International Journal of Hybrid Information Technology
Vol.9, No.3 (2016), pp. 291-300
http://dx.doi.org/10.14257/ijhit.2016.9.3.27

The Image Multi Feature Retrieval based on SVM Semantic Classification

Che Chang[1,2*], Yu Xiaoyang[1] and Bai Yamei[3]

1. Measuring and Control Technology and Instrumentations,Harbin University of Science and Technology, Harbin, China
2. School of Engineering,Harbin University, Harbin, China
3. School of Electronic and Information Engineering,Harbin Huade University Harbin, China
E-mail: chechang@163.com

Abstract

The image retrieval performance based on single feature is limited. For different kinds of images, it can not a better retrieval result. This paper raises image retrieval method based on weighted multi feature. In each kind of images, each feature precision is the weight evidence. On this basis, we research the existing semantic retrieval technology. Choosing the SVM classification theory which is more mature. Selecting parts of images as training set. Doing the training classification. From the research of different characteristics priority, it raises the image retrieval technology which synthesizes SVM and multi feature. From that, it can get a higher retrieval efficiency.

Keywords: *image retrieval, multi feature, training classification, SVM*

1. Introduction

In recent years, a large number of multimedia information such as image, text, video increase in geometric style. Everyday the growth in digital images is very huge. At present, one of the focus issues is how to search results which the user want quickly and effectively from a large number of images.

In order to solve the problems above, the earliest image retrieval system based on text was exploited. This system needs a large number of artificial workload. Making keys to all the images. With the emergence of massive images, the workload of manual image annotation is huge. It would make the image understanding indefinitely. In order to reduce the human factor in this kind of retrieval. Using image visual content features to retrieve images (CBIR) is being raised. This retrieval system focus on the human vision. The visual features are as the query condition. CBIR has a big advantage in understanding image content correctly. However, with the development of technology, the image retrieval technology could not meet the technical requirements. After all, the images visual feature could not represent the image connotation completely. To different kinds of query images, only one independent feature retrieval returned result could not meet people wills. It would make the retrieval result unreliably. In this paper, we use multi feature fusion to retrieve images. Choosing the basic visual features such as color, texture and shape. They can patch up the other drawback to a certain degree. Each feature is organic integration. By introducing decision theory and FFT algorithm, the retrieval performance is more easy and smooth. Besides, the semantic classification based on SVM can promote the theoretical level and pragmatic value of image retrieval.

ISSN: 1738-9968 IJHIT

2. Basic Principle of SVM

Support Vector Machine (SVM) is one of statistiacal learning theory. The theoretical basis of this method is closely. It could be used in the nonlinear function. This method uses kernel function expansion theorem. It can get the linear function which do not need the nonlinear mapping expression. Using this method can get the global minimum. The other algorithms only get the locally optimal solution.

2.1VC Dimension Theory

VC dimension is an important indicator which can reflect the learning ability of function set. If F is the function set, VC dimension is the elements number which the function in F could do the right category in R^n.

$$Z_n = \{(x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n)\} \tag{1}$$

For any $x_i, i = 1,2,\cdots,n$, Giving an arbitrary label, 1 or -1. Making up the set:

$$T = \{(x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n)\} \tag{2}$$

There is one function in F which could classify T accurately. f could be shown as:

$$f(x_i) = y_i, i = 1,2,\cdots,n \tag{3}$$

The VC dimension of F is n.

VC dimension describe the learning performance strength of function set F. If the VC dimension in function set is larger. The function set is high complexity and high-capacity. How to confirm the VC dimension of function set is an important problem in statistical theory.At present, it is hard to deal with.

2.2Minimization Principle of Structure Risk

The minimization principle of structure risk could ensure the classification accuracy. Besides, it can reduce the VC dimension, which can control the risk. It can be shown as:

$$R(f) \le R_{emp}(f) + \sqrt{\frac{8}{n}(h(\ln\frac{2n}{h}+1)\ln\frac{4}{\sigma})}, \tag{4}$$

f expresses the function which is in the decision function set. n expresses the elements number in the training set. h expresses the VC dimension of decision function candidate set F. $\sigma \in (0,1]$.

