Partial similarity of shapes using a statistical significance measure

Alexander M. Bronstein Michael M. Bronstein
Yair Carmon
Ron Kimmel
{bron@cs,mbron@cs,yairc@tx,ron@cs}.technion.ac.il

January 9, 2009

Abstract

Partial matching of geometric structures is important in computer vision, pattern recognition and shape analysis applications. The problem consists of matching similar parts of shapes that may be dissimilar as a whole. Recently, it was proposed to consider partial similarity as a multi-criterion optimization problem trying to simultaneously maximize the similarity and the significance of the matching parts. A major challenge in that framework is providing a quantitative measure of the significance of a part of an object.

Here, we define the significance of a part of a shape by its discriminative power with respect do a given shape database — that is, the uniqueness of the part. We define a point-wise significance density using a statistical weighting approach similar to the term frequency-inverse document frequency (tf-idf) weighting employed in search engines. The significance measure of a given part is obtained by integrating over this density.

Numerical experiments show that the proposed approach produces intuitive significant parts, and demonstrate an improvement in the performance of partial matching between shapes.

1 Introduction

Partial matching is associated with shape similarity for which dissimilar shapes may still consist similar parts. Problems that involve in partial matching often arise in computer vision, for example, when the acquired data to be matched is incomplete. Partial matching can be used to determine *partial similarity* between shapes, quantifying how different the shapes are, or *partial correspondence* determining the relation between the points on the shapes.

Two shapes are partially matching if they have significant similar parts. For example, a centaur and a horse have a similar part, the horse body, which makes them partially similar [15]. Thus, partial matching can be theoretically reduced to the problem of full shape matching by segmenting the shape into significant parts and trying to match these parts separately [24, 18, 19, 17, 3]. However, *significance* is often considered to be a semantic definition, and thus automatically finding such parts is not a well-defined. Many heuristic approaches have been proposed in the literature for shape decomposition (see e.g. [1]).

For two-dimensional shapes represented as curves, it is possible to employ technique from text matching. The boundary of a shape in polygonal representation can be thought of as an ordered sequence of characters (string) [11, 20]. Matching two boundaries boils down to matching two strings. Such approaches can handle partial similarity by finding similar sub-strings, e.g., using dynamic programming methods. However, string-based shape matching approaches rely on the fact that one can represent a shape as an ordered set (a string). While this is true for two-dimensional shapes (whose boundaries can be ordered), such a representation is impossible for more general, three-dimensional shapes.

Recently, Bronstein *et al.* proposed to consider partial similarity as a multi-criterion optimization problem trying to simultaneously maximize the similarity and the significance of the matching parts [7]. The solution of such a problem is a *Pareto optimum*, that is the set of all similarity-significance pairs for which there exist no two parts of the objects which are more similar and more significant at the same time.

The two major components in this framework are quantitative definitions of similarity and significance. Depending on the nature of the shapes and the class of transformations they may undergo, different similarity criteria should be employed. In particular, for partial matching of rigid shapes, a Hausdorff-type distance and a scheme similar to *iterative closest point* (ICP) methods [10, 2] can be used [5, 6]. For non-rigid shapes, one needs a deformation-invariant distance. Metric approaches considering the shapes as metric spaces with geodesic metrics and defining shape similarity as a distance between metric spaces [12, 21, 9, 8] were used in [7] as criteria for partial similarity of non-rigid shapes.

While one can employ standard and well-researched methods for formulating similarity criteria, a bigger challenge is finding a quantitative measure for the significance of a part of an object. The main problem stems from the fact that unlike similarity, which in many cases can have a strict definition, significance is rather a semantic notion and thus much more difficult to quantify. A naive approach proposed in [7] was to define significance as the area of the part. Intuitively, the larger the portion of the shape is, the more it is significant. Yet, it appears that such straightforward definition ignores the quality of the parts and in some cases tends to prefer multitude disconnected components with large total area over a single large part, often leading to meaningless results. As a remedy, claiming that "not only size matters," Bronstein [5, 6] proposed a regularization approach taking into consideration the boundary of the parts. The significance, according to this definition, is a combination of the part's area and the length of its boundary.

In this paper, we propose to define the significance of a part of the shape by its discriminative power with respect do a shape database — that is, the uniqueness of the part. We define point-wise *significance density* using a statistical weighting approach similar to the term frequency-inverse document frequency (tf-idf) weighting employed in search engines. The significance measure of a part is obtained by integrating over this density. This approach generalizes the significance definition used in [7].

The rest of this paper is organized as follows. In Section 2, we present the generic Paretian formulation of partial matching, and specifically address the cases of partial matching of rigid and nonrigid shapes. In Section 3, we define the significance distribution based on tf-idf weighting and in Section 4 describe its numerical computation. Section 5 is dedicated to implementation details of the algorithm. Section 6 shows experimental results. Finally, Section 7 concludes the paper.

2 Partial similarity

Let X and Y be two shapes we would like to compare. We say that X and Y are *partially matching* if there exist *parts* $X' \subseteq X$ and $Y' \subseteq Y$ which are *similar* and *significant*. In the following, for the sake of convenience, we will use the notions of *dissimilarity* and *insignificance* instead.

