

LOOKING BEYOND REGION BOUNDARIES: REGION-BASED IMAGE RETRIEVAL USING FUZZY FEATURE MATCHING

Yixin Chen James Z. Wang

The Pennsylvania State University, University Park, PA 16802, USA

ABSTRACT

For a region-based image retrieval system, performance depends critically on the accuracy of object segmentation. We propose a soft computing approach, unified feature matching (UFM), which greatly increases the robustness of the retrieval system against segmentation related uncertainties. In our retrieval system, an image is represented by a set of segmented regions each of which is characterized by a fuzzy feature (fuzzy set) reflecting color, texture, and shape properties. The resemblance between two images is then defined as the overall similarity between two families of fuzzy features, and quantified by a similarity measure, UFM measure. The system has been tested on a database of about 60,000 general-purpose images. Experiment results demonstrate improved accuracy, robustness, and high speed.

1. INTRODUCTION

In many application fields such as biomedicine, the military, education, commerce, entertainment, crime prevention and World Wide Web searching, large volume of data appear in image form. As a result, it happens quite often that a customer or operator need to find relevant information from an image database based on the contents of the images. This problem has been known as content-based image retrieval (CBIR) for more than a decade. The difficulties and complexities in quantifying “meanings” of images via feature set and designing efficient matching algorithms make CBIR a very challenging problem. There is a rich resource of prior work on this subject. For example, IBM QBIC [1], MIT Photobook [2], Virage System [3], Columbia VisualSEEK and WebSEEK [4], Stanford WBIIS [5], UIUC MARS [6], UCSB NeTra system [7], Berkeley Blobworld system [8], PicHunter system developed by Cox *et al.* [9], and PicToSeek system developed by Gevers *et al.* [10] are some of the systems that are most related to our work.

A central task in CBIR systems is similarity comparison. In general, the comparison is performed either globally using techniques such as histogram matching and color layout indexing, or locally based on decomposed regions (objects) of the images. A major drawback of the global histogram search lies in its sensitivity to intensity variations, color distortions, and cropping. The color layout indexing method is proposed to alleviate this drawback. The color layout of an image is essentially a low resolution representation of the image obtained by partitioning the image into

Yixin Chen is with the Department of Computer Science and Engineering (e-mail: yixchen@cse.psu.edu). James Z. Wang is with the School of Information Sciences and Technology and the Department of Computer Science and Engineering (e-mail: jwang@ist.psu.edu). An on-line demonstration is available at URL: <http://wang.ist.psu.edu/IMAGE/>

blocks and taking the average or dominating color of each block. In general color layout searching is sensitive to shifting, cropping, scaling, and rotation.

Recognizing these deficiencies and the fact that human beings are capable of this technically complex performance, researchers try to relate human perception to CBIR systems. In a human visual system, although color and texture are fundamental aspects of visual perceptions, human discernment of certain visual contents could potentially be associated with interesting classes of objects, or semantic meanings of objects in the image. Motivated by this intrinsic attribute of human visual perception, a region-based retrieval system applies image segmentation to decompose an image into regions, which correspond to objects if the decomposition is ideal. Since the retrieval system has identified objects in the image, it is relatively easy for the system to recognize similar objects at different locations and with different orientations and sizes.

The performance of a region-based image retrieval system depends critically on the accuracy of object segmentation. For general-purpose images such as the images in a photo library or the images on the World Wide Web, precise object segmentation is nearly as difficult as computer semantics understanding. To reduce the influence of inaccurate segmentation, Li *et al.* [11] propose an integrated region matching (IRM) scheme which allows for matching a region of one image to several regions of another image and thus decreases the impact of inaccurate segmentation by smoothing over the imprecision. The scheme is implemented in the SIMPLIcity system [12]. Nevertheless, the inaccuracies (or uncertainties) are not explicitly expressed in the IRM measure.

Semantically precise image segmentation by an algorithm is very difficult [13] [14]. However, a single glance is sufficient for human to identify circles, straight lines, and other complex objects in a collection of points and to produce a meaningful assignment between objects and points in the image. Although those points cannot always be assigned unambiguously to objects, human recognition performance is hardly affected. We can often identify the object of interest correctly even when its boundary is blurry. Based upon these observations, we hypothesize that, by softening the boundaries between regions, the robustness of a region-based image retrieval system against segmentation-related uncertainties can be improved.

