Emotion Recognition using Color and Pattern in Textile Images

Na Yeon Kim, Yunhee Shin, Youngrae Kim, and Eun Yi Kim
Dept. of advanced technology fusion, school of Internet and multimedia Eng. Konkuk Univ., Korea
{yeon0830, ninharsa, kissgood, eykim}@konkuk.ac.kr

***Abstract*—In this paper, a novel method is proposed using color and pattern information for recognizing some emotions included in a textile. Here we use 10 *Kobayashi emotion keywords*. Our method is composed of feature extraction and classification. For accurate emotion recognition, both color and pattern are extracted from a textile. Then the representative color prototypes of a textile are extracted using color quantization method and pattern is described by wavelet transform followed by some statistical terms. The extracted features are given to the neural network (NN)-based classifiers. The advantages of our method include the following: 1) it is a generalized method to accurately recognize emotions in textile images which used in various application domains; 2) it is a fully automatic method with no manual interaction. To prove these advantages, the experiments are performed on 389 textiles obtained from various application domains such as interior, fashion, and artificial ones. Our method shows the precision of 100% and the recall of 99%, regardless of the specific domain.**

***Keywords*—emotion recognition, textile indexing, wavelet transform, color quantization, neural networks**

I. INTRODUCTION

For a given product or object, predicting human emotions is very important in many business, scientific and engineering applications. In particular, human emotion recognition in textiles has been considerable attentions by pattern designer or interior designer, etc., as it can be applicable to automatically recommend some textiles to the customers in e-Business or psychotherapy [7]. Traditionally, the emotions included in some textiles have been manually analyzed by a few experts, which cause a huge amount of time and effort. To reduce the cost and time, automatic emotion recognition system should be developed. Thus, some works have been developed, which are summarized in Table 1.

From Table 1, we obtain some conclusions. First, there are three important application domain such as interior domain, fashion domain, and artificial domain. Second is that using compound features allows for the better performance than using unitary feature such as color or pattern. Third is that the fully automatic method has not yet developed to be used in various domains, although some semi-automatic method has been developed [5].

For a method to be used practically, two requirements should be satisfied: 1) considering the practical application domain of emotion recognition in textile images should proceed in an automatic way; 2) in order to be universal, the system should be capable of analyzing textiles regardless of used domain (e.g., interior, fashion, and artificial data). To develop the methods satisfying these requirements, it is important to describe the physical features in an effective manner and extract accurately.

Accordingly, we present a novel emotion recognition method using color and pattern information. Our goal is on developing the generalized fully automatic method to accurately recognize emotions included in textile images, regardless of the domain where the textiles used.

The proposed method is composed of feature extraction and classification. To describe the physical feature information in the textiles, the wavelet transform and color quantization are used. And the neural network is used as the classifier.

TABLE I. RELATED WORKS

Characteristics	Conventional Methods			
	T. Soen [3]	**E. Kim [2]**	**N. Kim [4]**	**S. Kim [5]**
Features	Texture	Color	Pattern	Color & Pattern
Extraction Methods	Fourier Transform	Raw color histogram	Wavelet transform	Row color histogram & wavelet transform
Classificication Methos	Regresstion model	Fuzzy system	Neural network	Neural network
# Database	Artificial data (13)	Interior data (179)	Artifical data (220) Fashion data (41)	Fashion data (41) Interior data (179)
Recognition Rate (%)	80%	78%	90% 62%	93% 92%

978-1-4244-1674-5/08 /$25.00 ©2008 IEEE CIS 2008

To assess the validity of our method, the experiments are performed on 389 textiles obtained from various application domains such as interior, fashion, and artificial ones. And the result of proposed method was compared with the other methods: Kobayashi and Kim's method [1, 4]. Then, our method shows the superior performance to others: it has the precision of 100% and the recall of 99%, regardless of the specific domain. These results confirmed that our method has the potential to be applied for various applications such as textile industry and e-Business.

This paper is organized as follows. Section 2 shows the system overview, and Section 3 shows the methods of feature extraction. And the NN-based classifier is described in Section 4. Section 5 shows experimental results, and the conclusion is followed.