The degree of *dissimilarity* of parts X and Y can be expressed by a nonnegative function $d : \Sigma_X \times \Sigma_Y \to \mathbb{R}_+$ (here Σ_X and Σ_Y denote the collection of all the parts of the shapes X and Y, respectively). In the case of non-rigid shapes, the similarity criterion must be insensitive to non-rigid deformations that the shapes can undergo. It was argued in [12, 21, 9, 8] that deformation-invariant shape matching can be performed using the intrinsic geometric properties of the shapes. More formally, this approach models a shape X as a *metric space* (X, d_X), where $d_X(x, x')$ is the *geodesic metric*, measuring the length of the shortest path between the points x, x' on X. Comparison of two shapes (X, d_X) and (Y, d_Y) thus boils down to the comparison of the metrics d_X, d_Y, which can be quantified as

$$d(X,Y) = \operatorname*{argmin}_{\substack{\varphi:X\to Y\\ \psi:Y\to X}} \int_{X\times X} |d_X(x,x') - d_Y(\varphi(x),\varphi(x'))|^2 \, d\mu_X \times d\mu_X + \int_{Y\times Y} |d_Y(y,y') - d_X(\psi(y),\psi(y'))|^2 \, d\mu_Y \times d\mu_Y .$$

This similarity criterion is similar to the stress function used in multidimensional scaling (MDS) problems [4, 9, 8] and also be related to the Gromov-Hausdorff distance [14] between metric spaces (X, d_X) and (Y, d_Y).

As the measure of insignificance, Bronstein *et al.* [7] used the *partiality* function

$$p(X') = \text{area}(X) - \text{area}(X') = \int_{X\setminus X'} d\mu_X . \tag{1}$$

Here μ_X denotes the area measure on X. In this formulation, partial matching can be stated as the problem of simultaneous minimization of d and p over pairs of all the possible parts,

$$\min_{X',Y'} (d(X',Y'), p(X') + p(Y')). \tag{2}$$

A solution of the multi-criterion optimization problem (2) is the set of parts (X^*, Y^*) achieving an optimal tradeoff between dissimilarity and partiality, in the sense that there exists no other pair of parts (X', Y') with both $d(X', Y') < d(X^*, Y^*)$ and $p(X') + p(Y') < p(X^*) + p(Y^*)$. Such a solution is called *Pareto optimal* and is not unique; it can be visualized as a curve in the (d, p) plane (referred to as the *Pareto frontier*). The value of d at the point with partiality p_0 on the Pareto frontier can be computed as

$$\min_{X',Y'} d(X',Y') \ \text{ s.t. } \ p(X') + p(Y') \leq p_0, \tag{3}$$

or, using a Lagrange multiplier λ,

$$\min_{X',Y'} d(X',Y') + \lambda\left(p(X') + p(Y')\right). \tag{4}$$

Problem (3) can be interpreted as fixing some value of area and trying to distribute this area over parts of the shape in such a way that the dissimilarity is minimized. While Bronstein *et al.* [7] show that in practical cases this approach produces semantically correct results, there is no guarantee that the parts obtained by such a matching procedure are always meaningful. In many cases, the solution of problem (3) manifests a tendency of finding multiple disconnected parts. The quality of the part X' can be measured using some *irregularity* function $r(X')$ [5, 6]. We are thus looking for the largest, most similar and most regular parts, giving rise to the following multi-criterion optimization problem,

$$\min_{X',Y'} (d(X',Y'), p(X') + p(Y'), r(X') + r(Y')). \tag{5}$$

The new formulation can be regarded as a regularized version of problem (3), and the use of the regularity term will give preference to parts with larger regularity even at the cost of smaller area or larger dissimilarity. Alternatively, we can interpret the aggregate $p(X') + p(Y') + (r(X') + r(Y'))$ as a new definition of significance, taking into consideration not only the size but also the regularity of the parts and rewrite (6) as

$$\min_{X',Y'} (d(X',Y'), p(X') + p(Y') + r(X') + r(Y')). \tag{6}$$

2.1 Fuzzy formulation

Though the above formulation allows us to interpret the meaning of partial shape matching, minimization over all the parts of shape X is a computationally intractable combinatorial problem. To handle this difficulty, Bronstein *et al.* [7] replaced the subset $X' \subseteq X$ with a *membership function* (or *fuzzy part*) $u : X \to [0,1]$. This way, a *crisp* part is equivalent to $u^{-1}([\theta,1])$ for some threshold $\theta \in [0,1]$. We denote by M_X the set of all membership functions on X which are Σ_X-measurable.

