Advanced Materials Research *Online: 2012-06-14*
ISSN: 1662-8985, Vols. 532-533, pp 1297-1302
doi:10.4028/www.scientific.net/AMR.532-533.1297

A Survey on Emotional Semantic Mapping in Image Retrieval

Zengrong Liu[1, a], Zhi Li[1, b], Xueli Yu[1, c]

[1.]College of Computer Science and Technology, Taiyuan University of Technology, Taiyuan Shanxi, China

[a]zr_liu@sxinfo.net, [b]lizhi_tyut@hotmail.com, [c]yugroup@126.com

Keywords: emotional semantic; emotional semantic mapping; image retrieval; feature extraction

Abstract. Emotion plays an important role in the human perception and decision-making process. Human comprehension and perception of images is subjective, and not merely rely on lower-level visual features. Semantic gap is regarded as the most important challenge of image retrieval. In this paper, we analyzed the emotional features as well as emotional semantic description of images, which comes from the image emotional semantics retrieval framework. And also the mapping ways and means were summarized from image visual features to emotional semantics. Finally, the disadvantages of emotional semantic mapping and developing tendency were discussed.

Introduction

With the rapid development of information technology in recent years, multimedia information has increased largely. Such information's storage, management, organization, retrieval have become a public problem. Since 1970s, image retrieval technology has developed into two basic retrieval methods: TBIR and CBIR. The advantage of TBIR techniques is that the retrieval result will surely be accurate when the image is labeled completely. However, the label work will not only be great, but also bring about subjectivity and uncertainty when the image database is extremely large, which will be difficult to retrieve technology. CBIR technology is to extract low-level visual features and then compute according to similarity. In most cases, human comprehension about the image is based on not just the visual features but also the image content. It can not be obtained directly through visual features of the images, which is defined as the semantic features of the image. It is usually uneasy to express the semantics of images by visual features. In some cases, there exists great semantic difference among those images even with very similar visual features. As a result, a disaster occurs in CBIR query. To avoid this situation, CBIR should have the ability to deal with high-level semantic features [1].

Image semantics include object semantic, spatial semantic, scene semantic, behavioral semantic and emotion semantic, among which the highest level is the emotion semantic. It is usually described by the adjectives that can well express human emotions [2]. It is very difficult to obtain image emotion semantic by low-level visual features just because there is no direct relation between high-level image semantics and low-level visual features, which is usually called the semantic gap. The semantic gap refers to the different expressing abilities between the image low-level features and the high-level richer semantics. The existence of the semantic gap leads to the difficulty that the low-level image features can extrapolate the high-level semantics. So we addressed these two levels in our study, focusing on how to create a mapping method from the low-level features to the high-level emotion semantic.

The framework of emotion semantic image retrieval

Present CBIR system is mostly realized according to the similarity matching queries of multi-dimensional physical features of the image. For the user's preferences, emotions and other subjective factors, it is not considered or less. The method of emotional image retrieval mapping is from the user's emotion, which is used for studies of the relation and mapping between high-level emotions semantic and low-level visual features so as to establish the interactive model of dynamic image emotion semantic retrieval system. The model is as follow (Fig. 1).

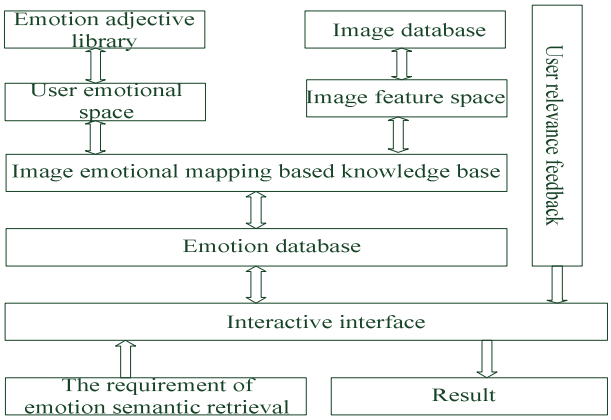

Fig.1 The Model of emotion semantics image retrieval

The description of image emotion semantic

Image semantics is abstract, vague and complex, which concerns cultural background, nationality, hobbies and other aspects. It is obviously not enough to describe the semantics just based on low-level features. And also the abstract description of the image need be done by high-level emotional semantics. In other words, some expressions about emotion need be used in the level of the emotional semantics.

