Generalized 3D Co-occurrence Matrix for Visual Information Retrieval*

LI Xue-long\textsuperscript{1}, LIU Zheng-kai\textsuperscript{1}, YU Neng-hai\textsuperscript{1}, YUAN Yuan\textsuperscript{1}
(1. Dept. EE&IS, University of Science and Technology of China, Hefei 230027, China)

TAO Da-peng\textsuperscript{2}, WEN Yong-gang\textsuperscript{3}, TANG Xiao-ou\textsuperscript{4}
(2. Electronic Optical Device Institute, Luoyang, Henan, China, 471000;
3. Department of Electronic Engineering, Massachusetts Institute of Technology, Boston, MA, U.S.;
4. Dept. IE, Chinese University of Hong Kong, Shatin, N.T., Hong Kong)

Abstract: A novel feature set based on generalized co-occurrence matrix is proposed for Content-based Image Retrieval (CBIR). Texture features provide important cue for the visual information retrieval and discrimination of image content, therefore a set of 3D co-occurrence matrix based texture features is proposed. The preliminary results based on the novel features are much better than traditional ones. Experiments are implemented on an image database including 25000 images.

Key words: Content-Based Image Retrieval; Image Database; Texture Features; 3D Co-occurrence matrix.

1 Introduction

Recently, Content-based image retrieval (CBIR) has gained much attention\cite{1-2}, and many systems\cite{3-7} have been built for this purpose. In those systems, low-level features are employed. Visual features, together with distance metric, are used as criteria to retrieve related images from the databases or Internet.

In CBIR, a challenge is to find efficient and effective features for image representation and similarity measurement, these features should be closely associated with the main contents of the image concerned. However, features in general meaning have not been found. Of all the proposed approaches based on low-level features, co-occurrence matrix\cite{8}, one of the fundamental image properties, is used extensively.

Many researchers applied texture in finding similarities between images. In the QBIC \cite{3}, Niblack et al. used coarseness, contrast and directionality; In the Photobook\cite{4}, Pentland et al. introduced a global signature based on a sum of weighted Gaussian to describe the texture. Manjunath and Ma\cite{10} used the means and standard deviations of Gabor transform coefficients for different scales and orientations. Li et al.\cite{11} used different spatial features like gray level differences, moments, autocorrelation functions fractals and Rober's gradient. All the methods have received great achievements, but sometimes the features do not work for color images, so it is demanding to find a novel feature for color visual texture information retrieval.

Actually, the color spatial information is important to characterize the images\cite{9}. However, most of previous research focuses on gray images, so a set of novel 3D co-occurrence matrices is defined in this paper, and the paper is organized as follows: algorithms are presented in Section 2; experiments and results are shown in Section 3; Finally, in Section 4, conclusions and future work are presented.

2 Algorithm

Co-occurrence matrix describes periodical information in an image. The feature is descriptive and easily computable based on gray-tone spatial dependency. But if the color images are translated into gray-level, color information will be lost, therefore a novel color co-occurrence matrix (3D) is built in the HSV color space, which is selected because it suits human’s visual perception:

\[
V = \frac{1}{7} (R + G + B) \quad S = 1 - \frac{3 \min(R, G, B)}{R + G + B} \quad H = \begin{cases} \theta & G \geq B \\ 2\pi - \theta & G < B \end{cases} \quad \theta = \cos^{-1} \left[ \frac{\frac{1}{3}[(R-G)+(R-B)]}{\sqrt{(R-G)^2 + (R-B)(G-B)}} \right]
\]

And then, each color channel is quantized into bins. Typically, the first thing we usually notice about a

* Manuscript retrieved: March 15, 2002 revised: April 27, 2002
Supported by: National Key Science-Technology Project of the Nine-Five-Year-Plan of China under Grant No. 96-B02-01-05
color is its hue, so we quantize H into 8 bins, while S and V are equally quantized into 4 bins. In HSV (3D) space, nine directions are used to describe the feature. Every pixel on the H plane is corresponding to nine neighbors on S plane and nine on V plane. From the Fig.1, it could be seen easily, pixel o, in H plane, has nine neighbor pixels on S plane and nine on V plane. There are nine directions in the matrix space. Each line’s center is o. The \((i, j, k)\) matrix entry represents the number of times pixel o has value \(j\) in H plane, pixel a has value \(i\) in V plane and pixel b has value \(k\) in S plane when pixel a and b are neighbors of pixel o and on the same direction. o varies across the entire H plane. Along the edges, the image is assumed to repeat itself, according to a torus topology. Based on these nine matrices, a large number of features will be defined. The results from previous studies in the 2D co-occurrence matrix are compared. From these definitions, corresponding features were built up. Assuming G is the gray level, we compare the traditional co-occurrence features to our novel 3D features.

