Final report (Option 1)

Course name: Multimedia Processing Technique (DD026_1594)

Project number: Option 1 - 9

Project name: Survey the paper: Content-based image retrieval using color difference histogram

Paper link: https://www.sciencedirect.com/science/article/pii/S0031320312002713

Student name: LeeYunji. 201912110

Github of project: https://github.com/yunzi125

1. Introduction

With the development of digital image processing technology, it has become imperative to find a method to efficiently search and browse images from large image collections. Prompted by market demand for search services, image retrieval has become an extremely active research area in the field of pattern recognition and artificial intelligence. Current image search technologies are typically based on low-level capabilities. To reduce the 'semantic gap' with high-level image search, researchers adopted machine learning techniques to derive high-level semantics. In this paper, we mainly focus on content-based image retrieval.

Much research is needed to use visual differences of two different colors in image search. To address this problem, this paper presents a novel feature representation method for content-based image retrieval, i.e., a chromatic difference histogram.

2. Related work

The visual system extracts information from the environment and converts it into neural code that makes it perceptible. CBIR technology is based on global and local functions and various algorithms have been designed to extract global and local features.

Because color histograms do not change for direction and size, they are easy to implement image retrieval and have been well studied and widely used in CBIR systems. However, it is difficult to characterize the spatial structure of the image. Therefore, several color descriptors have been proposed to utilize spatial information.

Texture functionality is also widely used in CBIR systems. Algorithms for texture analysis include Tamura texture feature, the Markov random field model, Gabor filtering, and local binary patterns. The MPEG-7 standard adopts three texture descriptors: the homogeneous texture descriptor, the texture browsing descriptor and the edge histogram descriptor.

In addition to color and texture features, shape features are also used in CBIR because humans can recognize objects solely based on their shapes. In MPEG-7, three shape descriptors are used for object-based image retrieval; these are the 3-D shape descriptor, region-based shapes derived from Zernike moments, and the curvature scale space (CSS) descriptor.

Several types of local descriptors have been reported in the literature, among which there is a technique called 'scale invariant characteristic transformation (SIFT). It is the most popular form of local feature representation and can tolerate certain levels of illumination changes, perspective distortions, and image transformations, and are very robust to occlusion.

3. The color difference histogram (CDH)

The perceptually uniform color difference between the color and the edge direction is very useful information. It plays an important role in analyzing and understanding image content in the human visual system. This descriptor combines the use of orientation, color and color difference and considers the spatial layout without the use of any image segmentation or learning processes.

3.1 L*a*b* Color space

R, G and B components are highly correlated, and therefore, chromatic information is not directly fit for use. CLE L*a*b* was designed to be perceptually uniform. CLE determines the difference between colors with high uniformity, and the difference between two color points can be measured by Euclidean distance. In the CLE L*a*b* color space, the L*, a* and b* components are computed obtained through a non-linear mapping of XYZ coordinates.

3.2 Edge orientation detection in L*a*b* color space

L*a*b* color space computationally efficient algorithm was adopted for edge direction detection. To minimize color information loss, we propose a method for gradient calculation. The core idea involves extending the concept of a gradient to the vector maximum rate of a scalar function f(x, y) at coordinates (x, y).

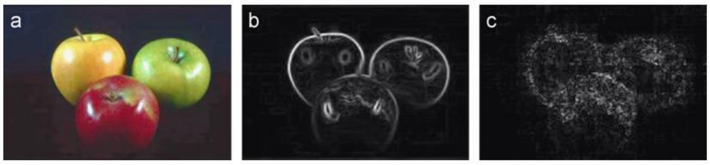

Fig. 1. Edge detection: (a) original full color image, (b) maximum gradient image and (c) minimum gradient image

3.3. Color quantization in the L*a*b* color space

To extract color information and simplify manipulation, color quantization needs to be implemented. The task of color quantization is to select and assign a limited set of colors for representing a given color image with maximum fidelity. The selection of color spaces for color quantization is an important step in many image search and object recognition algorithms. In this work, we use the L*a*b* color space, which is quantized into 90 colors.