In our multi-criterion problem, the optimization is now performed over fuzzy parts

$$\min_{u,v}(\tilde{d}(u,v), \tilde{p}(u) + \tilde{p}(v)), \tag{7}$$

where $\tilde{d} : M_X \times M_Y \to \mathbb{R}_+$ and $\tilde{p}$ denote the fuzzy versions of dissimilarity and insignificance, respectively. The fuzzy version of insignificance (8) is obtained straightforwardly as

$$\tilde{p}(u) \;\; = \;\; \int_X u(x) d\mu_X. \tag{8}$$

A fuzzy version of the intrinsic dissimilarity is obtained by weighting the stress by the membership functions,

$$\begin{aligned} \tilde{d}(u,v) = \min_{\substack{\varphi:X\to Y \\ \psi:Y\to X}} \quad & \int_{X\times X} \left|d_X(x,x') - d_Y(\varphi(x),\varphi(x'))\right|^2 u(x)u(x')d\mu_X \times d\mu_X \\ + \quad & \int_{Y\times Y} \left|d_Y(y,y') - d_X(\psi(y),\psi(y'))\right|^2 v(y)v(y')d\mu_Y \times d\mu_Y. \end{aligned}$$

Solution of problem (7) is carried out iteratively, by a two-step alternating optimization. In Step 1, optimization is performed over correspondences φ, ψ

$$\begin{aligned} (\varphi^*,\psi^*) = \operatorname*{argmin}_{\substack{\varphi:X\to Y \\ \psi:Y\to X}} \quad & \int_{X\times X} \left|d_X(x,x') - d_Y(\varphi(x),\varphi(x'))\right|^2 u^*(x)u^*(x')d\mu_X \times d\mu_X \\ + \quad & \int_{Y\times Y} \left|d_Y(y,y') - d_X(\psi(y),\psi(y'))\right|^2 v^*(y)v^*(y')d\mu_Y \times d\mu_Y. \end{aligned}$$

having parts u^*, v^* fixed. In Step 2, optimization is performed over the fuzzy parts u, v

$$\begin{aligned} (u^*,v^*) = \operatorname*{argmin}_{u,v} \quad & \int_{X\times X} \left|d_X(x,x') - d_Y(\varphi^*(x),\varphi^*(x'))\right|^2 u(x)u(x')d\mu_X \times d\mu_X \\ + \quad & \int_{Y\times Y} \left|d_Y(y,y') - d_X(\psi^*(y),\psi^*(y'))\right|^2 v(y)v(y')d\mu_Y \times d\mu_Y. \end{aligned}$$

with the correspondences φ^*, ψ^* from Step 1 fixed. The value of p_0 determines the required significance of the parts.

3 Statistical significance

Next, we would like to modify the definition of significance based on the relative importance of different parts. Intuitively, the most significant parts of a shape are those that distinguish the shape from the rest. In other words, significant parts discriminate between the shape (or shapes from a similar category) and the rest of the shapes. Obviously, this definition depends on the rest of the shapes. In a world dominated by horses, a human will stand out with distinctive parts such as hands, a head, and legs. While, in a world inhabited by humans, one would have to resort to more detailed features in order to distinguish between the different residents of our imaginary world.

This problem has an analogy to the problem of comparison of text documents. A standard technique used in search engines for detecting the similarity of two documents is comparing the histograms of the appearance of different words (*term frequency*, abbreviated as tf) in each document. On one hand, the more a word is repeated in a document (high tf), the more significant it is in the document. On the other hand, there are words that appear frequently in all documents – in English language, such words are articles "a" and "the", prepositions and connection words. Such words are said to have large *document frequency*, i.e., the appearance of a word in the entire database of documents.

What one would actually like to compare is the appearance of discriminative words, which are present in the current document and are absent in the rest of the documents. Intuitively, the more a word appears in a specific document, the more relevant it is, and the more it appears in other documents, the less relevant it becomes. For example, the term "partial" is significant in this document, while the word "the" is not, even though the latter appears more frequently. However, "the" appears frequently in many other documents as well, while "partial" probably not. Quantitatively, this can be measured by a ratio referred to as *term frequency-inverse document frequency*, or *tf-idf* for short [26]. More formally, given M documents and a vocabulary of T terms, let n_{ij} be the number of appearances of term i in document j. The frequency of term i in document j is defined as

$$\mathrm{tf}_{ij} = \frac{n_{ij}}{\sum_{t=1}^{T} n_{tj}} . \tag{9}$$

The *inverse document frequency* of term i in the database is defined as

$$\mathrm{idf}_i = \log\left(\frac{M}{\sum_{m=1}^{M} \mathbf{1}_{\{n_{im}>0\}}}\right) , \tag{10}$$

where $\mathbf{1}_{\{n_{im}>0\}}$ is an indicator equal to one if $n_{im} > 0$ and zero otherwise. An alternative definition, used in the case of a small database, is

$$\mathrm{idf}_i = \log\left(1 + \frac{M}{\sum_{m=1}^{M} \mathbf{1}_{\{n_{im}>0\}}}\right) . \tag{11}$$

This allows avoiding zero weights to infrequent terms. The significance, or weight, of term i in document j is given by

$$\mathrm{s}_{ij} = \mathrm{tf}_{ij}\mathrm{idf}_i \ , \tag{12}$$

known as *tf-idf weighting*, and is widely used and very successful in the field of information retrieval [27].

