A Bayesian Network Approach for Image Similarity

Yeni Herdiyeni, Rizki Pebuardi, Agus Buono
Dept of Computer Science, Bogor Agricultural University
Darmaga Campus, Wing 20 Level V, Bogor West Java, Indonesia
Email: yeni_herdi@ipb.ac.id

Abstract – This paper proposed Bayesian Network approach for image similarity measurement based on color, shape and texture. Bayesian network model can determine dominant information of an image using occurrence probability of image's characteristics. This probability is used to measure image similarity. Performance of the system is determined using recall and precision. Based on experiment, Bayesian network model can improve performance of image retrieval system. Experiment result showed that the average precision gain up of using Bayesian network model is about 8.28 %. The average precision of using Bayesian network model is better than using color, shape, or texture information individually.

Keywords: Bayesian network, histogram-162, edge direction histogram, co-occurrence matrix

I. INTRODUCTION

CBIR (content based image retrieval) is developed to retrieve images based on image's information such as color, shape, and texture. The main processes of CBIR are preprocessing, feature extraction, indexing, and retrieval. One of the most important processes in retrieval process is image similarity measurement process. Vailaya (1995) [6] used weighted function to measure image similarity which combine color and shape information. Meanwhile Osadebey (2006) [3] used weighted function to measure image similarity which combines texture, shape, and spatial information. In some condition, the use of weighted function gives imprecise result because weighted value is specified manually.

Rodrigues and Araujo (2004) [4] have developed a Bayesian network model to measure image similarity in a CBIR system. Bayesian network model applied probabilistic theory to measure image similarity. This model is used to combine color, shape, and texture information. This model can handle the weakness of weighted function because weighted value is specified automatically by the system based on image's information which is dominant.

If every image characteristic has occurrence probability in every database images, then similarity probability between query and database images are determinable. This matter can be modeled using Bayesian network model. On this model, image characteristics, query image, and database images are modeled as connected nodes that formed a Bayesian network model. This research implements and analyzes performance of Bayesian network model which have developed by Rodrigues and Araujo (2004) [4]. The use of Bayesian network in CBIR system is expected can improve relevancy of retrieval result. This research makes change in shape and texture extraction to improve retrieval result.

II. PROPOSED METHOD

The main purpose of this research is to implement and analyze performance of Bayesian network model in image similarity measurement. The steps of this research are: feature extraction, Bayesian network model development, similarity measurement, and retrieval evaluation. The steps are illustrated in Figure 1.

Feature Extraction

a Color feature extraction

Color feature extraction is done by determining color histogram using CCH (conventional color histogram). Firstly, the image is transformed to HSV (hue, saturation, value) because HSV is an intuitive color space in the sense that each component contributes directly to visual perception [4].

The transformation from RGB to HSV is accomplished through the following equations:

$$h = \begin{cases} \theta, & g \geq b \\ 2\pi - \theta, & g < b \end{cases}$$

$$\theta = cos^{-1}\left[\frac{\frac{1}{2}[(r-g)+(r-b)]}{\sqrt{(r-g)^2+(r-b)(g-b)}}\right]$$

$$s = 1 - \frac{3}{r+g+b}[min(r,g,b)]$$

$$v = \frac{1}{3}(r+g+b)$$

where (r, g, b) are colors in RGB color space and (h, s, v) are colors in HSV color space [1].

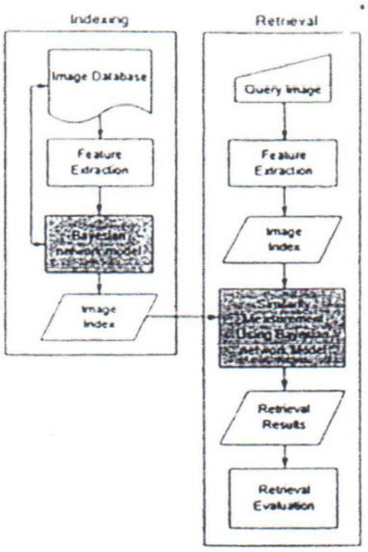

Figure 1 Proposed method.

Color quantization is done after RGB image transformed to HSV image. Color quantization is useful for reducing the calculation cost and efficient storage. Furthermore, it can eliminate color components which can be considered noises [4]. In this research, color quantization that will use is histogram 162 (HSV-162). Hue is quantized finer than saturation and value (18 bins) because human visual system is more sensitive to hue than saturation (3 bins) and value (3 bins).

