Shape Similarity Estimation using Ordinal Measures

Faouzi Alaya Cheikh, Bogdan Cramariuc, Mari Partio, Pasi Reijonen and Moncef Gabbouj

Tampere University of Technology,
P.O. Box 553, FIN-33101, Tampere, Finland

Abstract— In this paper we present a novel approach to shape similarity estimation based on ordinal correlation. The proposed method operates in three steps: object alignment, contour to multilevel image transformation and similarity evaluation. This approach is suitable for use in CBIR. The proposed technique produced encouraging results when applied on the MPEG-7 test data.

Index Terms— Content, Retrieval, Indexing, Image, Contour, Boundary, Shape, Ordinal, Correlation, Similarity, Measure.

I. INTRODUCTION

Generally shape representation can be based on its outer boundary or on the regions it contains. Characterizing the shape of an object by its boundary meets the way humans perceive objects. Since the human visual system itself concentrates on edges and ignores uniform regions [3]. This capability is hard-wired into our retinas. Connected directly to the rods and cones of the retina are two layers of the neurons that perform an operation similar to the Laplacian. This operation is called local inhibition and helps us to extract boundaries and edges [4].

Object shapes however, will have intrinsic intra-class variations. Moreover, object boundary deformation is expected in most imaging applications due to the varying imaging conditions, sensor noise, occlusion and imperfect segmentation.

Deformable models may be a promising approach for solving this problem due to their flexibility in object modeling. On the other hand, they are computationally very expensive to be used in a real time application, or even in a retrieval application where the user expects to have a response within few seconds after he puts his query. Hence these are not suitable for large databases where thousand of images are involved.

Therefore, simpler shape features have been used in several content-based indexing and retrieval (CBIR) systems, e.g. QBIC [10], MUVIS [1, 2]: high curvature points [5, 6, 7, 11], moments, morphological features (skeleton) and topological features.

This paper is introducing a novel boundary-based approach to shape similarity computation suitable for use in CBIR systems. The rest of the paper is organized as follows: Section 2 presents an overview of the proposed method, followed by a detailed description of each step. Assessments of experimental results using a subset of the MPEG-7 test data are presented in Section 3. In Section 4 conclusions are drawn.

II. THE PROPOSED METHOD

We are assuming in this paper that the shapes are already extracted from the gray level images and are stored in separate data files. The goal of this method is to compute similarity between any two shapes. The proposed method operates in three steps: alignment, boundary to multilevel image transformation and similarity evaluation.

To estimate the similarity between two objects shapes the boundaries are first aligned. The binary images containing the boundaries are then transformed into multilevel images; which are compared using the ordinal measure introduced in [9]. This ordinal measure estimates the similarity between the two shapes based on the correlation of their corresponding transform images. In the rest of this Section we give a detailed description of each one of the steps mentioned above.

A. Alignment

The alignment is performed by first detecting the major and minor axes of each shape, followed by reorientation of the shape in such a way that these axes are oriented in a standard way for all shape boundaries.

Once the major and minor axes are found, the boundary points {P1, P2, P3, and P4} that intersect with these axes are used to reorient/reposition the boundary as follows. The point that is closest to the center of mass among {P1, P2} is kept on the right. And the point among {P3, P4} closest to the major axe is kept on the top. We understand that this simple alignment process may not be enough in certain situations, but can be used to prove the validity of the proposed technique. In future work a more robust alignment algorithm may be used [12].

B. Boundary to multilevel image transformation

The shape is represented as a thin contour, C, in a binary image. This image is transformed into a multilevel (gray-scale) image G using a mapping function, such that the pixel values in G, $\{G_1, G_2, \ldots, G_n\}$, depend on the actual pixel position i and the positions of the contour pixels C_k for $k = 1, \cdots, p$, in the image G:

$$G_i = \phi(i, C_k : k = 1, ..., p), \text{ for } i = 1, ..., n. \quad (1)$$

Several transformations satisfy this requirement. For example any distance transformation or the transformations simulating the heat dissipation process.

