Horror Image Recognition Based on Emotional Attention

Bing Li[1], Weiming Hu[1], Weihua Xiong[2], Ou Wu[1], and Wei Li[1]

[1]National Lab of Pattern Recognition, Institute of Automation,CAS, Beijing, China.
[2]OmniVision Technologies, Sunnyvale, CA, USA.
bjtulb@gmail.com

Abstract. Along with the ever-growing Web, people benefit more and more from sharing information. Meanwhile, the harmful and illegal content, such as pornography, violence, horror etc., permeates the Web. Horror images, whose threat to children's health is no less than that from pornographic content, are nowadays neglected by existing Web filtering tools. This paper focuses on horror image recognition, which may further be applied to Web horror content filtering. The contributions of this paper are two-fold. First, the emotional attention mechanism is introduced into our work to detect emotional salient region in an image. And a top-down emotional saliency computation model is initially proposed based on color emotion and color harmony theories. Second, we present an Attention based Bag-of-Words (ABoW) framework for image's emotion representation by combining the emotional saliency computation model and the Bag-of-Words model. Based on ABoW, a horror image recognition algorithm is given out. The experimental results on diverse real images collected from internet show that the proposed emotional saliency model and horror image recognition algorithm are effective.

1 Introduction

In the past decades, the implosive growth of on-line information and resources has given us the privilege to share pictures, images, and videos from geographically disparate locations very conveniently. Meanwhile, more and more harmful and illegal content, such as pornography, violence, horror, terrorism etc., permeate in the Web across the world. Therefore, it is necessary to make use of a filtering system to prevent computer users, especially children, from accessing inappropriate information.

Recently, a variety of horror materials are becoming a big issue that enters into children's daily life easily. There are a number of psychological and physiological researches indicating that too much horror images or words can seriously affect children's health. Rachman's research [1] shows that horror information is one of the most important factors for phobias. To solve this problem, many governments also have taken a series of measures to ban horror films from children.

Over the last years, a number of high-quality algorithms have been developed for Web content filtering and the state of art is improving rapidly. Unfortunately, in contrast to either the pornographic content filtering [2] [3] [4] or Violent content filtering [5], there has been comparatively little work done on horror information filtering. In this paper, we aim to rectify this imbalance by proposing a horror image recognition algorithm which is the key ingredient of horror content filtering. Horror image recognition can benefit from image emotion analysis which lots of researchers have been paying more attention to recently. For example, Yanulevskaya et al [6] classify images into 10 emotional categories based on holistic texture features. Solli et al [7] propose a Color based Bag-of-Emotions model for image retrieval from emotional viewpoints. Without loss of generality, we also introduce emotional aspect into horror image recognition. In particular, we present a novel emotional saliency map based on emotional attention analysis which is then fed into a modified Bag-of-Words model, Attention based Bag-of-Words (ABoW) model, to determine images' emotional content. The experiments on a real image database show that our method can efficiently recognize those horror images.

2 Image and Emotion Introduction

Images not only provide visual information but also evoke some emotional perceptions, such as excitement, disgust, fear etc. Horror unexceptionally comes out when one experiences an awful realization or deeply unpleasant occurrence [8]. Published research in experimental psychology [9] shows that emotions can be measured and quantified, and that common emotional response patterns exist among people with widely varying cultural and educational backgrounds. Therefore, it is possible to establish a relationship between an image and the emotions.

Generally, a horror image includes two parts: the region with highly emotional stimuli, named emotional salient region (ESR) and a certain background region. The feeling of horror depends on the interaction between them. Therefore, accurate detection and separation of emotional salient region is important for horror image recognition. Intuitionally, visual saliency scheme may be a proper tool for this task. Unfortunately, existing visual saliency algorithms are incompetent to this task. To this end, we introduce emotional attention mechanism into our work and an emotional saliency map method is proposed.

3 Emotional Attention

It is well known that the primate visual system employs an attention mechanism to limit processing to important information which is currently relevant to behaviors or visual tasks. Many psychophysical and neurophysiologic experiments show that there exist two ways to direct visual attention: One uses bottom-up visual information including color, orientation, and other low level features, the other one uses those information relevant to current visual behaviors or tasks [10]. Recent psychophysical research further reveals that visual attention can be modulated and improved by the affective significance of stimuli [11]. When simultaneously facing emotionally positive, emotionally negative, and nonemotional pictures, human automatic attention is always captured by the following order: negative, both negative and positive, and nonemotional stimuli. Therefore, modulatory effects implement specialized mechanisms of "emotional attention" that might not only supplement but also compete with other sources of top-down control on attention.