Every image is represented with a vector that has 162 elements. Vector elements value represents number of image pixels that quantized to certain bins. On another word, image's vector represents image's color histogram. Finally, every image's vector is normalized.

b Shape feature extraction

Shape feature extraction is done using edge direction histogram. Firstly, the RGB image is transformed to grayscale image. Then, Sobel edge detector is applied to the image.

The direction is accomplished through the following equation:

$$\theta = tan^{-1}(Gy/Gx)$$

where Gx is matrix of gradient of image I that is estimated in x-direction and Gy is matrix of gradient of image I that is estimated in y-direction. Whenever the gradient in the x direction is equal to zero, the edge direction has to be equal to 90 degrees or 0 degrees, depending on what the value of the gradient in the y-direction is equal to. If Gy has a value of zero, the edge direction will equal 0 degrees. Otherwise the edge direction will equal 90 degrees [2].

The next step is determining the magnitude of the image strength that is accomplished through the following equation:

$$|G| = |Gx| + |Gy|$$

An image's pixel will be considered as an edge pixel if its magnitude value is bigger than specified threshold. Some experiments are done to determine threshold value which is optimized shape based image retrieval system. The threshold values which are used are 0.066, 0.08, 0.1, 0.12, 0.14, 0.16, and 0.18.

Number of bins that is used in this research is 72 bins (each bin consisting of a range 5^0). Every image is represented with a vector that has 72 elements. The next step is determining number of edge pixels that is related with 72 defined bins. Finally, every image's vector is normalized with the number of edge pixels in order to achieve scale invariant [6].

c Texture feature extraction

Texture feature extraction is done using co-occurrence matrix. In this research, texture information will be represented using energy, moment, entropy, maximum probability, contrast, correlation, and homogeneity.

The mathematics definition of them is accomplished through the following equation:

$$energy = \sum_{i,j} P(i,j)^2$$

$$inverse\ moment = \sum_{\substack{i,j \\ i \neq j}} \frac{P(i,j)^2}{|i-j|}$$

$$entropy = -\sum_{i,j} P(i,j)\ log(P(i,j))$$

$$maximum\ probability = max(P(i,j))$$

$$contrast = \sum_{i,j} |i-j|^2\ P(i,j)$$

$$correlation = \sum_{i,j} \frac{(i-\mu_i)(j-\mu_j)\ P(i,j)}{\sigma_i \sigma_j}$$

$$homogeneity = \sum_{i,j} \frac{P(i,j)}{1+|i-j|}$$

where $P(i,j)$ is element of the i-th rows and j-th columns of normalized co-occurrence matrix. μ_i is average value of the i-th rows and μ_j is average value of the j-th columns of matrix P. σ_i is standard deviation of the i-th rows and σ_j is standard deviation of the j-th columns of matrix P.

The first step to determine texture information of an image is determining its co-occurrence matrix. Co-occurrence matrix is determined on four directions (0^0, 45^0, 90^0, and 135^0). So that, every image will have four co-occurrence matrixes. The next step is determining energy, moment, entropy, maximum probability, contrast, correlation, and homogeneity for each co-occurrence matrixes. So that, each feature will have four value, each value for direction 0°, 45°, 90°, and 135°. Value of each feature is the average value of its four values to make texture information is rotation invariant. Texture information is represented with a vector that has seven elements. Finally, every image's vector is normalized.

This research uses some of grey level number on determining co-occurrence matrix in order to get optimal result. Numbers of grey level that are used are 8, 16, 32, and 64. Number of grey level that make texture image retrieval is optimal, is selected.

Bayesian Network Model

This research used Bayesian network model that was developed by Rodrigues and Araujo (2004) [4]. It will be used in indexing and retrieval process.

Color information of an image is representing by a vector that have 162 elements, shape information is representing by a vector that have 72 elements, and texture information is representing by a vector that have 7 elements. Every bin in above three vectors has occurrence probability in every database images. This is modeled in a network structure. Network model for texture information is illustrated in Figure 2.

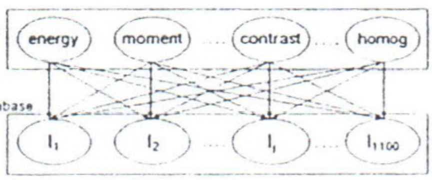

Figure 2 General Bayesian network model for CBIR using texture information.