As a result of this mapping the information contained in the shape boundary will be spread throughout all the pixels of the image. Computing the similarity in the transform domain will benefit of the rearrangement of the boundary information. We expect that there is no single optimal mapping; different mappings will emphasize different features of the contour. Which of these features is the most important is application and data dependent.

In this work we have implemented a mapping based on a simple geodesic metric. The metric is integer and its application is done through an iterative wave propagation process. The contour points are considered as seeds during the construction of the distance map. The distance map can be generated inside and/or outside the contour. The values can increase or decrease starting from the contour and can be limited.

Figure 2 presents an example of a distance map generated only inside the contour of a rat.

C. Similarity evaluation

The evaluation of image similarity is based on the framework for ordinal-based image correspondence introduced in [8]. Figure 3 gives a general overview of this region-based approach.

Suppose we have two images, X and Y, of equal size. In a practical setting, images are resized to a common size. Let $\{X_1, X_2, \ldots, X_n\}$ and $\{Y_1, Y_2, \ldots, Y_n\}$ be the pixels of image X and image Y, respectively. We select a number of areas $\{R_1, R_2, \ldots, R_m\}$ and extract the pixels from both images that belong to these areas. Consequently, R_j^X and R_j^Y contain the pixels from image X and Y, respectively, which belong to area R_j, with $j = 1, ..., m$.

The goal is to compare the two images using a region-based approach. To this end, we will be comparing R_j^X and R_j^Y, for each $j = 1, ..., m$. Thus, each block in one image is compared to the corresponding block in the other image in an ordinal fashion. Because our approach is an ordinal one only the ranks of the pixels are to be utilized. For every pixel X_k, we construct a so-called slice, which is defined as: $S_k^X = \{S_{k,l}^X : l = 1, ..., n\}$, where:

$$S_{k,l}^X = \begin{cases} 1, & \text{if } X_k < X_l \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

As can be seen, slice S_k^X corresponds to pixel X_k and is a binary image of size equal to image X. Slices are built in a similar manner for image Y as well.

With the goal of comparing regions R_j^X and R_j^Y, we first combine the slices from image X, corresponding to all the pixels belonging to region R_j^X. The slices are combined using the operation $OP_1(.)$ into a so-called metaslice M_j^X. More formally, $M_j^X = OP_1(\{S_k^X : X_k \in R_j^X\})$, $j = 1, ..., m$. Similarly, we combine the slices from image Y to form M_j^Y. It should be noted that the metaslices are equal in size to the original images and could be multi-valued, depending on the operation $OP_1(.)$. Each metaslice represents the relation between the region it corresponds to and the entire image.

The next step is a comparison between all pairs of metaslices M_j^X and M_j^Y by using operation $OP_2(.)$, resulting in the metadifference D_j. That is, $D_j = OP_2(M_j^X, M_j^Y)$, $j = 1, ..., m$. We thus construct a set of metadifferences $D = \{D_1, D_2, \ldots, D_m\}$. The final step is to extract a scalar measure of correspondence from set D, using operation $OP_3(.)$. In other words, $\lambda = OP_3(D)$. In [8] it was shown that this structure could be used to model the well-known Kendall's τ and Spearman's ρ measures.

The image similarity measure used in this paper is an instance of the previously mentioned framework. This measure has been analyzed more extensively in [9]. Following is a short description of the operations $OP_k(.)$ adopted for this measure. Operation $OP_1(.)$ is chosen to be the component-wise summation operation; that is, metaslice M_j is the summation of all slices corresponding to the pixels in block j or in other words, $M_j = \sum_{k: X_k \in R_j} S_k$.

Next, operation $OP_2(.)$ is chosen to be the squared Euclidean distance between corresponding metaslices. That is, $D_j = \left\| M_j^X - M_j^Y \right\|_2^2$.

Finally, operation $OP_3(.)$ sums together all metadifferences to produce $\lambda = \sum_j D_j$.

