area of the matrix blocks of five strategies with reciprocal 0 payoff can be found: this is related to the way we built  $S$ . In fact we chose to include for each model point 5 candidates in the data and they are mutually non compatible as they share the same source point and we are looking for a one-to-one match. In the top and bottom half of Fig. 3(d) we can see respectively



**Fig. 3.** Example of a rigid enforcing payoff and of the evolution of the matching process

model and data feature points at the beginning of the matching process. After just one round of replicator dynamics we see that many outliers have been peeled off from the initial set  $S$ , but still some wrong matches are present. After 100 iterations only a few matches have been retained, but it is easy to see that they are extremely coherent. Finally, in Fig. 3(h) and Fig. 3(i) we show the (sorted) population histogram respectively after 1 and 100 iterations. The first histogram shows that all the strategies are still played by a sizeable amount of the population, while after 100 iterations most of the consensus is held by the few surviving matches.

## 4 Experimental Results

In this section we study the behavior of the proposed surface registration technique with respect to different Surface Hashes and scales. In addition we evaluate both the performance of the proposed feature descriptor with other matches and the quality of the alignment obtained by comparison with other pipelines. The rigidity-enforcing payoff function used throughout the experiments is defined as

$$\pi((a_1, b_1), (a_2, b_2)) = \frac{\min(|a_1 - a_2|, |b_1 - b_2|)}{\max(|a_1 - a_2|, |b_1 - b_2|)} \quad (2)$$

where  $a_1$ ,  $a_2$ ,  $b_1$  and  $b_2$  are respectively the two model (source) and data (destination) points in the compared mating strategies. The initial set of strategies

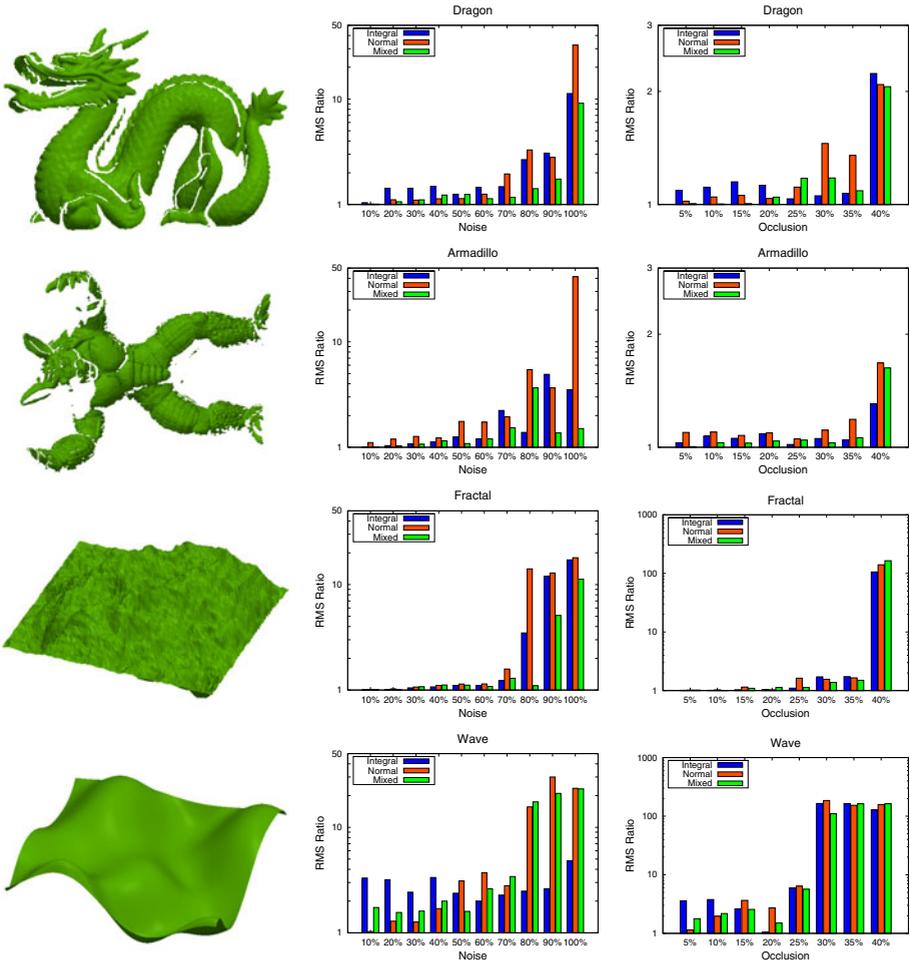
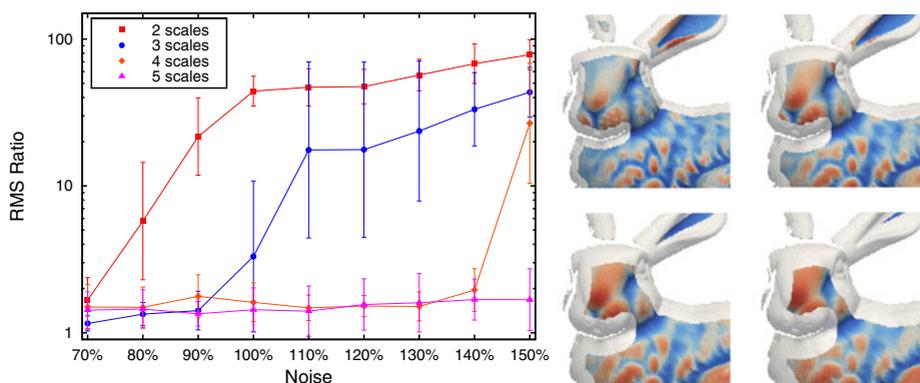