From formula 4, the empiric risk is on the right side of the first and the second:

$$\varphi(h,n,\sigma) = \sqrt{\frac{8}{n}(h(\ln\frac{2n}{h}+1)\ln\frac{4}{\sigma})}, \tag{5}$$

In formula 5, $\varphi(h,n,\sigma)$ is a decreasing function. When the number of training samples n is bigger,the $\varphi(h,n,\sigma)$ is smaller. If the training samples number is too small, we need to consider the confidence interval.

The structure risk is defined as the sum of empirical risk and confidence interval in STL. It needs to consider the confidence interval also. The structure risk minimization is shown as:

$$T = \{(x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n)\} \tag{6}$$

From that, $x_i \in R^n, y_i \in \{-1,1\}, i = 1,2,\cdots,n,$ The decision function set F was decomposed to the decision set $F(t)$ which is related to parameter t. When $\forall t_1 < t_2$, $F(t_1) < F(t_2)$.

3. SVM Classification

3.1SVM Dichotomies

SVM is from the development of linear separable case. Solving the situation of linear non separable. It can built the SVM model. It can also be linearized and solved. SVM could be expressed the question of convex optimization. Getting the global minimum of objective function. Its optimal hyperplane is shown as in Figure 1.

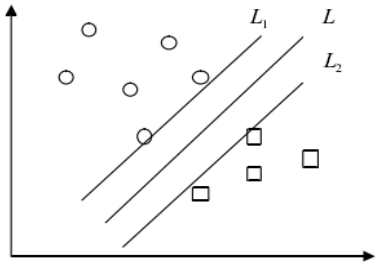


Figure 1. The Optimal Separating Hyperplane

The training set $T=\{(x_1,y_1),(x_2,y_2),\cdots,(x_n,y_n)\}$ would produce a lot of classifications. The optimal separating hyperplane $((\omega\cdot x)+b)$ should meet the following conditions.

$f(x_i)=y_i, i=1,2,\cdots,n$;

$\frac{2}{\|\omega\|}$ achieve optimal in all classification plane;

To support vector x, $f(x)=0$. To non support vectors x, $|f(x)|\geq 0$.

From the three conditions above, evaluating the optimal hyper plane f become a convex quadratic programming problem which can evaluate ω and b.

$$\min_{\omega,b}\frac{1}{2}\|\omega\|^2 \tag{7}$$

$$s.t.\ y_i((\omega\cdot x_i)+b)\geq 1, i=1,2,\cdots,n \tag{8}$$

Through a certain transformation, the training set is projected to high dimensional feature space. Then searching the optimal separating hyper plane in high dimensional space. The hyperplane do the corresponding transformation, which change into classification in the original space. As is shown in Figure 2.

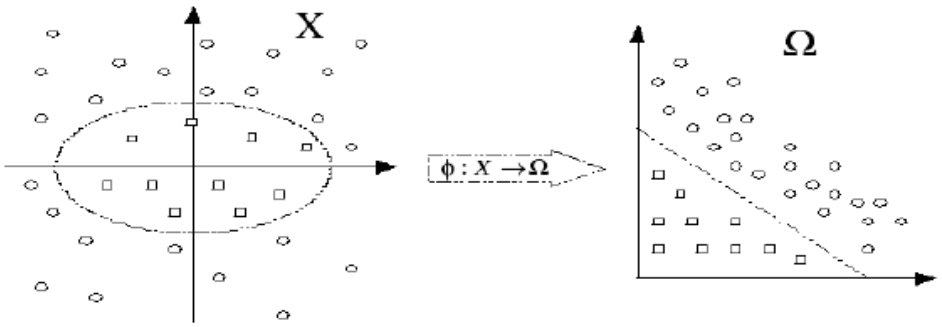


Figure 2. Nonlinear Mapping: the Spatial Data of X is Mapped onto Feature Space

The original feature space is mapped into Hibert space H, Φ is shown as:

$$\Phi : R^n \to H$$
$$X \to z = \Phi(x) \tag{9}$$

The optimal hyper plane in Hibert space H meet the following conditions:

$$(\omega \cdot z) + b = 0 \tag{10}$$

The optimal hyper plane which is transformation of the original space meet the following conditions:

