FINAL PROJECT
An Image Retrieval System for Cloth Images

Abstract

In this study, the design and implementation of an image retrieval system for cloth images is studied. This study proposes a general purpose image retrieval system and extends this idea to allow searching and retrieving images, which include cloth, based on the criteria that is determined by users. The criteria/query word in this project is image-based so that user searches images in database by feeding the system with image. General purpose image retrieval system is implemented with novel methods and an idea for cloth searching is proposed in this study. Experiments show that proposed method for image retrieval has promising results.

Introduction

This study proposes two systems, one is for general purpose image retrieval and other is for cloth-based image retrieval. General purpose image retrieval system is designed and implemented with a novel approach for histogram discrimination. Cloth-based search is proposed as design and not implemented in this study.

A general purpose image retrieval system has two operations. It indexes the image database and retrieves images associated with query image from database. Indexing the image database is realized by extracting characteristics information from each image in database. This information is called as fingerprint and it's a kind of summarization of the image. Efficient image retrieval systems keep this information as compact as possible so that its allocation on disk is feasible for millions of images in database and it's fast to evaluate the similarity of images in database with query image given by user. In other words, an efficient image retrieval system should consider computation in time/space.

Cloth-based image retrieval is a customized version of general purpose of image retrieval system. It indexes the cloth regions in the image instead of whole image by extracting fingerprints for those regions only. And, it compares fingerprints of cloth regions detected in the query image with fingerprints of each image in database to retrieve closer images.

There are three problems to be solved in this study. Images should be represented by an effective summarization method so that both indexing and retrieving are feasible. A similarity criteria/discrimination method should be decided for good precision/recall. And a cloth segmentation method should be determined to detect cloth regions in the images.

Problems mentioned above is solved by deciding on colour model, image fingerprint type, histogram discrimination method and designing a cloth segmentation method. It's also necessary to find so many in-cloth images so that performance of the systems can be evaluated with real-world data.

Related Works

Literature is reviewed for image retrieval systems and cloth segmentation separately. In [1], colour regions in image are extracted by back-projecting of binary colour sets. The method proposed here allows automatic extraction of regions. In [2], six different image retrieval approaches are examined for two different colour spaces: RGB and HSV. According to report, HSV colour space yields better results compared to RGB colour space. In addition, among six approaches such as histogram euclidean and quadratic distances, histogram intersection-based distance outperforms for comparison of the images. In [7], two methods for "Near Duplicate Image Detection (NDID)" are proposed: 1) Colour histogram (CH) and Locality Sensitive Hashing (LSH), 2) SIFT and MinHash. CH/LSH, which is only approach examined for this study, uses Opponent Colour Model with spatial representation [8], euclidean distance for descriptors and LSH for effective retrieving. [3] is a study on cloth segmentation, the method proposed here segments colours in the image into clusters by utilizing from an ant colony clustering algorithm. The performance of this method is evaluated with only one image and it doesn't give clue about the accuracy. In [4], the cloth segmentation is achieved in real-time by using hue and intensity histograms with spatial priors. The method proposed here has a 97% of F-score value for segmentation.

Method

Image retrieval system uses HSV as colour model for representation of intensities in the images. In [2], HSV is said to be better than RGB and this result is justified in the study. As image fingerprint, 64-bin HSV colour histogram is used so that each image is represented by a 64-dimensional feature vector. 64-dimensional vector is fast to create and hence indexing is realized in short time. Index file size is also reasonable because of shortness in feature vector. For 5120 images in database, it's observed that an index file with size of 2.7 MB is created in 4-5 minutes.

Three histogram discrimination methods are used and tested to intersect histograms of the images. At the beginning, intersection of corresponding bins in histograms of query and reference images are summed and then result is normalized by histogram magnitude whose number of samples is minimum.

$$\sum_R \sum_G \sum_B min(h(r,g,b), g(r,g,b)) / min(|h|, |g|)$$

where $|h|$ and $|g|$ are magnitudes of histograms h and g as number of samples in those images respectively. Assuming f_h is the image in database and f_g is the query word, after histogram intersection distance of histograms h and g for images f_h and f_g is evaluated and thresholded with a predefined value, decision on relativity of two images is made and f_h is retrieved based on this decision. This approach didn't yield satisfied results because it usually has high similarity results even for non-similar images in database. High similarity values are cause of taking minimum on both of nominator and denominator. By considering this, sum of intersection is normalized by histogram of reference image only as second discrimination method.

$$\sum_R \sum_G \sum_B min(h(r,g,b), g(r,g,b))/|h|$$

This approach had better results but it still lacks of accurate retrieval results. As a final method, sum of intersection as percentage is considered.

$$\sum_R \sum_G \sum_B min(h(r,g,b)/|h|, g(r,g,b)/|g|)$$

This approach had satisfied retrieval results. Because first two methods didn't consider differences in size of images. Query image may be much larger than reference image and this leads to abnormal increase of some intensities and intersection of those histogram bins is not reasonable. Intersection by percentage overcomes such issue.

Cloth segmentation is realized by locating face in the image and then predicting location of cloth according to face. For images in database, a 4x4 grid is opened on the image to divide it to 16 sub-images. Haar-Cascade face detector is applied to each sub-image to check whether or not one face, at least, exists in it. If this is the case then sub-image below the region the face exists is considered as cloth region. Here, assumption is that there is human body below the face and it's covered by cloth. Histogram of cloth region is extracted and indexed as fingerprint. For query image, a 4x4 grid is opened and cloth regions are detected by method as in database images. Once histogram for query image is extracted, it's compared with indexed histograms/fingerprints and similar results are retrieved.