One advantage of this approach over classical ordinal correlation measures is its' capability to take into account differences between images at a scale related to the chosen block size.

III. Experimental Results

The experiments were conducted on two sets of 20 images each. The two sets are from the MPEG-7 CE Shape/Motion test set B, which contains (1400 images 20 in each category).

To assess the performance of our technique the two test sets where chosen such that each contains four categories of objects.

The difference between silhouettes from any two categories of the first test set is obvious; see Figure 4. In Figure 6, this can be clearly noticed in the high levels away from the diagonal. The intra-category discrepancy is rather small. Those plateaus on the diagonal have lower levels, meaning that the shapes within a given category are well grouped.

All the images of the second test set, are side shots of animals in similar positions; see Figure 5. Therefore, the dissimilarity between the shapes in this set is more difficult to estimate. Especially, when a given category presents a clear difference between its members. This is the case of the deer-1 and elephant-12, see Figure 7. This difference is reflected by the larger values of the dissimilarity scores presented in Table 2.

IV. Conclusions

The proposed technique produced encouraging results when applied on the MPEG-7 test data. These results have been obtained by using some intuitively selected parameters for the generation of the mapping and the similarity evaluation. Better results are expected if the parameters are optimized.

Due to the large number of parameters involved in the generation of the distance map and in the similarity computations the behavior of the proposed technique is quite complex. Therefore, further study is required for the optimization of the results and the understanding of the method's behavior.

V. References

[1] Alaya Cheikh F., Cramariuc B., Reynaud C., Quinghong M., Dragos-Adrian B., Hnich B., Gabbouj M., Kerminen P., Mäkinen T. and Jaakkola H., "MUVIS: A System for Content-Based Indexing and Retrieval in Large Image Databases," Proceedings of the SPIE/EI'99 Conference on Storage and Retrieval for Image and Video Databases VII, Vol. 3656, pp.98-106, San Jose, California, 26-29 January 1999.

[2] Trimeche M., Alaya Cheikh F., Gabbouj M. and Cramariuc B., "Content-based Description of Images for Retrieval in Large Databases: MUVIS", X European Signal Processing Conference, Eusipco-2000, Tampere, Finland, September 5-8, 2000.

[3] Hildreth C., "The Detection of intensity changes by computer and biological vision systems", Comput. Vis. Graphics Image Proc.Vol. 22, pp. 1-27, 1983.

[4] Russ J. C., "The Image Processing Handbook", 3rd edition, CRC, Springer and IEEE Press inc., 1995.

[5] Teh C.H. and Chin R. T., "On the detection of dominant points on digital curves", IEEE Trans. PAMI, vol. 11, pp. 859-872, 1989.

[6] Koch M. W. and Kashyap R.L., " Using polygon to recognize and locate partially occluded objects", IEEE Trans. PAMI, vol. 9, pp. 483-494, 1987.

[7] Abbasi S., Mokhtarian F., and Kittler J., "Curvature Scale Space image in Shape Similarity Retrieval," Springer Journal of Multimedia Systems, 1999.

[8] Shmulevich I., Cramariuc B. and Gabbouj M., "A framework for ordinal-based image correspondence," X European Signal Processing Conference (EUSIPCO-2000), 5-8 September 2000, Tampere, Finland, pp. 1389-1392.

[9] Cramariuc B., Shmulevich I., Gabbouj M. and Makela A., "A new image similarity measure based on ordinal correlation," IEEE International Conference on Image Processing, Vancouver, BC, Canada, September 10 - 13, 2000.

[10] Niblack W. et al, "The QBIC project; querying images by content using color, texture and shape", SPIE, Vol. 1908, 1993.

[11] Quddus A., Alaya Cheikh F. and Gabbouj M., "Wavelet-Based Multi-level Object Retrieval in Contour Images," Proceedings of the Very Low Bit rate Video Coding (VLBV'99) workshop, pp.43-46, Kyoto, Japan, October 29-30, 1999.

