ADVERTISEMENT EVALUATION USING VISUAL SALIENCY BASED ON FOVEATED IMAGE

Zhiguo Ma\textsuperscript{1}, Laiyun Qing\textsuperscript{1,2}, Jun Miao\textsuperscript{1}, Xilin Chen\textsuperscript{1}

\textsuperscript{1}Key Laboratory of Intelligent Information Processing, CAS, Beijing 100190, China
\textsuperscript{2}Graduate School of the Chinese Academy of Sciences, CAS, Beijing 100039, China
\{zgma,lyqing,jmiao,xlchen\}@jdl.ac.cn

ABSTRACT

This paper proposes a novel approach to advertisement evaluation using automatic salient regions. The salient regions are detected using a predicting model, in which the predicted values are obtained by the foveated image. The saliency is defined as the difference between the input and its prediction. Then an advertisement is determined as attractive if the detected salient regions are overlapped with the interested regions of the advertisement. The experimental results on the advertisements data set are encouraging.

Index Terms— Advertisement Evaluation, Salient Region, Foveated Image

1. INTRODUCTION

The evaluation of an advertisement about its attractiveness is one of the most fundamental and important tasks in the field of advertising psychology. Previously works in this field focus mainly on the activity of the customer’s eyes captured by eye tracking devices [1, 2, 3]. Wedel and Pieters [1] divided an advertisement into three elements (brand, pictorial and text), and recorded the frequency of eye fixations falling into each element of the advertisement. They found that viewers are more likely to fixate on the pictorial element, followed by the text and the brand element. They also found that the memory of advertised brand is related to the frequency of eye fixations on the advertisement.

The eye tracking device records the actual activity of the viewer’s eyes in realtime. However, such expensive devices are not available in many cases, and it is also annoying to wear the device in most situations. Therefore some researchers concentrate on the visual saliency or attention of the advertisement stimuli to overcome the problem of using the eye tracking device in advertisement evaluation [4, 5, 6, 7].

Some computational models of visual saliency and attention has been investigated in recent years [8, 9, 10, 11]. Itti and Koch [8] proposed a computational saliency model based on feature integration theory to simulate the visual search mechanism of the HVS. The computational model integrates low level features such as color, intensity, orientation, and accounts for some characteristic of HVS such as center-surround differences, inhibition of return. The model achieves successful performance on task of searching target in complex natural scenes. Kienzle et al. [9] utilized machine learning techniques to build a saliency model from recordings of human eye fixations on natural images. Gao and Vasconcelos [10] implemented the center-surround mechanisms as a discriminant process based on the distributions between the local images patch its surrounding. Hou and Zhang [11] took the advantage of the spectral residual of an image in spectral domain in the task of detection of visual saliency in natural images.

In this paper, we propose a novel computational model to detect the salient regions in an image, and apply it to advertisement evaluation. The proposed saliency model exploits the foveated image characteristic of the retina [12, 13] and uses the foveated image as the prediction of the input. The saliency of a region is defined as the difference between the original image and its prediction in the region. The foveated image is updated by placing a new fixation on the center of the region with maximum saliency.

The rest of the paper is organized as follows. The details of the proposed saliency method are described in Section 2. The experimental results are shown in Section 3. Finally, conclusion and future work are given in Section 4.

2. SALIENCY BASED ON FOVEATED IMAGE

Visual psychophysics research has shown that human observers are able to obtain the outline of a novel image within a short glance [14]. Then this initial analysis of global structure guides the analysis of local details [15, 16]. The HVS will focus the eye on some salient regions and gain some extra information of these regions.

These observations inspire the proposed computational model (in Fig.1.), mimicking the behavior of the HVS. The input image is first blurred to reduce the impaction of the texture in detecting salient regions. The most blurred image without any fixations simulates the outline of the input image
and serves as the initial estimation of the input. The saliency
is defined as the difference between the input and its predic-
tion. The next fixation is determined as the center of the most
salient region. The information gained from the fixation is
computed using the foveated image, where the regions near
the fixation have higher resolution than those far away from
it. Then the foveated image is used as the estimation in next
iteration and this process continues until we accomplish the
specific visual task.

Some more details of the saliency and the foveated image
are given in followings.

2.1. Saliency

The image is divided into $N \times N$ regions. The saliency
of the region $R_{i,j}$ is defined as the difference between the input
$I$ and its estimation. In this implementation, the difference is
the mean square difference, and the estimation is the foveated
image $F$. The difference of each color channel is accumu-
lated. Therefore, the saliency of the region is:

$$Sal(R) = \sum_{x=x_0}^{x_0+w-1} \sum_{y=y_0}^{y_0+h-1} \sum_{c=1}^{C} [I(x,y,c) - F(x,y,c)]^2,$$

where $(i,j)$ is the index of the region, $(x_0, y_0)$ is the starts
coordinate of the region $R$, and $w$, $h$, $C$ are the width, height
and the number of color channels, respectively.