3.1 Emotional Saliency Map Framework

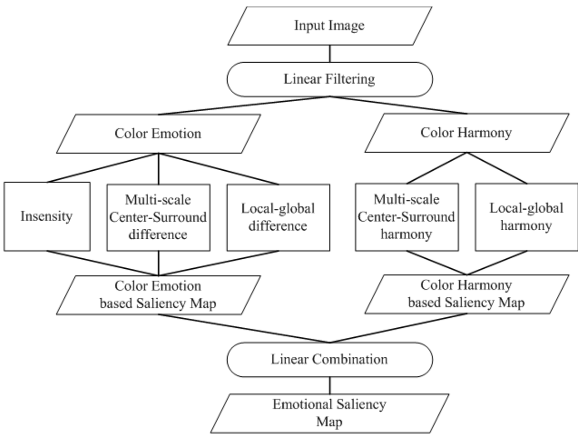

Fig. 1: Emotional saliency map computation Framework.

Motivated by emotion attention mechanism, we propose a new saliency map of image, called 'emotional saliency map', which is different from the traditional visual saliency map in three major aspects: (1) it is based on some high-level emotional features, such as color emotion and color harmony which contain more emotional semantics [12] [13]; (2) it takes both contrast information and isolated emotional saliency property into account which is inspired by lots of psychological experiments indicating that isolated pixels with various colors have different saliency to human visual system; and (3) it uses both contrast and color harmony value to represent the relationship between pixels. The emotional saliency map computation framework is given out in Fig.1, and we will discuss its implementation more details in the next section.

3.2 Emotional Saliency Map Implementation

Color Emotion based Saliency Map. As a high-level perception, emotion is evoked by many low-level visual features, such as color, texture, shape, etc. However, how these features affect human emotional perception is a difficult issue in current research. Color emotion refers to the emotion caused by a single color and has long been of interest to both artists and scientists. Recent research of Ou et al gives a 3D computational color emotion model, each dimension representing color activity (CA), color weight (CW), and color heat (CH) respectively. The transformation equations between color space and color emotion space are defined as follows [12]:

$$\begin{aligned} CA &= -2.1 + 0.06\left[(L^* - 50)^2 + (a^* - 3)^2 + \left(\tfrac{b^* - 17}{1.4}\right)^2\right]^{1/2} \\ CW &= -1.8 + 0.04(100 - L^*) + 0.45\cos(h - 100^o) \\ CH &= -0.5 + 0.02(C^*)^{1.07}\cos(h - 50^o) \end{aligned} \tag{1}$$

where (L^*, a^*, b^*) and (L^*, C^*, h^*) are the color values in CIELAB and CIELCH color spaces respectively. Given a pixel $I(x, y)$ at coordinates (x, y) of image I, we transform its RGB color to CIELAB and CIELCH color spaces, followed by another computation according to Eq (1), its color emotion value is denoted as $CE(x, y) = [CA(x, y), CW(x, y), CH(x, y)]$. We define single pixel's emotional saliency EI as:

$$EI(x, y) = \sqrt{[CA(x, y) - minCA]^2 + [CH(x, y) - minCH]^2} \tag{2}$$

where $minCA$ and $minCH$ are the minimal color activity value and minimal color heat value in the image respectively. This definition is based on the observation that pixels with high color activity (CA) and color heat (CH) are more salient, especially for those horror images; while color weight (CW) has less effects on emotional saliency.

Center-surround contrast is also very important operator in saliency models. Considering that the contrast in the uniformed image regions is close to 0, we propose a multi-scale center-surround operation for color emotion contrast computation:

$$\begin{aligned} EC(x, y) &= \sum_{l=1}^{L} \Theta_l(CE(x, y)), \\ \Theta_l(CE(x, y)) &= \sum_{i=x-d,}^{x+d} \sum_{j=y-d}^{y+d} |CE_l(x, y) - CE_l(i, j)| \end{aligned} \tag{3}$$

where Θ is center-surround difference operator that computes the color emotion difference between the central pixel (x, y) and its neighbor pixels inside a $d \times d$ window in l^{th} scale image from L different levels in a Gaussian image pyramid. $EC(x, y)$ is the multi-scale center-surround color emotion difference which combines the output of Θ operations. In this paper, we always set $d = 3$ and $L = 5$.