2. OUR APPROACH

Our system segments images based upon color and spatial variation features using k -means algorithm, a very fast statistical clustering method. To obtain these low-level image features, the system first partitions the image into blocks with 4×4 pixels. A feature vector, consisting of six features, is then extracted for each image block. Among the six features, three of them are the av-

erage color (in the LUV color space) of the corresponding block. The other three represent energy in the high frequency bands of the wavelet transform applied to the L component of the image block.

The k -means algorithm is used to cluster the feature vectors into several classes with every class corresponding to one region in the segmented image. Blocks in each cluster does not have to be neighboring blocks in the images because clustering is performed in the feature space. The k -means algorithm does not specify the number of clusters to choose. So we adaptively select the number of clusters, C , by gradually increasing C until a stop criterion is met. The average number of clusters for all images in the database changes in accordance with the adjustment of the stop criteria.

Classically, a region is represented either by a feature set (consisting of all features vectors corresponding to the region) or a prominent feature vector (for example the center of the feature set). The former incorporates all the information available in the form of feature vectors. But it is sensitive to the segmentation-related uncertainties, and increases the computational cost for similarity calculation. The latter could, to some extent, mitigate the influence of inaccurate segmentation. But much useful information is also submerged in the process of mapping a set of feature vectors to a single feature vector.

Representing regions by fuzzy features, to some extent, combines the advantages and avoids the drawbacks of both aforementioned region representation forms. In this representation form, each region is associated with a fuzzy feature (defined by a *membership function* mapping the feature space to the real interval $[0, 1]$) that assigns a value (between 0 and 1) to each feature vector. The value, named *degree of membership*, illustrates the degree of wellness that a corresponding feature vector characterizes the region, and thus models the segmentation-related uncertainties. In this way, more information is maintained in fuzzy feature representation because each region is represented by multiple feature vectors each with a value denoting its degree of membership to the region. Moreover, the membership functions of fuzzy sets naturally characterize the gradual transition between regions within an image. To that end, they characterize the blurring boundaries due to imprecise segmentation.

A direct consequence of fuzzy feature representation is the region-level similarity measure. Instead of using the Euclidean distance between two feature vectors, a fuzzy similarity measure is used to describe the resemblance of two regions. It is defined as the maximum value of the membership function of the intersection (the union and intersection operations on fuzzy sets are extensions of classical set operations) of two fuzzy features. It is always within $[0, 1]$ with a larger value indicating a higher degree of similarity between two regions. The value depends on both the Euclidean distance between the center locations and the grade of fuzziness of two fuzzy features. Intuitively, even though two fuzzy features are close to each other, if they are not “fuzzy” (i.e., the boundary between two regions is distinctive), then their similarity could be low. In the case that two fuzzy features are far away from each other, but they are very “fuzzy” (i.e., the boundary between two regions is very blurring), the similarity could be high. These correspond reasonably to the viewpoint of the human perception.

Attempting to provide a comprehensive and robust “view” of the similarity between images, the region-level similarities are combined into image-level similarity vectors which incorporate all the information of similarities between regions. The entries of similarity vectors are then weighted and added up to generate the UFM similarity measure that depicts the overall resemblance of images

in color, texture, and shape. The comprehensiveness and robustness of UFM measure can be examined from two perspectives namely the contents of similarity vectors and the way of combining them. Each entry of similarity vectors signifies the degree of closeness between a fuzzy feature in one image and all fuzzy features in the other image. Intuitively, an entry expresses how similar a region of one image is to all regions of the other image. Thus a region is allowed to be matched with several regions in case of inaccurate image segmentation which in practice occurs quite often. By weighted summation, every fuzzy feature contributes a portion to the overall similarity measure. This further reduces the sensitivity of UFM measure.

3. SYSTEM OVERVIEW

The system is implemented on a Pentium III 700 MHz PC running Linux operating system with a general-purpose image database including about 60,000 images. These images are stored in JPEG format with size 256×384 or 384×256 . Each image is associated with a keyword describing the main subject of the image. About 20 research groups around the world have downloaded our image database¹ and use as a testbed for benchmarking purposes.

In our current system all 60,000 images in the database are automatically classified into two semantic types: textured photograph, and non-textured photograph [15]. Feature extraction and image segmentation for all images in the database are performed off-line, which takes about 60 hours. On average, it takes about one second to segment and compute the fuzzy features for an image, and another 1.5 seconds to calculate and sort the UFM measures.

