論 文
7-1-3

대화형 감성기반 칼라영상 검색
(Interactive emotion-based color image retrieval)

엄경배(Kyoungbae Eum)[1] 박중수(Joongsoo Park)[2]

Abstract

Variable contents are extracted and used to improve the correctness of the retrieval in the content-based image retrieval. This way use the physical feature for the retrieval.
In this way of retrieval, the user has to know the basic physical features and spatial relationship of target images that he wants to retrieve. There are some restriction to reflect the user's intend. We need the retrieval system that reflect the user's intend. In this paper, we propose an emotion-based retrieval system. It is different from past emotion based image retrieval in point of view that it uses relevance feedback to estimate the users intend and it is easily combined with past content-based image retrieval system. The features and similarity measures are adopted from MPEG-7 color descriptors which are proper retrieval of large multimedia databases. We use wallpaper images for the experiment. The result shows that the system get successful result.

요 약

영상으로부터 여러 가지 내용을 추출하여 검색에 사용함으로써 내용기반 영상검색에서 검색의 정확도를 높이고 있다. 내용기반 영상검색 방법은 영상 검색을 위하여 물리적 속성을 이용한다. 즉 사용자가 검색을 원하는 영상의 물리적인 속성 또는 이들의 공간적인 배치등의 내용을 상당부분 알아야 검색이 가능하다는 제약을 수반한다. 그러므로, 이방법은 사용자의 의도를 반영하는데 제약이 있다. 따라서, 본 논문에서는 사용자의 의도를 반영하는 감성기반 검색 시스템을 제안하였다. 제안된 시스템은 사용자의 의도를 추정하기위해 relevance feedback을 이용한다는 점에 있어 기존의 내용기반 영상 검색 방법과 다르고, 제안된 시스템은 내용기반 영상검색 방법에 기반하고 있으므로 기존의 내용기반 검색방법과 통합이 용이하다. 속성들과 유사도측도들은 MPEG-7 칼라 기술자를 사용하였다. 벽지 영상을 이용하여 이같은 실험을 행하였고, 감성형용사들을 적용하여 DB에 저장된 벽지 영상들을 대상으로 검색한 결과 만족할 만한 결과를 얻을 수 있었다.

논문접수 : 2006. 1. 10
심사완료 : 2006. 2. 11

1) 정회원 : 군산대학교 컴퓨터정보과학과 교수
2) 정회원 : 우석대학교 전산정보학부 교수

1. Introduction

There has been an explosion in the amount and complexity of multimedia data with the advances in computer technologies and internet. In order to make use of this vast amount of data, efficient and effective techniques to retrieve multimedia information based on its content need to be developed.

To overcome the difficulties of the text-based approach, content-based image retrieval(CBIR) was proposed. The CBIR systems use the visual content of the images such as color, texture, and shape features. Recently, MPEG-7 standard that facilitates the CBIR was defined.[1]

In the CBIR, the query is an image itself and the retrieval process is the process that seeking the similar images to the given query image. In this way of retrieval, the user has to know the basic physical features of target images that he want to retrieve. It has some restriction not offering various retrieval method to users.

In this paper, we proposed an emotion-based retrieval system that retrieve the color images by using an emotion language such as "warm","soft","cozy" and "clean".

In the past emotion-based retrieval system, it is hard to combine with past CBIR because the retrieval is done through the index in the emotional feature space after transforming the physical features into emotional features. The emotion is very subjective and differ according to the users. The retrieval through the index don't sufficiently reflect on the users intend.[2]

Previously, main emotion language and its representative image are stored for the retrieval. The system seeks the similar images by using MPEG-7 color descriptor of the representative images, if an user choose and evaluate them.

The subjective evaluation of an user is reflected to the retrieval through the relevance feedback.[3,4] The filtering is possible based on it. We proposed the system which can retrieve the color images a user want. Proposed method is the emotion-based retrieval and it can be easily combined with traditional CBIR. It can be used for various retrieval images. It can overcome the difference of subjective emotion according to the different users.

To test the performance of proposed system, we tested about 1500 images and used an emotional language such as ""warm" ,"soft", "cozy" and "clean". The successful results are gotten in our experiments.

