An Approach to Optimal Text Placement on Images

Gautam Malu and Bipin Indurkhya

International Institute of Information Technology,
Hyderabad, India
gautam.malu@research.iiit.ac.in, bipin@iiit.ac.in

Abstract. In deciding where to place a text block on an image, there are two major factors: aesthetic of the design composition, and the visual attention that the text block naturally attracts. We propose a computational model to address this problem based on the principles of visual balance and the diagonal method of placing emphasis. A between-subject study with seven participants was conducted to validate our model with subjective ratings. Eight color photographs were used to generate a set of text-overlaid images as the stimuli. Participants rated the stimuli for aesthetic appeal on a seven-point likert scale. Results show that the participants preferred text-overlaid images generated by our method of text placement over random text placement.

Keywords: Computational aesthetics, Interface design, Visual Balance, Diagonal Method.

1 Introduction

Some previous studies in the field of computational aesthetics have focussed on layouts containing text blocks and images on a solid color background [1,2], but these studies do not address the aesthetics of text-overlaid images. Sandhaus *et al* [3] have dealt with the same but on an image background. These studies are concerned with the arrangement of multiple elements on a single background. Lai *et al.* 2010 [4] have proposed a computational model for overlaying a single text element on a background image. However, their approach only works for images with homogeneous backgrounds.

In this paper, we propose a computational model for optimal text placement on images using the principles of visual balance. We also used the diagonal method to improve the overall aesthetics of text-overlaid images.

Visual balance is a key principle in the study of design and composition [5,6]. In his book, *Art and Visual Perception,* Rudolf Arnheim [7] articulates visual balance in terms of perceived *visual weights*. In recent years, there have been several studies on computational modeling of visual balance using different kinds of visual weights: for example, contrast for grayscale images [2], visual saliency [8], and color contrast for colored images [4]. It is generally agreed that in a balanced composition the visual weight is equally distributed in every directions.

D. Harris (Ed.): EPCE/HCII 2013, Part I, LNAI 8019, pp. 68–74, 2013.

The diagonal method [9] suggests that the objects an artist wants to emphasize are often found at one of the bisecting diagonals of the frame. A bisecting diagonal is one that bisects a corner angle [Figure 1]. Arnheim [7] also stated that objects when placed at the diagonals appear *visually heavier* than any other location. Following this principle, text can be emphasized if placed at one of the bisecting diagonals.

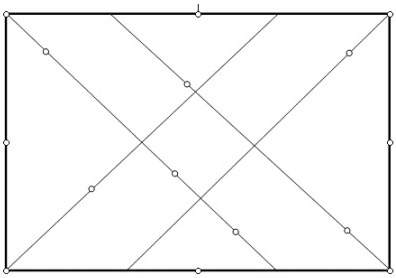


Fig. 1. Bisecting diagonals of a rectangle

2 Approach

We followed two approaches to position text on an image:

- Based on Visual Balance (VB)
- Based on Visual Balance and Diagonal Method (VB + DM)

2.1 Approach Based on Visual Balance (VB)

2.1.1 Visual Weight We defined the visual weight in terms of visual saliency values, as follows:

$$W(x,y) = \begin{cases} S(x,y), S(x,y) > T \\ 0, otherwise \end{cases} \tag{1}$$

Here, T is the threshold obtained using Otsu's method[10].

2.1.2 Saliency Algorithm We used a graph-based saliency model that also incorporates face detection along with low-level saliency features of color, intensity, and orientation [11]. Voila Jones algorithm was used for face detection [12]. This saliency algorithm does not incorporate the saliency of text; so we made a text conspicuity map by using delta functions at the center of the text block with 2D Gaussian with the standard deviation equal to the minimum of the sides of the text block. We added this conspicuity map to the existing saliency map. Now, the saliency map is the uniform linear combination of all five normalized conspicuity maps: Color(C), Intensity(I), Orientation(O), Face(F) and Text(T) [Figure 2].

$$S = \frac{1}{5}[N(\overline{C}) + N(\overline{I}) + N(\overline{O}) + N(\overline{F}) + N(\overline{T})] \tag{2}$$

