Ordinal-Measure Based Shape Correspondence

Faouzi Alaya Cheikh

Signal Processing Laboratory, Tampere University of Technology, P.O. Box 553, FIN-33101 Tampere, Finland
Email: faouzi@cs.tut.fi

Bogdan Cramariuc

Signal Processing Laboratory, Tampere University of Technology, P.O. Box 553, FIN-33101 Tampere, Finland
Email: crama@cs.tut.fi

Mari Partio

Signal Processing Laboratory, Tampere University of Technology, P.O. Box 553, FIN-33101 Tampere, Finland
Email: partio@cs.tut.fi

Pasi Reijonen

Signal Processing Laboratory, Tampere University of Technology, P.O. Box 553, FIN-33101 Tampere, Finland
Email: pasi.reijonen@tut.fi

Moncef Gabbouj

Signal Processing Laboratory, Tampere University of Technology, P.O. Box 553, FIN-33101 Tampere, Finland
Email: Moncef.Gabbouj@tut.fi

We present a novel approach to shape similarity estimation based on distance transformation and 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 shape classification, content-based image retrieval, and performance evaluation of segmentation algorithms. The two latter applications are addressed in this paper. Simulation results show that in both applications our proposed measure performs quite well in quantifying shape similarity. The scores obtained using this technique reflect well the correspondence between object contours as humans perceive it.

Keywords and phrases: shape, ordinal, correlation, content, retrieval, indexing, segmentation, performance.

1. INTRODUCTION

Shape representation techniques are generally characterized as being boundary-based or region-based. The former (also known as contour-based) represents the shape by its outline, while the latter considers the shape as being formed of a set of two-dimensional regions. The human visual system itself focuses on edges and ignores uniform regions [1, 2]. This capability is hardwired into the retina. 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 lateral inhibition and helps us to extract boundaries and edges. Therefore, in this paper we focus on this aspect of the shapes and not on the regions they may contain. Object contours 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.

Estimating the similarity between objects shapes can be described in a simplistic way in two steps: shape features extraction and feature comparison. Each of these two steps however represents a difficult problem by itself. Selecting a set of features to characterize a shape for a certain application is not easy, since one must take into consideration the variability of the shapes and the application domain specificity. Feature comparison can be understood as a way of quantifying the similarity/dissimilarity between their corresponding objects. This is a very difficult problem since it tries to mimic the human perception [2].

Several shape features have been proposed in the literature for shape characterization [3]. Many of these techniques however, cannot be used for content-based indexing and retrieval due to their complexity or because they have no
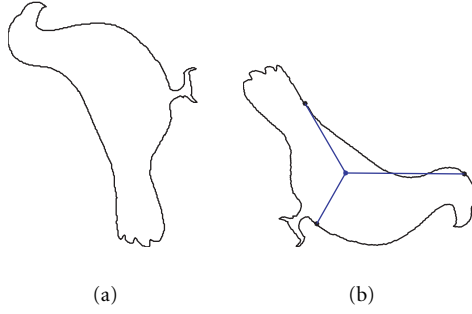
FIGURE 1: (a) The original bird contour before the alignment step, (b) the contour after alignment using three universal axes.

obvious counterpart in the human vision. Therefore, techniques based on simple and visually meaningful shape features have been used in several content-based indexing and retrieval (CBIR) systems, for example, QBIC [4], MUVIS [5, 6] such as high curvature points [7, 8, 9], polygonal approximation [10], morphological and topological features and others [3, 11].

In this paper, we introduce a novel boundary-based approach to shape similarity estimation. This technique is applied to two problems: shape-based image retrieval and performance evaluation of segmentation algorithms. 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. Experimental results for both applications are presented in Section 3, using a subset of the MPEG-7 shape test data and the segmentation masks obtained by the COST Analysis Model (COST AM) [12, 13] segmentation algorithm. In Section 4, conclusions are drawn.

2. THE PROPOSED METHOD

Images in the target applications are representing either: a single object outline or a segmentation mask. Therefore, we will not discuss how to obtain the contour or the segmentation masks. Our goal is to compute a similarity score between any two shapes or two segmentation masks. The proposed method operates in three steps: alignment, boundary to multilevel image transformation, and similarity evaluation. The alignment step is not needed in the case of the segmentation performance evaluation, since we are comparing segmentations masks corresponding to the same image.

Once the boundaries are aligned, the binary images containing the boundaries are transformed into multilevel images through distance transformation [3]. The obtained images are then compared using the ordinal correlation measure introduced in [14, 15]. 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 detailed description of each one of the steps mentioned above.

2.1. Object alignment based on universal axes

The alignment is performed by first detecting three universal axes [16] (those with the largest magnitude) for each shape, then orienting the shape in such a way that these axes are aligned in a standard way for all the objects to be compared. We use the same notation as in [16].

In this implementation of the universal axes determination algorithm we use the version number $\mu = 2l$. The steps of the alignment algorithm are detailed below.

Step 1. Translate the coordinate system so that the origin becomes the center of gravity of the shape S.

Step 2. Compute

$$
\begin{aligned}
\left| x_\mu^{(l)} + i y_\mu^{(l)} \right| &= \int\int_S \left(\sqrt{x^2 + y^2} \right)^\mu \left(\frac{x + iy}{\sqrt{x^2 + y^2}} \right)^l dx\,dy \\
&= \int\int_S r^\mu e^{il\theta} dx\,dy,
\end{aligned}
\tag{1}
$$

and using normalized counterpart (called Universal Axes (UA))

$$
\left| \tilde{x}_\mu^{(l)} + i \tilde{y}_\mu^{(l)} \right| = \frac{\left| x_\mu^{(l)} + i y_\mu^{(l)} \right|}{\int\int_S \left(\sqrt{x^2 + y^2} \right)^\mu dx\,dy}, \quad \text{for } l = 1, 2, 3.
\tag{2}
$$

Step 3. Compute the polar angle $\Theta_\mu \in [0, 2\pi]$ so that

$$
R_\mu e^{i\Theta_\mu} = \left| x_\mu^{(l_1)} + i y_\mu^{(l_1)} \right|
\tag{3}
$$

with R_μ being the magnitude of $|x_\mu^{(l_1)} + i y_\mu^{(l_1)}|$. l_1 is the number of axes needed to align an object.