On Figure 2, nodes in first level are feature of texture information. So that, first level is consist of seven nodes. If the network modeled color information, then first level of the network is consist of 162 nodes. Furthermore, if the network modeled shape information, then first level of the network is consist of 72 nodes. Nodes in second level of network are images in database (1100 images).

On the network model, occurrence probability value of image I_i in database which have characteristic C_j $(P(I_i|C_j))$ is the j-th element vector value of image I_i. For example, occurrence probability value of image I_{15} in database which have characteristic of moment is the second element vector value of I_{15}'s texture vector. This is use of Bayesian network model in indexing process.

Similarity Measurement

Bayesian network model which is developed, is not only used in indexing process, but also used to measure similarity between query image and database images. If query image is exist, then probability value of query image for every characteristic $P(Q|C_j)$ is determinable. So that, occurrence probability of every database image for a given query image $P(I_j|Q)$ is determinable. On another word, similarity value between query image and database images is determinable. Network model which is used in this process is illustrated in Figure 3.

On Figure 3, database is placed three times to avoid arcs crossing which cause visual noise.

Similarity value between query image and database images is accomplished through the general Bayesian network model equation:

$$P(I_j|Q) = \eta[1 - (1 - P(CC_j|CC))\ldots$$
$$\ldots x\,(1 - P(CS_j|CS))\; x\; (1 - P(CT_j|CT))]$$

$P(CC_j|CC)$, $P(CS_j|CS)$, and $P(CT_j|CT)$ are accomplished through cosine similarity:

$$sim(\vec{I},\vec{Q}) = \frac{I^t Q}{|I|\,|Q|} = \frac{\sum_{i=1}^{n} I_i\; x\; Q_i}{\sqrt{\sum_{i=1}^{n} Q_i^2}\; x\; \sqrt{\sum_{i=1}^{n} I_i^2}}$$

where I_i is the i-th characteristic of a database image, while Q_i is the i-th characteristic of query image. As much this function is near to one, as much the database image and query image are equals.

$P(CC_j|CC)$ represents similarity value between color vector of query image and color vector of database images, $P(CS_j|CS)$ for shape vector, and $P(CT_j|CT)$ for texture vector.

The general Bayesian network model equation is not only can be used to combine color, shape, and texture information, but also to determine similarity based on individual information (color, shape, or texture). Similarity measurement based on color can be done by setting $P(CS_j|CS) = 0$ and $P(CT_j|CT = 0)$. $P(I_j|Q)$ is accomplished through the following equation:

$$P(I_j|Q) = \eta[1 - (1 - P(CC_j|CC))]$$

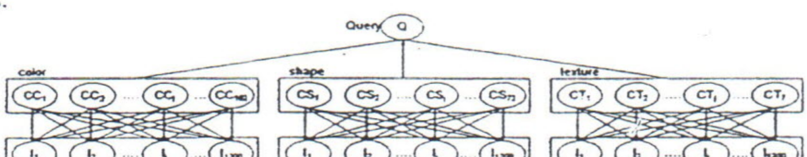

Figure 3 General Bayesian network model for CBIR which combine color, shape, and texture information.

This rule can be applied for similarity measurement based on shape or texture information. So that, similarity measurement based on shape information is accomplished through the following equation:

$$P(I_j|Q) = \eta[1-(1-P(CS_j|CS))]$$

Furthermore, similarity measurement based on shape information is accomplished through the following equation:

$$P(I_j|Q) = \eta[1-(1-P(CT_j|CT))]$$

The results from this step are similarity value between query image and database images. After similarity values are known, the database images are sorted based on its similarity value.

Evaluation

On this step, performance of the system is determined using recall and precision in order to determine the effectiveness of retrieval process. Every image in database is used as query and the relevant images are determined by manually counting number of images which has the same class with query image. Precision for each class is determined by averaging precision value of every image which has the same class. Finally, the table of recall against precision is constructed for each class and for general system.

III. EXPERIMENTAL RESULTS

Data

The research data consist of 1100 images that are manually classified into 10 classes. The classes are cars, lions, sunsets, textures, bears, elephants, arrows, landscapes, reptiles, and aircrafts. This data is taken from http://www.fei.edu.br/~psergio/MaterialAulas/Generalist1200.zip. Format of the images are TIF which have different size. Image examples can be seen in Figure 4.

