# An Algorithm of Image Classification Based on BP Neural Network

Zhiyong Xiong[1,2], Ke Chen[1,2], Caidong GU[1,2], Yinhong Liang[1,2], Fusheng Yu[1,2]

1 JiangSu Province Support Software Engineering R&D Center for Modern Information Technology Application in Enterprise
SuZhou, China
e-mail: xiongzhiyong@jssvc.edu.cn

2 Department of Computer Engineering SuZhou Vocational University
SuZhou, China
e-mail: cke@jssvc.edu.cn

***Abstract*—To improve the performance of image classification, we propose an image classification method based on BP neural network. Firstly, one image is segmented and clustered several visual objects. By means of the unit image library, we calculate one probability vector, which is composed by probability of every visual object of the image and every unit image of the unit image library. So a total feature vector of one image can been attain. By means of the BP neural network, which we construct, we attain a class about the image, thereby we can realize image classification. The end, Ground Truth Database is adopted experimental image library in this paper. The method attains good effect based on experimental result.**

***Keywords- unit image; content vector; feature vector; image classification; BP neural network***

## I. INTRODUCTION

In recent years, people more and more care for how to improve the level of image processing, and image classification becomes a hot research in the domain of image processing. So this paper discusses and studies the image similarity classification as the background.

## II. IMAGE CLASSIFICATION

At present the most existing methods of image classification, based on content of image, firstly generally extract low-level image features[1], then compute the degree of similarity of different images according to the low-level image features, finally finish image classification. But in the real word, people generally analyses one image, according to its' semantic to judge whether the image is needed or not. So it is a fact which the result of image classification is largely determined the analysis of image semantic. So in the process of the image selection, one image can be selected accurately according to probability, which is the probability of the image belonging to an object image classes according to the similarity of their semantics. Allowing for there are many kinds of image classifications, one image can be belonging to several image classifications, so images can be classified by the BP neural network.

The structure of this paper is as follows, section III builds the database of the image semantics, section IV explains the method for extracting the feature of image semantics, section V builds the BP neural network of image classification, section VI proposes a classified algorithm of images, section VII analyzes the experimental result of the BP neural network, and the final section is summary, which proposed farther works.

## III. KEYWORD LIBRARY PROCESSING

Because one image is a direct copy of some part of nature or reality, and or artificial creation, the content of one image conforms to an objective law of real world or a logical law of the human thinking. Based on this assumption, the paper considers that an image library is an image reproduction of one real environment Vi, the set $\Phi$ of all independent keywords from the environment Vi (ignoring these scene of images) constitutes a partition of the image library[2]. So the set $\Phi$ composes a keyword dictionary-W. Thus we use keywords of the keyword dictionary W to describe an image coming from the environment Vi.

In order to exactly establish relationship of images and semantic, this paper proposes a concept of unit image. A unit image' content is single and is belonged only to a keyword in a keyword dictionary according to image semantic, and every keyword of the keyword dictionary relates some unit images. So information of several unit images approximatively linearly combines the information of one real image.

### *A. Manual Classification*

If images of a sample image library originate from an environment Vi, so the image library can be used to classify unit images as a training set. Firstly the image library is classified T classes by professional people to transform N dimensions feature space of the image library into T dimensions classificatory feature space, which is accepted by experts. It is obvious that T is less then N. since every keyword of keyword dictionary can be appears in every image class, a keyword is figured out T frequencies. And a frequency, that a unit image belonging to one image class, equals to the frequency, that a keyword belonging to the same image class. The number $N_{ji}$ of images, which appear the keyword $w_i$ in the image class $C_j$, is figured out, and the number $N_i$ of images, which appear keyword $w_i$ in the image library C, is figured out. So the frequency can be calculated. The formula is as follow:

$$Freq(C_j|w_i) = \frac{N_{ji}}{N_i} = \frac{Freq(C_j w_i|C)}{Freq(w_i|C)} \quad (1)$$

978-1-4673-0089-6/12/$26.00 ©2012 IEEE

Note that the coefficient $C_j$ is the class j in the image library C, and coefficient $w_i$ is the $i$ keyword in the keyword library.

One image may be belonged to several classes when the image library is classified artificially. Then there is an obvious inequality $\sum_{j=1}^{t} Freq(C_j|w_i) \geq 1$ . So the probability $P(C_j|w_i)$ can be approximatively computed as follows:

$$P(C_j|w_i) \approx \frac{Freq(C_j|w_i)}{\sum_{j=1}^{t} Freq(C_j|w_i)} \qquad (2)$$

In this way we construct a probability matrix P about the image library C, this is as follow:

$$P=\begin{pmatrix} P(C_1|w_1) & P(C_1|w_2) & \cdots & P(C_1|w_n) \\ P(C_2|w_1) & P(C_2|w_2) & \cdots & P(C_2|w_n) \\ \cdots & \cdots & \cdots & \cdots \\ P(C_t|w_1) & P(C_t|w_2) & \cdots & P(C_t|w_n) \end{pmatrix} = \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_t \end{pmatrix} = (s_1 \ \ s_2 \ \ \cdots \ \ s_n) \qquad (3)$$

There is an obvious that every row vector $p_j(p_j \in P, j = 1,2,\cdots,t)$ of the matrix P represents a classification of the image library C. Because some of keywords always frequently appear in one classification of the image library C, these keywords is principal component of the classification. Every column vector $s_i(s_i \in P, i = 1,2,\cdots,n)$ of the matrix P represents a classification attribute of the image library C.