3.4. Feature representation

It is an important challenge that feature representation use color difference and take into account both the spatial information of color and edge orientation cues. Based on this idea, we propose a novel image feature representation method called chromatic difference histogram (CDH) for image retrieval.

CDH is combined into an integrated framework of directional and perceptual color information and both spatial layouts are considered. The proposed algorithm consists of two special histograms types, which are computed in a parallel manner under the background of colors and orientations.

4. Experiments and results

In this section, we demonstrate the performance of the proposed algorithm using two Corel datasets. In these experiments, we randomly selected 20 images from each category as query images. The performance was evaluated based on the average results of each query. For fair comparison, we selected

algorithms that were originally developed for image retrieval, such as the edge histogram descriptor (EHD) [9], color autocorrelograms (CAC) [8], and the multi-texton histogram (MTH) because these adopted edge orientations or color information for image representation without the use of image segmentation and model training.

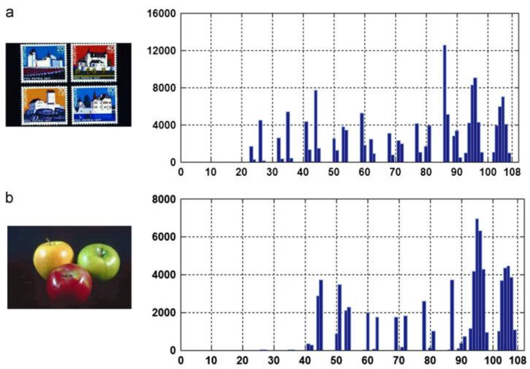

Fig.2. Two examples of CDH: (a) stamps and (b) fruit. The horizontal axis corresponds to the index values for edge orientation and color (where values in the range 1-90 denote color index values and values in the range 91-108 denote edge orientation index values). The vertical axis corresponds to perceptually uniform color difference values.

4.1 Datasets

There are data sets the Corel dataset, the Brodatz texture dataset, the OUTex texture dataset, the Coil-100 dataset, the ETH-80 dataset, the Caltech 101 dataset and the PASCAL VOC dataset, which are frequently used for various purposes in the field of image research.

4.2 Distance metrics

. In this paper, we extend the Canberra distance as a distance metric. For each template image in the dataset, an M-dimensional feature vector T=[T1, T2,⋯,TM] is extracted and stored in the database. Let Q=[Q1, Q2,⋯,QM] be the feature vector of a query image; then, the distance metric between them is simply calculated as follows:

$$D(T,Q) = \sum_{i=1}^{M} \frac{|T_i - Q_i|}{|T_i + u_T| + |Q_i + u_Q|}$$

4.3 Performance metrics

In our experiments, we use the Precision and Recall curves, a performance metric commonly used in information retrieval. Precision and Recall is defined as follows:

P(N) = I_N/N

R(N) = I_N/M

where IN is the number of images retrieved in the top N positions that are similar to the query image, M is the total number of images in the database that are similar to the query, and N is the total number of images retrieved.

4.4 Implementation details

In the RGB color space, the color difference between two colors is defined as follows:

$$D_{RGB}(P_1,P_2) = \sqrt{(\Delta R)^2 + (\Delta G)^2 + (\Delta B)^2}$$

where DR denotes the color difference between two points P1 and P2 in channel R, and so on. In this paper, the proposed algorithm is designed for color image retrieval. However, the MPEG-7 edge histogram descriptor (EHD) was originally designed for the retrieval of gray texture images.

4.5 Retrieval performance

In the experiments, we first demonstrate the reason why we adopted the L*a*b* color space for the proposed framework and confirm the final quantization number for color and edge orientation. Second, we demonstrate that the distance metric proposed in Eq. (21) is more suitable as the color difference histogram. Third, we will test the effect of the distance parameter on the color difference histogram. Finally, the retrieval performances are compared.