2. Organization

2.1 MPEG-7 color descriptor

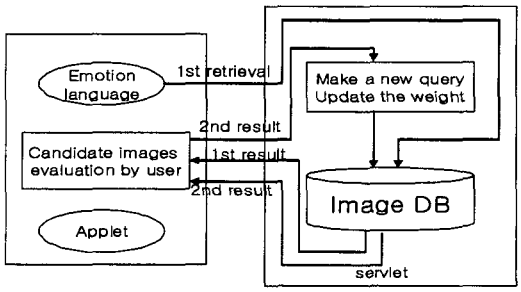
Through psychological experiments, Soen found that the physical features like a color and texture affected to the emotion of human-being.[5] Through the factor analysis, he knows the emotional features are affected by the physical features such as an average hue, saturation, intensity and the frequency components of the spatial distribution. Under the asumption that the information on the color and brightness can be represented by the scalable color descriptor(SCD) and the frequency component corresponding to color variation can be represented by the color structure descriptor(CSD). We used two

descriptor such as SCD and CSD of MPEG-7.[6]

2.2 Proposed system

Traditional emotion-based retrieval system used a converter to transform the physical features into the emotion features. The multiple regression, neural network, and adaptive fuzzy system were used as a converter. The converted emotional features are stored in DB by using the R-tree and clustering. Traditional emotion-based retrieval system is shown in [7].

Proposed method do not need a complicated converter and store the image DB as an index of emotional language in the emotional feature space. Proposed system shows the representative images which correspond to it, when getting the emotion language from an user. The representative images are evaluated by the user. The system makes a new query and updates the weight of each descriptor by above evaluation value. The system shows a new candidate images. The images for the experiment are classified according to the emotion language which the users often use. The classified images are evaluated on the degree of each emotion language. The system extracts the representative images and store it to DB. The representative images stored in above DB are shown to user in the 1st retrieval which use an emotion language as a query. The recommended images are evaluated by users again. It is transmitted to the system. The new query was made by this evaluation value. The query is represented by the color descriptor of MPEG-7. The queries used are the 16bin SCD and 32bin CSD. Proposed system consists of two part. The one is an applet.The other is a sublet. The former shows the candidate images and they are evaluated by the user. The latter gets the user's evaluation value and analyze it to know the user's intend. The candidate images are retrieved from the image DB by the user.



[Fig. 1] Proposed system

2.3 relevance feedback

It is easy to know user's intention through the user's evaluation in the relevance feedback.[3,4]
New query by user's emotion can be made, when tranmitting the value evaluated by user into the system. The individual difference between the users can be compensated. In the process of feedback, the retrieval and evaluation is processing at the same time. So, the retrieval time can be reduced.

Retrieval process by the relevance feedback is as follows.
(1st step) Every image is represented by the vector model having descriptor component.
(2nd step) Initialize the weights of each

descriptor.
(3rd step) Evaluate the representative images.
(4th step) Make a new query.
(5th step) Show them to user after retrieving candidate images by similarity measure.
(6th step) Evaluate the candidate images.
(7th step) Update the weights of each descriptor.
(8th step) Go to step 4th.

The vector models of image I by the SCD and CSD are as eq. (1)

$$I_1 = [d_{1,1}, \cdots, d_{1,16}],$$
$$I_2 = [d_{2,1}, \cdots, d_{2,32}] \qquad (1)$$

There are two kinds of weights. The weight of vector component d_{ij} is W_{ij}. Initial weight $W_{ij}^{\,0}$ is as eq.(2).

$$W_{ij}^{\,0} = \frac{1}{J_i} \qquad (2)$$

Where $J_1 = 16$ and $J_2 = 32$

The weights on SCD and CSD are represented by the W_1 and W_2. Initial weight is as eq.(3).

$$W_i = \frac{1}{2}, \quad i = 1, 2 \qquad (3)$$

For each of the retrieved objects, the user marks it as "highly relevant", "relevant","no opinion","nonrelevant", or "highly nonrelevant", according to his information need and perception subjectivity. Evaluation values are 0, 2, 3, 4, and 6.

The similarity measure between two images is as eq.(4).

$$S_i = \sum_{j=1}^{J_i} W_{ij} * |d_{ij} - d'_{ij}|, \quad i = 1, 2$$
$$S = \sum_{i=1}^{2} W_i * S_i \qquad (4)$$

The images of DB are retrieved by using a similarity measure such as eq.(4), if a query image is updated. The range of each descriptor value is different and the measures of SCD and CSD are different. Before updating the weight in the 7th step, we should nomalize the value of each descriptor. The former is an intra-normalization and the latter is an inter-normalization.[4]
One query image should be made for next retrieval, when the user evaluate the images as "highly relevant" and "relevant". We used the mean and weighted mean for it. We get the mean of each bin of evaluated images as "highly relevant" or "relevant" and used it as a bin value of new query. The weighted ratio of "highly relevant" and "relevant" is 1 to 2 in case of weighted mean.

