Content-Based Image Retrieval Using Adaptive Thinning and
Illumination Invariant Color Features

Keisuke KAMEYAMA, Houssem CHATBRI and Wataru MATSUMOTO*

(a) Adaptive Information Processing Group, Dept. of Computer Science,
Graduate School of Systems and Information Engineering, University of Tsukuba
1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8573, Japan
Email : kame@adapt.cs.tsukuba.ac.jp

Abstract

When implementing a Content-Based Image Retrieval (CBIR) system, various types of image features are used for relevance evaluation. The extracted features need to be selected according to the nature of the images and the requirements of the application, enabling a desirable evaluation of image relevance. This paper reports on image features and their treatment in CBIRs for binary sketch images and color photographs taken under varying illumination conditions.

Keywords  Content-based image retrieval (CBIR), Image matching, Thinning, Color invariants

1. Introduction

Content-based image retrieval (CBIR) is an alternative approach to the mainstream image retrieval methods based on the matching of metadata and text-based queries. The idea of finding relevant images using a submitted query image had been investigated during the 1990’s [7-9]. Nowadays, large-scale CBIR systems are available for public use [4]. Besides the massive cloud-based implementation of data storage and parallel search engine, the key technique that determines the quality of CBIR is how the relevance of two images is evaluated. A careful designing of the relevance function is required especially in special-purpose CBIR that deals with images of a specific class.

In the following, approaches of relevance evaluation for hand-drawn sketch images and images of common objects observed under varying illumination conditions, are reported.

2. Sketch image matching using adaptive thinning and Support Region Descriptor [5,6]

Sketch-Based Image Retrieval is the branch of CBIR concerned with the use of sketch images as queries to retrieve sketch or full color images. In sketch images, the shape information reflecting the stroke is the only available hint towards finding a relevant image. However, binary images including hand-drawn sketches tend to include noise as shown in Fig. 1. Therefore, stable extraction and characterization of the stroke shape regardless of the noise is essential. In order to achieve this goal, an adaptive thinning algorithm which balances the noise reduction and stroke preservation is used. Further, the skeletonized sketches are matched using histograms of local and global shape descriptors.

Fig. 1. Variations of noise in binary images. (a) Border noise. (b) Scratch. (c) Dither.
In the thinning stage, the sensitivity measure is introduced which reflects both the stroke preservation and noise reduction. Further, the sketch image is blurred by Gaussian filters having kernels of various scales, and blurred images are subsequently thinned. The optimal scale is selected by evaluating the sensitivity measure of the thinned images. This adaptive procedure produces an optimal result balancing the noise reduction and shape preservation.

Next, the thinned shape is characterized. For each pixel in the thinned sketch, “near” and “far” regions are defined according to the sketch size, and sketch pixels are counted in 8 directional bins relative to the pixel being characterized, both in near and far regions. Pixel counts in 8 directions are further mapped to an 8-bit binary pattern as shown in Fig. 2. For all the pixels in the thinned sketch, a pair of 8-bit patterns (for near and far regions, respectively) will be combined to make two histograms. The histogram pair is called the Support Region Descriptor (SRD), and it characterizes the sketch shape. When matching against other sketches, the discrepancies of SRD feature will be used as the distance between the two images.

An experiment was performed comparing the proposed shape descriptor and other known methods. An image set of 1431 images of handwritten alphabets, digits, math symbols and math expressions of 105 categories was used. The images were preprocessed by either a conventional thinning algorithm or the adaptive thinning method mentioned above to generate Datasets 1 and 2, respectively. In Fig. 3, the precision-recall curves are compared for the SRD, ERH[3], SC[1], and AP[2] descriptors. There, it is observed that the proposed SRD gave the best performances.

![Fig. 2. Calculation of SRD feature.](image1)

![Fig. 3. Matching performances](image2)

### 3. Image matching using illumination invariant color features [14]

When matching images of real-world objects observed under varying illumination conditions, it is important to evaluate and compare the actual color of the object. Geusebroek et al. have introduced numerous illumination invariant color features (color invariants) [10]. Among them, invariant $H$ has been applied to illumination invariant object matching using Colored-SIFT [11,12]. However, invariant $H$ has a disadvantage that pairs of opposite colors will give the same value of $H$, leading to an inability to characterize edges of opposite colors; a disadvantage in object matching.

The authors have proposed a novel invariant $H'$ as a function of two invariants $C$ and $H$. It solves the issue of $H$ by mapping opposite colors to different feature values as shown in Fig. 4.
A partial image matching experiment using invariants $H$ and $C$ by Geusebroek, the proposed invariant $H'$, luminance $L$ and Hue as image descriptors was performed. The images were chosen from the ALOI database [10], and the set included same objects observed under 8 different illumination conditions as shown in Fig. 6(a). The query images are partial images of the objects with variations such as scaling, rotation and severe illumination conditions (Fig. 6(b)-(d)). The procedure of image matching and retrieval is shown in Fig. 5. Point-to-point matching was solved by a nearest-neighbor match of Colored-SIFT features calculated using the image descriptors. The correspondences of regions were found using RANSAC[13]. Then, regions having positive normalized cross-correlation were included in the retrieved images. Then, the retrieval performance was measured using the F-measure. In this experiment, the difference in the retrieval performance according to the employed image descriptor was evaluated. The results are summarized in Fig. 6(e). For the scaled (Query A) and rotated (Query B) queries, luminance $L$ achieved better correspondence than the invariants, probably because there were less differences in the illumination intensity between the query and database images. When dark queries (Query C) were used, the advantage of using invariants $C$ and $H'$ became obvious. In contrast, the performance dropped significantly for $L$ and $H$. Improvement from $H$ to $H'$ shows that confining the equal-$H'$ color set to similar hue colors was a plus. Jointly using $L$ and other invariants generally improved performances. Performances for Hue was generally low. It may be due to its high variance against different illumination conditions in the dataset.

![Fig. 4. Sets of colors corresponding to different values of invariants $H$ and $H'$](image1)

![Fig. 5. Procedure of image matching](image2)
Fig. 6. (a) Database images taken under different illumination conditions, (b) Query-A (enlarged 200%), (c) Query-B (rotated), (d) Query-C (dark), and (e) retrieval performances for different image features.

References