$$(\omega \cdot \Phi(x)) + b = 0 \tag{11}$$

3.2SVM Multi-Classification

SVM algorithm can determine the data belong to positive class or anti class only. It is a two class classifier. But it is not so easy in practical application. It contains a lot of different categories generally, such as text classification, face recognition. In general conditions, we divide the area into many regions to distinguish the hyper plane. But it is theory of mind. The calculation quantity is very huge in practice. The workload is enormous in the real application. There are two ways to do the SVM multi-classification. One is to built a SVM classifier which has multi classification function, the other is to combine multiple SVM binary classifier. There are lots of combinations. They are several common methods following:

One to multi class classification method: It can only solve the problem of two class classification at any time. Such as both sides of the coin. If there are n kinds of different categories, 1 categories of samples is set to the obverse side of a coin, Other categories of samples are the reverse side of a coin. It would get the first two class classifier. After that, class 2 is changed to class 1 according to the first step. The others are unchanged. It would get the second two class classifier. Carring out that with the steps. It would get n class classifier. When making the quick sort, the classification would be overlap and unable to classify.

One to one class classification method: Each category built two kinds of class classifier beyond itself. From mathematical formula, if there are n categories, it built n(n-1)/2 classifiers. This method would not appear the inseparable phenomenon. When n is large, the number of classifier would be very big.

Directed acyclic graph classification method: Making 7 as an example. Classifying 1 to 7 with the classifier. If the classification result is 7, Classifying 2 to 7 with the right classifier. If the classification result is 2, Classifying 1 to 2 with the left classifier. Finally, it would get the right classification result. When the classification error is appearing, any classification process follows could not be corrected. The mistakes would always be continue.

4. Image Retrieval Experiments and Performance Analysis based on SVM Semantic Classification

4.1SVM Training Algorithm

The penalty function method and simple method are most common in quadratic programming method. When the samples number is enlarge, the complexity of the algorithm would be growing. The increasing trend is large. In order to overcome this shortcoming, lots of new training algorithms are appearance. The large scale primary problem is divided into some small-scale subproblem in most of the algorithms. From iterative solution, it get a good result. The decomposition algorithm steps are as follow:

If a is the subset size, to n samples, $\alpha_i = 0$;

Choosing a samples in the training set. Forming subset A;

Calculating quadratic programming question which is expresses by subset A from standards and procedures;

All relative dense data can use this model;

If the results are still divergence, again from the beginning step 3.

4.2 Experimental Procedure

In this paper, the framework all image classification system is shown as follows:

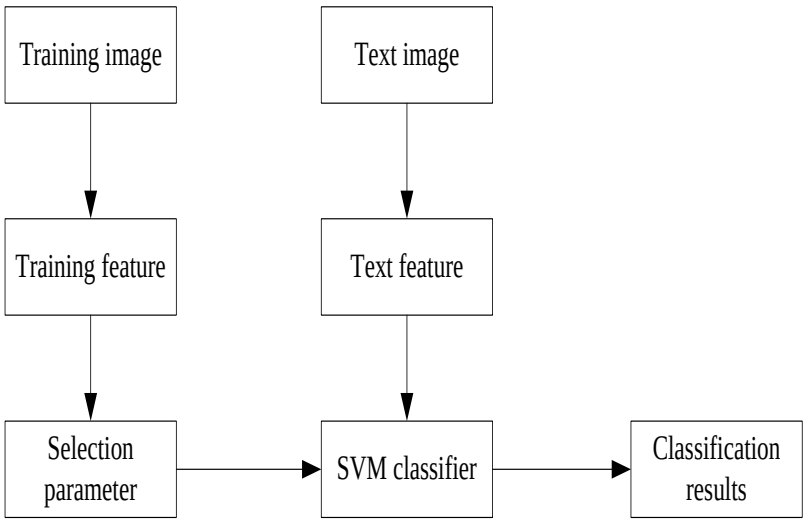


Figure 3. The Process of Image Classification System

The steps of this method are as follows:

The first step: Choosing 1000 images in image library. There are 100 images in each category. Choosing 80 images as training set in each category. The rest 20 images are used to test the retrieval performance.