### B. *Probability Matrix P Processing*

When the row number of the probability matrix P is far less then the column number of P, the matrix P can be cut out some of columns in the matrix P. If the mean square $\sigma_{s_i}$ of a column vector $s_i$ is very small, the column vector $s_i$ can be cut out from the matrix P, because of $P(C_j|w_i) \approx \frac{1}{n}\sum_{k=1}^{n} P(C_k|w_i), (j = 1,2,\cdots,n)$, in other words, the effect, which the keyword $w_i$ is classified the image library, isn't distinct. Based on the principle, all mean squares $\sigma_{s_i} (i = 1,2,\cdots,n)$ of column vectors are firstly computed in the matrix P, secondly these $\sigma_{s_i} (i = 1,2,\cdots,n)$ are sorted in descending order. A threshold $\theta$ is set to compare every of $\sigma_{s_i} (i = 1,2,\cdots,n)$ . If $\sigma_{s_i}$ is less than the threshold $\theta$, the column vector $s_i$ is be cut out from the matrix P. In this way, we reduce the dimension of the feature space about the problem of image classification. But the dimension of the feature space must be greater than the classification number of the image library, otherwise the image library can't be accurately classified.

## IV. IMAGE FEATURE EXTRACTION

The process of image feature extraction firstly computes the similarities between visual objects of a image and every unit images, and gets a linear combination of a keywords set. Because keywords were classified by means of learning mechanism or artificial annotation[3] in the keyword dictionary, the probabilities, what image classes do the image belong to, is computed.

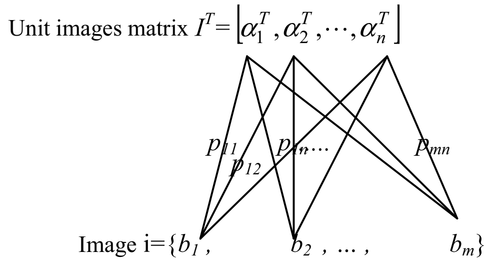


Figure 1. the Relationship of a Image's Blobs and Unit Images

### A. *Image Segmentation and Clustering*

We firstly separate and cluster the image i, which will be classified, then extract visual objects in the image i, and constitute a few blobs. So the image i is transformed into a form which is $i = \{b_1, b_2, \cdots, b_m\}$ (means the image i have m blobs). Then according to these blobs in the image i, we retrieve correlative[4] unit images in the unit image library based on the similar probability between every blobs of the image i and every unit image of the unit image library.

### B. *Generating Image's Content Vector*

After being separated and clustered, the image i can be represented a vector $i = \{b_1, b_2, ..., b_m\}$ . Because the feature information of every $b_i$ ( $b_i \in i$ , $b_i$ is the ith blob of the image i.) may be not complete, $b_i$ exists a similar probability with every keyword in the keyword dictionary W. So $b_i$ can be represented a probability vector $P_i = (P_{i1}, P_{i2}, \cdots, P_{in})$ . Next we normalize $P_i$ , that is to say $P_i \xrightarrow{normalize} p_i = (p_{i1}, p_{i2}, \cdots, p_{in})$ , among $p_{ij} = P_{ij} \Big/ \sum_{j=1}^{n} P_{ij}$ . According to the keyword dictionary W, the image i construct a similar probability vector p that which is as follow:

$$p = \left( \sum_{i=1}^{m} p_{i1}, \sum_{i=1}^{m} p_{i2}, \cdots, \sum_{i=1}^{m} p_{in} \right) \quad . \qquad (4)$$

There is an obvious equation $\|p\| = m$ , so (4) can be transformed as follows:

$$\frac{p}{\|p\|} = \left( \frac{\sum_{i=1}^{m} p_{i1}}{m}, \frac{\sum_{i=1}^{m} p_{i2}}{m}, \cdots, \frac{\sum_{i=1}^{m} p_{in}}{m} \right) \quad . \qquad (5)$$

*C. Generating Image's Feature Vector*

In fact, different visual objects of one image have different effects in the process of image classification. So there is a competitive relationship in these visual objects of one image, and different images, which have same image's content vector, can be classified different image classes. So we have to consider areas and locations of every blob, interrelationships of these blobs. Our method is as follow:

*1) Area coefficient*

$$\varepsilon_1(b_i)=\frac{n_i}{n} \tag{6}$$

Note that the coefficient $n_i$ is the pixel point number of the blob $b_i$, and coefficient $n$ is the pixel point number of the image.