3.1 Bags of features

While a document is a collection of words, a shape can be thought of as a collection of parts. [1] Parts are represented by local descriptors, capturing the geometry of the part and playing the role of words in text document. A descriptor is a vector associated with a region around a point on the shape. Instead of a human language vocabulary from which words are drawn in text documents, shapes can be described in "geometric vocabulary", a collection of descriptors corresponding to representative parts of which any shape can be composed. For example, representative parts can be convex and concave patches of different curvature.

Given a shape, one can find how often different parts from the vocabulary appear in it, and compare two shapes the way search engines compare documents using tf-idf weighting. This approach is widely used in shape retrieval applications [22]. More formally, given a set of M shapes, for each point x on shape X_j, we can define a descriptor $D(x)$, as will be described in the next section. The descriptors are represented in a finite vocabulary consisting of T elements (representative descriptors), such that we can represent the descriptor $D(x)$ by an index $i(x) \in \{1, ..., T\}$. The equivalent of term frequency can then be defined as

$$\mathrm{tf}_{ij} = \frac{\int_{X_j} \mathbf{1}_{i(x)=i} d\mu_{X_j}}{\int_{X_j} d\mu_{X_j}} . \tag{13}$$

[1] Unlike text documents which are ordered collections of words, there is no order between parts on a surface. However, since we are interested in the frequency of words, the order is not important.

In a similar manner, the inverse document frequency of descriptor i is defined as

$$\mathrm{idf}_i = \log\left(\frac{M}{\sum_{m=1}^{M} \mathbf{1}_{\{\int_{X_m} \mathbf{1}_{i(x)=i} d\mu_{X_m} > 0\}}}\right). \tag{14}$$

and the tf-idf weight is given by $\mathrm{s}_{ij} = \mathrm{tf}_{ij}\mathrm{idf}_i$.

3.2 Significance density

Using the search engine analogy, assume that we would like to determine which portions of a document are important for its comparison to other documents in the database. A document j is an ordered collection of words, $(i_1, ..., i_{K_j})$, where $i_n \in \{1, ..., T\}$. We could use the tf-idf weight in order to determine the *significance density* in the document,

$$\sigma_n = \frac{s_{i_n,j}}{\sum_{n=1}^{K_j} s_{i_n,j}}. \tag{15}$$

Thus, portions of the document containing terms with higher tf-idf weight will have higher density. This means that if we were to choose a portion of the document, we would prefer one with the largest total significance density.

Similarly, for a shape we can define the significance density as,

$$\sigma(x) = \frac{s_{i(x),j}}{\int_X s_{i(x),j} d\mu_X}. \tag{16}$$

Figure 2 shows a few examples of significance density for different shapes. One can see, for example, that a point on the back of the horse is less significant than a point on its hoof for discriminating a horse from the rest of the shapes in the database.

Let us now return to the trivial definition of significance based on the area of the parts,

$$p(X') = \int_{X \setminus X'} d\mu_X. \tag{17}$$

The underlying assumption in this definition is that every point of the shape has equal importance. However, as we have seen from the previous discussion, different point have different significance density $\sigma(x)$. Therefore, we can define a significance function taking this fact into consideration,

$$p(X') = \int_{X \setminus X'} (\lambda + (1-\lambda)\sigma(x)) d\mu_X, \tag{18}$$

where $0 \leq \lambda \leq 1$. For $\lambda = 1$, we obtain the old definition of significance based on part area.

4 Feature descriptors

Computation of descriptors that act as shape "words" is key for creating significance densities. The descriptors should be sufficiently rich to discriminate between different parts of shapes, and at the same time be invariant to different transformations that a shape can undergo. Since we are particularly interested in non-rigid shapes, our descriptors must be invariant to non-rigid deformations. A way to achieve it is to base the descriptors on intrinsic geometric properties.

In this paper, our descriptors are computed as local distributions of geodesic distances. Similar approaches have been used for shape recognition [23]. Such descriptors are suitable for non-rigid shapes (since geodesic distances are insensitive to inelastic deformations) as well as for rigid shapes, which can be considered a particular case thereof. More formally, given a shape X, for every point $x \in X$ we define a ball $B_R(x)$ of radius R around x and consider the geodesic distances $\{d_X(x', x''),\ \ x', x'' \in X\}$. Let $f(d) : [0, 2R] \to [0, 1]$ be the distribution of these distances, and $\hat{f}(d) = f(d)/2R$ be the normalized distribution. We define the descriptor $D(x)$ as an L-dimensional vector

$$D(x) \;\; = \;\; (d_1, ..., d_L), \tag{19}$$

where

$$d_l \;\; = \;\; \int_{(l-1)/L}^{l/L} \hat{f}(\delta) d\delta, \tag{20}$$

is the normalized distribution, sampled at L points. In other words, $D(x)$ is an L-bin histogram of geodesic distances. An example is shown in Figure 1.

The choice of the radius R determines the size of the neighborhood and the locality of the descriptor. It is reasonable to assume that for a fixed shape, R should be the same for all points. Additionally, we would like R to capture the same amount of detail in every shape, in order to make the descriptors comparable between different shapes. If the details of different shapes in the database appear on different scales, R should be shape-dependent. The simplest solution is to scale R by the geodesic diameter of the shape, $\mathrm{diam}(X) = \max_{x,x' \in X} d_X(x, x')$. Thus, two similar shapes at different scales would have similar environments. In our experiments, such a choice was reasonable for most shapes.