(a) The angular second-moment (ASM) is a measure of homogeneity of the image and is defined as

\[
\text{The counterpart of } f_{20} = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} P_{i,j,0} \text{ in 3D is } f_{20} = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} \sum_{k=0}^{G-1} P_{i,j,k,0}.
\]

(b) The contrast feature (CON) is a difference moment used to measure the contrast or the local gray level variation presented in an image

\[
\text{The counterpart of } f_{20} = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} \sum_{k=0}^{G-1} P_{i,j,k,0} \text{ in 3D is }
\]

\[
f_{20} = \frac{1}{G} \sum_{i=1}^{G} \sum_{j=1}^{G} \sum_{k=1}^{G} n^2 \sum_{|i-j|\leq n} P_{i,j,k,0} + \frac{1}{G} \sum_{i=1}^{G} \sum_{j=1}^{G} \sum_{k=1}^{G} n^2 \sum_{|j-k|\leq n} P_{i,j,k,0} + \frac{1}{G} \sum_{i=1}^{G} \sum_{j=1}^{G} \sum_{k=1}^{G} n^2 \sum_{|i-k|\leq n} P_{i,j,k,0}.
\]

(c) The correlation feature (COR) measures color linear dependencies in the image under HSV space.

\[
\text{The counterpart of } f_{20} = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} \sum_{k=0}^{G-1} P_{i,j,k,0} \text{ in 3D is } f_{20} = \frac{1}{G} \sum_{i=1}^{G} \sum_{j=1}^{G} \sum_{k=1}^{G} \mu_x \cdot \mu_y.
\]

(d) The entropy feature (ENT) measures complexity

\[
\text{The counterpart of } f_{20} = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} \sum_{k=0}^{G-1} P_{i,j,k,0} \cdot \ln(P_{i,j,k,0}) \text{ is } f_{20} = \frac{1}{G} \sum_{i=1}^{G} \sum_{j=1}^{G} \sum_{k=1}^{G} \ln(P_{i,j,k,0}).
\]

The vector \(F=[f_{10}, f_{20}, f_{30}, f_{40}]\) could represent the color images. In our experiments, we tested the features by following metrics, and the experimental results show that the Euclidean metric performs well for our features.

1. \(\sum_{i=1}^{\dim F} (F_i - F'_i)^2\)
2. \(\sum_{i=1}^{\dim F} (|\max(F_i, F'_i)| - |\min(F_i, F'_i)| - 1)\)
3. \(\sum_{i=1}^{\dim F} \max(F_i, F'_i)\)
4. \(\sum_{i=1}^{\dim F} |F_i / F'_i|\) (If \(F'_i \neq 0\) then \(F_i / F'_i\) is zero)

3 Experimental Results

iPHOTO system, our test bed, is built up under Visual Studio 7.0 / Matlab 6.0 /Windows 2000, and our image database contains 2,5000 pictures in JPEG format, based on it, a series of tests were made and hundreds of queries were tested, the some examples experiences and results are shown below.

![Fig 1: Neighbors of a pixel in color space](image)

![Fig 2: Results by 3D Co-occurrence matrix features](image)
In Fig. 2(b), the x coordinate stands for the amount of the pictures retrieved by iPHOTO, and the y coordinate is the correct number. The retrieval result shown in Fig. 2 (b) means that from No.1 to No.9 images are all contain bird or birds in them, and No.10 doesn't contains a bird. The same is others. Fig. 2 shows that in top 20 images, there are 13 images are right. Fig. 2 (a) shows the retrieval correctness ratio. The x-coordinate is the number of the pictures, which are retrieved by IPHOTO system. Therefore, from the results, it is easy to know that top 5’ correctness ratios are always 100%. About 80% to top 10’s correctness ratios, sometimes, the top 10’s correctness is 100%, such as the presented retrieval results.

Finally, the results by both the presented method (3D) and the traditional method (2D) are shown in Fig. 3 as statistic curves for the purpose of comparison. The x coordinate stands for the top N results, while the y coordinate means the right number among the total N result, and the two curves indicate that the developed scheme outperforms the other scheme in terms of the retrieval efficiency.

4 Conclusions

A new texture-based image retrieval algorithm in our CBIR system iPHOTO has been presented. In the proposed system, another dimension is added to the traditional 2D co-occurrence matrix to form a 3D co-occurrence matrix as to take the image spatial information into account. Based on our image feature matrix, the similarity of images is redefined. Experimental results show that the proposed algorithm outperforms the traditional ways.

References:

Authors: LI Xue-long (李学龙, 1976), Ph.D. Candidate in the department of Electronic Engineering and Information Science (EEIS) of the University of Science and Technology of China (USTC). He has received his B.S. (1998) and M.S. (2000) in USTC, his research focus on digital image processing, pattern recognition, machine Learning; LIU Zheng-kai (刘政凯, 1940), Professor and doctorate student supervisor, the chairman of the Research Committee in the department of EEIS of the USTC; YU Neng-hai (余能海, 1964), Professor in the department of EEIS of USTC; YUAN Yuan (袁媛, 1980) undergraduate students in the department of EEIS of the USTC; TAO Da-cheng (陶大鹏, 1977) Received his B.S. (1999) in Northwestern Polytechnical University, and his current research focus on Signal Processing; WEN Yong-gang (文勇刚, 1977), Ph.D. candidate in the department of Electronic Engineering of Massachusetts Institute of Technology; TANG Xiao-ou (汤晓鸥, 1966), professor in the Department of Information Engineering of The Chinese University of Hong Kong.

(Chinese abstract continued on page 48 )
The RF and IF Design of an LEO CDMA Transmission System

LAO Yuan-wei, ZHANG Zhao-yang, QIU Pei-liang
(Communication and Information System, Zhejiang University, Hangzhou 310027,China)

Abstract: In this paper, based on the characteristic of LEO transmission channel, we give out an analysis of the system's performance and a link transmission equation of CDMA of the reverse link as well. Then we put forward a type of CDMA transmission system and the implementation of RF and IF modules. The design and implementation of the key parameters and relative modules of the receiver are specially discussed here.

Key words: LEO; CDMA; RF; IF