II. Overview of the Proposed Method

The proposed method aims at classifying the textiles into some emotion groups. Here, we use 10 Kobayashi emotion keywords: {romantic, clear, natural, casual, elegant, chic, dynamic, classic, dandy, modern}. Then, for accurately mapping textile image to emotion groups, the proposed method uses the color and pattern information.

Fig. 1 shows the outline of proposed method, where it is composed of 10 neural network (NN)-based recognizers for the respective human emotions. For a given input image, the input image is firstly normalized to 64×64. And the normalized image is used as the input for NN-based recognizer. The recognizer for a specific emotion is composed of feature extraction and classification. The representative color prototypes of a textile are extracted using color quantization method and pattern is described by wavelet transform followed by some statistical terms. The NN-based classification, which produces the output value ranged from 0 to 1. Then, if the output value is bigger than 0.5, the recognizer decides that the image is positive for the corresponding emotion. Otherwise, the recognizer decides the textile image is negative for the corresponding one.

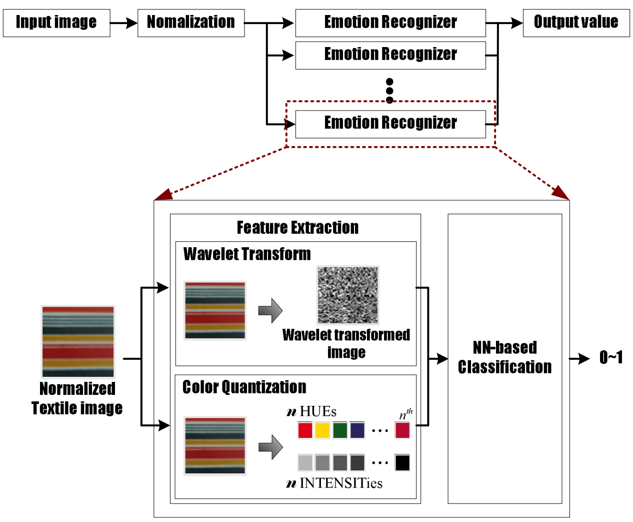


Figure 1. The outline of proposed method

III. Featrue Extraction

A. Color extraction

Kobayashi have investigated how the color affects human emotion based on the survey, and developed the color image scale to transpose the color image into some emotion keywords [1]. In color image scale, all the colors have been positioned along three axes: soft/hard, warm/cool, and clear/grayish. The coordinates in the color image scale are corresponding to the channels in the HIS color space: soft/hard to intensity, warm/cool to hue, and clear/grayish to saturation.

In this work, we represent a textile image in the HIS color space, thus firstly transform the RGB color image into HIS color image using the following equations.

$$H = \cos^{-1}\left\{ \frac{\frac{1}{2}[(R-G)+(R-B)]}{\sqrt{(R-G)^2+(R-B)(G-B)}} \right\} \tag{1}$$

$$S = 1 - \frac{3}{(R+G+B)}[\min(R,G,B)] \tag{2}$$

$$I = \frac{1}{3}(R+G+B) \tag{3}$$

Among three axes, Kobayashi highlights the importance of soft/hard and warm/cool, thus we use the color values from H-channel and I-channel. The values of each channel are ranged from 0 to 255. Generally to deal with 256 image scale requires a lot of computation cost and a poor clustering result. To reduce computing time and obtain the more meaningful results, we quantize the respective images on H or I channel into a few prototypes. The prototypes are found as local maxima in a smoothed color histogram of the input image, all other colors are then assigned to the nearest prototype. In most of images used in our experiments, this result in about peaks in the histogram, and the components that seem different to the human eye are quantized into different classes.

quant = quantized color prototype
start,end = start and end position of a color cluster
L = maximum of input feature value
max_value = max value initialization
NoOfQuant = current number of color prototype
oldNoOfQuant = previous number of color prototype
L = 256
max_value = -1
Given *hist*, input histogram, *hist* is normalized.
Initialize *NoOfQuant*.
While
- *oldNoOfQuant* ← *NoOfQuant*
- For *i* = 0, ... , *L*
  * Fine feature with maximum frequency.
  * If *hist*[*i*] > *max_value* then *max value* ← *i*
- For *i* =100, . . . , *max_value*
  * Given *i*, a local maxima in a smoothed color histogram of the input image.
  * For *j* = 0, ... , *L*
    - if *j* is near by the start of i_{th} color cluster, then *start* ← *j*
  * For *j* = 0, ... , *L*
    - if *j* is near by the end of i_{th} color cluster, then *end* ← *j*
  * *quant*[*NoOfQuant*] ← (*start* + *end*) / 2
- If *NoOfQuant* == *oldNoOfQuant* then Quantization is completed.

