POSTER IMAGE MATCHING BY COLOR SCHEME AND LAYOUT INFORMATION

Cheng-Yao Chen, Takayuki Kurozumi, and Junji Yamato

NTT Communication Science Laboratories, NTT Corporation
3-1, Morinosato Wakamiya, Atsugi-shi, Kanagawa, Japan 243-0198

ABSTRACT

In this paper, we demonstrate a novel poster image matching system for wireless multimedia applications. We propose a method that incorporates both color and layout information of the poster image to achieve a robust performance in poster image matching. We apply both color compensation and background separation to extract a poster from an image effectively. Based on our experiment, we show that even under the effects of lighting, image rotation, scaling, and occlusion, our system can still maintain high recall and precision. We also show that our system can recognize the correct image from a database which contains several poster images with similar features. Finally, the promising performance of the poster image matching encourages us to further enrich the information retrieval for wireless environment.

1. INTRODUCTION

Information retrieval by cell phone through wireless network has become a trend in recent years. However, typing text queries by cell phone keypads is not comfortable and convenient for everyone. An alternative is querying by images. Image querying has been extensively studied since the last two decades. It turns out to be more efficient and natural than text queries in several cases. With the advances of recent technology, the digital camera becomes more and more popular within the cell phone. Taking pictures by the camera from the cell phone also becomes an inevitable usage. As a result, we can expect that information retrieval by mobile phone camera images will be needed soon.

Posters are used everywhere for movies, activities, announcements, etc. Sometimes if people are interested in the event that the poster is about, all they can do now is remember the event and look it up later by text. This is not efficient and convenient. Here we hope to design a poster image matching system that allows users to take pictures when they see a poster on the wall, on a TV screen or any other displays, and learn more information about the event. By sending poster images to the server through wireless network, we can provide more details and follow up messages to users. The most important step of succeeding in this application is the accuracy of recognizing the poster from an image and finding of the correct match from the poster database.

The key of a successful information retrieval by image query is image matching or image indexing. Generally speaking, image matching can be divided into three steps, image feature selection, matching of image features, and evaluation of similarity distance. Image feature selection basically includes color [1], shape [2], and texture [1]. Matching of image features determines how to compare the selected image features, including direct pixel comparison, sub-image comparison [1], histogram [3], singular value decomposition [1], and probability distribution [4]. Methods for evaluating similarity include Euclidean distance, Hausdorff distance [5], sum of square difference, and Mahalanobis distance [2]. The selection of methods in each step depends on the characteristics of target images. As a result, we have to specify the properties of a poster and find a set of features which characterizes posters best.

Posters have their own special features. First, they usually have a dominant background color. If a poster is created carefully, this background color should be easily distinguished from foreground objects. Second, there are always some complicated parts within a poster, and those parts are usually text descriptions. Those can be good blocks to describe the layout of the poster. Third, a poster usually has a rectangular shape. This feature can help when we try to separate the poster from an image with mixed objects.

2. METHODOLOGY

The system block diagram is shown as Figure 1. Our system can be separated by five stages, color compensation, background subtraction, color analysis, layout information analysis, and similarity evaluation. First, we use a log-shape nonlinear color adjustment to compensate the poster image color shifting due to the environment color shifting. Continuing that, we locate the poster within an image by background subtraction. Following that, we summarize the color scheme of the poster image with spatial information embedded. Fourth, we describe the layout information for each poster by a layout histogram which is rotation- and scaling- free. Finally, we define the similarity between the query image and the database image by combination of criteria in the similarity evaluation stage. Details of each stage will be given in the following subsections.
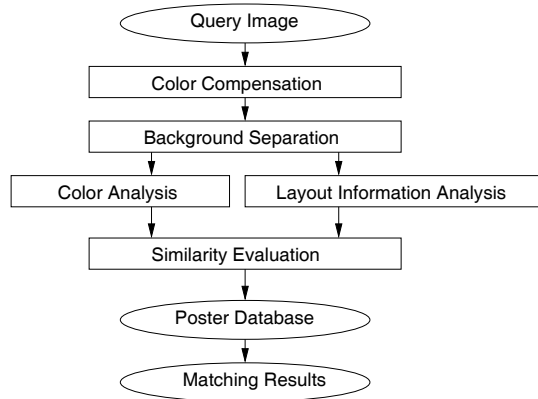
ICME 2006

