Color Features for Image Fingerprinting

Marios A. Gavrielides, Elena Sikudova, Dimitris Spachos, and Ioannis Pitas

AIIA Laboratory, Computer Vision and Image Processing Group,
Dept. of Informatics, Aristotle University of Thessaloniki,
Box 451 Thessaloniki, GR-54124, Greece

Abstract. Image fingerprinting systems aim to extract unique and robust image descriptors (in analogy to human fingerprints). They search for images that are not only perceptually similar but replicas of an image generated through mild image processing operations. In this paper, we examine the use of color descriptors based on a 24-color quantized palette for image fingerprinting. Comparisons are provided between different similarity measures methods as well as regarding the use of color-only and spatial chromatic histograms.

1 Introduction

Image fingerprinting (or perceptual hashing) refers to the extraction of a unique description of an image that would be resilient to transformations, in an analogous manner to human fingerprints. Detecting transformed versions of images could be used to fight piracy of such material.

Image fingerprinting differs from image *watermarking* in the sense that watermarking involves embedding information into the image, whereas fingerprinting, as defined here, involves descriptors extracted from the image content. Fingerprinting can be used to search for those copies of an image that have already been circulating in the internet with no watermarks embedded on them.

In order for such a system to be successful, it has to be robust against a number of frequent attacks, have good discriminating ability to avoid the retrieval of false alarms, provide efficient storage of extracted image descriptors that would be used for image matching and an efficient search mechanism that would compare the image description of the query image to those in a database of image descriptors. A number of approaches [1], [2], [3] have been presented in the literature.

In this paper, we examine the use of color-based descriptors for an image fingerprinting system and its robustness to most common attacks. We demonstrate the effect of various color descriptors including color-only and color-spatial information, reduced number of colors, and different types of histogram-similarity measures.

2 Color-Based Fingerprint Extraction

The fingerprint extraction procedure involves the quantization of the image colors and the calculation of color histograms based on the resulting colors. We used a quantization method based on a pre-defined color palette known as the Gretag Macbeth Color

G. Antoniou et al. (Eds.): SETN 2006, LNAI 3955, pp. 494–497, 2006.

Checker, which was designed to match human perception of colors. The Macbeth chart is a standard used to test color reproduction systems and it consists of 24 colors, scientifically prepared to represent a variety of different naturally occurring colors. The procedure for the generation of the color chart is reported in [4]. We created a color palette based on the Macbeth xyY values found in [5] in Table G.10.

Color histograms have been used extensively in CBIR systems due to the simplicity of their calculation and their robustness to common image transformations. Many types of histograms exist in the literature falling mainly into two categories: those based only on the quantized colors and those incorporating information on the spatial color distribution. We examined the use of both kinds of color histograms. The first one was the *normalized color-only histogram*, providing probability of the occurrence of a certain color in an image. The normalized color histogram depends only on the color properties of an image without providing any information on the spatial distribution of colors. In order to examine any advantages of using histograms incorporating color-spatial information for image fingerprinting, we also experimented with the *spatial chromatic histogram* proposed by Cinque et.al. [6]. The spatial chromatic histogram descriptor gives information on color presence, and color spatial distribution.

3 Color-Based Fingerprint Matching

Image matching between color histogram descriptors depends on the choice of similarity measures, so we investigated the use of three different measures given below to match the normalized color histograms. A fourth measure was used for the spatial chromatic histogram. The matching measures used were *Scaled* L_1*-norm* and L_2*-norm* distance, defined as $\mathrm{d}_{L_1}(H_1, H_2) = 1 - 0.5 * \sum_{i=1}^{C_p} |H_{1_i} - H_{2_i}|$ and $\mathrm{d}_{L_2}(H_1, H_2) = 1 - \frac{1}{\sqrt{2}} * \sqrt{\sum_{i=1}^{C_p} (H_{1_i} - H_{2_i})^2}$.

The above two measures are scaled versions of the L_1-, and L_2-norms which have been previously used for matching color histograms.

Scaled Histogram Intersection defined as $\mathrm{d}_{HI}(H_1, H_2) = \sum_{i=1}^{C_p} \min(H_{1_i}, H_{2_i}) * (1 - |H_{1_i} - H_{2_i}|)$ is a modified version of the Histogram Intersection measure. Only colors present in the image contribute to this metric.

Finally, in order to compare the spatial chromatic histograms between images I_1, I_2, we used the *spatial chromatic distance* defined in [6].

4 Results and Discussion

A database of 450 art images of variable sizes, provided by the Bridgeman Art Library, was used to evaluate the method. The following set of 20 transformations was applied to each image: Scaling (25,50,75,125,150 and 200%), Rotation (10°, 20°, 30°, 90°), Cropping - (both sides by 10, 20 and 30%), Compression - (JPEG with 25, 50, 75 quality factor), Blurring (median with 3x3, 5x5,7x7 masks), and Combination of attacks (Rotation 10°, cropping 10%, resizing to 25%, median filtering 5x5 and compression with quality factor 50). The images were resized using nearest neighbor interpolation. For the image rotation, it has to be noted that a black frame was added around the image, thus producing an additional source of degradation, except for the case of 90°. The set of 450 original images defined the Original Image Set and the resulting set of 9000 transformation images defined the Transformed Image Set.