Figure 2. Outline of color quantization algorithm.

The outline of the used quantization method is shown in Fig. 2, and Fig. 3 shows the results of color quantization.

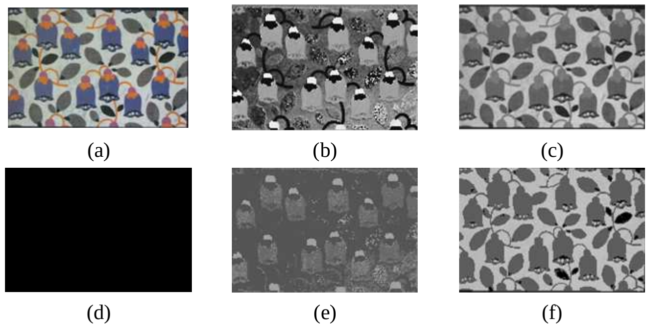


Figure 3. The results of color quantization: (a)original image, (b)hue conversion, (c)intensity conversion, (d)saturation conversion, (e)quantized hue(Q=5), (f)quantized intensity(Q=5)

B. *Pattern extraction*

Generally, a pattern can be described as a combination of texture and edge. Therefore, we use a wavelet transform.

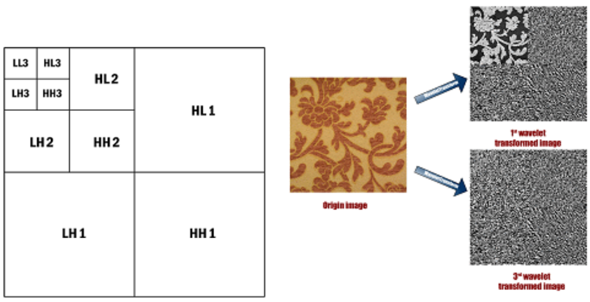


Figure 4. Wavelet transformed image.

The wavelet transform provides successive approximations to the image by down-sampling and have the ability to detect edges during the high-pass filtering. The wavelet transform decomposes to 4 sub-blocks as LL, LH, HL, and HH as shown in Fig. 4. LL is involving the textual content, while the other sub-blocks are involving edge information for vertical, horizontal and diagonal orientations.

In our method, the LL level is again decomposed into 4 sub-blocks using wavelet transform. This process is iterated P times, so that the sub-blocks of $3P$+1 are created. Then, from each block the following parameters are calculated [6].

$$M(I) = \frac{1}{N^2} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} I(i,j) \tag{4}$$

$$\mu_2(I) = \frac{1}{N^2} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (I(i,j) - M(I))^2 \tag{5}$$

$$\mu_3(I) = \frac{1}{N^2} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (I(i,j) - M(I))^3 \tag{6}$$

Given $N \times N$ image, $M(I)$ represents average value, and $\mu_2(I)$ and $\mu_3(I)$ represent 2nd and 3rd momentums. As these parameters are extracted for a block, total of $3(3P+1)$ parameters are created after P_{th} wavelet transforms. All of them are used as the input of the classifier that recognizes the pattern information in some textiles.

IV. NN-BASED CLASSIFICATION

In this paper, we use multilayer perceptron (MLP) as a classifier. The network is composed of input layer, hidden layer, and output layer. The adjacent layers are fully connected.

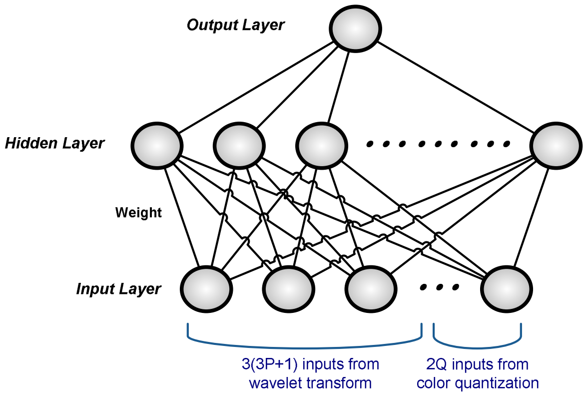


Figure 5. The NN-based classifier.