Fig. 1. System Block Diagram

2.1. Color Compensation

Color compensation is very important for image processing since shifts of color can result in totally different images. As we mentioned before, we will use color information as one of our similarity metrics so we can not ignore this effect. There are three sources of color shifting, one is from the environment lighting, another is from the display characteristics, and the other is from the camera characteristics. All the effects mentioned above are non-linear. Since it is difficult to separate each factor, we apply a single gamma correction to characterize the non-linear color changing process by the following equation.

$$P_{modified} = \alpha + \log(P_{original} + 1)/\beta$$

where, $P_{original}$ and $P_{modified}$ are the pixel value of each channel respectively. α and β are the curve fitting parameters. We first use a least square error to determine the curve parameters by training on all the images within the database. Each query image is then modified based on the curve we obtained by the training results of the database. In this way, we can achieve a closer color distribution from the poster image database.

2.2. Background Separation

Background separation is one of the most important factors for the performance of the poster image matching. However, there is no known effective algorithm that works well for all cases without preliminary knowledge about the target [6]. Here we will not try to propose a global solution for all kinds of images, but instead we will introduce a good solution for poster images.

In order to effectively separate the poster region from a given image, we have to make some assumptions. First, we assume that the poster region is relatively more complicated than the whole image. This complexity can be from either text or colorful representation of a poster. Second, the shape of the

poster is close to rectangular. Third, there is only one poster image and it is larger than other objects in the image. Here we do not deal with multiple posters within an image, and we assume that when users are trying to get the information for a poster, they will capture the poster as the main subject within the image.

Based on the assumptions, we first perform a corner detector for the whole image with a four-level pyramid. Then we divide the image into sub-regions, and for each sub-region we calculate the average density of the corner. We treat this average density as the index for the complexity. Then we model all the blocks into a Gaussian distribution, and select blocks which have a density greater than two standard deviations from the mean. We consider those blocks are more complicated than other regions and could be possible candidates of the poster image. Then we project those blocks to the corresponding blocks in the original size image. After identifying possible poster image blocks, we perform morphological operations to re-link broken fragments and fill in holes. Continuing that, we find all the connected components and only pick up the largest block as the rough region of the poster image. With this rough region, we then further refine the boundary of the poster by three metrics.

First, we move each boundary inwards and outwards, and calculate the new average detail density for the whole region. Second, we again move each boundary inwards and outwards, and calculate the average color error in hue domain compared to the original rough poster region. Third, we perform Canny edge detection with non-maximum suppression in eight ways, and locate the four longest edges in pair near the original boundaries. Then we apply a weighted threshold for those three metrics to determine the final poster boundary. Modifying boundaries in this recursive way is based on the features we observe from posters.

2.3. Color Analysis

Huang et al [7] proposed a color correlogram algorithm which incorporates both color and spatial information with a single representation. Since color is not the only information that we apply for the matching, we only use one-pixel-distance correlogram representation to save processing power and keep the feature size compact. We first convert our poster image from RGB to HSV color space. Then we quantize the hue value of each pixel to one of the 18 steps. There is a trade-off between larger and smaller number of steps. Since we hope to get a higher recall rate by color, we use a smaller number. It results in a 18 by 18 co-occurrence matrix. We only choose 18 hue values for our application because it is efficient enough based on the experiment.