As an example of the disadvantages of this method, imagine two shapes of dogs with bodies of the same size, one with a short tail and the other with a long one. While it would make sense to use same value R for both shapes, the long tail of the second dog will cause it to have a significantly larger diameter, and thus a significantly higher value of R. Finding a simple and robust method to estimate the correct R for every shape remains an open question.

4.1 Vocabulary construction

Once we computed the descriptors for all the shapes in the database, we need to construct a vocabulary containing representative descriptors that would allow us to describe parts of all the shapes in the database. A straightforward way to achieve this goal is using vector quantization of the set of all descriptors. This is a method widely used in creating vocabularies of visual words in computer vision applications [28].

Here, we use a variation of the K-means algorithm for vector quantization. We use the *Earth Mover's Distance* (EMD) between the distributions. The EMD can be intuitively thought of as the amount of work required to transform one distribution into another. The EMD has been shown to be a successful perceptual metric in many applications [25]. For one-dimensional distributions, EMD boils down to computing the L_1 distance between cumulative distributions.

Vector quantization clusters the descriptors into a fixed number T of regions that best represent the data. The centroids $\{v_1, ..., v_T\}$ of these regions serve as words or indices in our vocabulary. Given a descriptor $D(x)$, we define its index in the vocabulary as the closest centroid,

$$i(x) = \underset{k=1,...,T}{\operatorname{argmin}} \|v_k - D(x)\|. \tag{21}$$

5 Numeric computation

In the discrete setting, we represent the shapes as triangular meshes. We assume that shape X has M vertices $\{x_1, ..., x_M\}$, and shape Y has N vertices $\{y_1, ..., y_N\}$. The geodesic distances $d_X(x_i, x_j)$ and $d_Y(y_i, y_j)$ on the shapes are computed using the *fast marching method* [16]. We discretize the measure μ_X as an M-dimensional vector $\mathbf{a}_X = (a_1, ..., a_M)$, where a_i is $1/3$ of the sum of areas of the triangles at vertex x_i. The measure μ_Y is discretized in the same way.

The meshes are subsampled by selecting subsets $\{x_1, ..., x_m\}$ and $\{y_1, ... y_n\}$ of vertices, for example, using *farthest point sampling* [13]. The descriptors are computed at the vertices of the subsampled mesh, using the geodesic distances between all the vertices. To compensate for possible non-uniform sampling, the contribution of the distance $d_X(x_i, x_j)$ to the distribution is weighted by the product of the areas corresponding to those points, $a_i a_j$. The significance density of shape X is represented as an m-dimensional (n-dimensional for Y, respectively) vector denote by $\mathbf{s}_X$.

In discrete formulation, the correspondence $\varphi : X \to Y$ can be expressed by specifying m points $\{y'_1, ..., y'_m\}$ in Y that correspond to the points $\{x_1, ..., x_m\}$. The partiality function will be discretized as the vector $\mathbf{w} = (w_1, ..., w_m)^{\mathrm{T}}$. The components of $\mathbf{w}$ relate to the partiality of the sampled vertices $\{x_1, ..., x_m\}$ as well as their corresponding points $\{y'_1, ..., y'_m\}$. Thus, the discrete version of the optimization problem over correspondences (Step 1) reads,

$$\min_{\{y'_1,...,y'_m\}} \sum_{i,j=1}^{m} w_i w_j a_i a_j \left| d_X(x_i, x_j) - d_Y(y'_i, y'_j) \right|^2 . \tag{22}$$

where the optimization is preformed over all sets of m points on the mesh Y. While distance terms of the form $d_X(x_i, x_j)$ can be pre-computed using the fast marching algorithm [16], the terms $d_Y(y'_i, y'_j)$ have to be interpolated, since the $\{y'_1, ..., y'_m\}$ do not necessarily coincide with the vertices of Y. Problem (22) can be solved using the *generalized multidimensional scaling* (GMDS) algorithm [9].

Assuming a fixed correspondence, we now discretize the second part of the alternating minimization. Denote by $\mathbf{E}_\varphi$ the $m \times m$ local distortion matrix,

$$[\mathbf{E}_\varphi]_{ij} = \left| d_X(x_i, x_j) - d_Y(y'_i, y'_j) \right|^2 . \tag{23}$$

The dissimilarity term is the simple quadratic form,

$$\mathbf{w}^{\mathrm{T}} \mathbf{A} \mathbf{E}_\varphi \mathbf{A} \mathbf{w} , \tag{24}$$

where $\mathbf{A} = \mathrm{diag}(\mathbf{a})$. In order to discretize the significance terms, let $\mathbf{P}_X$ be an $M \times m$ matrix interpolating function on m sampled vertices to the entire mesh X. Similarly, define $N \times m$ matrix $\mathbf{P}_Y$, that will be constructed separately for every correspondence. Let $\mathbf{u} = \mathbf{P}_X \mathbf{w}$ and $\mathbf{v} = \mathbf{P}_Y \mathbf{w}$ denote the interpolated membership functions. In vector notation, the discretized significance term becomes $\mathbf{s}_X^{\mathrm{T}} \mathbf{A}_X \mathbf{u} + \mathbf{s}_Y^{\mathrm{T}} \mathbf{A}_Y \mathbf{v}$, where $\mathbf{A}_X = \mathrm{diag}(\mathbf{a}_X)$ and $\mathbf{A}_Y = \mathrm{diag}(\mathbf{a}_Y)$. Putting all