[12] Shen D., and Ip H.H.S., "Optimal Axes for Defining the Orientations of Shapes," Electronic Letters, Vol. 32, No. 20, pp. 1873-1874, September 1996.

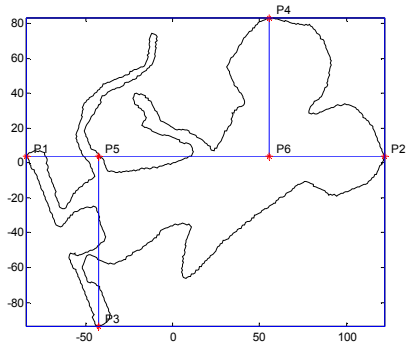

Figure 1. Boundary points used for the alignment.

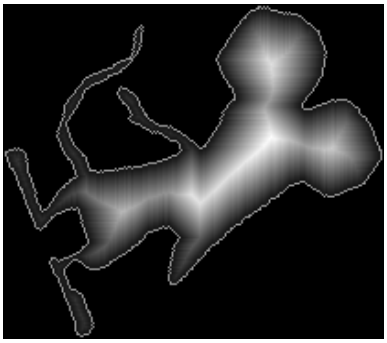

Figure 2. The object boundary overlaid on top of the distance map.

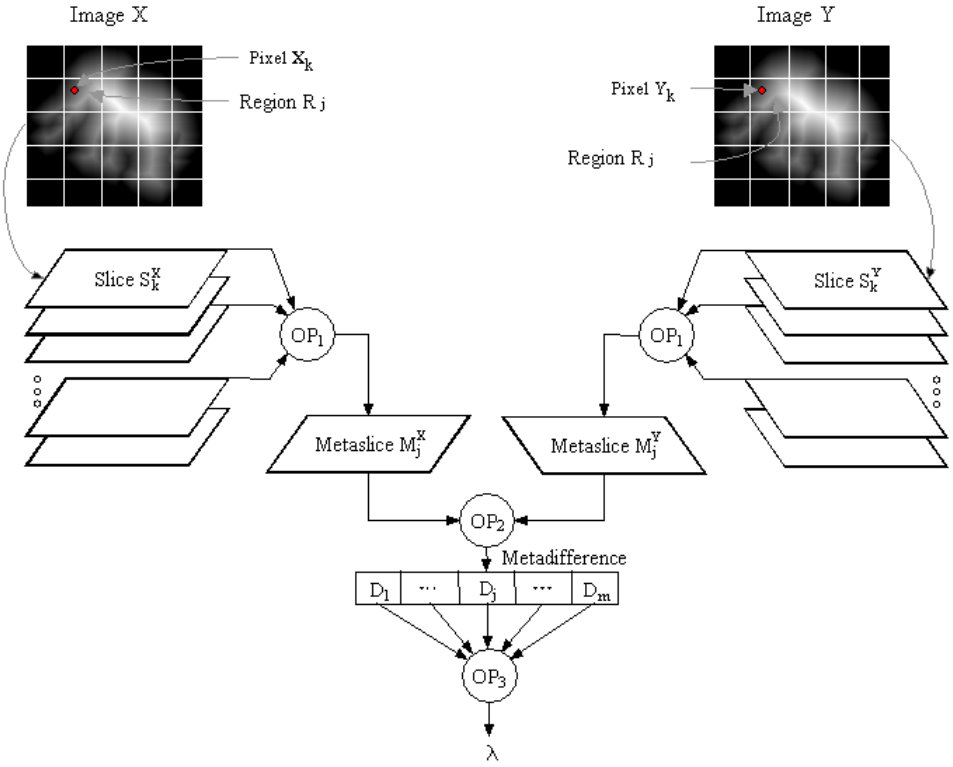

Figure 3. A general framework for ordinal-based image correspondence.