Another factor is emotional difference between the local and global $EL(x, y)$, as defined in Eq(4), where $\overline{CE}$ is the average color emotion of the whole image.

$$EL(x, y) = \left|CE(x, y) - \overline{CE}\right| \tag{4}$$

Finally, we define the color emotion based emotional saliency map $CS(x, y)$ as Eq.(5), where $Norm()$ is the Min-Max Normalization operation in an image.

$$CS(x, y) = \frac{1}{3}[Norm(EI(x, y)) + Norm(EC(x, y)) + Norm(EL(x, y))] \quad (5)$$

Color Harmony based Saliency Map. Judd et al [13] pointed out: "When two or more colors seen in neighboring areas produce a pleasing effect, they are said to produce a color harmony." Simply speaking, color harmony refers to the emotion caused by color combination. A computational model $HA(P^1, P^2)$ to quantify harmony between two colors P^1, P^2 is proposed by Ou et al [13].

$$HA(P^1, P^2) = H_C(P^1, P^2) + H_L(P^1, P^2) + H_H(P^1, P^2) \quad (6)$$

where $H_C()$, $H_L()$, and $H_H()$ are the chromatic harmony component, the lightness harmony component, and the hue harmony component, respectively. Since the equations for three components are very complex, we will skip their details due to space limitation, the reader can refer to [13]. $HA(P^1, P^2)$is harmony score of two colors. The higher the score of $HA(P^1, P^2)$ is, the more harmonious the two-color combination is. Similarly, we define a multi-scale center-surround color harmony as well:

$$\begin{aligned} HC(x, y) &= \sum_{l=1}^{L} \Psi_l(x, y), \\ \Psi_l(x, y) &= \sum_{i=x-d,}^{x+d} \sum_{j=y-d}^{y+d} HA(P_l(x, y), P_l(i, j)) \end{aligned} \quad (7)$$

where Ψ is a center-surround color harmony operator that computes the color harmony between the central pixel's color $P(x, y)$ and its neighbor pixel's color $P(i, j)$ in a $d \times d$ window in l^{th} scale image from L different levels in a Gaussian image pyramid. Next, we define the color harmony score between a pixel and the global image as:

$$HL(x, y) = HA(P(x, y), \overline{P}) \quad (8)$$

where $\overline{P}$ is average color of whole image. The high color harmony score indicates the calm or tranquility emotion [14]. In other words, the higher the color harmony score is, the weaker the emotional response is. Therefore we can transform color harmony score to an emotional saliency value using a monotonic decreasing function. Because the range of color harmony function $HA(P^1, P^2)$ is $[-1.24, 1.39]$ [13], a linear transformation function $T(v)$ is given as:

$$T(v) = (1.39 - v)/2.63 \quad (9)$$

Finally, the color harmony based emotional saliency is computed as:

$$HS(x, y) = \frac{1}{2}[T(HC(x, y) + T(HL(x, y))] \quad (10)$$

Emotional Saliency Map. Color emotion based emotional saliency map and color harmony based emotional saliency maps are linearly fused into emotional saliency map:

$$ES(x, y) = \lambda \times HS(x, y) + (1 - \lambda) \times CS(x, y) \quad (11)$$

where λ is fusing weight and $ES(x, y)$ indicates the emotional saliency of the pixel at location (x, y). The higher the emotional saliency is, the more likely this pixel captures the emotional attention.

4 Horror Image Recognition based on Emotional Attention

Bag-of-Words (BoW) is widely used for visual categorization. In the BoW model, all the patches in an image are represented using visual words with same weights. The limitation of this scheme is that a large number of background image patches seriously confuse image's semantic representation. To address this problem, we extend BoW to integrate emotional saliency map and call it Attention based Bag-of-Words (ABoW). The framework is shown in Fig.2.