Fig. 4. Comparison of different descriptors using real and synthetic objects

$S$  was built by including all the pairs composed by a feature point in the model and the 5 feature points in the data with the nearest descriptor.

#### 4.1 Sensitivity to Noise, Occlusion, and Scale of the Descriptor

The performance of different descriptors was tested for various levels of noise and occlusion applied to two surfaces obtained from real range scans (“armadillo” and “dragon” from Stanford) and two synthetic surfaces designed to be challenging for coarse registration techniques (“fractal” and “wave”). The noise is a positional Gaussian perturbation on the point coordinates with its level ( $\sigma$ ) expressed in terms of the percentage of the average edge length, while the occlusion denotes the percentage of data and model surfaces removed. The RMS



**Fig. 5.** Effect of scale on the matching accuracy

Ratio in the charts is the ratio of the root mean square error (RMS) obtained after registration and the RMS of the ground truth alignment. The Normal and Integral Hashes were calculated over 3 levels of scale and the “Mixed” Hash is simply the juxtaposition of the previous two. In Fig. 4 we see that all the descriptors obtain good results with real ranges and the registration “breaks” only with very high levels of noise (on the same order of magnitude of the edge length). It is interesting to observe that the Mixed Hash always obtains the best performance, even with high level of noise: This higher robustness is probably due to the orthogonality between the Normal and Integral Hashes. The behavior with the “fractal” synthetic surface is quite similar, by contrast all the descriptors seem to perform less well with the “wave” surface. This is due to the lack of distinctive features on the model itself, which indeed represents a challenge for any feature based registration technique. The performance obtained with respect to occlusion is similar: all the descriptors achieve fairly good results and are resilient to high levels of occlusion (note that 40 percent occlusion is applied both to data and model). Overall the Mixed Hash appears to be consistently more robust. Since we found that the descriptors calculated over 3 levels of scale break at a certain level of noise, we were interested in evaluating if their performance can be improved by increasing their dimension. In Fig. 5 we present the results obtained with different levels of scale for the Mixed Hash. The graphs show the average over all the surfaces and the associated RMS. It is interesting to observe that by reducing the scale level the technique becomes less robust, whereas its performance increases dramatically when the number of scales increases. With a scale level of 5 our approach can deal even with surfaces subject to Gaussian positional noise of  $\sigma$  greater than the edge length. Unfortunately this enhanced reliability comes with a drawback: by using larger levels of scale the portion of boundary that cannot be characterized grows. In the right half of Fig. 5 the shrinking effect is shown for scale levels from 2 to 5.