The classifier receives the $3(3P+1)+2Q$ inputs, where P is the level of wavelet transform and Q is the number of quantized colors. The selection of P and Q is very important factor to affect the performance. From various experiments, we obtained the fact that it is important to adjust the balance between the pattern feature and color feature. That is, $(3P+1) \approx 2Q$. The pairs can be one of {(3, 5), (5, 8), (7, 11), …}. The larger P and Q are, the larger the computational cost is. Accordingly, we select (P, Q) to (3, 5).

The hidden node in the classifier plays a role as discriminant function, so that it is determined by the experiment.

The output value of classifier is normalized to 0~1. And if the output value is bigger than 0.5, the classifier decides that the image includes the corresponding emotion. Otherwise, the classifier decides the textile has not the correlation for the emotion.

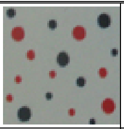
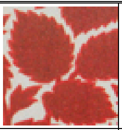
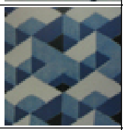
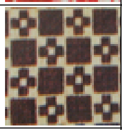
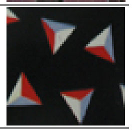
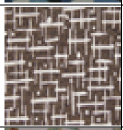
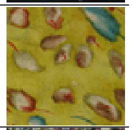
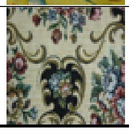
To train the classifier, our method uses feature (I, d), where I is textile image, and d is emotion value manually annotated by human experts in the survey. The input layer receives the wavelet transformed values and color quantized values of 64×64 textile image. The output value of hidden node is obtained from the dot product of the vector of input values and the vector of the weights connected to the hidden node. It is then presented with the output nodes. The weights are adjusted by training with a back-propagation algorithm in order to minimize the sum of squared error during the training session.

V. EXPERIMENTAL RESULTS

The proposed method was implemented using MS Visual C++ from Window XP environment. The parameters of NN were fixed as follows: error rate is fixed to 0.001, momentum to 0.5, and iteration to 5000. In the stage of NN training, the number of hidden node was determined to 120, which was not changed in testing.

For experiments, we collected 389 textile images from three application domains such as interior, fashion, and artificial ones. The artificial data were collected from the Pattern-Book[1]. And the fashion and interior data were collected from the Dongdaemoon textile shopping mall in Seoul. These data samples are shown in Table 2.

TABLE II. THE EXPERIMENTAL DATA

	Data domain	Count	Samples		
Training Data	Artificial domain	120			
Test Data	Artificial domain	100			
	Fashion domain	41			
	Interior domain	128			

Out of artificial data, 120 data were used for the NN training, while the remainders were used for the test. And a total of fashion and interior data were used for only the test, to prove the generalization of our method.

To evaluate the performance our method, the ground truth of 389 data are required. For this, the survey was conducted on 20 peoples. On emotions, the pollees put one out of {-1,0,1} on each textiles: -1 represents negative emotion; 1 represents positive emotion; 0 represents image does not have relation with emotion. The accumulated sums from 20 peoples are given to a textile. Thus, the integer is from -20 to 20 for each textile. Then, fuzzy rules were applicable for accumulated sum of each emotion [4]. Through this process, we classified the textile images into 10 emotion groups, and annotated them using their groups.

The advantages of our method include the following: 1) it is a generalized method to accurately recognize emotions in textile images which used in various application domains; 2) it is a fully automatic method with no manual interaction. To prove these advantages, the proposed method was compared with other methods. Here, two methods were adapted: Kim et al.'s and Kobayashi's method: 1) Kim et al.'s method developed the NN-based recognizer using only the pattern information [4]. 2) Kobayashi's method uses only color information [1]. Then for the quantitatively comparison two measures were used such as precision and recall. These are defined as follows:

$$precision(\%) = \frac{\#\ of\ correctly\ detected\ textile\ image}{\#\ of\ detected\ textile\ image} \times 100 \tag{7}$$