The second step: From the groundwork and preparation above. Getting the feature vector such as color, texture and shape. Doing the artificial classification according to the classification of image database. Comparing the influence of image classification from different characteristics. Increase the efficiency of image classification further.

The third step: Lots of mature software about SVM algorithm appear in the market. The LIBSVM software system updates all the time. It can be wrote by Matlab language. LIBSVM is easy to use. This paper choose latest version in LIBSVM official website.

The forth step: Normally, the classifying quality of RBF kernel function is better. So, SVM algorithm use RBF kernel function to do the image classification. The acquiescent kernel function in LIBSVM system is as the classification algorithm parameters. Training the classification model.

The fifth step: After classification, retrieving thought normal procedure.

The sixth step: The precision ratio is as the evaluation criterion in this classification and retrieval experiments.

4.3Experimental Results and Analysis

In the fusion experiments, it returns the first 10 most similar images as experimental results to output. Selecting the native images as query images to query. There are 100 images related to the sample images in image library. The precision ration of image retrieval algorithm based on DS theory five feature fusion is 0.30. Removing the lowest single characteristics precision Hu matrix (0.11). The precision ratio which the rest four kinds of characteristics do the DS theory integrating retrieval is 0.60. The African image

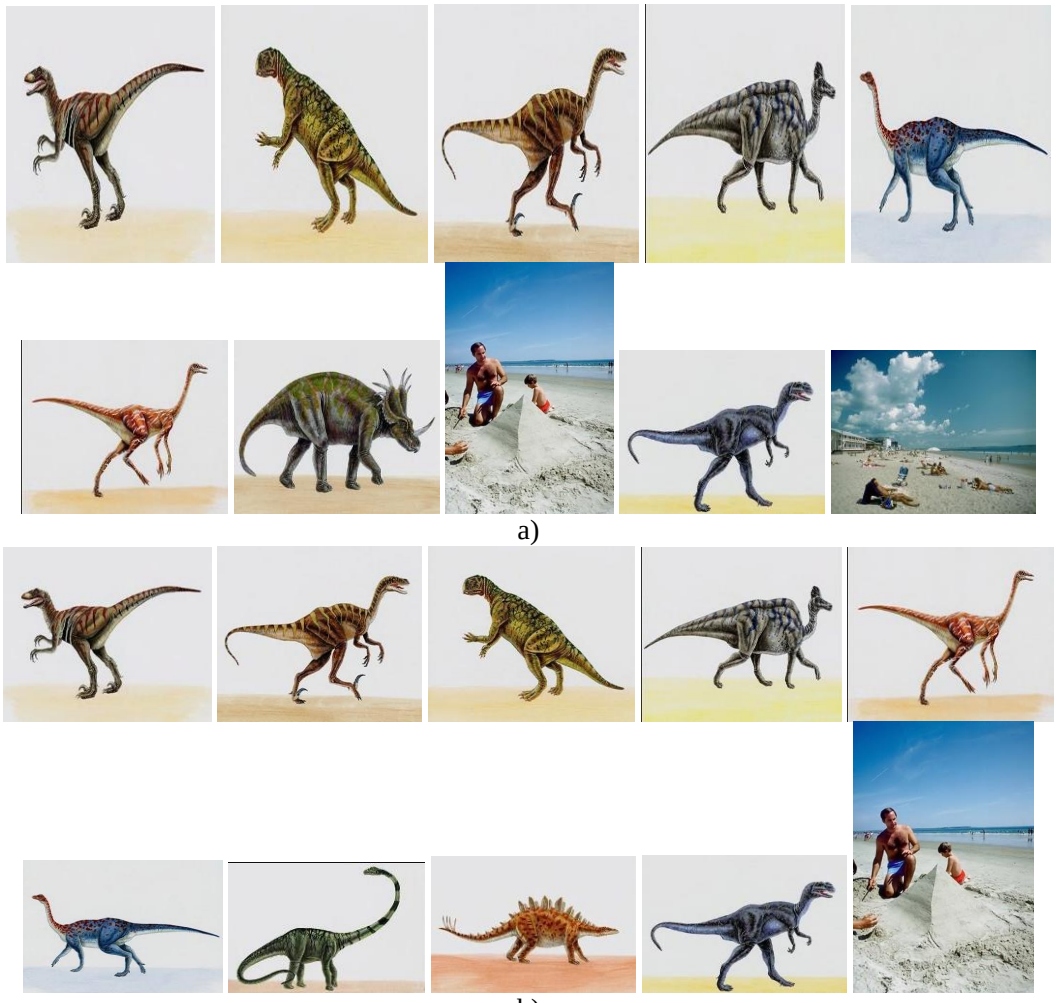
retrieval based on five characteristics DS fusion is shown in Figure 4a. The African image retrieval based on improved DS fusion is shown in Figure 4b.