Step 4. Compute the directional angles of the l_1 universal axes of the shape S as follows:

$$
\theta_j = \frac{\Theta_\mu}{l_1} + (j - 1)\frac{2\pi}{l_1}, \quad \text{for } j = 1, 2, \dots, l_1.
\tag{4}
$$

In our implementation we used $l_1 = 3$, see Figure 1, since for $l_1 = 2$ the two universal axes orientation will verify $\theta_2 = \theta_1 + \pi$. Therefore, they cannot be used alone to determine if an object is flipped around the direction they define or not.

Step 5. Once the three universal axes are determined, rotate the contour so that the most dominant UA (UA with the largest magnitude) will be aligned with the positive x-axis, see Figure 1.

Step 6. Then, if the y-component of the second most dominant UA is positive, flip the contour around the x-axis.

To illustrate the alignment performance we applied it to the set of contours in Figure 2. The results of the alignment are presented in Figure 3. It can be noticed that this alignment scheme solved both problems of rotation and mirroring.
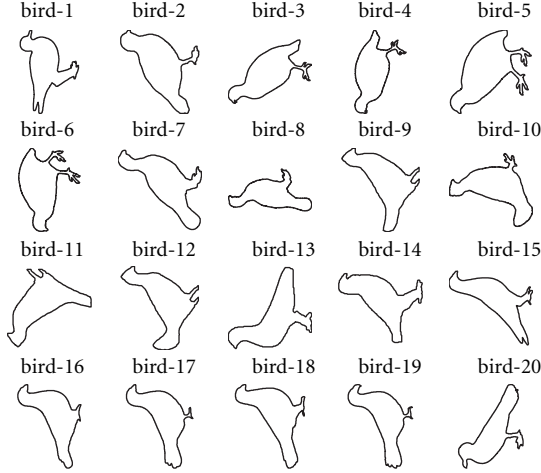
bird-1 bird-2 bird-3 bird-4 bird-5

bird-6 bird-7 bird-8 bird-9 bird-10

bird-11 bird-12 bird-13 bird-14 bird-15

bird-16 bird-17 bird-18 bird-19 bird-20

FIGURE 2: Bird contours from the MPEG-7 shape test set B.

bird-1 bird-2 bird-3 bird-4 bird-5

bird-6 bird-7 bird-8 bird-9 bird-10

bird-11 bird-12 bird-13 bird-14 bird-15
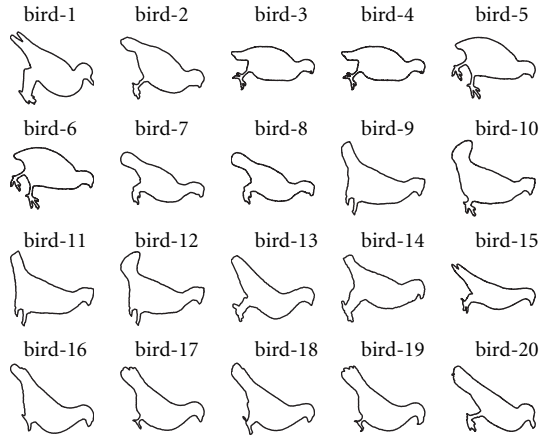
bird-16 bird-17 bird-18 bird-19 bird-20

FIGURE 3: The contours in Figure 2 after alignment.

2.2. Boundary to multilevel image transformation

Let S be a shape represented by its contour C in a binary image. The binary image is transformed into a multilevel (grayscale) image G using a mapping function ϕ, such that the pixel values in G, $\{G_1, G_2, \ldots, G_n\}$, depend on their relative position to the contour pixels $C_1, C_2, \ldots, C_p$:

$$G_i = \phi(C_k : k = 1, 2, \ldots, p), \quad \text{for } i = 1, 2, \ldots, n, \quad (5)$$

where C_k is the position of the contour pixel k in the image G. It should be observed that several transformations satisfy this requirement, including any distance transform [3].

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 from the boundary information redundancy in the new image. We expect that there is no single optimal mapping; different mappings will emphasize different features of the contour. We distinguish however two special cases:

- the first is defined as follows:

$$G_i = \begin{cases} V_0 - d(P_i, C), & \text{if } d(P_i, C) < T_h, \\ 0, & \text{otherwise}, \end{cases} \quad (6)$$

for $i = 1, 2, \ldots, n$ and $V_0 > 0$ is the value assigned to the pixels on the boundary. The larger the distance $d(P_i, C)$ from the contour points is, the smaller the new pixel value will be. This mapping function emphasizes the details on the boundary;

- the second mapping is when $V_0 = 0$ and $G_i = d(P_i, C)$, using a geodesic distance [17]. When this mapping is applied inside the contour only, the emphasis is on the shape skeleton which is a very important feature of a shape, see Figure 4. Applying the distance mapping inside and outside the contour can lead to a better evaluation of segmentation results. One can even assign different weights inside and outside of the contours.

In this work, we implemented the second mapping based on the geodesic distance. The metric is integer and its application is done through an iterative wave propagation process [18]. 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, as stated earlier. The values can increase or decrease starting from the contour and can be limited. The pixel values in the distance map can be therefore written as follows:

$$G_i = |V_0 \pm d(P_i, C)|, \quad \text{for } i = 1, 2, \ldots, n, \quad (7)$$

where, V_0 is the value on the contour and $d(P_i, C)$ is the distance from any point P_i in the image to the contour C.

Figure 4a represents an example of a distance map generated only inside the contour of a bird contour. Figure 4b shows a 3D visualization of this distance map.

2.2.1 Similarity evaluation

The evaluation of image similarity is based on the framework for ordinal-based image correspondence introduced in [14]. Figure 5 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 images X and 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. Let R_j^X and R_j^Y be the pixels from images X and Y, respectively, which belong to areas R_j, with $j = 1, 2, \ldots, m$.