The use of color-based descriptors for the application of image fingerprinting was evaluated using Receiver-Operator Characteristic (ROC) analysis. Specifically, the evaluation consisted of taking the color descriptors from each of the 450 images in the Original Image Set and matching them against the descriptors from each of the 9000 images in the Transformed Image Set. Matches were determined by applying a threshold on the similarity measures and identifying those images with measures higher than the threshold. The well-known measures *True Positive Fraction* (TPF or sensitivity) and *False Positive Fraction* (FPF) were used. By sweeping the threshold and averaging the measures of TPF and FPF over all images in the Original Image Set, ROC curves were calculated.

The ROC curves for the four similarity measures described in section 4 are plotted in Figure 1. It can be seen from the graph, that the normalized color histogram with the similarity measures scaled L_1-norm and scaled histogram intersection show the best performance whereas the quadratic histogram measure shows the worst performance. The quadratic histogram measure incorporates information regarding the distance of colors in the color space, which might be more useful if the query was for images of *similar* color, as opposed to exact matches. It can also be seen that the spatial chromatic histogram shows slightly worse performance compared to the color-only histogram for this experiment. The spatial information could prove useful when two images have exactly the same colors but in different locations. However, we did not design a database having those requirements since the goal was to examine the robustness of color-based descriptors over transformation changes in a general database.

We also examined the robustness of color-based descriptors for specific transformations. The ROC curves, taken using the color-only histogram with the normalized histogram intersection measure, that have been evaluated (but will not be presented here due to space limitations), demonstrate the invariance of color-based descriptors to resizing, to JPEG compression, to median filtering, and to rotation of 90°, keeping in

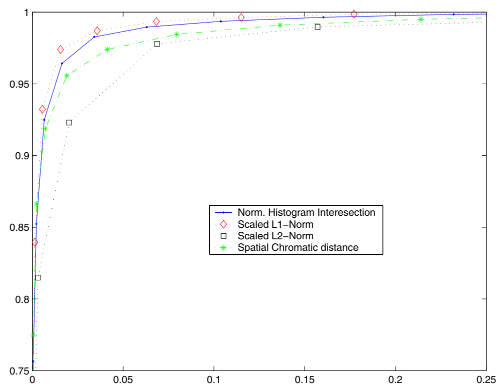

Fig. 1. ROC curves comparing the different similarity measures for matching of color histograms

mind that for that particular angle no black frame was added to the image. As expected, the performance dropped for higher degrees of cropping and for rotations where a black frame is added. However, even for cropping of 30% both sides, a sensitivity of 90% is achieved at only about 2.3% false positive rate.

The experimental results demonstrate the robustness of color-based descriptors for the application of image fingerprinting. The results were very good keeping in mind that the Transformed Image Set included images that were badly deteriorated.

In the future, we plan to examine the performance of the algorithm when the database includes similar images. For such a query, spatial descriptors might prove more useful. Moreover, we will address other attacks, such as illumination changes.

5 Conclusion

In this manuscript, we presented an image fingerprinting system that was designed to retrieve transformed versions of a query image from a large database. We examined the use of color-only and spatial chromatic histograms and the effect of different similarity measures. The results on a database of 450 original images and their 9000 transformation images, showed the robustness of color-based descriptors under high degrees of common attacks and are very encouraging for the use of this system for image fingerprinting.

Acknowledgement

This work has been partly supported by EU and Greek national funds under Operational Programme in Education and Initial Vocational Training II through the Archimedes project Retrieval and Intellectual Property rights protection of Multidimensional Digital Signals (04-3- 001/4).

References

1. M. Schneider and S-F Chang, "A Robust content based digital signature for image authentication," in *Proc. 1996 International Conference on Image Processing (ICIP 1996)*, pp. 227-230, 1996.
2. R. Venkatesan and S.-M Koon and M. H. Jakubowski and P. Moulin, "Robust Image Hashing," in *Proc. of the 2000 International Conference on Image Processing (ICIP 2000)*, 2000.
3. M. Johnson and K. Ramchandran, "Dither-based secure image hashing using distributed coding," in *Proc. 2003 International Conference on Image Processing (ICIP 2003)*, pp. 495-498, 2003.
4. C.S. McCamy and H. Marcus and J.G. Davidson, "A color-rendition chart," *Journal of Applied Photographic Engineering*, vol. 2, no. 3, pp. 95-99, 1976.
5. A.S. Glassner, *Principles of Digital Image Synthesis*. Morgan-Kaufmann Publishers Inc., San Francisco, USA, 1995.
6. L. Cinque and G. Ciocca and S. Levialdi and A. Pellicano and R. Schettini, "Color-based image retrieval using spatial-chromatic histograms," *Image and Vision Computing*, vol. 19, pp. 979-986, 2001.