Compared with other region-based retrieval systems, such as [8], the query interface, which is shown in Figure 1, for our system is very simple. It provides a CGI-based Web access interface and a



Figure 1: The query interface.

CGI-based outside query interface.

The Web access interface is designed for query images which are in the database. The system provides a *Random* option that will give a user a random set of images from the database to start with. Users can also enter the ID of an image as the query. If the user moves the mouse on top of a thumbnail shown in the window, the thumbnail will be automatically switch to its region segmentation with each region painted with corresponding color. This feature is important for partial region matching since the user may choose a subset of the regions of an image to form a query rather than

¹Available at URL: <http://wang.ist.psu.edu/IMAGE>

using the entire query image. The outside query interface allows the user to submit any image on the Internet as a query image to the system by entering the URL of the image. Our system is capable of handling any image format from anywhere on the Internet and reachable by our server via the HTTP protocol. The image is downloaded and processed by our system in real-time. The high efficiency of our image segmentation and matching algorithm made this feature possible.

4. EXPERIMENTAL RESULTS

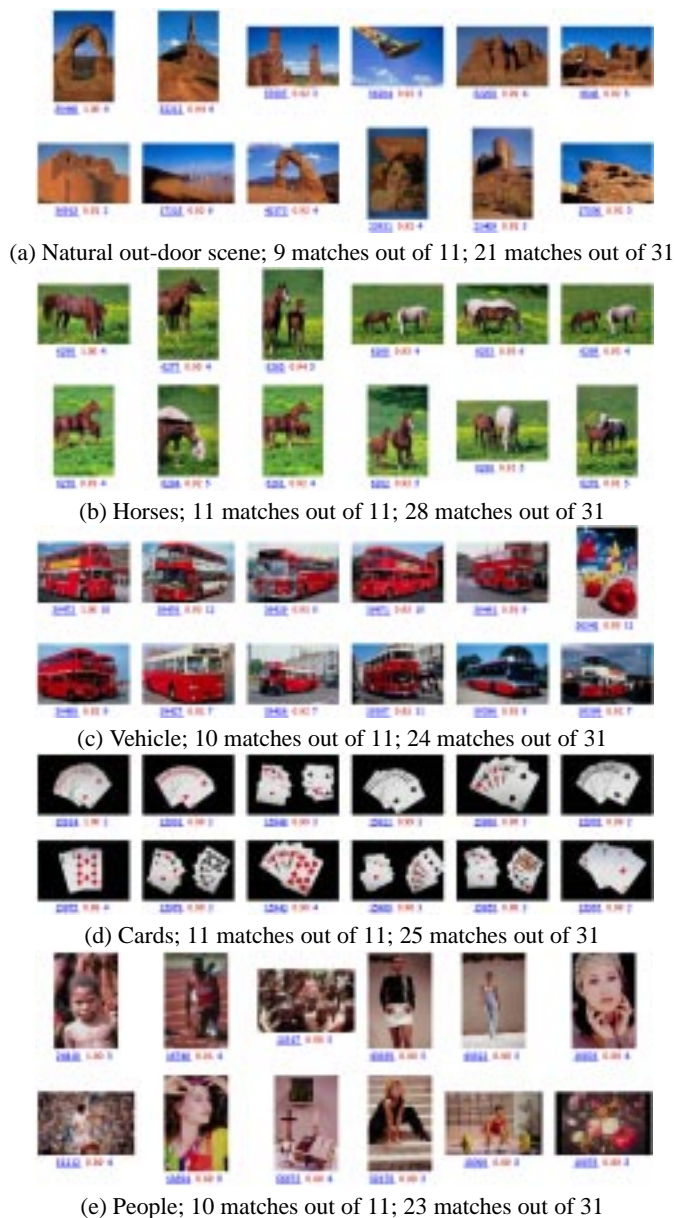


Figure 2: The query image is the upper-left corner image of each block of images.

To qualitatively illustrate the accuracy of the system over the 60,000-image COREL database, we randomly pick 5 query images with different semantics, namely natural out-door scene, horses, vehicle, cards, and people. For each query example, we manually exam the precision of the query results depending on the relevance of the image semantics. Due to space limitation, only top 11 matches to each query are shown in Figure 2. We also provide the number of relevant images among top 31 matches. More matches can be viewed from the on-line demonstration site. For each block of images, the query image is on the upper-left corner. There are three numbers below each image. From left to right they are: the ID of the image in the database, the value of the UFM measure between the query image and the matched image, and the number of regions in the image.