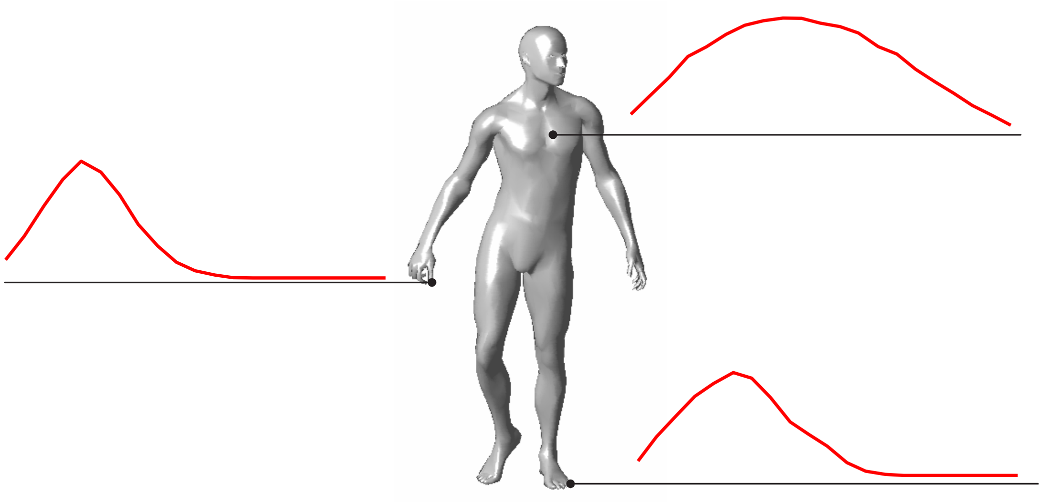

Figure 1: Examples of shape descriptors.

the terms together and substituting for $\mathbf{u}$ and $\mathbf{v}$, the second part of the alternating minimization algorithm reads,

$$\min_{0 \leq \mathbf{w} \leq 1} \quad \mathbf{w}^{\mathrm{T}} \mathbf{A} \mathbf{E}_{\varphi} \mathbf{A} \mathbf{w}$$
$$\text{s.t.} \quad \frac{1}{2} \left(\mathbf{s}_X^{\mathrm{T}} \mathbf{A}_X \mathbf{P}_X + \mathbf{s}_Y^{\mathrm{T}} \mathbf{A}_Y \mathbf{P}_Y \right) \mathbf{w} \geq p_0 \,. \tag{25}$$

6 Results

In the following experiment, we used shapes from the TOSCA dataset[2]. Each shape was sampled at 2500 points. As described in Section 4, descriptors were computed with $L = 21$ and $R = 10\%$ of the shape diameter (see Figure 1 for a few examples of descriptors). Geodesic distances were computed using fast marching [16]. Clustering was performed using the Statistical Learning Toolbox[3], with 1500 descriptors from every shape (at points chosen with farthest point sampling) to obtain a vocabulary of size $T = 300$.

Figure 2 shows examples of the significance density obtained from tf-idf weights. Distinctive features such as hand, paws and faces have high density. The resulting

[2] http://tosca.technion.ac.il
[3] http://web.mit.edu/dhlin/www/softwares/index.html

significance measures are generally consistent along various poses of the shapes.

Figure 3 shows a matching between the shape of a centaur and a human. Once can observe that a meaningful correspondence is obtained.

7 Conclusions

In this paper, we defined the significance of regions by their ability to distinguish the shape to which they belong from other shapes in a database. We implemented this definition using descriptors based on the intrinsic shape geometry, and the tf-idf ranking common in the field of data-retrieval. We showed how to integrate the resulting significance maps into existing partial matching frameworks in order to enhance their performance.

8 Acknowledgements

This research was supported by the Israel Science Foundation (grant No. 623/08).

References

[1] R. Basri, L. Costa, D. Geiger, and D. Jacobs. Determining the similarity of deformable shapes. *Vision Research*, 38:2365–2385, 1998.

[2] P. J. Besl and N. D. McKay. A method for registration of 3D shapes. *IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI)*, 14:239–256, 1992.

[3] O. Boiman and M. Irani. Similarity by composition. In *Int'l Conf. Neural Information Processing Systems (NIPS)*, 2006.

[4] I. Borg and P. Groenen. *Modern multidimensional scaling - theory and applications*. Springer, 1997.

[5] A. M. Bronstein and M. M. Bronstein. Not only size matters: regularized partial matching of nonrigid shapes. In *Proc. Nonrigid Shape Analysis and Deformable Image Registration (NORDIA)*, 2008.