Image retrieval system is implemented with Python [9] and SimpleCV [10] is utilized for image processing operations. There is no "Database Management System (DBMS)" used to store reference images, instead, all reference images are stored in file system as directories. Indexing is realized in an index file so that index is portable. Performance evaluation is realized by recall, precision and F-score.

Dataset

Two datasets used in [4] are considered for testing the system implemented in this study. One is called as Soton dataset [5] and other is called as “Images of Groups” dataset [6]. "Images of Groups" dataset is decided to be enough for testing of the system. There are 4992 images with group of people in "Images of Groups" dataset. It's very suitable for cloth retrieving system.

In addition to "Images of Groups" dataset, 4x7=28 images are added for testing of image retrieval system. These images are created by using 4 images from "Images of Groups" dataset as query and reference images and modifying them as cropping, blurring, text-inserting, bordering, resizing in both directions.

test3.jpg

test4.jpg/test5.jpg

test8.jpg

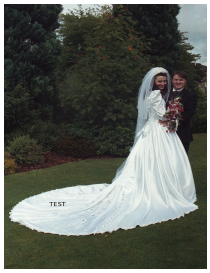
Figure 1. Query images

Modified versions of one query image can be reviewed below:

Blur

Border

Crop



Text-insert

Original

Resizes

Figure 2. Modifications on query images

Results

Performance in this study is evaluated by “recall” and “precision” measures. Recall is the ratio of the number of detected frames by the algorithm to total number of shots in the video. Precision gives how much of the detection by the algorithm is correct. The formulas for these measures are as below:

$$Recall = D/(D + D_M)$$
$$Precision = D/(D + D_F)$$

where D is the number of correctly detected shot transition, D_M is the number of missed detection and D_F is the number of false detection by the algorithm.

For query of test3.jpg, blur and border images are missed; recall is 5/7=0.71, precision is 5/5=1 and F-score is 10/12=0.83. For query of test4.jpg, only text-insert image is missed; recall is 6/7=0.85, precision is 6/35=0.17 and F-score is 2/7=0.28. For query of test5.jpg, blur and border images are missed; recall is 5/7=0.71, precision is 5/5=1 and F-score is 10/12=0.83. For query of test8.jpg, text-insert image is missed; recall is 6/7=0.85, precision is 6/6=1 and F-score is 12/13=0.92. Recall vs. precision plot can be reviewed below:

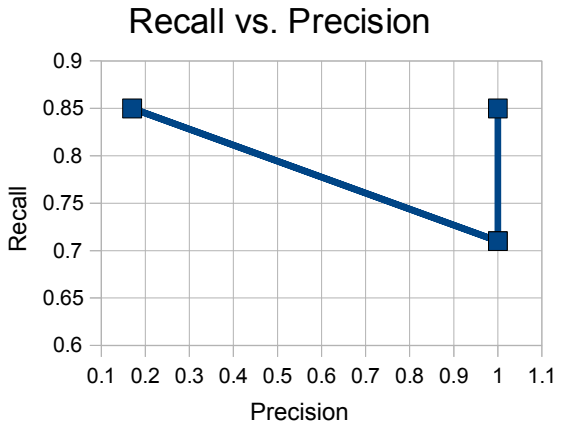


Figure 3. Recall vs. Precision

Conclusions

In this study, a general purpose image retrieval system is designed and implemented. Image retrieval system is customized so that cloth images can be searched based on cloth colours. Cloth-based search is an extension of this study and the idea is proposed in the paper. This study proposes novel histogram discrimination and cloth segmentation methods. Image retrieval system has satisfactory recall results and very good precision results except one case.

References

[1] J. R. Smith and S. F. Chang, “Tools and techniques for color image retrieval,” IST/SPIE-Storage and Retrieval for Image and Video Databases IV, San Jose, CA, 2670, 426-437, 1996.
[2] Sangoh Jeong, “Histogram-Based Color Image Retrieval”, Project Report, 2001.
[3] Xinrong Hu, Naixue Xiong, Shuqin Cui, Wang Hui, Jianchun Wang, “A Color Clustering Algorithm for Cloth Image,” APSCC, 1500-1505, 2008.
[4] George A. Cushen, Mark S. Nixon, “Real-Time Semantic Clothing Segmentation,” ISVC (1), 272-281, 2012.
[5] Seely, R.D., Samangooei, S., Lee, M., Carter, J.N., Nixon, M.S.: The University of Southampton Multi-Biometric Tunnel and introducing a novel 3D gait dataset. In: BTAS, IEEE (2008) 1-6.
[6] A. Gallagher, T. Chen, “Understanding Groups of Images of People,” IEEE Conference on Computer Vision and Pattern Recognition, 2009.
[7] Ondrej Chum, James Philbin, Michael Isard, Andrew Zisserman, “Scalable near identical image and shot detection,” Proceedings of the 6th ACM International Conference on Image and Video Retrieval, Amsterdam, Netherlands, pp.549-556, July 09-11, 2007.
[8] S. Lazebnik, C. Schmid and J. Ponce, “Beyond bag of features: Spatial pyramid matching for recognizing natural scene categories,” In Proc. CVPR, 2006.
[9] Python, http://www.python.org/
[10] SimpleCV, http://simplecv.org/