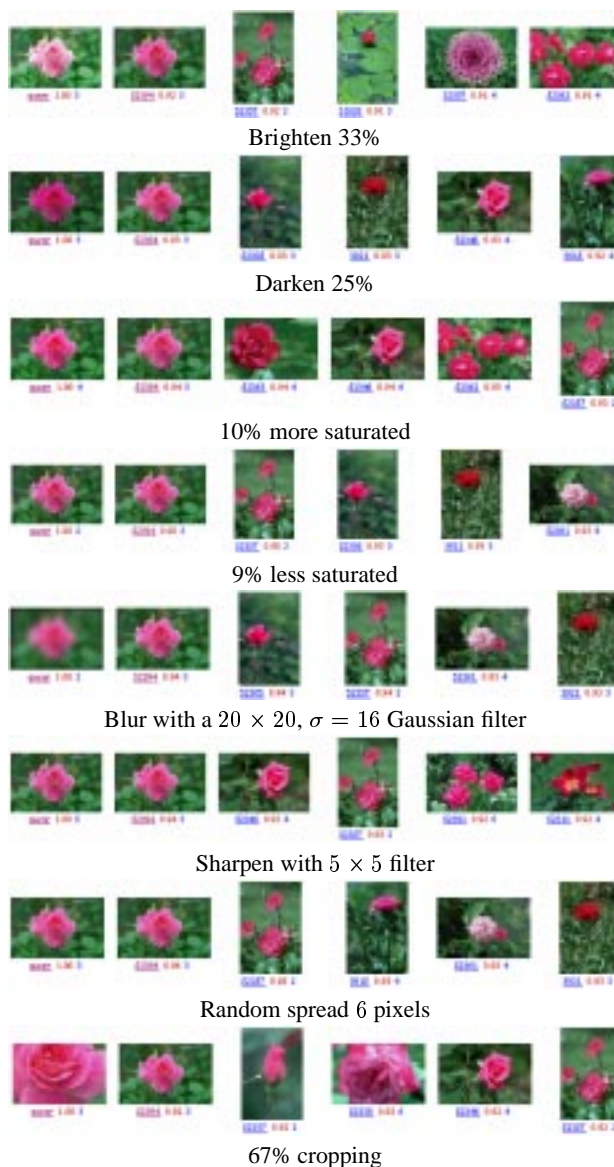


Figure 3: The robustness of the system to image alterations. The query image is the first image in each example.

The robustness of system is also tested with respect to image alterations such as intensity variation, sharpness variation, color distortion, shifting, and cropping. Figure 3 shows some query results using the 60,000-image COREL database. The query image is the left image for each group of images. In this example, the first retrieved image is exactly the unaltered version of the query image for all tested image alterations.

The system is also compared with the IRM [11] and EMD-based color histogram [16] approaches and demonstrates improved accuracy, robustness, and high speed [17].

5. CONCLUSIONS AND FUTURE WORK

We have developed UFM, a novel soft computing approach for region-based image retrieval. It is a scheme of measuring the similarity between two images based on their fuzzified region representations. The fuzzification of feature vectors brings in two crucial improvements on the region representation of an image,

- Fuzzy features naturally characterize the gradual transition between regions (blurry boundaries) within an image.
- Compared with a feature vector in a conventional region representation, a fuzzy feature incorporates more information of the region since more feature vectors are included inside with different degrees of membership. In addition, each feature vector usually belongs to multiple regions with different degrees of membership.

Based on this fuzzy feature representation, the similarity measure of two images is defined as the overall similarity between two classes of fuzzy features. The UFM measure we have developed has two major advantages:

- Compared with retrieval methods based on individual regions, the UFM approach reduces the adverse effect of inaccurate segmentation, and make the retrieval system more robust to image alterations.
- Compared with overall similarity approaches, such as that proposed in [11], the UFM is better in extracting useful information under the same uncertain conditions.

Experiments have shown the accuracy of the scheme and the robustness to image segmentation and image alterations.