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 $j = 1, 2, \ldots, m$. Thus, each block in image X is compared to the corresponding block in image Y in an ordinal fashion. The ordinal comparison of the two regions means that only the ranks of the pixels are utilized. For every pixel X_k, we construct a so-called slice $S_k^X = \{S_{k,l} : l = 1, 2, \ldots, n\}$,
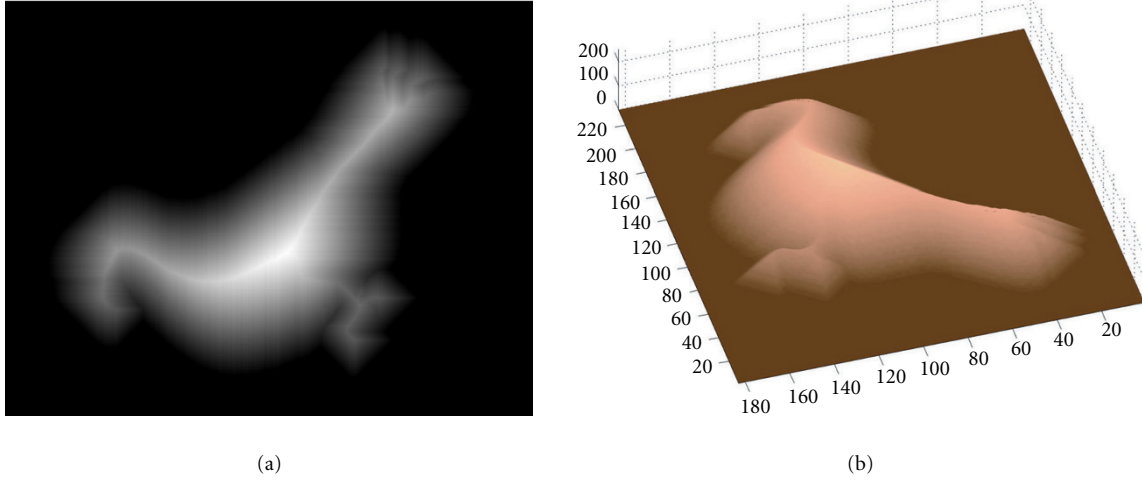
(a)

(b)

FIGURE 4: (a) The distance map generated for the "bird-19" contour in Figure 1, (b) a 3D view of this distance map.

where

$$S_{k,l}^X = \begin{cases} 1, & \text{if } X_k < X_l, \\ 0, & \text{otherwise.} \end{cases} \tag{8}$$

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.

To compare 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 $OP_1(\cdot)$ into a metaslice M_j^X.

Figure 6 shows an illustration of the slices and metaslices creation for a 4×4 image and blocks of 2×2. The four slices S_1, S_2, S_5, and S_6 shown in this figure are computed for the four pixels in block B_1. The operation used in this illustration to create the metaslice M_1 is $OP_1(\cdot) = \sum(\cdot)$.

More formally, $M_j^X = OP_1(S_k^X : X_k \in R_j^X)$ for $j = 1, 2, \ldots, m$. Similarly, we combine the slices from image Y to form M_j^Y for $j = 1, 2, \ldots, m$. It should be noted that the metaslices are equal in size to the original images and could be multivalued, depending on the operation $OP_1(\cdot)$. 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, $D_j = OP_2(M_j^X, M_j^Y)$, $j = 1, 2, \ldots, m$. We thus construct a set of metadifferences $D = \{D1, D2, \ldots, Dm\}$. The final step is to extract a scalar measure of correspondence from set D, using operation $OP_3(\cdot)$. In other words, $\lambda = OP_3(D)$. It was shown in [14] that this structure could be used to model the well-known Kendall's τ and Spearman's ρ measures [19].

The image similarity measure used in this paper is an instance of the previously mentioned framework. This measure has been analyzed more extensively by Cramariuc et al. [15]. Following is a short description of the operations $OP_k(\cdot)$, $k = 1, 2, 3$ adopted for this measure. Operation $OP_1(\cdot)$ 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 $OP_2(\cdot)$ is chosen to be the squared Euclidean distance between corresponding metaslices. That is, $D_j = ||M_j^X - M_j^Y||_2^2$. Finally, operation $OP_3(\cdot)$ sums together all metadifferences to produce $\lambda = \sum_j D_j$, for $j = 1, 2, \ldots, m$. Small values of λ mean similar objects.

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.

3. EXPERIMENTAL RESULTS

The proposed technique is applied to two important problems: content-based retrieval of shape images and performance evaluation of segmentation algorithms. The experiments performed are presented and their results analyzed in the rest of this section.