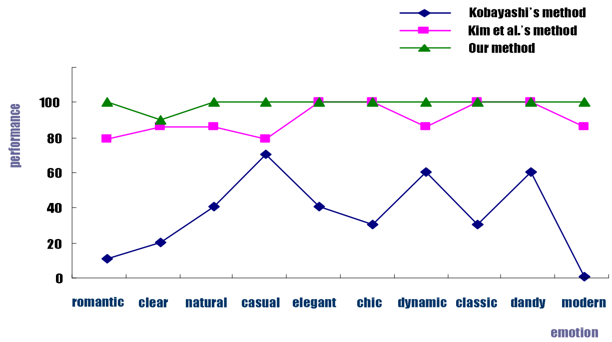
$$recall(\%) = \frac{\#\ of\ correctly\ detected\ textile\ image}{\#\ of\ textile\ image} \times 100 \tag{8}$$

A. Application to artificial domain

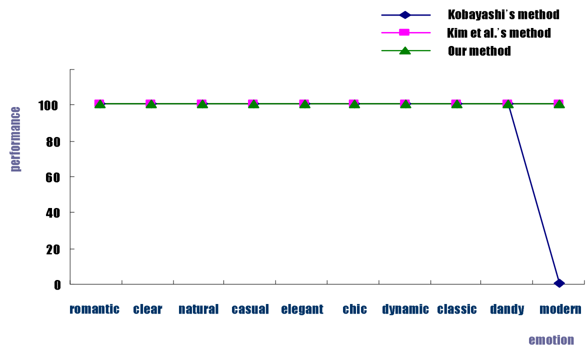
Firstly, our method was evaluated with the artificial data. These data are mainly composed of unitary patterns and simple color combination. Table 3 summarized the performance of our method. As can be seen in Table 3, our method produced the precision of 100% and the recall of 99% on average.

TABLE III. THE PERFORMANCE OF ARTIFICIAL DOMAIN

Emotion type	Number of data	Number of detected data	False alarm	False dismissal	Recall (%)	Precision (%)
romantic	20	20	0	0	100	100
clear	20	18	1	1	90	100
natural	20	20	0	0	100	100
casual	20	20	0	0	100	100
elegant	20	20	0	0	100	100
chic	20	20	0	0	100	100
dynamic	20	20	0	0	100	100
classic	20	20	0	0	100	100
dandy	20	20	0	0	100	100
modern	20	20	0	0	100	100
Average	200	198	1	1	99	100



(a)



(b)

Figure 6. The performance comparison for artificial domain: (a)recall, (b)precision.

---

[1] Meller, Susan, "Textile designs : 200 years of European and American patterns for printed fabrics organized by motif, style, color, layout", Harry N. Abrams, 1991.

Fig. 6 shows a comparison of the three methods in terms of their recall and precision when processing the artificial data. Fig. 6(a) shows the results relative to recall and Fig. 6(b) shows a comparison among the three methods in terms of precision. Kobayashi's method shows the precision of 90% and the recall of 36% on average. Kim et al.'s method shows the precision of 98% and the recall of 90% on average. Among three methods, Kim et al's method and proposed method have almost perfect performance.

B. *Application to fashion domain*

To fully demonstrate the effectiveness of our method, it was tested with data which not to be used in training. For this, three methods have been evaluated with fashion data and interior data. The fashion data is less complex than the interior data, but they are more complex than the artificial data in terms of color combination and the variety of pattern.

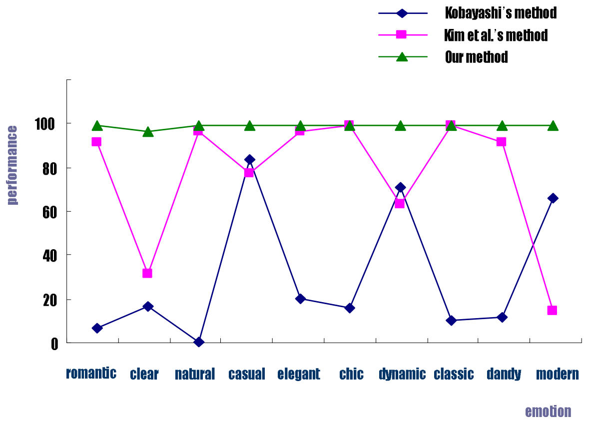
The results to the fashion data have been shown in table 4. As the table indicates, proposed method shows the precision of 100% and the recall of 99.74% on average. The performance of our method is slightly improved, when comparing with its performance in artificial domain. This is unexpected result because the fashion data were not used in training.