Currently, we are working on statistical modeling based image comparison, statistical feature-space selection, feature clustering scheme for region-based retrieval, and biomedical applications. We will continue to distribute our 60,000-image test database as a benchmark testbed for image retrieval research.

6. ACKNOWLEDGMENTS

This work was supported in part by the US National Science Foundation, the PNC Foundation, and SUN Microsystems. The authors would like to thank Jia Li for valuable discussions.

7. REFERENCES

[1] C. Faloutsos, R. Barber, M. Flickner, J. Hafner, W. Niblack, D. Petkovic, and W. Equitz, "Efficient and Effective Querying by Image Content," *J. Intell. Inform. Syst.*, 3(3-4):231–262, July 1994.

[2] A. Pentland, R.W. Picard, and S. Sclaroff, "Photobook: Content-Based Manipulation for Image Databases," *Int. J. Comput. Vis.*, 18:233–254, 1996.

[3] A. Gupta and R. Jain, "Visual Information Retrieval," *Commun. ACM*, 40(5):70–79, May 1997.

[4] J.R. Smith and S.-F. Chang, "VisualSEEK: A Fully Automated Content-Based Query System," *Proc. ACM Multimedia'96*, pages 87–98, 1996.

[5] J.Z. Wang, G. Wiederhold, O. Firschein, and X.W. Sha, "Content-Based Image Indexing and Searching Using Daubechies' Wavelets," *Int. J. Digital Libraries*, 1(4):311–328, 1998.

[6] S. Mehrotra, Y. Rui, M. Ortega-Binderberger, and T.S. Huang, "Supporting Content-Based Queries over Images in MARS," *Proc. IEEE Int. Conf. on Multimedia Computing and Systems*, pages 632–633. Ottawa, Ont., June 1997.

[7] W.Y. Ma and B.S. Manjunath, "NeTra: A Toolbox for Navigating Large Image Databases," *Proc. IEEE Int. Conf. Image Processing*, pages 568–571, 1997.

[8] C. Carson, M. Thomas, S. Belongie, J.M. Hellerstein, and J. Malik, "Blobworld: A System for Region-Based Image Indexing and Retrieval," *3rd Int. Conf. on Visual Information Systems*, June 1999.

[9] I.J. Cox, M.L. Miller, T.P. Minka, T.V. Papatomas, and P.N. Yianilos, "The Bayesian Image Retrieval System, PicHunter: Theory, Implementation, and Psychophysical Experiments," *IEEE Trans. Image Processing*, vol. 9, no. 1, pp. 20–37, 2000.

[10] T. Gevers and A. W. M. Smeulders, "PicToSeek: Combining Color and Shape Invariant Features for Image Retrieval," *IEEE Trans. Image Processing*, vol. 9, no. 1, pp. 102–119, 2000.

[11] J. Li, J.Z. Wang, and G. Wiederhold, "IRM: Integrated Region Matching for Image Retrieval," *Proc. 8th ACM Int. Conf. on Multimedia*, pages 147–156. Los Angeles, CA, October 2000.

[12] J.Z. Wang, J. Li, and G. Wiederhold, "SIMPLiCity: Semantics-Sensitive Integrated Matching for Picture Libraries," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 23, no. 9, 2001.

[13] W.Y. Ma and B.S. Manjunath, "EdgeFlow: A Technique for Boundary Detection and Image Segmentation," *IEEE Trans. Image Processing*, vol. 9, no. 8, pp. 1375–1388, 2000.

[14] S.C. Zhu and A. Yuille, "Region Competition: Unifying Snakes, Region Growing, and Bayes/MDL for Multiband Image Segmentation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 18, no. 9, pp. 884–900, 1996.

[15] J. Li, J.Z. Wang, and G. Wiederhold, "Classification of Textured and Non-Textured Images Using Region Segmentation," *Proc. 7th Int. Conf. on Image Processing*, pp. 754–757, Vancouver, BC, September 2000.

[16] Y. Rubner, L.J. Guibas, and C. Tomasi, "The Earth Mover's Distance, Multi-Dimensional Scaling, and Color-Based Image Retrieval," *Proc. ARPA Image Understanding Workshop*, pages 661–668, New Orleans, LA, May 1997.

[17] Y. Chen, J.Z. Wang, and J. Li, "FIRM: Fuzzily Integrated Region Matching for Image Retrieval," *Proc. 9th ACM Int. Conf. on Multimedia*, Ottawa, Canada, September 2001.