3. Experiment

The wall paper images are used for the proposed retrieval system. The MPEG-7 color descriptors of wallpaper images are extracted and stored to DB by Visual C++. In our proposed system, we used the emotional language such as "warm","soft","cozy","clean" and "refined".

The 1st retrieval is done by an emotional language. The user evaluates the candidate images. After that, a new query is made. The 2nd retrieval is done by a new query. The relevance feedback is applied to the retrieval since the 2nd one. The new query is made by the user's evaluation value. We used the clustering method to consider various physical features. Proposed retrieval method is implemented by the JAVA.

To test the performance of proposed retrieval system, the relevant rate measure and penalty measure are used. The relevant ratio is the ratio of the number of "highly relevant" and "relevant" images against the number of displayed images. When the system shows the retrieval result to user, the images are displayed by the order as follows : "highly relevant", "relevant", "no opnion", "nonrelevant", and "highly nonrelevant". The penalty measure give a penalty score to them, when this images are not in its order. This is similar to the performance measure of MPEG-7.[1] The good retrieval result can be obtained, when having small penalty score and large relevant rate.

The retrieval results on the 6 emotional languages are shown in Table 1. These results are averaged by the 20 subjectives.

<Table 1> The retrieval result by the emotional languages

	1st	2nd	3rd	4th
	relevant rate	relevant rate	relevant rate	relevant rate
	penalty	penalty	penalty	penalty
soft	50.0	58.0	75.0	91.7
	6	6	4	2
comfotable	41.0	66.7	83.3	83.3
	7	6	4	4
cozy	41.6	75.0	83.3	91.7
	5	4	4	2
clean	33.3	41.6	66.6	75.0
	5	6	5	5
refined	50.0	75.0	83.3	83.3
	7	4	4	4
vigorous	50.0	66.6	75.0	91.7
	8	5	4	4

4. Conclusion

In this paper, we proposed the emotion-based retrieval system to reflect the user's emotion. Traditional emotion-based retrieval system retrieve the images after converting the physical features to the emotional features. But, proposed method do not need the complicated converter. Proposed method can fast convert the user's emotional feature to physcal feature through the relevance feedback and retrieve the images by compensating the individual difference of user's emotion. We tested about 1,500 wallpaper images and verified it well reflect the user's emotion through the experiment. Proposed system can be used as the product retrieval system based on emotion in various area such as the wallpaper retrieval, a papered floor retrieval, and textile design

retrieval.

In the future research, we will make more robust system by including the texture feature. The fine retrieval is possible when increasing the number of bins. Then, the retrieval speed is dropped. So, we should develope the retrieval algorithm to reduce computational complexity.

References

[1] B.S. Manjunath, J.R. Ohm, V.V. Vasudevan and A. Yamada(2001), Color and Texture Descriptors, IEEE Trans. On Circuits and Systems for Video Technology, No. 6.

[2] J. Um(2000), A study on the emotion-based recognition and retrieval of color pattern, Ph. D. Thesis, Chonbuk National Univ..

[3] Y. Rui, T. S. Huang and S. Mehrotra(1997), Content-based image retrieval with relevance feedback in MARS, Proc. IEEE Int. Conf. Image Processing, pp.815-818.

[4] Y. Rui and T. S. Huang(1998), Relevance Feedback: A power Tool for Interactive Content-based Image Retrieval, IEEE Transactions on Circuits and Systems for Video Technology, Vol.8, No.5, pp.644-655.

[5] N. Kawamoto and T. Soen(1993), "Objective evaluation of color design 2," Color Res. Appl., Vol. 12.

[6] "MPEG-7 Visual Experimentation Model(2001), ISO/IEC/JTC1/SC29/WG11, Doc. N4063.

[7] J. Um, J. Lee(1997), Color pattern retrieval by using emotional language, Journal of KFIS, Vol. 7 No. 3.

엄경배

1990 : 전북대학교 대학원 전자공학과 졸업(공학박사)

1995-1996 : 캐나다 토론토대학교 컴퓨터 및 전기공학과 Post Doc.

1999 : 일본 동경대학 생산기술연구소 객원연구원

1989-현재 : 군산대학교 컴퓨터정보과학과 교수

관심분야 : 영상처리 및 멀티미디어처리, 감성공학

박중수

1980 : 공주대학교 사범대학 수학교육과 졸업(학사)

1983 : 성균관대학교 대학원 수학과 졸업(이학석사)

1987 : 미국 오하이오대학 박사과정수료

1993 : 전북대학교 대학원 수학과 졸업(이학박사)

1988-현재 : 우석대학교 전산정보학부 교수

관심분야 : 영상처리, 정보검색, 응용수학