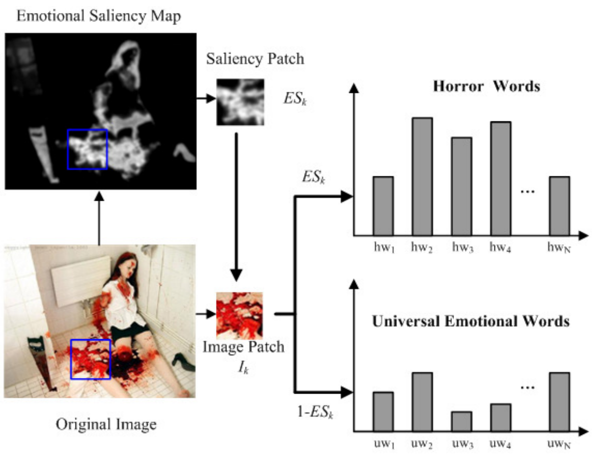

Fig. 2: Framework for Attention based Bag-of-Emotions.

4.1 Attention based Bag-of-Words

In the ABoW model, two emotional vocabularies, Horror Vocabulary (HV) and Universal Vocabulary (UV), and their histograms are defined to emphasize horror emotion character. The former one is for image patches in ESR while the latter one is for background. The combination of these two histograms is used to represent the image's emotion. The histogram for each vocabulary is defined as:

$$Hist(w_m, VOC) = \sum_{k=1}^{M} p(I_k \in w_m) \Big/ \sum_{k=1}^{M} \sum_{m=1}^{|VOC|} p(I_k \in w_m) \tag{12}$$

where $VOC \in \{HV, UV\}$, w_m is the mth emotional word in VOC. $p(I_k \in w_m)$ is the probability that the image patch I_k belongs to the word w_m. Because the words in the two vocabularies are independent. Consequently, $p(I_k \in w_m)$ can be defined as

$$p(I_k \in w_m) = p(I_k \in VOC) p(I_k \in w_m | VOC) \tag{13}$$

The $p(I_k \in VOC)$ is the probability of selection VOC to represent I_k. Because the image patches in ESR is expected to be representd by the HV histogram, we assume that

$$p(I_k \in HV) = p(I_k \in ESR) \propto ES_k; \quad p(I_k \in UV) = p(I_k \notin HV) \tag{14}$$

where ES_k is the average emotional saliency of image patch I_k and computed by the emotional saliency computation model. The higher ES_k is, the more possibility I_k belongs to ESR. Consequently, Eq (13) can be rewritten as:

$$p(I_k \in w_m) = \begin{cases} p(I_k \in hw_m) = ES_k \times p(I_k \in hw_m | HV) \\ p(I_k \in uw_m) = (1 - ES_k) \times p(I_k \in uw_m | UV) \end{cases} \tag{15}$$

where hw_m and uw_m are the emotional words in HV and UV respectively. The conditional probability $p(I_k \in w_m | VOC)$ is defined by the multivariate Gaussian function.

$$p(I_k \in w_m | VOC) = \frac{1}{((2\pi)^{D/2} |\Sigma|^{1/2})} \exp\{-\frac{1}{2}(F_k - c_m)^T \Sigma^{-1} (F_k - c_m)\} \tag{16}$$

where F_k is a D dimensional feature vector of I_k; C_m is a feature vector of emotional word w_m, Σ is the covariance matrix, and $|\bullet|$ is the determinate. The nearer the distance between F_k and C_m is, the higher the corresponding probability $p(I_k \in w_m | VOC)$ is.

4.2 Feature Extraction

In the ABoW framework, feature extraction for each image patch is another key step. To fill the gap between low level image features and high level image emotions, we select three types of features for each image patch, Color ,Color Emotion and Weibull Texture.

Color Feature: We consider the image color information in HSV space. Two feature sets are defined as: $[f_1^k, f_2^k, f_3^k] = [H_k, S_k, V_k]$ and $[f_4^k, f_5^k, f_6^k] = [|H_k - \overline{H}|, |S_k - \overline{S}|, |V_k - \overline{V}|]$. The former one is the average of all the pixels in the patch, and the latter one is defined as the difference between average value of the patch and that of the whole image.

Texture Feature: Geusebroek et al [15] report a six-stimulus basis for stochastic texture perception. Fragmentation of a scene by a chaotic process causes the spatial scene statistics to conform to a Weibull-distribution.

$$wb(x) = \frac{\gamma}{\beta}\left(\frac{x}{\beta}\right)^{\gamma-1} e^{-(\frac{x}{\beta})^{\gamma}} \tag{17}$$

The parameters of the distribution can completely characterize the spatial structure of the texture. The contrast of an image can be represented by the width of the distribution β, and the grain size is given by γ, which is the peakedness of the distribution. So $[f_7^k, f_8^k] = [\gamma_k, \beta_k]$ represents the local texture feature of the patch I_k and the texture difference between I_k and whole image is $[f_9^k, f_{10}^k] = [|\gamma_k - \bar{\gamma}|, |\beta_k - \bar{\beta}|]$.