*2) Location coefficient*

$$\varepsilon_2(b_i)=\exp\left(-\ r_i^2 \Big/ 2\cdot\sum_{j=1}^{m} r_j^2\right) \tag{7}$$

Note that the coefficient $r_i$ is a Euclidean distance between the barycenter of the blob i and the barycenter of the image.

*3) Interrelationship coefficient*

$$\varepsilon_3(b_i)=\exp\left(-\frac{\sum_{j=1}^{m} L_{i,j}^2}{2\cdot\sum_{i=1}^{m}\sum_{j=1}^{m} L_{i,j}^2}\right) \tag{8}$$

Note that the coefficient $L_{i,j}$ is a Euclidean distance between the barycenter of the blob i and the barycenter of the blob j.

Therefore, an image's content vector can be transformed an image's feature vector as follows:

$$i=\frac{p}{\|p\|}=\left(\frac{\sum_{i=1}^{m}\varepsilon_1(b_i)\cdot\varepsilon_2(b_i)\cdot\varepsilon_3(b_i)\cdot p_{i1}}{\|p\|},\frac{\sum_{i=1}^{m}\varepsilon_1(b_i)\cdot\varepsilon_2(b_i)\cdot\varepsilon_3(b_i)\cdot p_{i2}}{\|p\|},\ldots,\frac{\sum_{i=1}^{m}\varepsilon_1(b_i)\cdot\varepsilon_2(b_i)\cdot\varepsilon_3(b_i)\cdot p_{in}}{\|p\|}\right) \tag{9}$$

## V. CONSTRUCTING IMAGE'S CLASSIFIER

According to above analysis, we construct one N-T-T three-layer BP neural network. In the BP neural network, the matrix P is initialized the weight between the input layer and the hidden-layer, a T dimensional matrix W is initialized the weight between the hidden-layer and the output layer. Firstly the number $N_k$ of images, which belong to the k class in the training samples, is figured out, and the number $N$ of images is figured out in the training samples. So a component $w_{jk}$ of matrix W equal to $\frac{N_k}{N},(k=1,2,\cdots,t)$. One N dimensional inputting vector x equal to one image's feature vector, that is to say $x^T=i$ .The BP algorithm is as follows:

$$y_j=f(net_j)=\frac{1}{1+e^{-net_j}}\text{ , among } net_j=\sum_{i=1}^{n}(p(C_j|w_i)\cdot x_i)$$

$$(j=1,2,\cdots,t) \tag{10}$$

$$z_k=g(net_k)\text{ , among } net_k=2\cdot y_k\cdot w_{kk}-\sum_{j=1}^{t} y_j\cdot w_{jk}$$

$$(k=1,2,\cdots,t) \tag{11}$$

$$g(\bullet)=f(\bullet) \tag{12}$$

*A. Computing the Probability of Un-classified Image*

After the BP network studied by means of the training image samples, the probability of un-classified image can be computed by the formula as follows:

$$P(C_k|image)=\frac{z_k}{\sum_{i=1}^{t} z_i},k=1,2,\cdots,t \tag{13}$$

*B. Feedback Mechanism*

After completing classification, we adjust the matrix P and W by means of the feedback mechanism as follows:

$$w_{jk}=\frac{\Phi\cdot w_{jk}+P(C_k|image)}{\Phi+1},k=1,2,\cdots,t;j=1,2,\cdots,t; \tag{14}$$

Note that the coefficient $\Phi$ is the number of the image library.

$$P(C_k|w_i)=\frac{\Phi_k\cdot P(C_k|w_i)+P(C_k|image)\cdot\frac{\sum_{j=1}^{m}\varepsilon_1(b_j)\cdot\varepsilon_2(b_j)\cdot\varepsilon_3(b_j)\cdot p_{ji}}{\|p\|}}{\Phi_k+\frac{\sum_{j=1}^{m}\varepsilon_1(b_j)\cdot\varepsilon_2(b_j)\cdot\varepsilon_3(b_j)\cdot p_{ji}}{\|p\|}} \tag{15}$$

$$,k=1,2,\cdots,t;i=1,2,\cdots,n;$$

Note that the coefficient $\Phi_k$ is the number of the category k in the image library.

## VI. CLASSIFIED ALGORITHM OF IMAGE

Based on the analysis of last section, we propose a classified algorithm of image as follow:

- Step1: we separate and cluster the image i to get the content vector $\{b_1,b_2,...,b_m\}$ about image i.
- Step2: we calculate the probability of $\{b_1,b_2,...,b_m\}$ basing on unit image library, and get the expression of (4) about the image i, and give the linear expression of formula 6 about the image i.
- Step3: using the BP neural network, we calculate probabilities, which the image I belong to every class of the image library, and sort these probabilities in descending order.