Fig. 7 shows a comparison of the three methods in terms of their recall and precision when processing the fashion data. Fig. 7(a) shows the results relative to recall and Fig. 7(b) shows a comparison among the three methods in terms of precision.

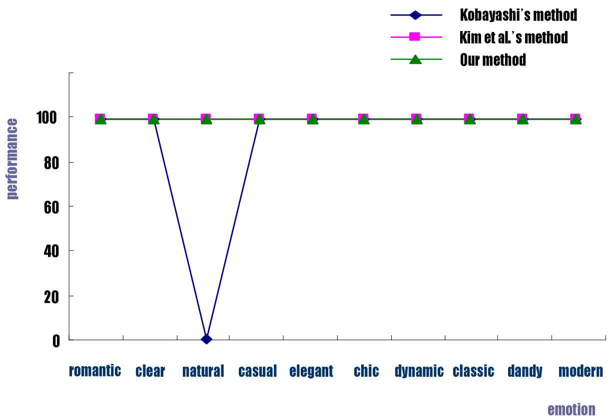
TABLE IV. THE PERFORMANCE OF FASHION DOMAIN

Emotion type	Number of data	Number of detected data	False alarm	False dismissal	Recall (%)	Precision (%)
romantic	37	37	0	0	100	100
clear	37	36	1	0	97	100
natural	34	34	0	0	100	100
casual	36	36	0	0	100	100
elegant	33	33	0	0	100	100
chic	34	34	0	0	100	100
dynamic	36	36	0	0	100	100
classic	34	34	0	0	100	100
dandy	38	38	0	0	100	100
modern	35	35	0	0	100	100
Average	354	353	1	0	99.7	100

Kobayashi's method shows the precision of 90% and the recall of 30% on average. Kim et al.'s method shows the precision of 100% and the recall of 62% on average. Note the performance of Kim et al.'s method. Its performance was rapidly declined when comparing with one of artificial data. This declination can be explained by the following reason: Kim et al.'s method used only the fashion data when training, so that it fails to accurately recognize the emotions in the data collected from other domain. On the contrary, there is no difference in the performance of our method. From this experiment, we can guess that our method can be universally used in variety of data.



(a)



(b)

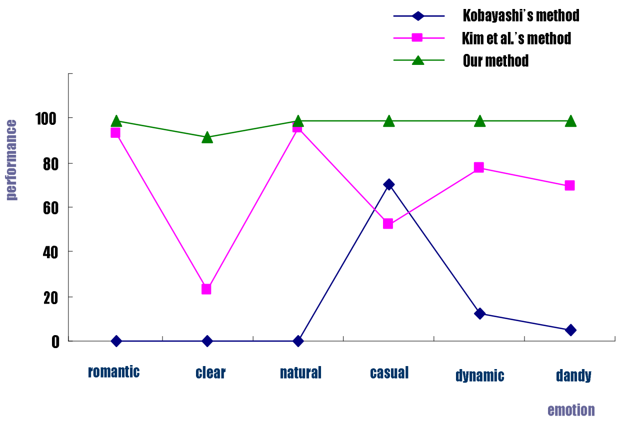
Figure 7. The performance comparison for fashion domain: (a)recall, (b)precision.

C. *Performance in interior domain*

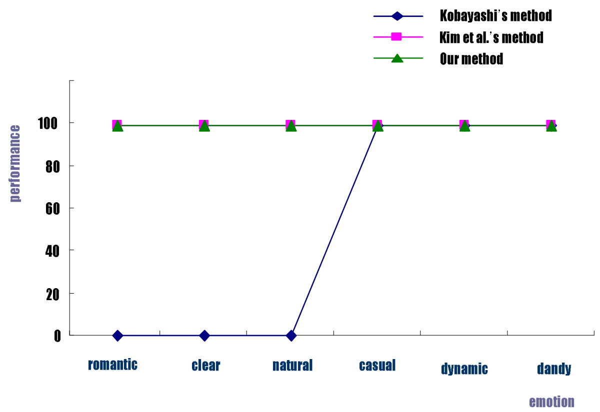
Our method was also evaluated with the interior data. Their data have more complicated color combination and pattern than fashion data. The results to the interior data have been shown in table 5. As the table indicates, proposed method shows the precision of 100% and the recall of 98.83% on average. Despite of image complexity, our method shows steadily performance.

