Visual Information Retrieval Based on Shape Similarity

Jong-Seung Park

Dept. of Computer Science and Engineering, University of Incheon,
177 Dohwa-dong, Nam-gu, Incheon, 402-749, Republic of Korea
jong@incheon.ac.kr

Abstract. An effective and fast shape description and retrieval method is presented for huge image databases. As a shape representation for deformable objects, a multi-scale skeleton representation is proposed in order to preserve the consistency of the skeletons and to reduce the effect of the structural changes. Incorrect matches due to the boundary noise in a segmentation process are avoided by including multiple coarse skeletons of different scales. A fast computational method for the similarity of skeletons is also proposed by using the moment invariants. Experimental results on animal databases showed that the proposed method gives prominent accuracy in retrieval.

1 Introduction

The indexing and retrieval of digital photographs is becoming relevant for many applications, including in multimedia libraries, art galleries and museum archives, in picture and photograph archiving and communication, and in medical and geographic databases. The problem of content-based image retrieval in an image archiving system is to obtain a list of images from a huge database which are most similar to the query description. Two critical issues of a content-based image retrieval system are feature extraction for indexing and similarity computation for retrieval. Feature extraction is for automatic characterization of an image. Based on the similarity measure, relevant images are retrieved. Most of the previous image retrieval systems were based on the image feature analysis such as colors, textures, and regions. Such features do not represent shape properties of a query object quite well hence irrelevant images are frequently retrieved. Recently, there has been works to handle shape features effectively [1][2][3].

In this paper, to improve the retrieval effectiveness of a content-based image retrieval system, a shape-based similarity comparison method is presented. A new multi-scale skeleton-based invariant feature representation is proposed as a shape representation. The method of multi-scale skeletons is not a new one [4][5]. In contrast to their approaches, our method extracts invariant features from the multiple skeletons and keeps only the feature vectors to the database. The moment invariants are used to compute the similarity of two images.

Z. Chen et al. (Eds.): ICADL 2004, LNCS 3334, pp. 458–461, 2004.

2 Multi-scale Skeleton Representation

There are two different processes for the image archiving system: image archival process and image retrieval process. The image archival process is related to registration of collected images to the database. For the registration of an image to the archive system we first obtain the region segmentation data using our specialized segmentation utility. From the segmentation data, multiple skeletons are computed according to the level of details for the boundary representation. Also a feature vector is computed for each skeleton representation. The image and the set of feature vectors are saved to the database in the image registration process.

The image retrieval process is related to the search and retrieval of similar images to the given query description. The user should provide a region boundary of the object to be searched in the query image. Then, the system constructs a skeleton structure for the specified region and also computes the feature vectors for the structures. The feature vectors for the structures are compared to the feature vectors stored in the database and the most similar results are retrieved.

Skeleton representation is a natural way of shape description especially for deformable objects such as human beings, animals, fishes and insects. Beside its naturalness, the shape can be reconstructed from the skeleton representation by taking an inverse skeleton transform. The major drawback of skeleton representation is that it is sensitive to noise. For an example, approaches of medial axis may cause spurious branches and shape distortions for the jagged boundaries. To prohibit the sensitivity to boundary noises we propose a method of skeleton structure by employing a multi-scale representation. The multiple skeletons describe coarse boundaries as well as fine details.

The skeleton of an object shape is defined as the locus of the centers of the maximal disks that are contained within the shape. Among several methods of skeleton extraction, we use a fast two-pass algorithm [6] to compute the skeleton using a distance transform. By limiting the minimal size of maximal disks multiple skeletons with different scales are obtained. For each skeleton an invariant feature vector is computed and the set of feature vectors is registered to the database together with the image.

3 Invariant Features for Similarity Measure in Retrieval

Moment invariants are useful measures for 2-D shape matching. Moments are defined on a continuous image intensity function. For a discrete binary image, a simple approximation is possible using summation operation. For each region, a 7-D feature vector of seven moments is computed where each moment is invariant to translation, scale changes and rotation for the case of continuous functions. The invariants are still strictly invariant under image translation and are approximately invariant under rotation and scale changes due to sampling, digitizing, and quantizing of the continuous image for digital computation [7].

In the case of articulated objects such as human beings or animals, the shape transformation due to the motion of articulations causes the shape matching to an original shape to fail and the system may regard them as a different object. In our multi-level structure representation, the shape deformation of a deformable object due to a small articulation movement is less affected since scale reduced skeletons could

be matched. The similarity between two feature vectors is computed as the Euclidean distance. Note that some values of feature vectors are quite small or vary a lot. So they need to be normalized for comparison.

4 Experimental Results and Conclusion

We tested the proposed method on several databases. Fig. 1 shows intermediate data for the invariant feature computation. Fig. 2 shows the multi-scale skeletons and their shape boundaries. For a query image the system retrieves the most relevant eight images from the database images. Fig. 3 shows two examples of image retrieval from a fish image database which was originally used in the work of Mokhtarian et al. [2]. The left-most images are the query images and the remains are the retrieved images in order of similarity measurements. The query processing time is less than one millisecond in most cases for the database of 1,100 animal images. For each query, there are only vector length operations for 1,100 feature vectors and one array sorting operation.