3.1. Shape similarity estimation

The shape similarity estimation experiments were conducted on two sets of 20 images. The two sets are taken from the MPEG-7 CE Shape test set B, which contains 1400 images grouped in 70 categories. These test sets are chosen in such a way as to assess the performance of our technique in estimating the object similarity within a single category (intra-category) and between contours from different categories (inter-category). Therefore, the first test set contains all the samples in the bird category of the MPEG-7 Shape test set B, see Figure 2. While, the second set contains 20 objects taken from four different categories, see Figure 7. In both experiments, the similarity score λ is computed for all the pairs
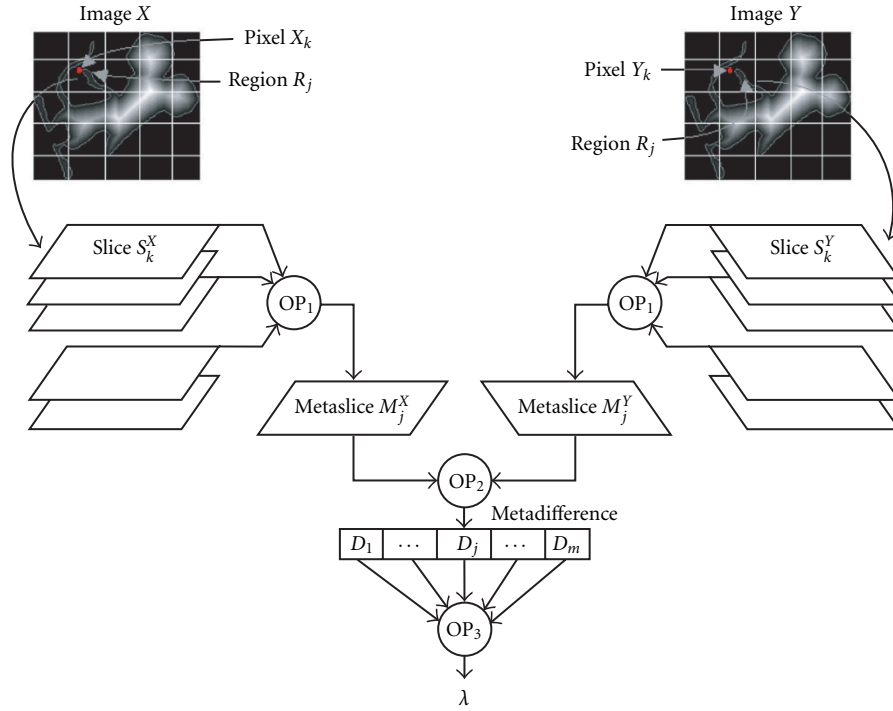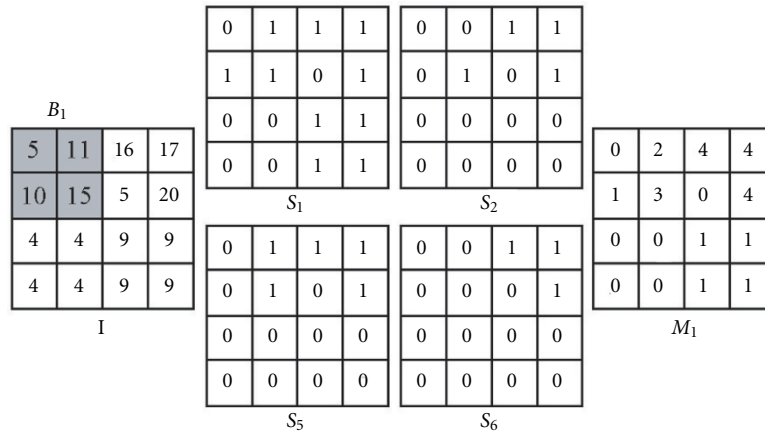
FIGURE 5: The general framework for ordinal correlation of images.

FIGURE 6: Example of slices and metaslice for a 4×4 image using blocks of 2×2.

of shapes in the set. The similarity scores obtained are presented in Tables 1 and 2. All the scores are multiplied by 10^3 when they are presented in the tables and the figures. The distance maps were generated inside the objects only with $V_0 = 50$. This setting emphasizes the shape skeleton and gives less importance to contour pixels. The distance transformed images are resized to 32×32 pixels and blocks of size 4×4 were used. Larger images can be used if more precision is needed, this would imply the creation of more slices and therefore more computational power would be needed.

Figure 8 represents a surface plot of the similarity scores in Table 1. It shows that within the same category, the scores have small values, which means that they are quite similar according to our measure. It is worth noticing that the scores on the diagonals are zero which means that each object is identical to itself, so there is no bias in the similarity scores. It is worth noticing that the scores obtained between the "bird-3," "bird-4," "bird-5," "bird-6," and the rest of the birds in this category are larger than the rest of the scores. This can be explained by the fact that these four birds have much shorter tails and have a more circular contours compared the rest of the birds. The similarity scores are low between themselves, moreover, the scores for the pairs (bird-3, bird-4), (bird-5, bird-6), (bird-7, bird-8), and (bird-9, bird-11) are very small. By visual inspection one can verify that each pair of con-

TABLE 1: Similarity scores for the contours in Figure 3. These scores are multiplied by 10^3.

Object	bird-1	bird-2	bird-3	bird-4	bird-5	bird-6	bird-7	bird-8	bird-9	bird-10	bird-11	bird-12	bird-13	bird-14	bird-15	bird-16	bird-17	bird-18	bird-19	bird-20
bird-1	0	32	75	74	80	81	46	44	52	44	52	49	64	49	53	63	60	51	64	59
bird-2	32	0	71	70	79	79	39	38	60	59	62	61	65	56	53	68	58	52	62	55
bird-3	75	71	0	0	48	48	80	77	88	84	85	87	102	53	71	117	112	105	116	106
bird-4	74	70	0	0	48	48	80	76	88	83	85	87	102	53	71	116	112	104	116	106
bird-5	80	79	48	48	0	1	101	98	105	89	102	98	120	76	95	125	122	112	125	121
bird-6	81	79	48	48	1	0	101	98	106	90	103	99	120	76	95	126	122	113	126	121
bird-7	46	39	80	80	101	101	0	7	39	42	42	40	42	54	28	53	42	38	46	43
bird-8	44	38	77	76	98	98	7	0	38	39	39	38	44	52	23	58	47	41	51	47
bird-9	52	60	88	88	105	106	39	38	0	25	7	17	32	51	33	51	44	39	49	60
bird-10	44	59	84	83	89	90	42	39	25	0	22	12	45	53	32	55	48	43	53	60
bird-11	52	62	85	85	102	103	42	39	7	22	0	16	36	49	31	56	49	45	55	63
bird-12	49	61	87	87	98	99	40	38	17	12	16	0	37	55	31	52	46	41	50	61
bird-13	64	65	102	102	120	120	42	44	32	45	36	37	0	60	42	41	35	37	40	57
bird-14	49	56	53	53	76	76	54	52	51	53	49	55	60	0	47	86	80	75	86	83
bird-15	53	53	71	71	95	95	28	23	33	32	31	31	42	47	0	67	56	52	59	61
bird-16	63	68	117	116	125	126	53	58	51	55	56	52	41	86	67	0	17	22	19	44
bird-17	60	58	112	112	122	122	42	47	44	48	49	46	35	80	56	17	0	18	8	39
bird-18	51	52	105	104	112	113	38	41	39	43	45	41	37	75	52	22	18	0	17	35
bird-19	64	62	116	116	125	126	46	51	49	53	55	50	40	86	59	19	8	17	0	38
bird-20	59	55	106	106	121	121	43	47	60	60	63	61	57	83	61	44	39	35	38	0

FIGURE 7: Contours of test set 2 after alignment.