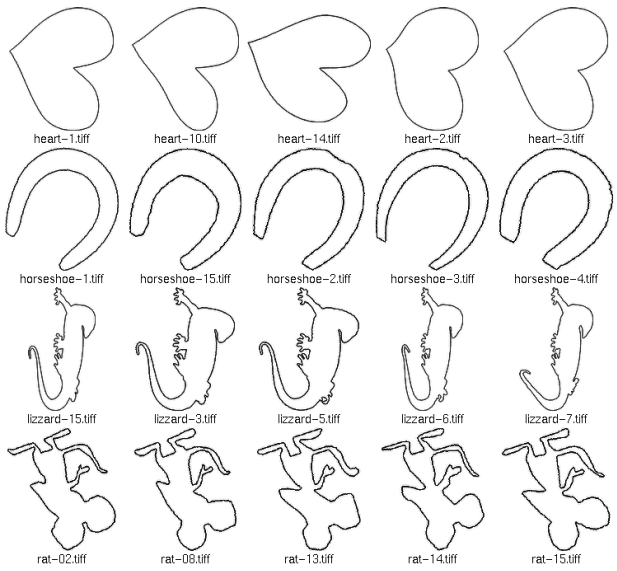

Figure4. Test set used in Experiment 1.

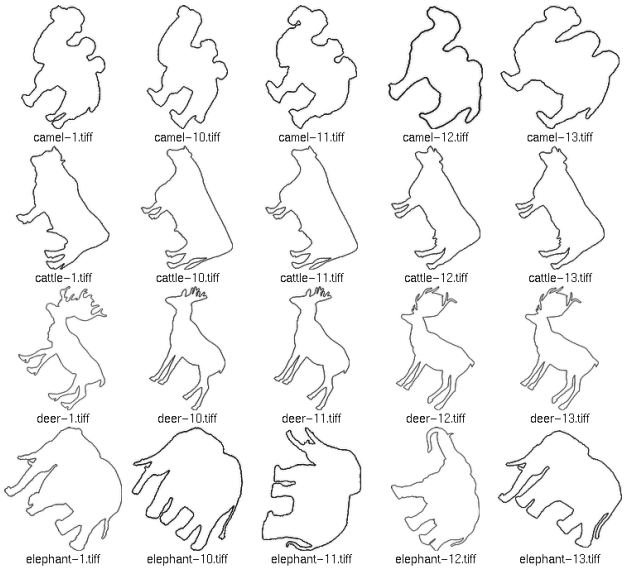

Figure 5. Test set used in Experiment 2.

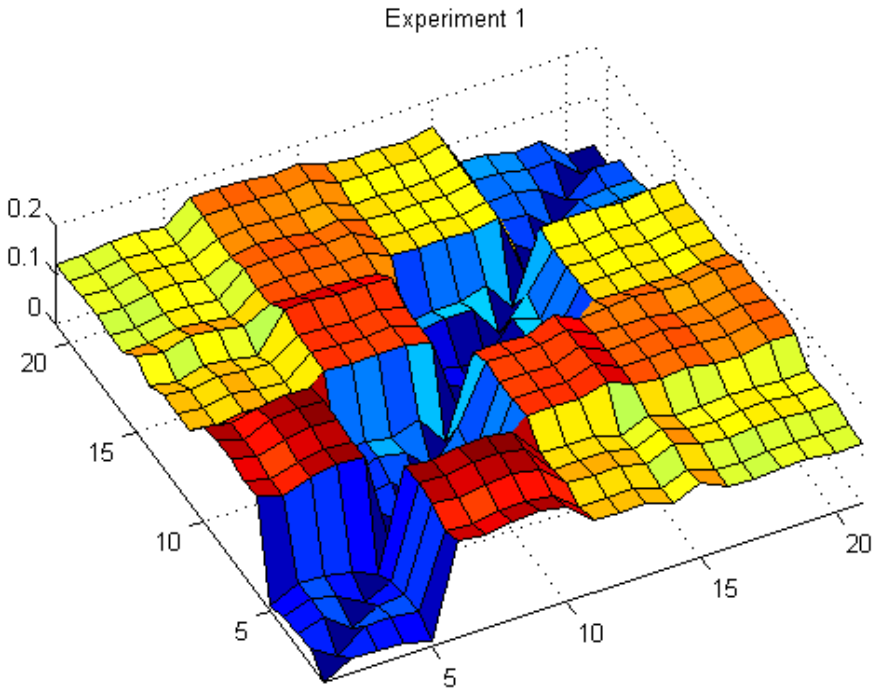

Figure 6. Surface plot of the scores obtained with the shape test set shown in **Figure 4**.