Based on the definition of the correlogram, the diagonal terms for distance one contain information close to the color histogram and the off-diagonal terms are related to the color edge pairs. Those off-diagonal terms will give us some fur-
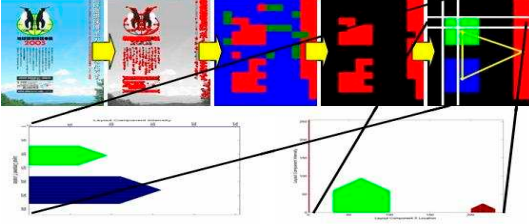
Fig. 2. Sample Operation of Layout Information Analysis

ther information of the whole color scheme which is crucial to enhance the precision of the color. Here we store both the diagonal and off-diagonal terms to serve as our main color features, and also the rank order of the diagonal terms as supplementary color features.

2.4. Layout Information Analysis

Since posters usually have text layout, it will be useful if we can describe the layout information and serve as one metric for the matching. We again first perform the corner detection to capture distinguished parts and connect blocks within the poster region. Then we select the largest three blocks and record their size, location, average color within the block. Here we propose a layout histogram to translate the geometry and size information of the layout to a one dimension representation.

The layout histogram transformation is designed as follows. First, we divide the poster into stripes both horizontally and vertically. Second, whenever there is a layout component within the stripes we have a histogram bin based on the location and the average color. The height of the bin is determined by the normalized distance to the center of the layout component and the average color of it. The closer the distance to the center, the higher the intensity. If there is no layout component, then the intensity of the bin at that location is zero. All the calculation of the layout histogram is scaling free because all the intensities are normalized to the size of the poster. Also since part of the poster can be occluded, we assume that relative long side and short side information is not changed by the occlusion, and then we can set up the length of the histogram by the short side. For rotation free checking, we just compare two settings, one uses top-left origin and the other uses the right-bottom origin. Assuming there is no exactly two reverse layout settings for two posters, we select the smallest one, and is thus rotation-free. This is especially useful for a up side down poster. The illustration of layout information analysis is shown as Figure 2.

2.5. Similarity Evaluation

We have two feature histograms now for us to calculate the similarity measure for each query image. Since we have multiple features and it is not easy to discover higher perceptual

Fig. 3. Sample Variations for Poster Image Database

relationship between them, we decide to use a linear distance calculation to define the final similarity. In order to fairly treat those features the same, all the features are normalized to their relative scales. Each entry in the color correlogram is divided by the total number of pixels times 4 because we calculate the relationship with eight neighbors. For layout information, all the size of the block is normalized to the size of the poster captured. Then the similarity measure is based on the following equation.

$$Similarity = D(color) \times C_{color} + D(layout) \times C_{layout}$$

Where C_{color} and C_{layout} are fixed first for all posters but then adjusted based on the matched order of dominant color and number of matched stripes (top left or right bottom) in the previous stage respectively. Then we find the best match of this similarity criterion by an adaptive threshold according to the image quality calculation from the previous steps.

3. EXPERIMENT RESULTS

The poster database is composed of 24 major image styles where they use the same background in pair, and for each style we have 4 color variations and 2 layout variations. This results in 24 x 4 x 2 = 192 images. This database is challenging since there are several similar features between posters. Moreover we randomly capture 10,000 images from video clips to serve as confusing data. Figure 3 shows the variations of one style example.

In order to comprehensively test our poster image matching system, we simulate many practical and difficult query images for our testing image database. We use four types of cameras, two are cell phone cameras, one is regular digital camera, and one is digital video recorder. We also simulate two types of lightings, fluorescent and pan yellow. We show our posters on five types of display, LCD, plasma, CRT, paper, and projector. There are three orientations for the poster, 0, 45, and 90 degree. The occlusion of test includes two types of images, normal and 25% occlusion from the longer side. This setting adds one more variation. Two sizes of the poster are also tested, large (occupy the whole image) and normal. It again adds one more variation. All the mentioned settings are tested with 12 types of the posters. Thus, the total number of the testing images is 12 x 4 x 2 x 5 x (3+1+1) = 2400. Sample testing images are shown as Figure 4.