Emotion Feature: The emotion feature is also extracted based on color emotion and color harmony theories. The average color emotion of image patch I_k, and the difference between I_k and whole image are $[f_{11}^k, f_{12}^k, f_{13}^k] = [CA_k, CW_k, CH_k]$ and $[f_{14}^k, f_{15}^k, f_{16}^k] = [|CA_k - \overline{CA}|, |CW_k - \overline{CW}|, |CH_k - \overline{CH}|]$. We use the color harmony score between average color P_k of I_k and average color $\overline{P}$ of the whole image to indicate the harmony relationship between them,$[f_{17}^k] = [HA(P_k, \overline{P})]$.

4.3 Emotional Histogram Construction

The feature vectors are quantized into visual word for two vocabularies, HV and UV, by k-means clustering technique independently. For HV, horror words are carried out from those interesting regions with manual annotation in each horror training image. All image patches in both horror images and non-horror images in the training set are used to construct words of UV. Then each image is represented by combining the horror word histogram $Hist(w_m, HV)$ and the universal emotional word histogram $Hist(w_m, UV)$, as $[Hist(w_m, HV), Hist(w_m, UV)]$. And Support Vector Machine (SVM) is applied as the classifier to determine whether any given image is horror one or not.

5 Experiments

To evaluate the performance of our proposed scheme, we conducted two separate experiments, one is about emotional attention, and the other one is about Horror image recognition.

5.1 Data Set Preparation and Error Measure

Because of lacking open large data set for horror image recognition, we collected and created one. A large number of candidate horror images are collected from internet. Then 7 Ph.D students in our Lab are asked to label them from one of the three categories: Non-horror, A little horror, and Horror. They are also required to draw a bounding box around the most emotional salient region (according to their understanding of saliency) of each image. The provided annotations are used to create a saliency map $S = \{s_X | s_X \in [0,1]\}$ as follow:

$$s_X = \frac{1}{N}\sum_{n=1}^{N} a_X^n \tag{18}$$

where N is the number of labeling users and $a_X^n \in \{0,1\}$ is a binary label give by nth user to indicate whether or not the pixel X belongs to salient region. We select out 500 horror images from these candidates to the create horror image set, each labeled as 'Horror' by at least 4 users.

On the other hand, we also collect 500 non-horror images with different scenes, objects or emotions. Specially, the non-horror images includes 50 indoor images, 50 outdoor images, 50 human images, 50 animal images, 50 plant images, and 250 images with different emotions (adorable, amusing, boring, exciting, irritating, pleasing, scary, and surprising) which are downloaded from an image retrieval system ALIPR(http://alipr.com/).

Finally, all these 500 horror images, along with their emotional saliency S, and 500 non-horror images are put together to construct a horror image recognition image set (HRIS) used in the following experiments.

Given the ground truth annotation s_X and the obtained emotional saliency value e_X with different alogrithms, the precision, recall, and F_α measure defined in Eq.(19) are used to evaluate the performance. As many previous work [16], α is set as 0.5 here.

$$\begin{array}{l} Precsion = \sum_X s_X e_X / \sum_X e_X ; Recall = \sum_X s_X e_X / \sum_X s_X \\ F_\alpha = \frac{(1+\alpha)\times Precsion \times Recall}{\alpha \times Precsion + Recall} \end{array} \tag{19}$$

5.2 Experiments for Emotional Saliency Computation

To test the performance of emotional saliency computation, we need to select optimal fusing weight λ in Eq(11) firstly. The $F_{0.5}$ value change curve with λ on the 500 horror image set is shown in Fig.3(A). It tells us that the best $F_{0.5}$ value, 0.723, is achieved when we set $\lambda = 0.2$. So λ is always set to be 0.2 in the followings. From the optimal λ selection, we can find that the emotional saliency map with only color emotion based saliency map ($\lambda = 0$) is better than that with only color harmony based saliency map ($\lambda = 1$). But the color harmony based saliency map also can improve the color emotion based saliency map. Both of them are important for the final emotional saliency map.