TABLE V. THE PERFORMANCE OF INTERIOR DOMAIN

Emotion type	Number of data	Number of detected data	False alarm	False dismissal	Recall (%)	Precision (%)
romantic	19	19	0	0	100	100
clear	40	37	3	0	93.02	100
natural	36	36	0	0	100	100
casual	17	17	0	0	100	100
dynamic	52	52	0	0	100	100
dandy	60	60	0	0	100	100
Average	224	221	3	0	98.83	100



(a)



(b)

Figure 8. The performance comparison for interior domain: (a)recall, (b)precision.

Fig. 8 shows a comparison of the three methods in terms of their recall and precision when processing the interior data. Fig. 8(a) shows the results relative to recall and Fig. 8(b) shows a comparison among the three methods in terms of precision. Kobayashi's method shows the precision of 50% and the recall of 14.8% on average. Kim et al.'s method shows the precision of 100% and the recall of 69.5% on average.

On average, our method achieved a recall that was 84.03% higher than the Kobayashi's method and 30.33% higher than the Kim et al.'s method. Kim et al.'s method produced a recall more about 4.5 times than Kobayashi's method. When applied to the interior data, there are dramatically decreasing in performance of both Kobayashi's method and Kim et al's method. However, our method shows the performance of above 98%, regardless of the data types.

In summary, the experiments for various domains show that our method has the superior performance to other two methods in all of domains. Kobayashi's method showed the lowest performance among three methods, and Kim et al.'s method has good performance in only the data used for the training. Therefore, these results show that our method can be generally used in a variety of application domain. Moreover, as it require the processing time of 10 frames/sec., our method can be easily applied in the commercial products or e-Business.

VI. Conclusion

In this paper, we proposed a generalized emotion recognition method using color and pattern information. Our method is component two modules: feature extraction and classification. To describe the color and pattern information from a textile, the color quantization and the wavelet transform are used. And the neural network is used as the classifier.

To assess the validity of the proposed method, it was applied to recognize the human emotions in three textile domains: fashion, interior and artificial data book. And then our method produced the precision of 100% and the recall of 99% on average. This result confirmed that our method has the potential to be applied for various applications such as textile industry and e-Business.

Acknowledgement

This work was supported by the Korea Science and Engineering Foundation (KOSEF) grant funded by the Korea government (MOST). (No. R01-2007-000-20997-0)

References

[1] Shigenobu Kobayashi, "COLOR IMAGE SCALE", Kodansha, 1991.

[2] Eun Yi Kim, Soo-jeong Kim, Hyun-jin Koo, Karpjoo Jeong, Jee-in Kim, "Emotion-based Textile Indexing using Colors and Texture," Lecture Notes in Computer Sciences, vol. 3613, pp. 1077-1080, 2005.

[3] T. Soen, T. Shimada, and M. Akita, "Objective Evaluation of Color Design Ⅱ", Color Res. Appl., Vol. 12, pp. 187-194, 1987.

[4] Na Yeon Kim, Yunhee Shin, Soo-jeong Kim, Jee-in Kim, Karpjoo Jeong, Hyunjin Koo, Eun Yi Kim, "Emotion Recognition System using Neural networks in Textile images", Jorunal of KIISE: Software and Applications, Vol. 34, No. 9, 2007.

[5] Soo-jeong Kim, "An Emotion Recognition Framework for Textile Images", Thesis for the Degree of Doctor of Philosophy in Konkuk graduate school, 2007.

[6] Huiping Li, David Doermann, and Omid Kia, "Automatic Text Detection and Tracking in Digital Video", IEEE Transactions on Image Processing, Vol. 9, No. 1, January 2000.

[7] http://www.multitherapy.com.