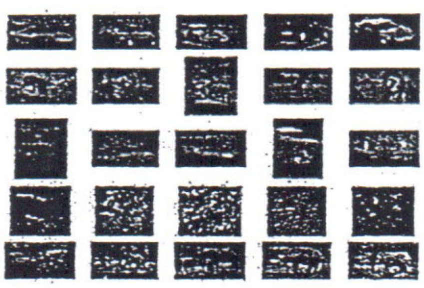

Figure 4 Image examples for cars, lions, sunsets, textures, and elephants class.

Image Preprocessing

Image preprocessing is done to eliminate border of original image. The border is eliminated using cropping operation in order to prevent Sobel edge detector get the wrong image edge.

Feature Extraction

a Color feature extraction

The result of color feature extraction for all images is a matrix 162 x 1100, because there are 1100 images in database and each image is represented by a vector which has 162 elements.

b Shape feature extraction

Sobel edge detector operation is done to all images in database to determine its edge direction histogram. Each image is represented by a vector which has 72 elements. The result of this process is a matrix 72 x 1100, because there are 1100 images in database.

c Texture feature extraction

The result of texture feature extraction for all images is a matrix 7 x 1100, because there are 1100 images in database and each image is represented by a vector which has 7 elements. The vector's elements are energy, moment, entropy, maximum probability, contrast, and correlation.

Retrieval Results

Retrieval result using color information can be seen in Figure 5.

Figure 5 Retrieval result using color information.

Retrieval result using shape information can be seen in Figure 5.

Figure 6 Retrieval result using shape information.

Retrieval result using texture information can be seen in Figure 7

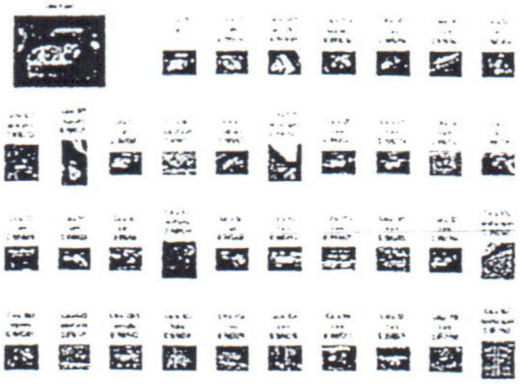

Figure 7 Retrieval result using texture information.

Then, Retrieval result using Bayesian network model can be seen in Figure 8.

Figure 8 Retrieval result using Bayesian network model.

Evaluation

Recall and precision value is calculated to determine the effectiveness of retrieval process. All images in each class are used as query. Precision value of each class is determined by averaging precision of all images in that class in order to measure performance of Bayesian network model which is developed.

This section will present precision value for some examples class. The classes are cars, lions, sunsets, textures, and elephants. The precision values which are presented are precision value for CBIR system that is using information color, shape, and texture individually and its combination in a Bayesian network model.

Cars class consists of 176 images. In this class, the highest average precision value is obtained using Bayesian network model. Furthermore, average precision value of using shape information is higher than using color or texture information. As we can see in Figure 4, images in cars class have shape which is almost same. The average precision values using color, shape, texture, and Bayesian network model are 0.299, 0.3942, 0.3435, and 0.3996 respectively.

Lion class consists of 103 images. In this class, the highest average precision value is obtained using color information. As we can see in Figure 4, images in lion class have color information which is almost same visually. In this class, the average precision value of using Bayesian network model is few lower than using color information, but still better than using shape or texture information. However, precision value of using Bayesian network model is higher than using color information for 0.1, 0.2, 0.3, and 1 recall. The average precision values using color, shape, texture, and Bayesian network model are 0.3729, 0.2706, 0.23, and 0.355 respectively.

Sunset class consists of 102 images. In this class, the highest average precision value is obtained using Bayesian network model. Furthermore, average precision value of using texture information is higher than using color or shape information. As we can see in Figure 4, images in sunset class have texture information which is almost same. Precision value of using color and shape information are lower than using texture information because its color and shape information are few different. The average precision values using color, shape, texture, and Bayesian network model are 0.2814, 0.2583, 0.3356, and 0.4074 respectively.

Texture class consists of 175 images. In this class, the highest average precision value is obtained using Bayesian network model. Furthermore, average precision value of using texture information is higher than using color or shape information. As we can see in Figure 4, images in texture class have texture which is almost same. The average precision values using color, shape, texture, and Bayesian network model are 0.3376, 0.3385, 0.4384, and 0.4582 respectively.