A. Emotional Feature Description Based on User's Experience

The emotional features of image generally refer to the visual features that can arouse people's senses, impressions, emotions, and other subjective experience. The image usually implicates large amount of information, because different shape, color, texture and different combinations of space shape will give different psychological reactions, such as happy, disgust, anger and so on. Careful study of the visual features of the image, and that of the law according to the physiologic, psychological and emotional response caused, as well as the appropriate way to study how to extract these features, quantitative description of these features are based on the premise of emotional semantic retrieval.

- The color has certain stability, which is the most intuitive features of an image. Color can wake up emotions, and lead to people's associations of ideas. It is a fact that color associates with a certain emotion. The expression methods about color features used commonly are color sets, color matrix and color histogram.
- Texture is an important and difficult feature to describe, which is still no the exact definition given upon now. Texture features include roughness, regularity, line similarity, concavo-convexity, direction, and contrast. Texture analysis methods are commonly used grayscale co-matrix, KL transform, texture spectrum analysis, wavelet transform. The emotional affect from texture is not as strong as color. However, the psychological and emotional effects can not be ignored because of the contrast and spatial frequency concerned. At present, both Tamura texture features and Gabor wavelet transform are consistent with human visual perception, which can be used to analyze the role of texture in the aspect of emotion.
- Shape has its own aesthetic values, which is another essential feature besides color and texture. Shape features can stimulate people's emotions, and different shapes express different visual effects and thoughts and feelings. Common feature representation method includes boundary-based methods, region-based methods, and Fourier descriptors, matching the finite element method and wavelet transform etc.

B. The Way of Eemotional Semantics Description

How many kinds of emotional states human has? There is no definite answer for this question. There is an old saying in China: "Seven Emotions and six Desire" which shows different versions in modern psychology. In fact, there are two theoretical systems available nowadays: that the emotion includes basic emotions and compound emotions; the compound emotions are composed of basic emotions. This so-called "dimensions of emotion" thinking is psychology. In the classification of

basic emotions, the theory of six basic facial expressions presented by Ekman has been extensively applied, which covers happiness, sadness, surprise, fear, anger and disgust [3]. Human emotions are integrated complexly from six basic facial expressions. Dimensions of thinking means that the basic emotions are used as a vector and these vectors create an emotional space through appropriate methods to represent the image semantics. Usually people use adjectives to express their feelings. By emotional semantic description we mean how to deal with and express the user's psychological activity, using the adjectives by computer.

The establishment of emotional space is as following steps [4]:

- Collect adjectives to express emotion and create adjectives library to express emotion;
- Do emotional semantics quantitative experiments, select image samples, classify the emotions, and evaluate by the user, create a user emotion database;
- Analyze the redundant information of the emotional adjective library by factor analysis methods, finally create the emotional space.

Image emotional semantic mapping

Let an image database D= $\{ I_1, I_2 \dots I_p \}$ (where I_m is a database of images, m = 1, ..., p), an emotional semantic space E=$\{ W_1, W_2 \dots W_q \}$ (where W is vector in emotional space), by the emotional semantics mapping, it is indicated that given an image ,the most match emotional semantic description can be found from the emotional space ; given an emotional semantics, the image set included the emotional semantics W can be found from image database. The mapping ability learns from training set T= $\{ I_1, W_1 \} \dots (I_r, W_r)$.