The saliency matrix in Fig. 1 stores the saliency value of
every region of the $N \times N$ regions. It means that

$$Saliency\ matrix(i,j) = Sal(R_{i,j}).$$

The fixations are generated progressively based on the
saliency of the regions. In each iteration, the region with max-
imum value in saliency matrix is selected as the next salient
region. The center of the salient region is added as a new

2.2. Foveated Imaging

We simulate the foveated imaging phenomenon in HVS with
different Gaussian smoothing operator applied in different re-
gions of the image. The window widths and variances of the
Gaussian smoothing operators increase with the distance be-
tween the current pixel and the closest fixation.

The foveated image with multiple fixations is constructed
as following:

1. Compute the distances between the current pixel and all the fixations;
2. Find the closest fixation;
3. Select corresponding window width and variance of Gaussian smoothing operator according to the distance between the current pixel and its closest fixation;
4. Employ the gaussian smoothing operator with the given window width and variance on original image;
5. Select the pixel value of the blurred image with the same coordinate as the value of the current pixel.

The foveated image without any fixations serves as the
initial estimation of the input. We employ the Gaussian
smoothing operator on the original image with variant win-
dow widths and variances and store the blurred images, to
avoid the repeatedly Gaussian smoothing.

A sample image generated by the foveated imaging
method with one fixation is shown in Fig. 2.
Table 1. Results of advertisement evaluation

<table>
<thead>
<tr>
<th>Labeler</th>
<th>Our method</th>
<th>Saliency Toolbox</th>
</tr>
</thead>
<tbody>
<tr>
<td>Labeler 1</td>
<td>102</td>
<td>94</td>
</tr>
<tr>
<td>Labeler 2</td>
<td>109</td>
<td>109</td>
</tr>
<tr>
<td>Labeler 3</td>
<td>100</td>
<td>93</td>
</tr>
<tr>
<td>Labeler 4</td>
<td>106</td>
<td>99</td>
</tr>
</tbody>
</table>

3. EXPERIMENTS

We use the data set from [11] to evaluate the performance of the proposed method on saliency detection. Some other 132 classic advertisement images collected from Internet are used in advertisement evaluation. Four students annotate the product and its trademark in each advertisement as salient regions. Each advertisement has three salient region at most. The effectiveness of an advertisement is measured by examining the frequency of fixations falling in these salient regions.

3.1. Experimental Settings

The input image is blurred to reduce the impaction of the texture in detecting salient regions. The blurring is obtained by convolution it with a Gaussian filter to reduce the high frequencies in texture regions. The window width and the standard deviation are set to 9, 1.5, respectively.

The parameter $N$ is set to 10 and the image is divided into $10 \times 10$ regions. The standard deviations $\sigma$ and window widths $w$ in foveated imaging are similar with our previous work [13].

We discretized the distances between the current pixel and its closest fixation to 5 values, which are 12, 24, 48, 96, 192, respectively. The window width of each Gaussian filter are 1, 5, 9, 13, 17, respectively. The corresponding $\sigma$ are calculated as

$$\sigma(i) = \left[ \frac{w(i)}{2 - 1} \right] \times 0.3 + 0.5, \ i \in [1, 5].$$

3.2. Results on Saliency Detection

Some of salient regions detected by the proposed method on data set from [11] are shown in Fig. 3. The results are compared with the those of labeled by humans and those of [17]. The fixations hit the 53% white regions and 73% the gray region. The results show that the method can detect most of the salient regions.

3.3. Results on Advertisement Evaluation

The experiential results on the advertisement data set are given in Table 1. The number of images be hit by the first three fixations are listed. It can be seen that the performance of our method is close to that of the human beings.

Some examples evaluated by the proposed method are shown in Fig. 4. Fig.4(a) are some examples of attractive advertisements. Fig.4(b) shows 4 images in all the 11 images that hit by neither our method nor the Saliency Toolbox using annotation of Labeler 1. We can see that in these images, one can hardly distinguish the advertised product from its background. The advertisements preferred by our method (Fig. 4 (c) ) are those with high color contrast compared with the Saliency toolbox (Fig. 4 (d)).

4. CONCLUSION AND FUTURE WORK

This paper proposes an approach to advertisement evaluation using salient regions based on foveated imaging. The proposed method achieves satisfying results on the task of advertisement evaluation by taking advantage of the foveated image feature of HVS. Future works include integration of other low level features such as orientation and intensity, and saliency on different scales.

Acknowledgements

This research is partially sponsored by NSFC (Nos. 60673091 and 60702031), Hi-Tech R&D Program of China (No. 2006AA01Z122) and Natural Science Foundation of Beijing (No. 4072023).

5. REFERENCES

Fig. 4. Example of attractive and non-attractive images. (a) Attractive images detected in both method; (b) Non-attractive images missed in both method; (c) Images that are detected in our method, but missed by Saliency Toolbox; (d) Images that are detected in Saliency Toolbox, but ignored by our method.