FIGURE 8: The similarity scores for the bird contours in Figure 3, dark blue cells mean most similar contours.
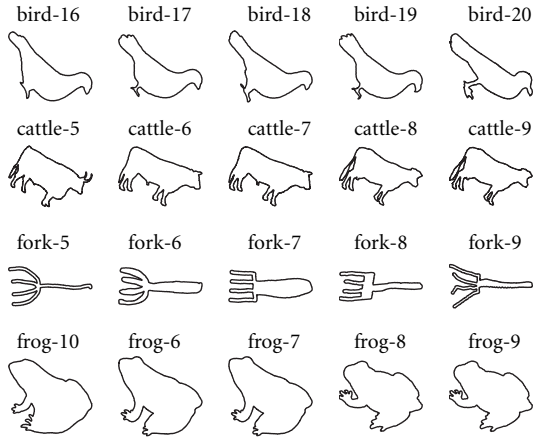
tours represent the same bird contour rotated or rotated and scaled. Therefore, we can safely say that our measures have a 0.5% error, which can be explained by the small contour variation introduced by rotation and the size reduction of the distance maps. Lower error can be obtained by increasing the size of the distance map images and reducing the block size used for the metaslices creation.

Dark blue regions in Figure 8 represent very low scores (close to zero), which shows that there are quite many objects in this category which are very similar or even identical.

To find out which are the most similar contours to a given contour in Figure 7 we sort the scores on the raw corresponding to this contour 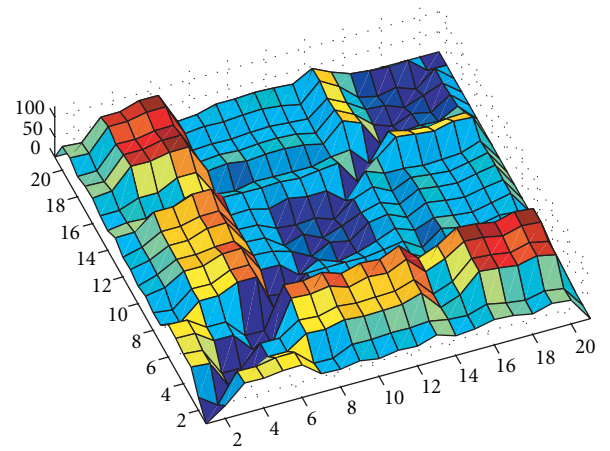in Table 2. Using Figure 9, one can easily estimate which are the most similar objects within this category, based on the clustered dark blue cells. Figure 9 shows that similarity scores between subjects from the same category are low, while those obtained for subjects from different categories are relatively high. Therefore, sorting the scores in ascending order will yield the most similar object first.

The inter-category scores obtained by our similarity estimation technique are larger than the inter-category ones. Therefore, this technique can be used as a shape classification technique. Moreover, it is sensitive to intra-category shape variations thus it can be used in a content-based retrieval system.

TABLE 2: Similarity scores for the contours in Figure 7. These scores are multiplied by 10^3.

Objects	bird-16	bird-17	bird-18	bird-19	bird-20	cattle-5	cattle-6	cattle-7	cattle-8	cattle-9	fork-5	fork-6	fork-7	fork-8	fork-9	frog-10	frog-6	frog-7	frog-8	frog-9
bird-16	0	39	63	39	101	282	322	321	351	348	358	368	361	379	401	196	209	210	303	302
bird-17	39	0	58	22	94	273	306	304	342	341	337	345	339	363	377	199	206	208	297	297
bird-18	63	58	0	58	83	256	293	291	325	322	325	330	323	346	364	182	185	187	285	284
bird-19	39	22	58	0	74	298	332	331	369	366	360	368	363	388	402	215	227	228	318	318
bird-20	101	94	83	74	0	304	333	334	370	367	365	370	363	387	392	234	233	234	330	330
cattle-5	282	273	256	298	304	0	73	71	75	73	247	221	184	206	244	162	152	153	136	138
cattle-6	322	306	293	332	333	73	0	1	52	53	246	204	168	198	234	195	192	193	143	145
cattle-7	321	304	291	331	334	71	1	0	51	52	247	206	169	199	236	193	191	192	141	144
cattle-8	351	342	325	369	370	75	52	51	0	5	244	207	172	191	228	206	199	200	141	143
cattle-9	348	341	322	366	367	73	53	52	5	0	244	209	173	190	232	202	196	197	140	142
fork-5	358	337	325	360	365	247	246	247	244	244	0	59	118	84	63	373	347	346	356	359
fork-6	368	345	330	368	370	221	204	206	207	209	59	0	62	74	77	355	337	338	326	329
fork-7	361	339	323	363	363	184	168	169	172	173	118	62	0	81	123	320	304	305	287	290
fork-8	379	363	346	388	387	206	198	199	191	190	84	74	81	0	79	351	331	331	319	321
fork-9	401	377	364	402	392	244	234	236	228	232	63	77	123	79	0	379	354	352	339	340
frog-10	196	199	182	215	234	162	195	193	206	202	373	355	320	351	379	0	50	51	132	129
frog-6	209	206	185	227	233	152	192	191	199	196	347	337	304	331	354	50	0	1	138	138
frog-7	210	208	187	228	234	153	193	192	200	197	346	338	305	331	352	51	1	0	139	139
frog-8	303	297	285	318	330	136	143	141	141	140	356	326	287	319	339	132	138	139	0	2
frog-9	302	297	284	318	330	138	145	144	143	142	359	329	290	321	340	129	138	139	2	0
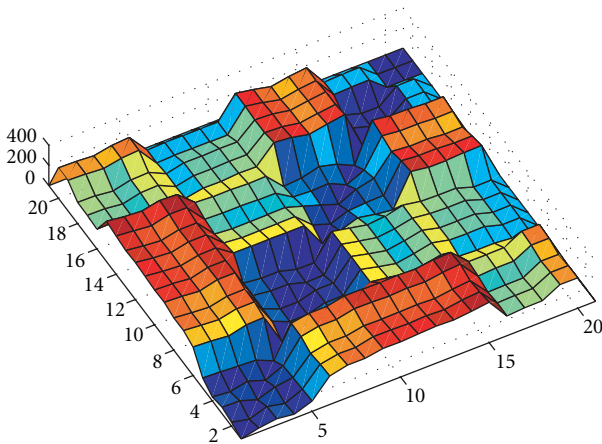
FIGURE 9: Similarity scores obtained for the contours in test set 2 presented in Figure 7.