The emotion semantic mapping method of image was first based on linear, such as regression analysis, multivariate analysis, and quantity theory and related computing method etc. However, the mapping methods based on linear could not describe the relationship accurately because of the complexity from image features and emotional relationships,. With the present advancement of non-linear mapping method in pattern recognition, such as artificial neural network (ANN), Hidden Markov Model (HMM), Gaussian mixture model (GMM), support vector machine (SVM) etc., the mapping method of image emotion semantic has also made good progress.

A. Image Emotional Semantic Mapping Based on Linear Method

Linear classification method is a way of categorizing the sample through classification feature vectors of sample located in a feature space which can be classified by a straight line (in two-dimensional space) or a hyperplane (in multi-dimensional space).

Kato had used linear regression method through the evaluation of user to the visual features classification of training image set and similarity agree to deduce the mapping relationship between visual features of image with user emotion, so as to extract the semantic in 1993 [5].

B. Image Emotional Semantic Mapping Based on Artificial Neural Network

Neural network is put forward on the basis of the research harvests in the modern neuroscience and it is one of the most active branches concerning computational intelligence and machine learning [6]. Neural network is the human brain function integration of simplification, abstraction and simulation, which consists of a large number of computing units with nonlinear mapping abilities - neurons. Neurons are connected by weighted value and the information is distributed in a complex network through the connection of the weight coefficients which can acquire knowledge and solve problems by way of training. In the process of emotion mapping, the first task is to construct a sub-network which has the architecture of MLP with the ability to extract the feature vector to itself. Train for each emotion a sub-network, and make decision of emotion using the outputs in each sub-network.

In 1997, Takashi Hayashi and Masafumi Hagiwara use neural networks to establish the relation between the impression words and image visual features to achieve the extraction of emotional semantics [7]. Shangfei Wang etc. use radial basis function neural network, whose inputs are the sample images features and outputs are the sample images coordinates in the emotional space. After learning, the weights memorize the connection between image features and the user's emotion. So the image from the feature space is mapped to emotional space, in which perceptual retrieval would be realized and good experimental results would be acquired [8].

C. Image Emotion Semantic Mapping Based on SVM

Support vector machine (SVM, Support Vector Machine) put forward by Vapnik and CorTes in 1995, is a machine learning method based on VC dimension theory in the area of statistical learning theory as well as structural risk minimization principle, which is used to find the best compromise in the model of the complexity and learning ability based on the limited sample of information. It aims to get the optimal solution under the existing information rather than the sample tending to infinity, so it is a learning method to obtain the best generalization ability. It has shown many unique advantages in solving the problem of the small sample, nonlinear and high dimensional pattern recognition.

Hualin Wan and Morshed U. Chowdhury, using the integrated texture, edge and color histogram as image feature vectors, together with radial basis function as kernel function of SVM, implemented the semantics classification of images [9]. Yu etc. used a Gaussian-kernel-function SVM algorithm in order to identify four kinds of emotional states, the average recognition rate reaching 74.3%. Besides, compared with the neural network algorithm, SVM has better robustness .

In image emotional mapping, low-level feature is a high dimension. The emotion annotations labels in training set with its subjective and quantitative reasons make it impossible to be carried out on a large scale. Therefore, SVM based on structural risk minimization principle is superior to the traditional learning methods based on empirical risk minimization principle. SVM can better settle the mapping problem of limited sample high-dimensional model, and it has better prediction performance.

D. Image Emotional Semantic Mapping Based on Fuzzy Theory

Fuzzy Theory established by Professor Zadeh in 1965, developed based on mathematics of fuzzy set theory, mainly including fuzzy set theory, fuzzy logic, fuzzy reasoning, fuzzy control and other aspects. Fuzzy objects have not explicitly connotation and extension, so it is not clear. However, in some specific conditions, it will lead to a more specific meaning by using some vague concept to describe a particular object. This ability can make qualitative thinking and judgment quantitative so as to adapt to the computer's processing, which represents the human experience, common sense. In addition, the model of human sense and language expressions can also be created in order to implement emotional semantic mapping.