We also compare our emotional saliency computation method with other existing visual saliency algorithms, Itti's method (ITTI) [17], Hou's method (HOU) [18], Graph-based visual saliency algorithm (GBVS) [19], and Frequency-tuned Salient Region Detection algorithm (FS) [20]. The precision (Pre), recall (Rec) and $F_{0.5}$ values for each method are calculated and shown in Fig.3 (B). In

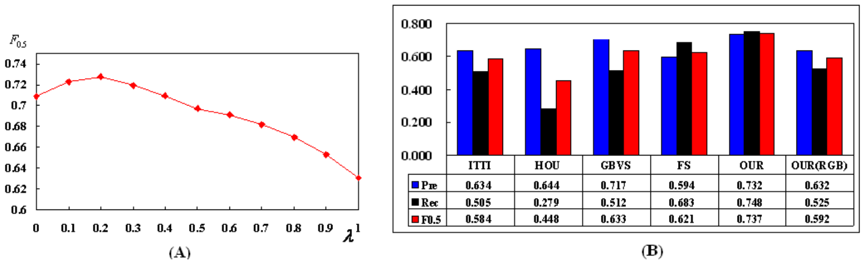

Fig. 3: (A)$F_{0.5}$ change curve as a function of λ. (B) Comparison with other methods.

addition, in order to validate the effectiveness of color emotion, we also apply our saliency framework on the RGB space, denote as 'Our(RGB)' in Fig.3. The output of our methods are $Pre = 0.732$, $Rec = 0.748$, and $F_{0.5} = 0.737$ respectively, showing that it outperforms all other visual saliency computation methods. This is due to the fact that visual saliency computation are solely based on local feature contrast without considering any pixel's salient property itself and some emotional factors, so neither of them can detect emotional saliency accurately. In contrast, emotional factors that can integrate the single pixel's salient value are added in our solution, so it can improve performances. In addition, the proposed method also outperforms Our(RGB), which indicates the color emotion and color harmony theories are effective for emotional saliency detection.

5.3 Horror Image Recognition

In this experiment, an image is divided into 16×16 image patches with overlapping 8 pixels. We have discussed many features in section 4.2. Now we need to find the feature or feature combination that best fits horror image recognition at hand. We divide these features into 3 subsets: Color Feature ($[f_1 - f_6]$), Texture Feature ($[f_7 - f_{10}]$) and Emotion Feature ($[f_{11} - f_{17}]$). Seven different selections (or combinations) are tried based on the ABoW model. In order to simplify the parameter selection, we set the word numbers in HV and UV equal, $N_H = N_U = 200$. We equally divide the image set into two subsets A and B. Each subset includes 250 horror images and 250 non-horror images. We then use A for training and B for test and vice versa. The combined results of the two experiments are used as the final performance. The Precision, Recall and F_1 values, which are widely used for visual categorization evaluation, are adopted to evaluate the performances of each feature combination. Different from the computation of Precision and Recall in emotional saliency evaluation, the annotation label here for horror image recognition is binary, '1' for horror image and '0' for non-horror images. The experimental results are shown in Table 1.

Table 1 shows that the Emotion feature outperforms the other single features, meaning that the color emotion and color harmony features are more useful for horror image recognition. Regarding combinational features, combination of the color, texture and emotion features performs best, with $F_1 = 0.769$.

Table 1: Comparison with different feature combinations based on ABoW.

Feature	*Prec*	*Rec*	F_1
Color	0.700	0.742	0.720
Texture	0.706	0.666	0.685
Emotion	0.755	0.742	0.748
Color + Texture	0.735	0.772	0.753
Color + Emotion	0.746	0.780	0.763
Texture + Emotion	0.769	0.744	0.756
Color + Texture + Emotion	**0.768**	**0.770**	**0.769**

In order to check the effectiveness of the emotional saliency map, we also construct the combinational histogram without emotional saliency for horror image recognition, in which the emotional saliency value for each image patch is set as 0.5. That is to say, the weights for the horror vocabulary and for the universal vocabulary are equal. The same feature combination (**Color + Texture + Emotion**) is used for horror image recognition both with and without emotional saliency. In addition, we also compare our method with Emotional Valence Categorization (EVC) algorithm [12] proposed by Yanulevskaya and color based Bag-of-Emotions (BoE) model[13]. The experimental results are shown in Table 2.

Table 2: Comparison with other emotional categorization methods.