[6] A. M. Bronstein and M. M. Bronstein. Regularized partial matching of rigid shapes. In *Proc. European Conf. Computer Vision (ECCV)*, pages 143–154, 2008.

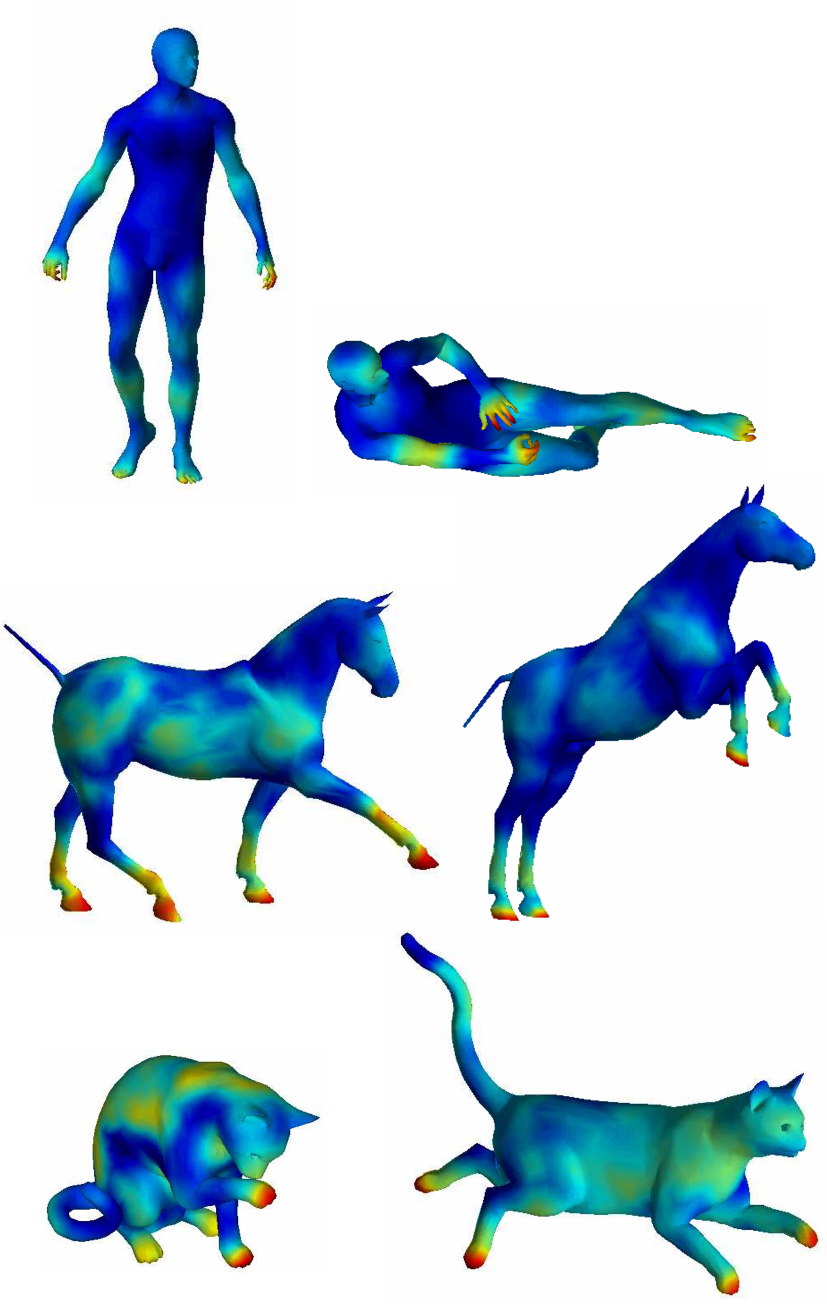

Figure 2: Six shapes from the database, colored by significance density.