a)

b)

Figure 4. African Image Retrieval based on DS Fusion

Selecting the dinosaur images as query images to query. There are 100 images related to the sample images in image library. The precision ration of image retrieval algorithm based on DS theory five feature fusion is 0.80. Removing the lowest single characteristics precision color correlogram (0.11) and gray level co-occurrence matrix (0.11). The precision ratio which the rest three kinds of characteristics do the DS theory integrating retrieval is 0.90. The dinosaur retrieval based on five characteristics DS fusion is shown in Figure 5a. The dinosaur image retrieval based on improved DS fusion is shown in Figure 5b.



a)

b)

Figure 5. Dinosaur Image Retrieval based on DS Fusion

Choosing Corel image database to do the image retrieval based on semantic classification also. Selecting the native images as sample images to query. There are 100 images related to the sample images in image library. In this paper, displaying the experimental results of the first 10 images. The precision ration of image retrieval algorithm based on SVM semantic classification is 0.90. It is shown as in Figure 6.

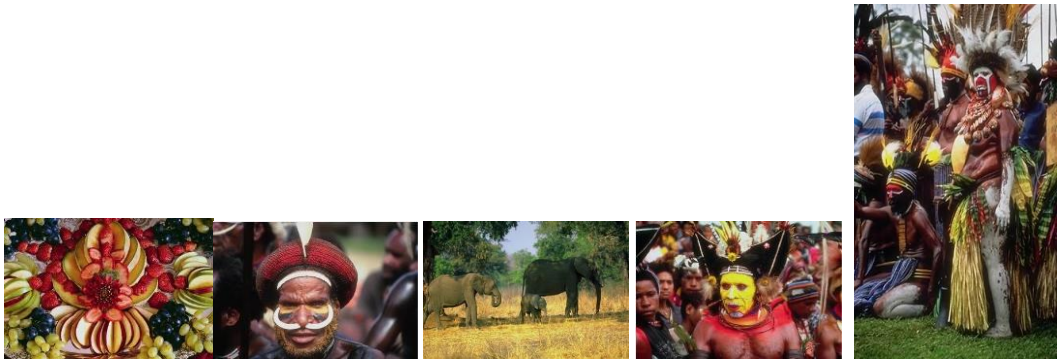


Figure 6. Image Retrieval Results for Indigenous based on SVM Semantic Classification

Selecting the dinosaur images as sample images to query. There are 100 images related to the sample images in image library. In this paper, displaying the experimental results of the first 10 images. The precision ration of image retrieval algorithm based on SVM semantic classification is 0.95. It is shown as in Figure 7.

Figure 7. Image Retrieval Result for Dinosaur based on SVM Semantic Classification

Selecting the snowberg images as sample images to query. There are 100 images related to the sample images in image library. In this paper, displaying the experimental results of the first 10 images. The precision ration of image retrieval algorithm based on SVM semantic classification is 0.75. It is shown as in Figure 8.