Jinsub established the mapping from color characteristics to emotional semantics with the help of self-adaptive fuzzy system, and compared with the mapping from color to the emotions by neural network [10]. Isomoto etc. used fuzzy set to represent the color, shape and other characteristics, and based on membership function to represent the emotional impression words .

E. Image Emotional Semantic Mapping Based on Interactive Genetic Algorithm

The GA (genetic algorithm) proposed by J. Holland in 1975, is a computation model by simulating natural selection, survival of the fittest etc. genetic mechanism in biological evolution and a method by simulating search the optimal solution in natural evolution. The IGA (Interactive genetic algorithm) replaces the traditional GA fitness function designed to improve the convergence speed and search capabilities by means of an interactive assessment of the individual user.

Sung-Bae Cho uses IGA in his emotional image retrieval system. In the evolutionary process, the value of fitness function is determined by the user, then the algorithm search in the feature space through selection, crossover, and mutation etc. genetic operation. User gives the value of fitness function to evaluate the similarity between search results and the user subjective mental space, which will introduce the subjective feelings to the evolution of GA so as to achieve the subjective psychological space and objective feature space mapping. Though the method by user assessment replacing fitness-function has obvious advantages, it also brings out uncertainty for mapping because user emotions are affected by many factors, which is difficult to evaluate the multiple individuals maintain stability.

F. Other Mapping Methods

The algorithms of image emotional mapping are not limited to what mentioned above. A common pattern recognition method, such as Bayesian classification, principal component analysis, hidden Markov model and the quadratic discriminate mapping method. Human emotion as subjectivity, abstractness and complexity, while the extraction methods also affect the

understanding of human emotions on the image, so a simple mapping method often does not achieve a better mapping from image to emotion, and sometimes combination of several mapping methods, coupled with user participation is necessary to achieve higher accurate mapping.

Future study of emotional semantic mapping and retrieval about image

With the efforts of the researchers, there is some progress done in the study of the image emotional semantic mapping and retrieval. However, it is not enough to improve the efficiency and effect of emotional semantic retrieval only by emotion semantic mapping. In fact, there are still many problems to solve.

A. Processing Low-level Features

Effective low-level visual features are extracted to express emotion semantic of image. At present, the researchers have studied low-level visual features of image such as color, texture and shape, but there are still some problems, such as high dimension vectors, imperfect feature extraction algorithm, immature image segmentation algorithm etc., which have affected the extraction of emotional semantics. Image emotion semantic is expressed by different visual feature. The images and regions that have similar emotion semantic should have similar global or local visual feature, which is the premise of study of emotion semantic mapping. Emotion semantic mapping need more low-level features that are fused in order to overcome influence from the light, rotation and other factors. The mapping method of multi-feature fusion can express accurately emotion and improve the accuracy of emotional semantic retrieval. It is a very interesting research subject that more valuable low-level features are extracted by stable, faster image segmentation and integrated to achieve more accurate emotional semantic.

B. Improving Emotional Semantic Mapping

In existing mapping methods of emotion semantic, the samples are trained by a prior knowledge and then recognized to get emotion semantic by machine learning. These methods have their own advantages, but the disadvantage is equally clear. Ideal method of mapping method should can be learned automatically and accumulate gradually knowledge to annotate precisely a new image with the premise on minimization interaction of human-computer. Therefore, new ideas are proposed and achieved by effective algorithms. All in all, it is a valuable research subject that mapping mechanism of emotion semantic is perfected to improve mapping precision.

C. Strengthening Emotional Semantic Description

Emotion semantic description of existing is still at the exploratory stage, and there is no uniform manner. At present, according to the different applications, the emotion space is realized through simple relatively psychology experiment. The emotional adjectives are usually selected based on the experience or some psychological investigation in specific applications. These experiences and results are concerned with district, culture, education, individual aesthetic etc., with some differences of common or personality, which are just reduced through psychology or other experiment. On the other hand, the distance between emotions semantics should be defined. The emotional semantics are quantified and calculated the distance to evaluate the semantic similarity between two emotions. The related work is still strengthened.