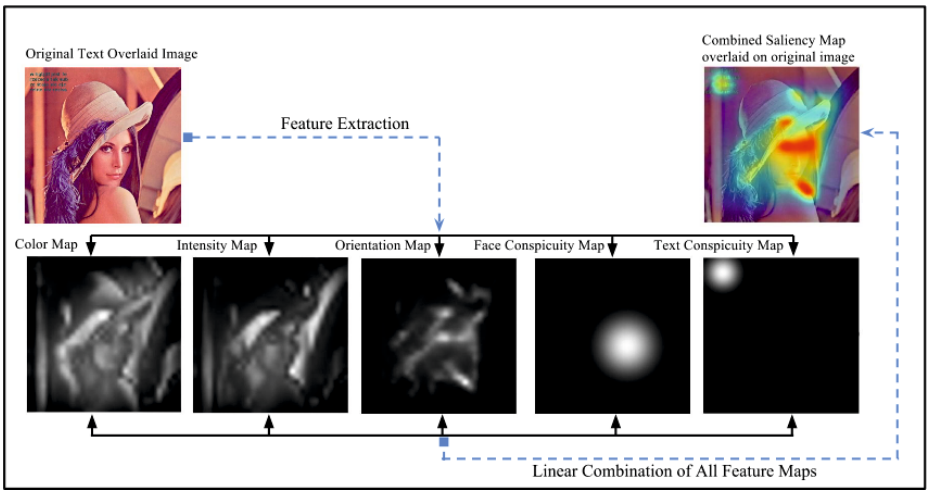


Fig. 2. Modified saliency model: An image is processed through standard color, orientation and intensity multi-scale channels [13], along with face detection channel and text detection channel. All five maps are normalized to the same range, and added with equal weights to a final saliency map.

2.1.3 Text Placement The center of mass (x_c, y_c) was computed from $W(x, y)$, as follows:

$$x_c = \frac{\sum_{j=1}^{h} \sum_{i=1}^{w} W(i,j) \times i}{\sum_{j=1}^{h} \sum_{i=1}^{w} W(i,j)} \tag{3}$$

$$y_c = \frac{\sum_{i=1}^{w} \sum_{j=1}^{h} W(i,j) \times j}{\sum_{i=1}^{w} \sum_{j=1}^{h} W(i,j)} \tag{4}$$

Here, w= Width of the image h= Height of the image

For a balanced composition, the center of mass should be at the minimal distance from the center of the frame. We find the quadrant in which the center of mass was located.To minimize the distance between the center of mass and the center of the image, the text block should be placed in a quadrant opposite to the quadrant in which the center of mass was located. We gridded this quadrant, with each grid cell being equal to the size of the text block. The text block was placed at each grid cell and the center of mass was calculated again. We calculated the Manhattan distance between the center of mass and the center of the image.The grid cell with the least distance was selected for placement of the text block.[Figure 3]

2.2 Visual Balance and Diagonal Method (VB+DM) Approach

This method was an extension of the visual balance method. Diagonal method states that the salient objects should be placed at one of the bisecting diagonals,

but it does not specify the diagonal or the location of the diagonal. After computing the text location by VB method (section 2.1), the text block was moved perpendicularly to the bisecting diagonal of the nearest corner. [Figure 3]

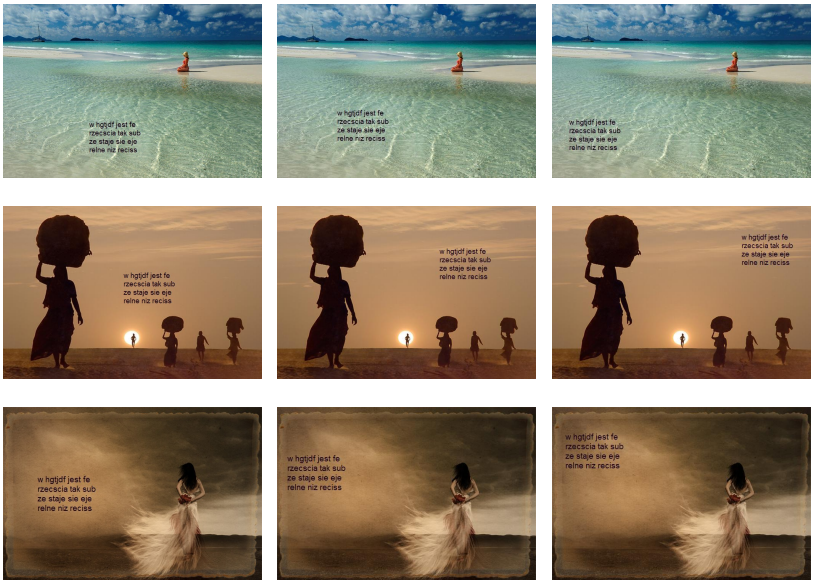
Fig. 3. Overview of the approach: (a) Input Image, (b) Saliency map, (c) Visual weights and center of visual weight (d) Text placement using VB model (red line indicates the bisecting diagonal), (e) Text placement using VB+DM method; red lines are for illustration only