Method	*Pre*	*Rec*	F_1
EVC	0.760	0.622	0.684
BoE	0.741	0.600	0.663
Proposed (Without Emotional Saliency)	0.706	0.698	0.702
Proposed (With Emotional Saliency)	**0.768**	**0.770**	**0.769**

The results show that the proposed horror image recognition algorithm based emotional attention outperforms all other methods. The comparison between with and without emotional saliency gives out that the emotional saliency is effective for horror image recognition. The potential reason lies in that the EVC and BoE describe image's emotion only from global perspective, but the horror emotional is usually generated by image's local region. This proves again that the emotional saliency is useful to find the horror region from images.

6 Conclusion

This paper has proposed a novel horror image recognition algorithm based on emotional attention mechanism. Reviewing this paper, we have the following conclusions: (1) This paper has presented a new and promising research topic, Web horror image filtering, which is supplement to existing Web content filtering. (2) Emotional attention mechanism, an emotion driven attention, has been introduced and a initial emotional saliency computation model has been proposed. (3)We have given out an Attention based Bag-of-Words (ABoW) framework, in which the combinational histogram of horror vocabulary and universal vocabulary has been used for horror image recognition. Experiments on real images have shown that our methods can recognize those horror images effectively and can be used as a new basis for horror content filtering in the future.

Acknowledgement. This work is partly supported by National Nature Science Foundation of China (No. 61005030, 60825204, and 60935002), China Postdoctoral Science Foundation and K. C. Wong Education Foundation, Hong Kong.

References

1. Rachman, S.: The conditioning theory of fear acquisition. Behaviour Research and Therapy **15** (1997) 375–378
2. Forsyth, D., Fleck, M.: Automatic detection of human nudes. IJCV **32** (1999) 63–77
3. Hammami, M.: Webguard: A web filtering engine combining textual, structural, and visual content-based analysis. IEEE T KDE **18** (2006) 272–284
4. Hu, W., Wu, O., Chen, Z.: Recognition of pornographic web pages by classifying texts and images. IEEE T PAMI **29** (2007) 1019–1034
5. Guermazi, R., M.Hammami, A.Hamadou: Webangels filter: A violent web filtering engine using textual and structural content-based analysis. In: Proc. of ICDM. (2008) 268–282
6. Yanulevskaya, V., van Gemert, J.C., Roth, K.: Emotional valence categorization using holistic image features. In: Proc. of ICIP. (2008) 101–104
7. Solli, M., Lenz, R.: Color based bag-of-emotions. In: Proc. of CAIP. (2009) 573–580
8. http://en.wikipedia.org/wiki/Horror_and_terror.
9. Lang, P.J., Bradley, M.M., Cuthbert, B.N.: International affective picture system (iaps): Technical manual and affective ratings (1999) Tech. Rep., Gainesville, Centre for Research in Psychophysiology.
10. Sun, Y., Fisher, R.: Object-based visual attention for computer vision. AI **146** (2003) 77–123
11. Vuilleumier, P.: How brains beware: neural mechanisms of emotional attention. TRENDS in Cognitive Sciences **9** (2005) 585–594
12. Ou, L., Luo, M., Woodcock, A., Wright, A.: A study of colour emotion and colour preference. part i: Colour emotions for single colours. Color Res. & App. **29** (2004) 232–240
13. Ou, L., Luo, M.: A colour harmony model for two-colour combinations. Color Res. & App. **31** (2006) 191–204

14. Fedorovskaya, E., Neustaedter, C., Hao, W.: Image harmony for consumer images. In: Proc. of ICIP. (2008) 121–124
15. Geusebroek, J., Smeulders, A.: A six-stimulus theory for stochastic texture. IJCV (**62**)
16. Liu, T., J.Sun, Zheng, N.N., Tang, X., Shum, H.Y.: Learning to detect a salient object. In: Proc. of CVPR. (2007) 1–8
17. Itti, L., Koch, C., Niebur, E.: A model of saliency-based visual attention for rapid scene analysis. IEEE T PAMI **20** (1998) 1254–1259
18. Hou, X., Zhang, L.: Saliency detection: A spectral residual approach. In: Proc. of CVPR. (2007) 1–8
19. Harel, J., Koch, C., Perona, P.: Graph-based visual saliency. In: Proc. of NIPS. (2006) 545–552
20. Achanta, R., Hemami, S., Estrada, F., Ssstrunk, S.: Frequency-tuned salient region detection. In: Proc. of CVPR. (2009) 1597–1604