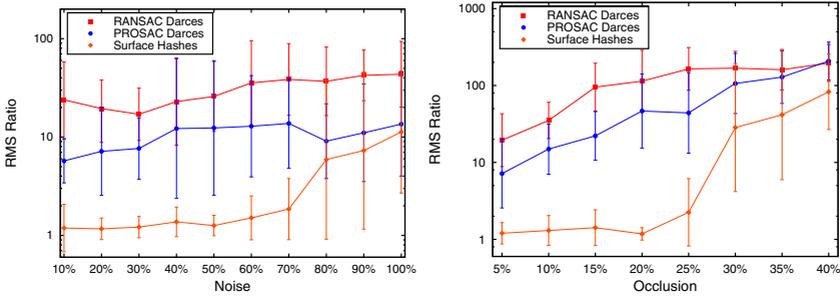


Fig. 6. Comparison of the performance obtained with different matchers

## 4.2 Comparisons with Other Matchers

Our goal in this set of experiments is to study if Surfaces Hashes can be used successfully with matchers alternative to the Matching Game described in Section 3.2. Specifically, we compared our full pipeline with standard DARCES [17] and with a DARCES variant that adopts PROSAC instead of plain RANSAC to take advantage of our descriptors as prior. To this end, we sorted the initial correspondence hypotheses by descriptor similarity and operated a PROSAC-like selection starting from an initial set of high-ranked matches and enlarging it progressively. In Fig. 6 we show the results of this test. As expected, RANSAC-based DARCES yields the worst results. Our PROSAC based variant obtains slightly better average registrations, but, the additional information provided by the descriptors is not distinctive enough to boost this technique to performance levels of the Matching Game that relies only on the global rigidity constraints.

## 4.3 System-Level Comparisons

Since our alignment approach does not need any initial estimate of the motion between surfaces, it can be classified as a coarse registration technique. For this reason we found appropriate to compare it with other widely used coarse registration methods. To this extent, we chose to use the Spin Images based approach proposed by Johnson [9] and the MeshDOG/MeshHOG combination suggested by Zaharescu [11]. The latter was selected because it adopts short descriptors very similar to the one proposed in this paper. In Fig. 7 we see that both techniques perform worse than the one based on Surface Hashes, even at low noise and occlusion levels. Surprisingly MeshDOG/MeshHOG obtains the worst results, probably because of the combination of a weak descriptor with a greedy matcher. Finally, we used the coarse registrations obtained with each approach to initialize a fine registration made with a best-of-breed ICP variant similar to the one proposed in [18]. Point selection is based on Normal Space Sampling [19], and point-surface normal shooting is adopted for finding correspondences, distant mates, candidates with back-facing normals, or matings established on the boundary of the mesh are rejected. In the leftmost plot of Fig. 8 we histogram the frequency of RMS ratio intervals obtained after the coarse registration. The histogram is based on bins of exponentially increasing

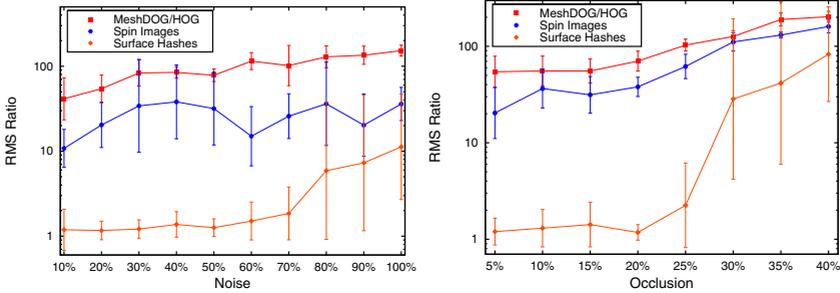


Fig. 7. Comparison of the performance obtained with different coarse techniques

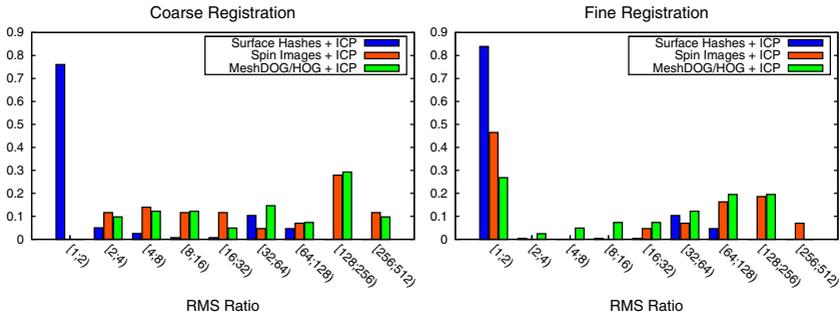


Fig. 8. Comparison of the performance between complete pipelines

size. In the rightmost chart the distribution change after a full round of ICP refinement can be seen. We can observe that while ICP is able to correct some wrong registrations with lower RMS Ratio, our approach still reaches the optimal alignment with a frequency that is almost double of the one obtained by the closest competitor. Regarding the computational complexity, it should be noted that the algorithm is quadratic in the number of strategies and thus the number of feature correspondences. Nevertheless, the initial interest points selection and the correspondences filtering by means of the descriptors, allow us to keep the computational time within a few seconds in all of our experiments.