FIGURE 10: The segmentation performance scores for frames 4–49 of the sequence "Erik," for COST AM versions 5.0 and 5.1.

3.2. Segmentation quality evaluation

The objective evaluation of the performance of segmentation algorithms is an important problem [20, 21, 22]. Even when a reference mask is available, comparing two segmentation masks is still a difficult problem. Several factors make such evaluation difficult, among the most important factors is the difficulty to discriminate between many small distributed error segments and few larger error segments.

Our shape correspondence technique proposed in Section 2, discriminates easily between the two cases of segmentation errors. The geodesic distance transformation is applied inside each segment of the mask. Therefore, small regions yield small distances inside them and therefore will generate pixels with low gray values.
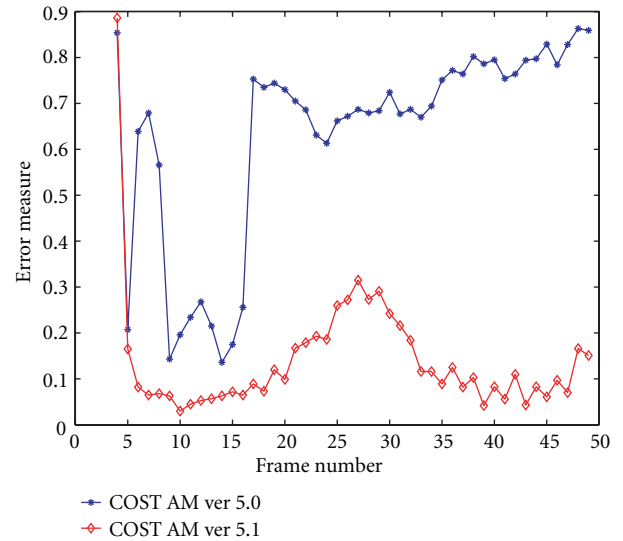
In this experiment, the segmentation masks resulting from the COST AM versions 5.0 and 5.1 [12, 13], are compared against a reference mask. The plot in Figure 10, shows the segmentation performance scores obtained by our technique, for the frames 4–49 of "Erik" sequence. The plots in Figures 11 and 12, show quantitative measures of the errors in number of pixels from both COST AM versions 5.0 and 5.1, respectively. Three different numbers are computed for each frame:

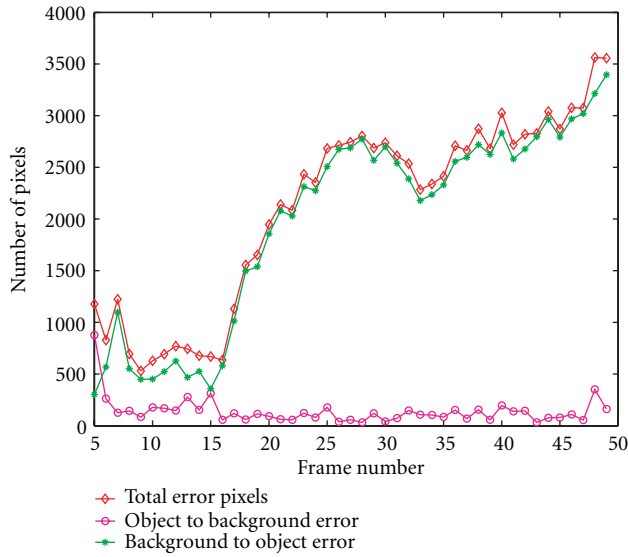- number of pixels of the background segmented as foreground pixels,

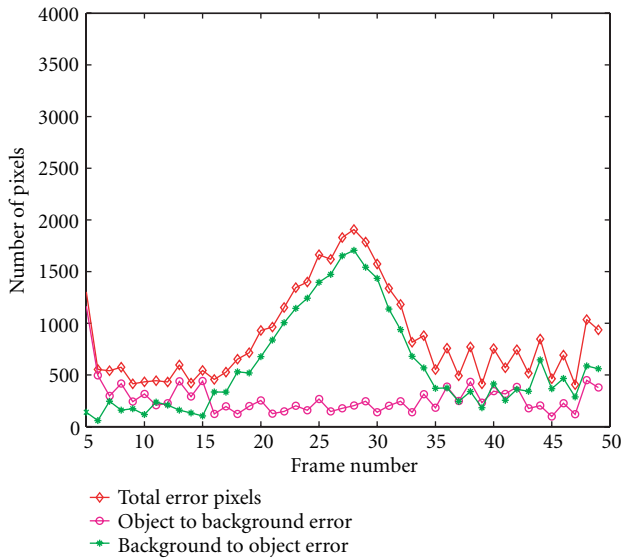FIGURE 11: Plot of the segmentation errors, for COST AM 5.0.

FIGURE 12: Plot of the segmentation errors, for COST AM 5.1.

- number of pixels of the foreground segmented as background pixels,
- sum of the two previous numbers.