Table 1. Scores obtained for the test set number one.

Files	heart-1	heart-10	heart-14	heart-2	heart-3	horseshoe-1	horseshoe-15	horseshoe-2	horseshoe-3	horseshoe-4	lizzard-15	lizzard-3	lizzard-5	lizzard-6	lizzard-7	rat-02	rat-08	rat-13	rat-14	rat-15
heart-1	0	0.31	0.27	0.24	0.11	1.94	1.89	1.98	2.01	1.91	1.45	1.45	1.45	1.28	1.51	1.24	1.21	1.26	1.27	1.22
heart-10	0.31	0	0.42	0.42	0.35	1.78	1.73	1.83	1.83	1.79	1.38	1.39	1.39	1.19	1.4	1.23	1.19	1.23	1.24	1.22
heart-14	0.27	0.42	0	0.29	0.34	1.97	1.92	1.98	2.03	1.93	1.48	1.48	1.47	1.33	1.57	1.37	1.32	1.37	1.36	1.31
heart-2	0.24	0.42	0.29	0	0.24	2.05	1.99	2.07	2.1	2	1.4	1.4	1.4	1.26	1.5	1.31	1.29	1.35	1.34	1.3
heart-3	0.11	0.35	0.34	0.24	0	2.02	1.96	2.04	2.07	1.97	1.36	1.36	1.36	1.18	1.41	1.24	1.22	1.28	1.27	1.24
horseshoe-1	1.94	1.78	1.97	2.05	2.02	0	0.36	0.34	0.37	0.48	1.77	1.77	1.76	1.92	1.69	1.62	1.66	1.55	1.55	1.64
horseshoe-15	1.89	1.73	1.92	1.99	1.96	0.36	0	0.41	0.6	0.51	1.73	1.74	1.72	1.93	1.66	1.67	1.69	1.6	1.59	1.66
horseshoe-2	1.98	1.83	1.98	2.07	2.04	0.34	0.41	0	0.46	0.41	1.73	1.72	1.71	1.88	1.64	1.65	1.69	1.58	1.57	1.63
horseshoe-3	2.01	1.83	2.03	2.1	2.07	0.37	0.6	0.46	0	0.63	1.76	1.75	1.74	1.87	1.63	1.59	1.61	1.5	1.52	1.58
horseshoe-4	1.91	1.79	1.93	2	1.97	0.48	0.51	0.41	0.63	0	1.61	1.6	1.59	1.83	1.57	1.62	1.65	1.57	1.56	1.62
lizzard-15	1.45	1.38	1.48	1.4	1.36	1.77	1.73	1.73	1.76	1.61	0	0.2	0.19	0.56	0.37	1.39	1.36	1.39	1.34	1.42
lizzard-3	1.45	1.39	1.48	1.4	1.36	1.77	1.74	1.72	1.75	1.6	0.2	0	0.05	0.66	0.46	1.35	1.33	1.36	1.32	1.37
lizzard-5	1.45	1.39	1.47	1.4	1.36	1.76	1.72	1.71	1.74	1.59	0.19	0.05	0	0.66	0.47	1.36	1.34	1.36	1.31	1.38
lizzard-6	1.28	1.19	1.33	1.26	1.18	1.92	1.93	1.88	1.87	1.83	0.56	0.66	0.66	0	0.62	1.33	1.3	1.34	1.31	1.36
lizzard-7	1.51	1.4	1.57	1.5	1.41	1.69	1.66	1.64	1.63	1.57	0.37	0.46	0.47	0.62	0	1.42	1.41	1.42	1.38	1.47
rat-02	1.24	1.23	1.37	1.31	1.24	1.62	1.67	1.65	1.59	1.62	1.39	1.35	1.36	1.33	1.42	0	0.37	0.45	0.46	0.56
rat-08	1.21	1.19	1.32	1.29	1.22	1.66	1.69	1.69	1.61	1.65	1.36	1.33	1.34	1.3	1.41	0.37	0	0.44	0.51	0.54
rat-13	1.26	1.23	1.37	1.35	1.28	1.55	1.6	1.58	1.5	1.57	1.39	1.36	1.36	1.34	1.42	0.45	0.44	0	0.38	0.41
rat-14	1.27	1.24	1.36	1.34	1.27	1.55	1.59	1.57	1.52	1.56	1.34	1.32	1.31	1.31	1.38	0.46	0.51	0.38	0	0.45
rat-15	1.22	1.22	1.31	1.3	1.24	1.64	1.66	1.63	1.58	1.62	1.42	1.37	1.38	1.36	1.47	0.56	0.54	0.41	0.45	0