## 5 Conclusions

In this paper we introduced a novel surface registration technique that uses very simple descriptors to create several weak correspondence hypotheses that are further optimized by a robust game-theoretic matcher. A theoretical result exposed the correspondence between optimal alignments and evolutionary equilibria, and the approach was validated on a wide range of experiments showing its greater robustness with respect to noise and occlusion in comparison with other well-known techniques.

## Acknowledgement

We acknowledge the financial support of the Future and Emerging Technology (FET) Programme within the Seventh Framework Programme for Research of the European Commission, under FET-Open project SIMBAD grant no. 213250.

## References

1. Harris, C., Stephens, M.: A combined corner and edge detector. In: Proc. Fourth Alvey Vision Conference, pp. 147–151 (1988)
2. Marr, D., Hildreth, E.: Theory of edge detection. Royal Soc. of London Proc. Series B 207, 187–217 (1980)
3. Matas, J., Chum, O., Urban, M., Pajdla, T.: Robust wide-baseline stereo from maximally stable extremal regions. *Image and Vision Computing* 22, 761–767 (2004); *British Machine Vision Computing 2002* (2002)
4. Mikolajczyk, K., Schmid, C.: An affine invariant interest point detector. In: Heyden, A., Sparr, G., Nielsen, M., Johansen, P. (eds.) *ECCV 2002, Part I. LNCS*, vol. 2350, pp. 128–142. Springer, Heidelberg (2002)
5. Lowe, D.: Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision* 20, 91–110 (2003)
6. Herbert Bay, T.T., Gool, L.V.: Surf: Speeded up robust features. In: Leonardis, A., Bischof, H., Pinz, A. (eds.) *ECCV 2006. LNCS*, vol. 3951, pp. 404–417. Springer, Heidelberg (2006)
7. Mikolajczyk, K., Schmid, C.: A performance evaluation of local descriptors. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 27, 1615–1630 (2005)
8. Chua, C.S., Jarvis, R.: Point signatures: A new representation for 3d object recognition. *International Journal of Computer Vision* 25, 63–85 (1997)
9. Johnson, A.E., Hebert, M.: Using spin images for efficient object recognition in cluttered 3d scenes. *IEEE Trans. Pattern Anal. Mach. Intell.* 21, 433–449 (1999)
10. Pottmann, H., Wallner, J., Huang, Q.X., Yang, Y.L.: Integral invariants for robust geometry processing. *Comput. Aided Geom. Des.* 26, 37–60 (2009)
11. Zaharescu, A., Boyer, E., Varanasi, K., Horaud, R.P.: Surface feature detection and description with applications to mesh matching. In: *CVPR* (2009)
12. Albarelli, A., Rota Bulò, S., Torsello, A., Pelillo, M.: Matching as a non-cooperative Game. In: *ICCV. IEEE Computer Society, Los Alamitos* (2009)
13. Weibull, J.: *Evolutionary Game Theory*. MIT Press, Cambridge (1995)
14. Albarelli, A., Rodolà, E., Torsello, A.: A game-theoretic approach to fine surface registration without initial motion estimation. In: *CVPR* (2010)
15. Fischler, M.A., Bolles, R.C.: Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Commun. ACM* 24, 381–395 (1981)
16. Chum, O., Matas, J.: Matching with prosac - progressive sample consensus. In: *CVPR*, Washington, DC, USA, pp. 220–226. IEEE Computer Society, Los Alamitos (2005)
17. Chen, C.S., Hung, Y.P., Cheng, J.B.: Ransac-based darces: A new approach to fast automatic registration of partially overlapping range images. *IEEE Trans. Pattern Anal. Mach. Intell.* 21, 1229–1234 (1999)
18. Turk, G., Levoy, M.: Zippered polygon meshes from range images. In: *SIGGRAPH 1994: Proc. of the 21st annual conference on Computer graphics and interactive techniques*, pp. 311–318. ACM, New York (1994)
19. Rusinkiewicz, S., Levoy, M.: Efficient variants of the icp algorithm. In: *Proceedings of the Third Intl. Conf. on 3D Digital Imaging and Modeling*, pp. 145–152 (2001)