Elephant class consists of 98 images. In this class, the highest average precision value is obtained using shape information. As we can see in Figure 4, images in elephant class have shape information which is almost same visually. In this class, the average precision value of using Bayesian network model is few lower than using shape information, but still better than using color or texture information. Due to the color and texture information of this class are few different, the precision value of using color or texture information are lower than using shape information. The average precision values using color, shape, texture, and Bayesian network model are 0.2461, 0.3930, 0.2330, and 0.3182 respectively.

Generally, the average precision values for another five classes have same tendency with the classes which have been presented. Table 1 presents average precision values for all images in database. All of images in database are used as query.

Table I Recall precision value of all images in database

Recall	Color	Shape	Texture	Bayes	Improve-ment (%)
0	1.0000	1.0000	1.0000	1.0000	0.0000
0.1	0.4008	0.3346	0.3561	0.4861	17.5591
0.2	0.3448	0.2859	0.3136	0.4223	18.3532
0.3	0.3058	0.2580	0.2798	0.3777	19.0406
0.4	0.2754	0.2407	0.2564	0.3386	18.6793
0.5	0.2568	0.2221	0.2363	0.3050	15.8142
0.6	0.2395	0.2062	0.2134	0.2704	11.4136
0.7	0.2237	0.1914	0.1929	0.2415	7.3467
0.8	0.2076	0.1735	0.1724	0.2122	2.1676
0.9	0.1924	0.1522	0.1525	0.1825	-5.4027
1	0.1586	0.1249	0.1267	0.1392	-13.8753
Average	**0.3277**	**0.2900**	**0.3002**	**0.3614**	**8.2815**

According to Table I, the highest average precision value is obtained using Bayesian network model. Furthermore, average precision value of using color information is higher than using shape or texture information. These show that Bayesian network model can improve precision values. Bayesian network model which is combine color, shape, and texture information, is better than using color, shape, or texture information individually. The improvement of precision values means improvement of the relevancy of retrieved images.

According to Table I, the average precision of using color information is better than using shape or texture information. The average precision of using Bayesian network model achieved the best result. The average precision gain up of using Bayesian network model is about 8.2815 %. Generally, performance of using Bayesian network model is better than of using color, shape, or texture information individually.

Comparative performance of retrieval based on color, shape, texture, and Bayesian network model is shown in Figure 9.

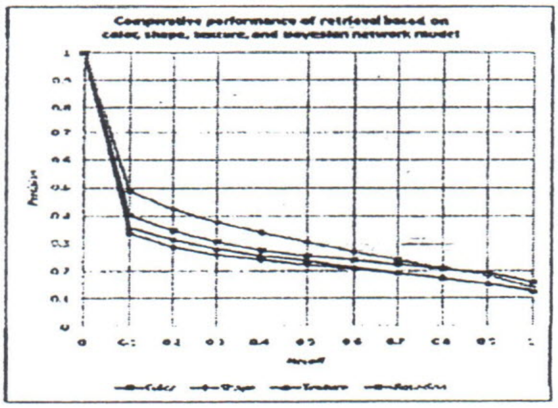

Figure 9 Comparative performance of retrieval based on color, shape, texture, and Bayesian network model.

IV. CONCLUSIONS

This paper proposed Bayesian Network approach for image similarity based on color, shape and texture. Color information is extracted using histogram 162, shape information is extracted using edge direction histogram, and texture information is extracted using co-occurrence matrix.

Bayesian network model determine weighted value automatically based on image's information which is dominant. Generally, Bayesian network model have good performance. The average precision of all images in database that are using Bayesian network model are better than using color, shape, and texture information individually. The average precision gain up of using Bayesian network model is about 8.28%.

REFERENCES

[1]. Gonzalez RC, et al. 2004. Digital Image Processing Using Matlab. New Jersey : Pearson Prentice Hall.

[2]. Green B. 2002. Edge Detection Tutorial http://www.pages.drexel.edu/%7Eweg22/edge.html [28 Juli 2008]

[3]. Osadebey ME. 2006. Integrated Content-Based Image Retrieval Using Texture, Shape and Spatial Information [thesis]. Umea : Department of Applied Physics and Electronics, Umea University.

[4]. Rodrigues PS & Arnaldo de Albuquerque Araujo 2004. A Bayesian Network Model Combining Color, Shape and Texture Information to Improve Content Based Image Retrieval Systems. LNCC, Petropolis, Brazil.

[5]. Vailaya A & Anil Jain. 1995. Image Retrieval Using Color and Shape. Michigan : Michigan State University.

[6]. Vailaya A. 1996. Shape-Based Image Retrieval [thesis]. Michigan : Michigan State University.