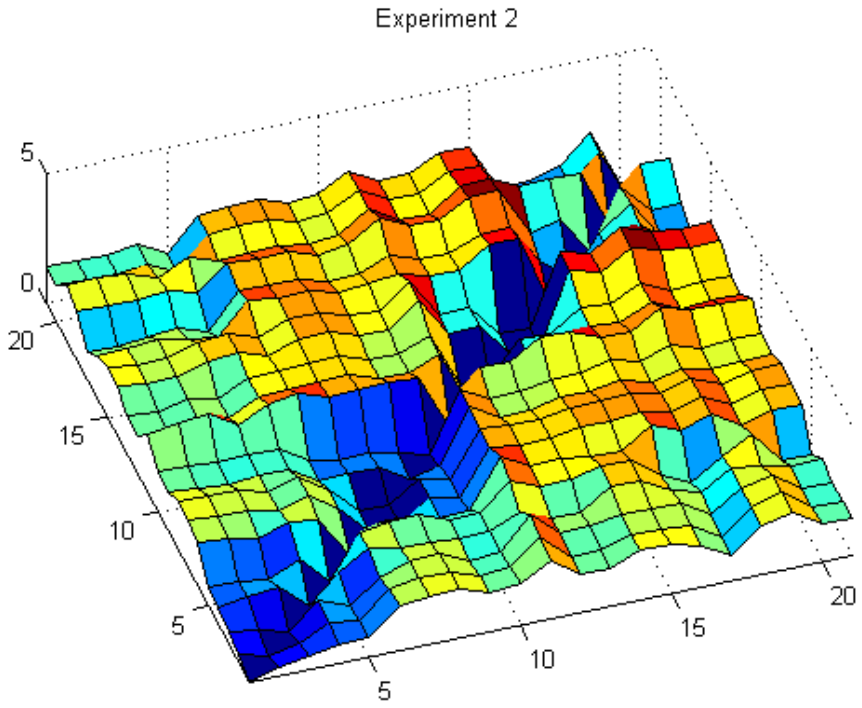

Figure 7. Surface plot of the scores obtained with the shape test set shown in **Figure 5**.

Table 2. Scores obtained for the test set number two.