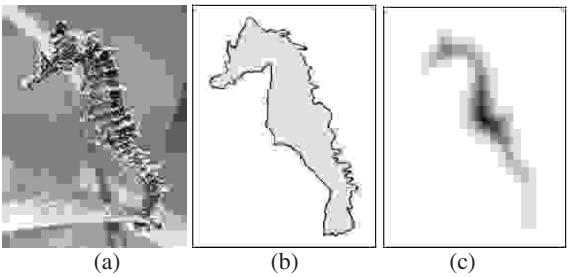

Fig. 1. (a) A sea horse image; (b) the region boundary; (c) the distance image

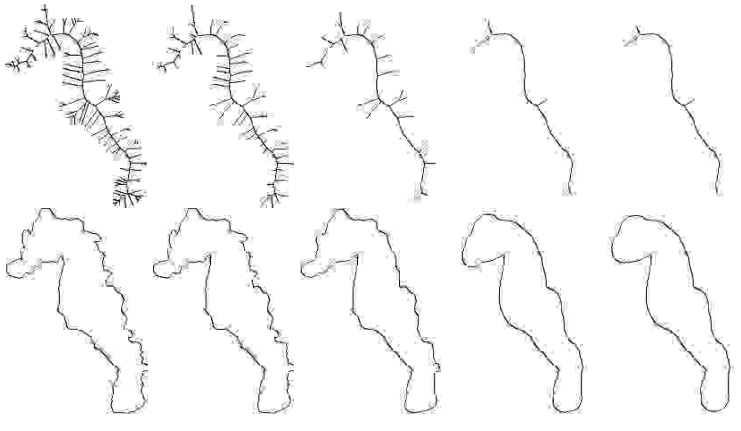

Fig. 2. Multi-scale skeletons (upper row) and corresponding region boundaries (lower row)

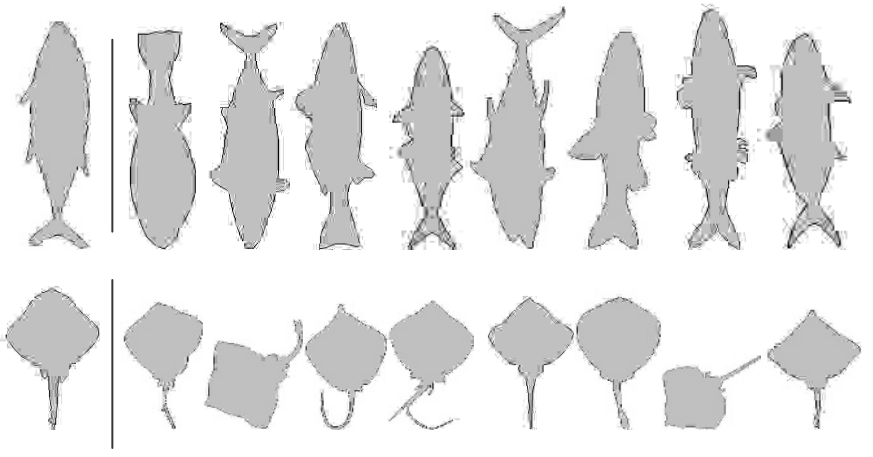

Fig. 3. Examples of image retrieval using the multi-scale skeleton feature vectors

This paper presented a shape-based image retrieval scheme which represents an object as a set of multi-scale skeleton structures. Invariant features which were obtained from multiple skeleton representations are insensitive to small boundary deformations. Hence the method overcomes the shape deformation of animated object. As future works of our research, we are developing a motion-based image segmentation algorithm which extracts objects in motion regardless of the complexity of the environment where the object is located in.

References

1. Sclaroff, S., Pentland, A.: Model matching for correspondence and recognition. IEEE Trans. Pattern Analysis and Machine Intelligence 17 (1995) 545–561
2. Mokhtarian, F., Abbasi, S., Kittler, J.: Robust and efficient shape indexing through curvature scale space. Proc. British Machine Vision Conf. (1996) 545–561
3. Celenk, M., Shao, Y.: Rotation, translation, and scaling invariant color image indexing. Storage and Retrieval for Image and Video Databases VII. SPIE 3656 (1999) 623–630
4. Ogniewicz, R.: Skeleton-space: a multiscale shape description combining region and boundary information. CVPR (1994) 746–751
5. Telea, A., Sminchisescu, C., Dickinson, S.: Optimal Inference for Hierarchical Skeleton Abstraction. IEEE ICPR (2004)
6. Arcelli, C., di Baja, G. S.: Euclidean skeleton via centre-of-maximal-disc extraction. Image and Vision Computing 11 (1993) 163–173
7. Teh, C.H., Chin, R.T.: On digital approximation of moment invariants. CVGIP 33 (1986) 583–598