- Step4: we select the top n class names in the probabilities sort, and submit user.
- Step5:the BP network studies the classified result, and adjust the classified probability matrix P and the competitive matrix W base on above feedback mechanism.

## VII. Algorithm Analysis

### A. *Efficiency of algorithm*

According to the existing knowledge about image classification of image library, the algorithm can complete classification of images. The accuracy of classification effect is depended on the matrix P, the matrix W and the accuracy of image's content vector. The algorithm's shortcoming can't discover new class of image. But the algorithm suits to solve image classification in a known environment, and the training of the matrix P and W is high efficiency and requires a short time.

### B. *The Experimental results*

These images are separated into two sets in Ground Truth Database, and these images are artificially classified sever classes. According to the algorithm, one class is the training set to train the BP neural network, which is constructed in the paper, other is the simulation set to test the BP neural network. The experimental results are as follow:

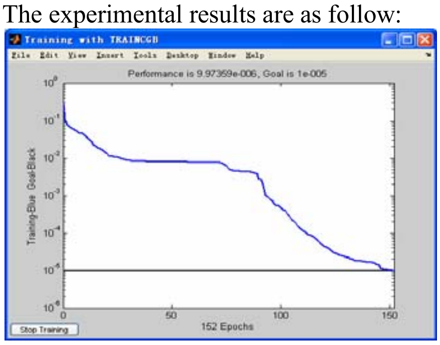


Figure 2. Training Error Curve.

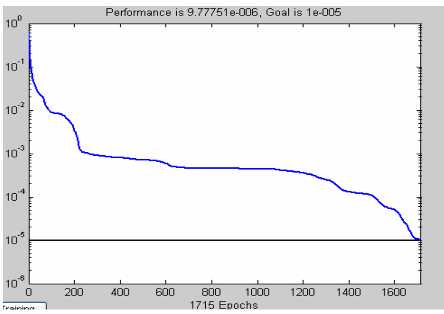


Figure 3. Simulating Error Curve.

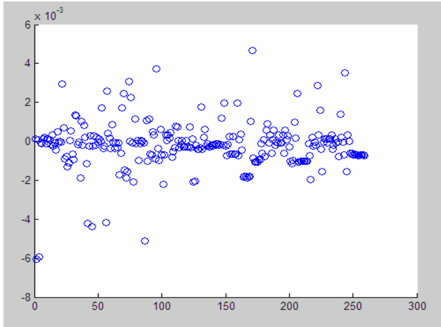

Figure 4. the Classification Error of Images in Training Set.

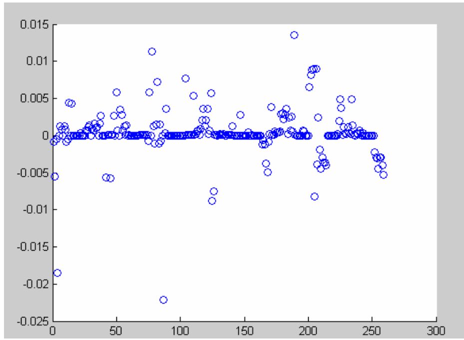

Figure 5. the Classification Error of Images in Simulating Set.

## VIII. Conclusions and Future Work

The algorithm is developed according to the hypothesis in the first section. First of all, the algorithm is feasibility in theory, and its' precision is depended on the accuracy of image keywords annotation. Based on the artificial classification of unit image library, the algorithm can achieve an expected result of image classification. If we hope to improve the performance of the algorithm, we have more work to do, for example: how to improve the accuracy of image annotation, how to induct one of machine learning in the algorithm. These are the focus of our work in the future.

## Acknowledgment

This work was partially supported by the Open Foundation of JiangSu Province Support Software Engineering R&D Center for Modern Information Technology Application in Enterprise under Grant No. SX201002, and this research was supported by the Innovation Team Grant of SuZhou Vocational University NO.3100124.

[1] J. Jeon, V. Lavrenko and R. Manmatha Automatic Image Annotation and Retrieval using CrossMedia Relevance Models July 28–August 1, 2003

[2] Shaowu Peng, Leyuan Liu, Xiong Yang and Nong Sang Visual knowledge representation based on image grammar and annotated database Application Research of Computers Vol. 26 No. 2 Feb. 2009

[3] Chong Pan and Hongbin Zhu Application of improved k-means algorithm in image annotation and retrieval. Computer Engineering and Applications 2010 46(4) : 183-185

[4] Jun Luo and Hang Kuang Content-based image retrieval using combination features. Computer Engineering and Applications 2009 45(1):153-155