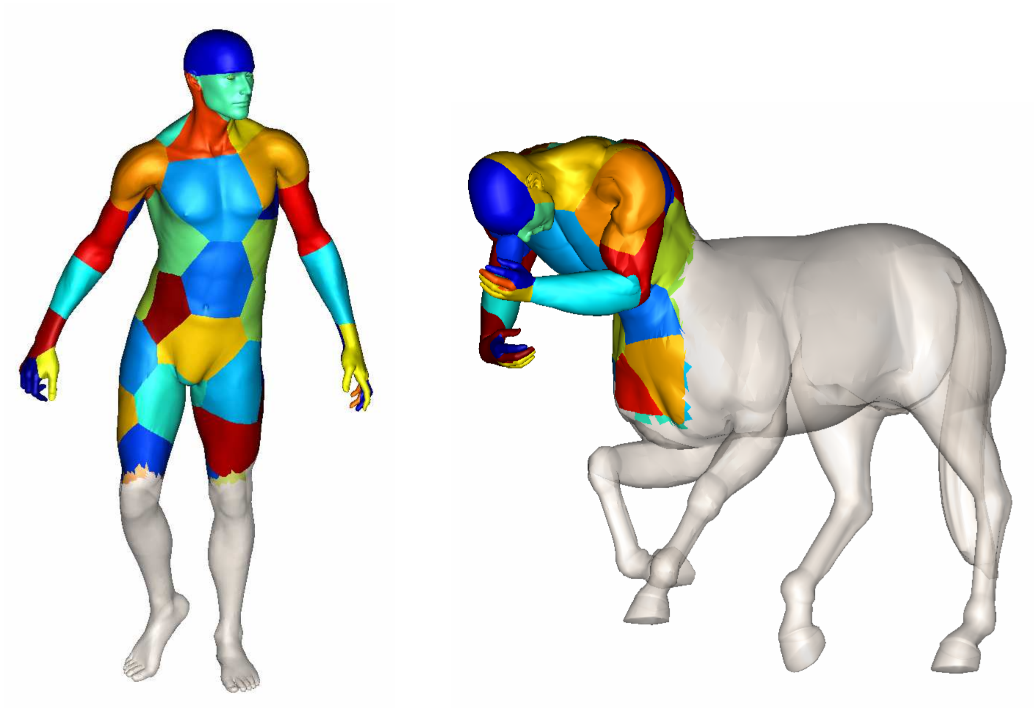

Figure 3: Matching between human and centaur shapes. Shown in different colors are corresponding Voronoi regions.

[7] A. M. Bronstein, M. M. Bronstein, A. M. Bruckstein, and R. Kimmel. Partial similarity of objects, or how to compare a centaur to a horse. *IJCV*, 2008. to appear.

[8] A. M. Bronstein, M. M. Bronstein, and R. Kimmel. Efficient computation of isometry-invariant distances between surfaces. *SIAM J. Scientific Computing*, 28(5):1812–1836, 2006.

[9] A. M. Bronstein, M. M. Bronstein, and R. Kimmel. Generalized multidimensional scaling: a framework for isometry-invariant partial surface matching. *Proc. National Academy of Science (PNAS)*, 103(5):1168–1172, 2006.

[10] Y. Chen and G. Medioni. Object modeling by registration of multiple range images. In *Proc. Conf. Robotics and Automation*, 1991.

[11] E. Arkin E, P. L. Chew, D. Huttenlocher, K. Kedem, and J. Mitchel. An efficiently computable metric for comparing polygonal shapes. *Trans. PAMI*, 13:209–216, 1991.

[12] A. Elad and R. Kimmel. On bending invariant signatures for surfaces. *IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI)*, 25(10):1285–1295, 2003.

[13] T. F. Gonzalez. Clustering to minimize the maximum intercluster distance. *Theoretical Computer Science*, 1985.

[14] M. Gromov. *Structures métriques pour les variétés riemanniennes*. Number 1 in Textes Mathématiques. 1981.

[15] D. Jacobs, D. Weinshall, and Y. Gdalyahu. Class representation and image retrieval with non-metric distances. *IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI)*, 22:583–600, 2000.

[16] R. Kimmel and J. A. Sethian. Computing geodesic paths on manifolds. *Proc. National Academy of Sciences (PNAS)*, 95(15):8431–8435, 1998.

[17] K. Kupeev and V. L. Brailovsky. A reinforced random algorithm for a partial contour perceptual similarity problem. *Pattern Recognition Letters*, 19:287–297, 1998.

[18] K. Kupeev and H. Wolfson. On shape similarity. In *Proc. Int'l Conf. Pattern Recognition (ICPR)*, pages 227–237, 1994.

[19] K. Kupeev and H. Wolfson. A new method of estimating shape similarity. *Pattern Recognition Letters*, 17:873–887, 1996.

[20] L. J. Latecki, R. Lakaemper, and D. Wolter. Optimal Partial Shape Similarity. *Image and Vision Computing*, 23:227–236, 2005.

[21] F. Mémoli and G. Sapiro. A theoretical and computational framework for isometry invariant recognition of point cloud data. *Foundations of Computational Mathematics*, 5:313–346, 2005.

[22] P. Min, M. Kazhdan, and T. Funkhouser. A comparison of text and shape matching for retrieval of online 3D models. Number 3232 in Lecture Notes in Computer Science. Springer.

[23] R. Osada, T. Funkhouser, B. Chazelle, and D. Dobkin. Shape distributions. *TOG*, 21(4):807–832, 2002.

[24] A. Pentland. Recognition by parts. In *Proc. Int'l Conf. Computer Vision (ICCV)*, pages 612–620, 1987.

[25] Y. Rubner, C. Tomasi, and L. J. Guibas. The earth mover's distance as a metric for image retrieval. *International Journal of Computer Vision*, 40(2):99–121, November 2000.

[26] G. Salton. *Automatic text processing: the transformation, analysis, and retrieval of information by computer*. Addison-Wesley, 1989.

[27] G. Salton and C. Buckley. Weighting approaches in automatic text retrieval. *Information Processing and Management*, 24(5):513–523, 1988.

[28] J. Sivic and A. Zisserman. Video Google: Efficient visual search of videos. In *Toward Category-Level Object Recognition*, volume 4170 of *Lecture Notes in Computer Science*, pages 127–144. Springer, 2006.