It may be seen that the CDH algorithm in the L*a*b* color space performs the best quantization number for the color and edge direction. Increasing the number of quantizations for colors reduces the performance of CDH because too many noise features may be obtained to improve the technology.

Table 1
The average retrieval precision and recall of the CDH with different quantization numbers for color and edge orientation using the Corel-5K dataset in the $L^*a^*b^*$ color space.

The quantization number for color	The quantization number for edge orientation											
	Precision (%)						Recall (%)					
	6	12	18	24	30	36	6	12	18	24	30	36
180	54.12	55.36	56.60	56.96	57.13	57.13	6.49	6.64	6.79	6.84	6.85	6.85
160	49.41	51.73	51.80	51.64	51.16	50.45	5.93	6.21	6.22	6.20	6.14	6.05
90	54.81	56.58	57.23	56.87	57.10	56.70	6.58	6.79	6.87	6.83	6.85	6.80
45	52.46	53.50	53.07	52.85	52.26	52.01	6.30	6.42	6.37	6.34	6.27	6.24

Table 2
The average retrieval precision and recall of the CDH with different quantization numbers for color and edge orientation using the Corel-5K dataset in the *RGB* color space.

The quantization number for color	The quantization number for edge orientation											
	Precision (%)						Recall (%)					
	6	12	18	24	30	36	6	12	18	24	30	36
128	51.59	53.32	53.57	53.94	53.81	53.75	6.19	6.40	6.43	6.47	6.46	6.45
64	46.89	52.94	52.87	52.81	53.16	52.65	5.63	6.35	6.35	6.34	6.38	6.32
32	50.05	52.20	51.66	51.09	50.87	50.49	6.01	6.26	6.20	6.13	6.11	6.06
16	46.60	48.40	48.18	48.21	48.04	47.86	5.59	5.81	5.78	5.79	5.77	5.74

Table 3
The average retrieval precision and recall of the CDH with different quantization numbers for color and edge orientation using the Corel-5K dataset in the HSV color space.

The quantization number for color	The quantization number for edge orientation											
	Precision (%)						Recall (%)					
	6	12	18	24	30	36	6	12	18	24	30	36
192	51.21	53.38	54.03	54.48	55.11	55.34	6.15	6.41	6.48	6.54	6.61	6.64
128	52.80	54.35	55.19	55.80	56.01	56.18	6.34	6.52	6.62	6.70	6.72	6.74
108	52.20	53.03	53.77	54.34	54.52	54.81	6.26	6.36	6.45	6.52	6.54	6.58
72	52.58	53.09	53.91	54.23	54.42	54.82	6.31	6.37	6.47	6.51	6.53	6.58

Table 4
The average retrieval precision and recall of CDH with different distance or similarity metrics.

Dataset	Performance	Distance or similarity metrics							
		Our distance metric	Canberra	x^2 statistics	L_1	L_2	Histogram intersection	Cos Correlation	Jeffrey divergence
Corel-5K	Precision (%)	57.23	56.40	56.21	53.06	53.06	24.28	45.01	45.96
	Recall (%)	6.87	6.76	6.74	6.37	6.37	2.93	5.40	5.52
Corel-10K	Precision (%)	45.24	44.30	44.47	42.05	42.05	16.48	36.34	35.08
	Recall (%)	5.43	5.32	5.34	5.05	5.05	2.01	4.36	4.22

The HSV color space is widely used in image retrieval and object recognition and achieves good performance. In the proposed framework, the HSV color space can also provide much better results than the RGB color space.

In practice, color histograms based on a given color space may perform very well. However, the proposed algorithm is entirely different from the existing color histograms, in that it does not merely focus on color information.

As seen from Tables 1-3, increasing the equalization number of color and edge orientation does not always enhance the description power. Based on the results shown in Table 1 and to balance the retrieval precision and vector dimensions, we set the final quantization numbers for color and edge orientation in the proposed algorithm to 90 and 18, respectively.