Table 1 summarizes the performance of our testing queries with each step separated. The complete matching solution includes color compensation, background separation, color

Fig. 4. Sample Images for Testing Image Database

analysis, and layout information analysis. In order to see the effects of each step, we also compare the performance when only one step is taken away. We achieve an average 85% recall and 86% precision. The average cumulative classification(ACC) rate of top 5 candidates is 93%. This performance would be efficient enough since showing 5 candidates on a small screen is still manageable. We found it difficult to tune the threshold for recall and precision since the image quality and condition varies a lot between query images. However, we observed that changing the similarity weighting of color information will affect the recall rate more while changing the weighting of layout information will affect the precision rate more. Based on the experiments results, we can also observe that background separation plays the most important role in the whole matching process. Without the background separation the matching will fail in general. As for the performance of similarity evaluation, the combined method outperform the others (color only and layout only). When a lot of image retrieval systems are based on the conventional color histogram, we can expect our system will outperform at least more than 20% in both recall and precision. Moreover, we can find out that without the layout information, the recall rate doesn't drop substantially, but the precision is seriously degraded. This matches the characteristics of our poster database, since we have several posters sharing the same background. However, increasing the weighting for layout information is risky because it is based on small details which are not detected robustly compared to general color information.

Some cases do not have good matching performance and the reasons can be divided into three categories. First, the image quality is very poor. For example, when capturing from CRT, there will be large dark stripes in the image due to the differences of sampling frequency of cameras. Second, 45 degree orientation images are not well separated from the background, and thus contain too much confusing information for both color and layout. Third, the lost information of 25% occlusion images sometimes results in lower precision rate.

4. CONCLUSION

In this paper, we introduce a novel poster image matching system. We demonstrate that after our color compensation and background separation, the query image features will be more robust to the effects of lighting, and mixed objects in the scene. We also show that our color and layout similarity evaluation are effective enough to distinguish similar posters and confusing data from the database. Future work includes a

Table 1. Matching Performance Analysis with Comparisons When without One Step Each Time

Method	Recall (%)	Precision (%)	Cumulative Classification Rate(%)
w/o Color Compensation	54.32	55.14	63.12
w/o Background Separation	26.42	31.71	34.09
Color Only	65.23	49.55	73.44
Layout Only	53.74	67.10	69.15
Complete Solution	85.13	86.21	93.02

better solution for background separation with rotation effect better alleviated, and a more detailed but still robust layout description to distinguish even smaller differences, e.g. font difference. Moreover, matching performance of real posters captured from movies, announcements, and activities will be evaluated. Although there are still refinements needed to be done, the promising results encourage us to extend the current wireless poster information retrieval to the next level.

5. REFERENCES

[1] C. Carson, M. Thomas, S. Belongie, J. M. Hellerstein, and J. Malik, "Blobworld: A system for region-based image indexing and retrieval," *Int'l Conf. on Visual Information and Information Systems*, pp. 509–517, 1999.

[2] G. C. H. Chuang and C. C. J. Kuo, "Wavelet descriptor of planar curves: Theory and applications," *IEEE Trans. on Image Processing*, vol. 5, no. 1, pp. 56–70, 1996.

[3] R. Rickman and J. Stonham, "Content-based image retrieval using colour tuple histograms," *SPIE Int'l Conf. on Storage and Retrieval for Image and Video Databases*, pp. –, 1996.

[4] B. Moghaddam, C. Nastar, and A. Pentland, "A bayesian similarity measure for direct image matching," *ICPR*, pp. 350–358, 1996.

[5] S. H. Kim and R. H. Park, "An efficient algorithm for video sequence matching using the modified hausdorff distance and the directed divergence," *IEEE Trans. on Circuits and System for Video Technology*, vol. 12, no. 7, pp. 592–596, 2002.

[6] Y. Rui, T. S. Huang, and S. F. Chang, "Image retrieval: Current techniques, promising directions, and open issues," *Journal of Visual Communication and Image Representation*, vol. 10, pp. 39–62, 1999.

[7] J. Huang, S. R. Kumar, M. Mitra, W. J. Zhu, and R. Zabih, "Image indexing using color correlograms," *IEEE CVPR*, pp. 762–768, 1997.