Figure 8. Image Retrieval Result of Snow Mountain based on SVM Semantic Classification

Three kinds of image retrieval algorithm results are shown as Table 1:

Table 1. Retrieval Performance Comparison of Three Kinds of Multi-Feature Fusion Algorithm

Precision ratio	Weighted fusion retrieval	Weighted fusion retrieval based on DS theory	SVM semantic classification retrieval
Native	0.75	0.60	0.90
Dinosaur	0.73	0.90	0.95
Snowberg	0.50	0.60	0.75
Average precision	0.66	0.70	0.87

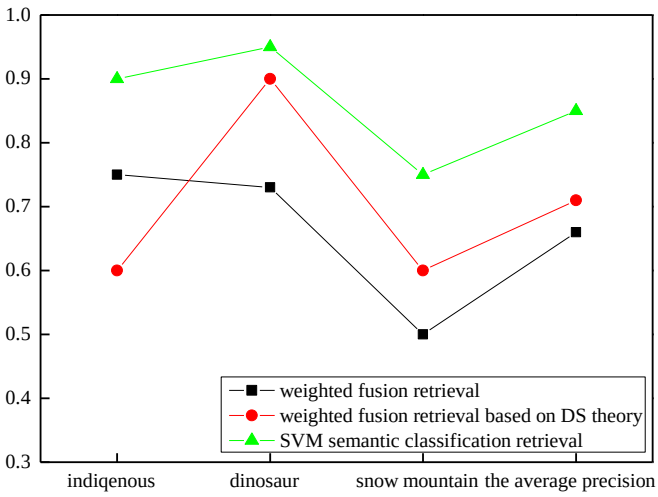


Figure 9. Precision Comparison of Three Kinds of Multi-Feature Fusion Algorithm

From the experimental result, the image retrieval algorithm based on SVM semantic classification is better than multi feature fusion algorithm based on content in retrieval average accuracy. Some images are very similar to the query images in image library, but they are not the target images actually. It would make the detecting error. The cause of the error is that the retrieval only adapt the color, texture and shape. It needs some manual annotation in the semantic classification training.

5. Conclusion

This paper introduces the concept of semantic retrieval. It raises multi feature image retrieval algorithm based on SVM semantic classification. Choosing LIBSVM algorithm to do the SVM training and classification. From the experiment, the average precision of image retrieval algorithm based on semantics is better than the image retrieval algorithm based on content. There is a ceratain degree of advantage.

References

[1] M. Flickner, H. Sawhney, W. Niblack, *et al.*, "Query by image and video content: the QBIC System", IEEE Computer, vol. 28, no. 9, **(1995)**, pp. 23-32.
[2] R. Yong, "3D Measurement Method by Combining RGB Gray Code with Trapezoidal Phase-shifting", Journal of Harbin University of Science and Technology , **(2011)**.
[3] C. Beecks, S. Kirchhoff and T. Seidl, "On stability of signature-based similarity measures for content-based image retrieval", Multimedia Tools and Applications, vol. 71, no. 1, **(2014)**, pp. 349-362.
[4] M. Flickner and H. Sawhney, "Query by image and video conterit", The QBIC System IEEE Computer, vol. 28, no. 9, **(1995)**, pp. 23-32.
[5] J. Assfalg, A. Del Bimbo and P. Pala, "Three-dimensional interfaces for querying by example in content based image retrieval", Visualization and Computer Graphics, vol. 8, no. 4, **(2002)**, pp. 305-318.

[6] Y. Fang, K. Yamada, Y. Ninomiya, *et al.*, "A Shape-independent Method for Pedestrian Detection with Far-infrared Images", IEEE Transactions on Vehicular Technology, vol. 53, no. 6, **(2004)**, pp. 1679-1697.
[7] B. Wu and R. Nevatia, "Detection of Multiple, Partially Occluded Humans in a Single Image by Bayesian Combination of Edgelet Part Detectors", IEEE International Conference on Computer Vision, vol. 1, **(2005)**, pp. 90-97.
[8] Z. Lin and L. S. Davis, "A Pose-invariant Descriptor for Human Detection and Segmentation", European Conference on Computer Vision, **(2008)**, pp. 423-436.
[9] M. Andriluka, S. Roth and B. Schiele, "People-tracking-by-detection and People detection-by-tracking", IEEE Conference on Computer Vision and Pattern Recognition. Anchorage, AK, **(2008)**.
[10] P. Srivastava, N. T. Binh and A. Khare, "Content-Based Image Retrieval Using Moments of Local Ternary Pattern", Mobile Networks and Applications, vol. 19, no. 5, **(2014)**, pp. 618-625.