The proposed distance metric can be considered as the improved Canberra distance. The Canberra distance is mostly used for data that are scattered around the origin. And here, using the weighting parameters improves performance.

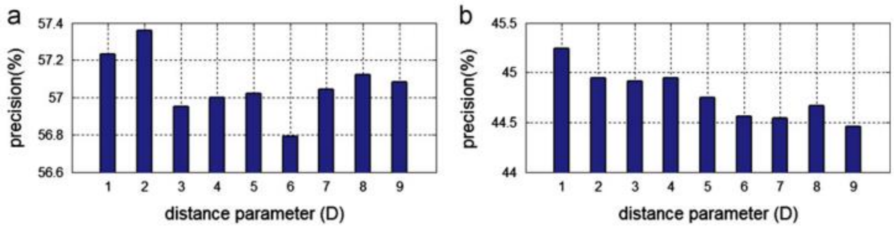

Fig. 3. The average retrieval precision of the CDH algorithm for various values of the distance parameter D: (a) the Corel-5K dataset and (b) the Corel-10K dataset.

Various distance parameter values were tested to test the effect of the distance between two points on perceptually uniform color differences. The results show that the perceptually uniform color difference between two adjacent pixels has the best discrimination.

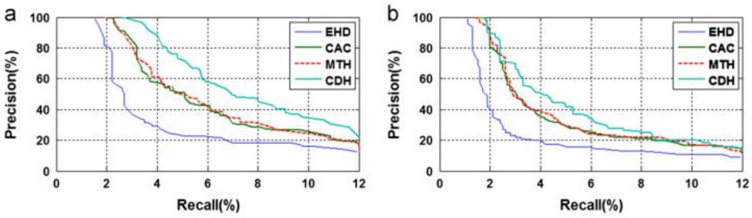

Fig. 4. The precision and recall curves of the EHD, CAC, MTH and CDH algorithms: (a) the Corel-5K dataset and (b) the Corel-10K dataset.

The average precision and recall curves are plotted in Fig. 4.. If the average retrieval precision and recall are higher, the curves will go gar from the original of coordinate.

Table 5
The average retrieval precision and recall results using the two Corel datasets.

Datasets	Performance	Method			
		EHD	CAC	MTH	CDH
Corel-5K	Precision (%)	39.46	49.05	49.84	57.23
	Recall (%)	4.74	5.89	5.98	6.87
Corel-10K	Precision (%)	32.31	40.94	41.44	45.24
	Recall (%)	3.88	4.92	4.97	5.43

It can be seen from Table 5 and Fig. 4 that the proposed algorithm outperforms the EHD, CAC and MTH algorithms. The vector dimensions of the EHD and CAC algorithms are 240 and 256, respectively, higher than that of the proposed algorithm. Only the vector dimension of the MTH algorithm is lower than that of the CDH algorithm.

In addition, experiments with stamp query images and fruit query images using Corel 10k datasets demonstrate that the proposed algorithm has discrimination against color, texture, and edge features. However, it should be taken into account that not all queries in the dataset can provide such high search accuracy.

The proposed algorithm analyzes the perceptually uniform color difference between neighboring colors and edge orientations based on two special histogram types in L*a*b* color space and overcomes the disadvantage of MTH, which discards perceptual color information.

This algorithm can represent the perceptually uniform color difference between colors and edge orientations, and contains the spatial information between them. Therefore, this algorithm can provide better performance than that of MTH and EHD.

Based on the synthetic analysis of retrieval precision and vector dimension, the CDH algorithm performs better than the EHD, CAC and MTH algorithms.

5. Conclusion

In this paper, we propose chromatic difference histogram (CDH) as a novel image feature representation method used to describe image features for image retrieval. This histogram calculates perceptually uniform color differences between two points below different backgrounds with respect to color and edge orientation, rather than simply calculating the number or frequency of pixels.

This method is similar to the way humans perceive colors. And it does not require image segmentation, learning processes, or clustering implementation. Experiments show that it is much more efficient than conventional image feature descriptors originally developed for content-based image retrieval.