Files	camel-1	camel-10	camel-11	camel-12	camel-13	cattle-1	cattle-10	cattle-11	cattle-12	cattle-13	deer-1	deer-10	deer-11	deer-12	deer-13	elephant-1	elephant-10	elephant-11	elephant-12	elephant-13
camel-1	0	0.36	0.51	0.68	0.74	1.75	1.87	1.88	1.50	1.57	2.23	1.47	1.48	1.92	1.92	1.60	1.01	1.87	2.14	1.45
camel-10	0.36	0	0.32	0.83	0.68	1.61	1.79	1.80	1.41	1.49	2.46	1.42	1.44	1.75	1.75	1.67	1.00	1.88	2.18	1.36
camel-11	0.51	0.32	0	0.97	0.53	1.36	1.76	1.76	1.28	1.40	2.55	1.61	1.63	1.81	1.81	1.75	1.13	1.82	2.12	1.34
camel-12	0.68	0.83	0.97	0	1.09	1.77	1.49	1.50	1.36	1.44	1.89	1.42	1.42	2.14	2.14	1.56	1.25	1.92	1.71	1.51
camel-13	0.74	0.68	0.53	1.09	0	1.61	1.79	1.80	1.36	1.43	2.42	1.95	1.96	2.23	2.23	1.37	0.83	1.62	2.25	0.97
cattle-1	1.75	1.61	1.36	1.77	1.61	0	1.05	1.05	0.73	0.78	2.62	2.01	2.03	1.86	1.86	2.57	2.15	2.53	1.89	2.17
cattle-10	1.87	1.79	1.76	1.49	1.79	1.05	0	0.01	0.74	0.54	2.23	2.07	2.08	2.29	2.29	2.38	2.13	2.45	1.94	2.24
cattle-11	1.88	1.80	1.76	1.50	1.80	1.05	0.01	0	0.74	0.54	2.23	2.08	2.09	2.29	2.29	2.39	2.14	2.45	1.95	2.25
cattle-12	1.50	1.41	1.28	1.36	1.36	0.73	0.74	0.74	0	0.32	2.33	1.70	1.71	1.94	1.94	2.10	1.81	2.09	1.81	1.88
cattle-13	1.57	1.49	1.40	1.44	1.43	0.78	0.54	0.54	0.32	0	2.24	1.68	1.69	1.90	1.90	2.19	1.87	2.30	1.91	2.00
deer-1	2.23	2.46	2.55	1.89	2.42	2.62	2.23	2.23	2.33	2.24	0	2.14	2.14	2.79	2.79	2.33	2.30	3.01	2.68	2.61
deer-10	1.47	1.42	1.61	1.42	1.95	2.01	2.07	2.08	1.70	1.68	2.14	0	0.03	1.22	1.22	2.03	1.91	2.49	1.96	2.05
deer-11	1.48	1.44	1.63	1.42	1.96	2.03	2.08	2.09	1.71	1.69	2.14	0.03	0	1.27	1.27	2.01	1.91	2.47	1.95	2.04
deer-12	1.92	1.75	1.81	2.14	2.23	1.86	2.29	2.29	1.94	1.90	2.79	1.22	1.27	0	0	2.73	2.40	3.12	2.80	2.62
deer-13	1.92	1.75	1.81	2.14	2.23	1.86	2.29	2.29	1.94	1.90	2.79	1.22	1.27	0	0	2.73	2.40	3.12	2.80	2.62
elephant-1	1.60	1.67	1.75	1.56	1.37	2.57	2.38	2.39	2.10	2.19	2.33	2.03	2.01	2.73	2.73	0	0.91	1.24	2.27	0.74
elephant-10	1.01	1.00	1.13	1.25	0.83	2.15	2.13	2.14	1.81	1.87	2.30	1.91	1.91	2.40	2.40	0.91	0	1.46	2.36	0.84
elephant-11	1.87	1.88	1.82	1.92	1.62	2.53	2.45	2.45	2.09	2.30	3.01	2.49	2.47	3.12	3.12	1.24	1.46	0	2.29	1.13
elephant-12	2.14	2.18	2.12	1.71	2.25	1.89	1.94	1.95	1.81	1.91	2.68	1.96	1.95	2.80	2.80	2.27	2.36	2.29	0	2.15
elephant-13	1.45	1.36	1.34	1.51	0.97	2.17	2.24	2.25	1.88	2.00	2.61	2.05	2.04	2.62	2.62	0.74	0.84	1.13	2.15	0