For illustration we present the colored segmentation masks for frames 15 and 20 from "Erik" sequence in Figures 13, 14, 15, and 16. The frame pixels are colored as follows:

- black represents the background,
- white is the region where the reference and estimated masks overlap,
- green represents the areas of the background segmented as part of the object,
- purple represents the regions from the object merged with the background.

FIGURE 13: The colored segmentation error of frame 15 from the sequence "Erik," segmented using COST AM 5.0.

FIGURE 14: The colored segmentation result of frame 15 from the sequence "Erik," segmented using COST AM 5.1.

FIGURE 15: The colored segmentation result of frame 20 from the sequence "Erik," segmented using COST AM 5.0.
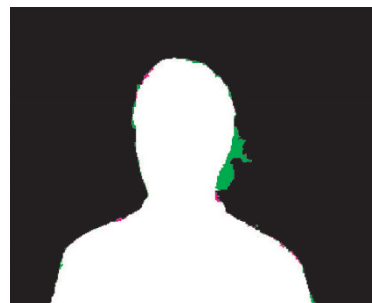

FIGURE 16: The colored segmentation result of frame 20 from the sequence "Erik," segmented using COST AM 5.1.

It can be easily seen that our segmentation performance scores in Figure 10, correlate very well with the variation of the total number of pixel errors. Moreover, it reflects the variation in both types of segmentation errors. Our measure inherently resolves the case of many small errors and the case of a single large error region.

4. CONCLUSIONS AND FUTURE WORK

In this paper, we proposed a contour correspondence measure, based on distance transformation and ordinal correlation. The similarity scores obtained are in line with the visual perception of the similarity between shapes. We showed that the proposed measure can be used in two applications: shape similarity estimation in the context of content-based image retrieval, and for performance evaluation of segmentation algorithms. Simulation results were presented for both applications using images from the MPEG-7 shape test set B for the first application, and 50 frames from "Erik" sequence for the second. The proposed technique produced encouraging results in both experiments. Further study is needed to optimize the proposed technique in order to select appropriate parameters for the application at hand. Finally, further analysis of the behavior of the proposed technique may identify new applications.

REFERENCES

[1] C. Hildreth, "The detection of intensity changes by computer and biological vision systems," in *Proc. Computer Vision, Graphics, and Image Processing*, vol. 22, pp. 1–27, 1983.

[2] T. V. Papathomas, "Special issue on visual-perception: guest editorial," *International Journal of Imaging Systems and Technology*, vol. 7, no. 2, pp. 63–64, 1996.

[3] L. da F. Costa and R. M. Cesar Jr., *Shape Analysis and Classification: Theory and Practice*, CRC Press, Boca Raton, Fla, USA, 2001.

[4] M. Flickner, H. Sawhney, W. Niblack, et al., "Query by image and video content: The qbic system," *IEEE Computer Magazine*, vol. 28, no. 9, pp. 23–32, 1995.

[5] C. F. Alaya, B. Cramariuc, C. Reynaud, et al., "Muvis: a system for content-based indexing and retrieval in large image databases," in *Proc. SPIE/EI '99 Conference on Storage and Retrieval for Image and Video Databases VII*, vol. 3656, pp. 98–106, San Jose, Calif, USA, January 1999.

[6] M. Trimeche, C. F. Alaya, M. Gabbouj, and B. Cramariuc, "Content-based description of images for retrieval in large databases: Muvis," in *Proc. X European Signal Processing Conference*, Tampere, Finland, September 2000.

[7] C. H. Teh and R. T. Chin, "On the detection of dominant points on digital curves," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 11, no. 8, pp. 859–872, 1989.

[8] S. Abbasi, F. Mokhtarian, and J. Kittler, "Curvature scale space image in shape similarity retrieval," *Springer Journal of Multimedia Systems*, vol. 7, no. 6, pp. 467–476, 1999.

[9] A. Quddus, C. F. Alaya, and M. Gabbouj, "Wavelet-based multi-level object retrieval in contour images," in *Proc. International Workshop on Very Low Bit Rate Video Coding*, pp. 1–5, Kyoto, Japan, October 1999.

[10] M. W. Koch and R. L. Kashyap, "Using polygon to recognize and locate partially occluded objects," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 9, no. 4, pp. 483–494, 1987.

[11] J. C. Russ, *The Image Processing Handbook*, CRC, Springer and IEEE Press, 3rd edition, 1999.

[12] A. Alatan, L. Onural, M. Wollborn, R. Mech, E. Tuncel, and T. Sikora, "Image sequence analysis for emerging interactive multimedia services—the European COST 211 framework," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 8, no. 7, pp. 802–813, 1998.

[13] M. Gabbouj, G. Morrison, C. F. Alaya, and R. Mech, "Redundancy reducation techniques and content analysis for multimedia services—the European COST 211quat action," in *Proc. Workshop on Image Analysis for Multimedia Interactive Services*, Berlin, Germany, 31 May–1 June 1999.

[14] I. Shmulevich, B. Cramariuc, and M. Gabbouj, "A framework for ordinal-based image correspondence," in *Proc. X European Signal Processing Conference*, Tampere, Finland, September 2000.

[15] B. Cramariuc, I. Shmulevich, M. Gabbouj, and A. Makela, "A new image similarity measure based on ordinal correlation," in *Proc. International Conference on Image Processing*, vol. 3, pp. 718–721, Vancouver, BC, Canada, September 2000.

[16] J. C. Lin, "The family of universal axes," *Pattern Recognition*, vol. 29, no. 3, pp. 477–485, 1996.

[17] P. J. Toivanen, "New geodesic distance transforms for grayscale images," *Pattern Recognition Letters*, vol. 17, no. 5, pp. 437–450, 1996.

[18] I. Ragnelmam, "Neighborhoods for distance transformations using ordered propagation," *Computer Vision, Graphics and Image Processing*, vol. 56, no. 3, pp. 399–409, 1992.

[19] M. Kendall and J. D. Gibbons, *Rank Correlation Methods*, Edward Arnold, New York, 5th edition, 1990.

[20] P. Villegas, X. Marichal, and A. Salcedo, "Objective evaluation of segmentation masks in video sequences," in *Proc. Workshop on Image Analysis for Multimedia Services (WIAMIS '1999)*, Berlin, Germany, May 1999.