D. Performance Evaluation

An effective performance evaluation is necessary for a new study or technologies and its application. Objective performance evaluation is a right guidance for a new study or technology. For retrieval, the general assessment is performance and efficiency. Precision ratio and recall ratio are evaluation criteria for image emotional semantics. Because of the subjective, abstract and complex of emotional semantics, different people have different emotional reactions to the same feature of image. How to objectively judge a emotion semantic retrieval system is lack of uniform criteria. How to study and evaluate objectively the criteria is a well worth subject.

Conclusion

Emotion is a physiological and psychological phenomenon. It is difficult to describe and define emotion accurately. By the research of image emotion semantic, it is attempted to find a model which can be identified by computer and accurately describe and express the emotional reaction when images occur. The research relates to the physiology, psychology, visual perception, cognitive science, and many other domains and requires the cooperation of many subjects.

For the emotional mapping, the feature extraction and mapping algorithm selection is very important. In the past few years, some progress has been made in the research of this field. However, it is still not a common standard. In future, computer vision, intelligent science, psychology, information retrieval and other achievements should be utilized in emotional mapping to achieve greater progress.

Acknowledgment

We gratefully acknowledge the support from National Natural Science Foundation of China (No.60873139), Shanxi Natural Science Foundation (2008011040), and State Key Laboratory of Virtual Reality Technology and Systems (SKVR-KF-09-04).

References

[1] Vasconcelos N, “From pixels to semantic spaces: advances in content-based image retrieval.”computer, 2007,40(7):20-26.

[2] R. W. Picard. Synthetic emotion. IEEE Computer Graphics and Applications. 2000,20(1):52-53

[3] P. Ekman, W. V. Friesen, “Manual for the Facial Action Coding System.”Consulting Psychologists Press, 1977.

[4] K. Yoshida, T. Kato, T. Yanaru, “Image Retrieval System Using Impression Words Systems, Man, and Cybernetics.” 1998 IEEE International Conference on [C]. 1998, 3(11-14): 2780–2784.

[5] Yoshida K, Kato T, Yanaru T. Image Retrieval System Using Impression Words Systems, Man, and Cybernetics [A]. 1998 IEEE International Conference on [C]. 1998, 3(11-14): 2780–2784.

[6] Simon Haykin, Translated by Ye Shi Wei, Shi Zhong Zhi, “Neural Networks: A Comprehensive Foundation, 2nd Edition.” China Machine Press ,2004(in Chinese).

[7] T. Hayashi, M. Hagiwara, “An image retrieval system to estimate impression words from images using a neural network.” 1997-IEEE International Conference on Systems, Man, and Cybernetics Computational Cybernetics and Simulation, IEEE, New York, NY, 1997(1):150-155.

[8] Shangfei Wang, “A Kansei Image Retrieval Based on Emotion Model. “ Journal of Circuits and Systems,2003,8(6):48-52.

[9] Hualin Wan, “Image Semantic Classification by Using SVM.” Journal of Software, 2003, 14(11):1891-1899.

[10] Jinsub Um. A Study of the Emotional Evaluation Models of Color Patterns Based on the Adaptive Fuzzy System and the Neural Network[J]. COLOR research and application,2002, 27(3):208-216

Materials Science and Information Technology II
10.4028/www.scientific.net/AMR.532-533

A Survey on Emotional Semantic Mapping in Image Retrieval
10.4028/www.scientific.net/AMR.532-533.1297

DOI References
[2] R. W. Picard. Synthetic emotion. IEEE Computer Graphics and Applications. 2000, 20(1): 52-53. 10.1109/38.814561
[10] Jinsub Um. A Study of the Emotional Evaluation Models of Color Patterns Based on the Adaptive Fuzzy System and the Neural Network[J]. COLOR research and application, 2002, 27(3): 208-216. 10.1002/col.10052