[21] R. Mech, "Objective evaluation criteria for 2D-shape estimation results of moving objects," in *Proc. Workshop on Image Analysis for Multimedia Services*, pp. 23–28, Tampere, Finland, May 2001.

[22] R. Mech and F. Marqu s, "Objective evaluation criteria for 2D-shape estimation results of moving objects," EURASIP Journal of Applied Signal Processing, Special Issue: Image Analysis for Multimedia Interactive Services, 2002.

Faouzi Alaya Cheikh received his B.S. degree in electrical engineering in 1992 from École Nationale d'Ingénieurs de Tunis, Tunisia. He received his M.S. degree in electrical engineering (Major in Signal Processing) from Tampere University of Technology, Finland, in 1996. Mr. Alaya Cheikh is currently a Ph.D. candidate and works as a Researcher at the Institute of Signal Processing, Tampere University of Technology, Tampere, Finland. From 1994 to 1996, he was a Research Assistant at the Institute of Signal Processing, and from 1997 he has been a Researcher with the same institute. His research interests include nonlinear signal and image processing and analysis, pattern recognition and content-based analysis and retrieval. He has been an active member in many Finnish and European research projects among them Nobless esprit, COST 211 quat, and MUVI. He served as Associate Editor of the EURASIP Journal on Applied Signal Processing, Special Issue on Image Analysis for Multimedia Interactive Services. He serves as a reviewer to several conferences and journals. He co-authored over 30 publications.

Bogdan Cramariuc received his M.S. degree in electrical engineering in 1993 from Polytechnica University of Bucharest, Faculty of Electronics and Telecommunications, Bucharest, Romania. Mr. Cramariuc is currently a Ph.D. candidate and works as Researcher for the Institute of Signal Processing at Tampere University of Technology, Tampere, Finland. From 1993 to 1994 he worked as Teaching Assistant at the Faculty of Electronics and Telecommunications at the Polytechnica University of Bucharest. During this period he has also been involved as Researcher with Electrostatica S.A., a national research institute in Bucharest, Romania. Since 1995 he has been with the Institute of Signal Processing at Tampere University of Technology, Tampere, Finland. His research interests include signal and image analysis, image segmentation, texture analysis, content-based indexing and retrieval in multimedia databases, mathematical morphology, computer vision, parallel processing, data mining, and artificial intelligence. Mr. Cramariuc has been an active member in several Finnish and European projects, such as Nobless, Esprit and MUVI. He served as Associate Editor of the EURASIP Journal on Applied Signal Processing, Special Issue on Image Analysis for Multimedia Interactive Services.

Mari Partio was born 1979 in Finland. Autumn 1998 she started her M.S. studies at Tampere University of Technology in department of Electric Engineering. She is majoring in Signal Processing and her minor is Software Engineering. Since summer 2000 she has been working as a research assistant at the Institute of Signal Processing. Her research interests include content-based image and video retrieval.

Pasi Reijonen was born 1976 in Finland. Started at Tampere University of Technology in Automation Degree Program 1996 and received M.S. degree 2001. Major: Electronics. Minor: Industrial Management, Automation and Control Engineering. For the academic year 1999–2000, he was as an Erasmus student in England (University of Leeds). Currently he is working as a researcher at Tampere University of Technology at Institute of Signal Processing. His research interests are Shape Analysis for Content-Based Image Retrieval.

Moncef Gabbouj received his B.S. degree in electrical engineering in 1985 from Oklahoma State University, Stillwater, and his M.S. and Ph.D. degrees in electrical engineering from Purdue University, West Lafayette, Indiana, in 1986 and 1989, respectively. Dr. Gabbouj is currently a professor and Head of the Institute of Signal Processing of Tampere University of Technology, Tampere, Finland. From 1995 to 1998 he was a professor with the Department of Information Technology of Pori School of Technology and Economics, Pori, and during 1997 and 1998 he was on sabbatical leave with the Academy of Finland. From 1994 to 1995 he was an associate professor with the Signal Processing Laboratory of Tampere University of Technology, Tampere, Finland. From 1990 to 1993 he was a senior research scientist with the Research Institute for Information Technology, Tampere, Finland. His research interests include nonlinear signal and image processing and analysis, content-based analysis and retrieval and mathematical morphology. Dr. Gabbouj is the Vice-Chairman of the IEEE-EURASIP NSIP (Nonlinear Signal and Image Processing) Board. He is currently the Technical Committee Chairman of the EC COST 211quat. He served as associate editor of the *IEEE Transactions on Image Processing*, and was guest editor of the European journal *Signal Processing*, special issue on nonlinear digital signal processing (August 1994). He is the chairman of the IEEE Finland Section and past chair of the IEEE Circuits and Systems Society, Technical Committee on Digital Signal Processing, and the IEEE SP/CAS Finland Chapter. He was also the TPC Chair of EUSIPCO 2000 and the DSP track chair of the 1996 IEEE ISCAS and the program chair of NORSIG' 96, and is the technical program chair of EUSIPCO 2000. He is also member of EURASIP AdCom. Dr. Gabbouj is the Director of the International University Program in Information Technology and member of the Council of the Department of Information Technology at Tampere University of Technology. He is also the Secretary of the International Advisory Board of Tampere International Center of Signal Processing, TICSP. He is a member of Eta Kappa Nu, Phi Kappa Phi, IEEE SP and CAS societies. Dr. Gabbouj was co-recipient of the Myril B. Reed Best Paper Award from the 32nd Midwest Symposium on Circuits and Systems and co-recipient of the NORSIG 94 Best Paper Award from the 1994 Nordic Signal Processing Symposium. He is co-author of over 150 publications. Dr. Gabbouj was the prime investigator in two ESPRIT projects, 1 HCM project and several Tempus projects. He also served as Evaluator of IST proposals, and Auditor of a number of ACTS and IST projects on multimedia security, augmented and virtual reality, image and video signal processing.