3 User Evaluation

To evaluate the performance of our method, we conducted user evaluation study in which we compared both models against random placement of text blocks as well as with each other.
Seven university students (three females and four males; mean age: 23.2; age range 20 -26) took part in the evaluation study. None of the participant had any formal art education. Art naive participants were selected to avoid any possible background effect of expertise in art.
Eight colored photographs were selected from `http://photo.net/`. All photpgraphs were rated at least six on a seven point scale by Photo.net users. Three variants of each photograph were prepared by placing a text block on them using the following methods [Figure4]:

- random placement excluding the salient regions and a 10 pixel margin on all sides.
- VB Approach (section 2.1).
- VB+DM Approach (section 2.2).

The only difference between the variants was the location of the text. The text block was devoid of any semantic meaning, as we are not considering the semantics of the text while finding its optimal position.
Each participant was shown all three variants of each photograph simultaneously and was asked to rate each variant on a seven-point scale for aesthetic quality. The photographs and variants were presented in random order. Their responses were normalized per image per participant between 0 and 1. A total of 168 (3x8x7) responses were gathered.



(a) (b) (c)

Fig. 4. Examples of the stimuli used in user evaluation study (a) Random placement of text,(b) Based on VB model, (c) Based on VB+DM model

4 Results and Discussion

A one-way between-subject ANOVA test revealed that the mean aesthetic scores were significantly different for all three models at $p < 0.01$ level[$F(2, 165) = 28.455, p = 0.000$]. Post-hoc comparison using Tukey HSD test indicted mean aesthetic scores for both VB model ($M = 0.3343, SD = 0.0745$) and VB+DM model($M = 0.3891, SD = 0.0771$) are significantly better than random placement ($M = 0.2741, SD = 0.0896$). The results also indicated that the mean aesthetic score for VB+DM model is significantly higher than VB model at $p < 0.01$ level.

Taken together, these results suggest that our approach using both visual balance and diagonal method performed better than random placement of text on images. It is noteworthy that the diagonal method further significantly increased the overall aesthetics of the text-overlaid image.

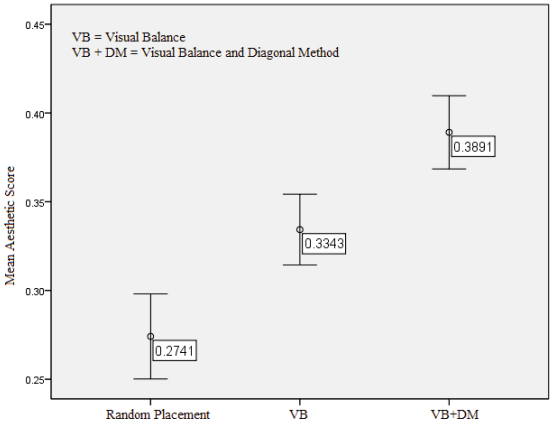


Fig. 5. Mean aesthetic scores for all three methods

5 Applications and Future Work

There are many websites that allow users to share their photographs in public domain, for example *Photo.net*.One example is a portal for online greeting cards (e-cards), where the user supplies a message which is overlaid on the selected image. Existing systems simply place the text message on the top of the image and send it as an e-card. In our model, the text color is defined by the user, but it could also be decided automatically using different color harmony schemes[14]. Different color evokes different emotional reactions [15,16]; so according to the emotional content of the message different color harmonies could be used.

References

1. Ngo, D.C.L., Teo, L.S., Byne, J.: A mathematical theory of interface aesthetics. Visual Mathematics 2 (2000)
2. Lok, S., Feiner, S., Ngai, G.: Evaluation of visual balance for automated layout (2004)
3. Sandhaus, P., Rabbath, M., Boll, S.: Employing aesthetic principles for automatic photo book layout. In: Lee, K.-T., Tsai, W.-H., Liao, H.-Y.M., Chen, T., Hsieh, J.-W., Tseng, C.-C. (eds.) MMM 2011, Part I. LNCS, vol. 6523, pp. 84–95. Springer, Heidelberg (2011)
4. Lai, C.Y., Chen, P.H., Shih, S.W., Liu, Y., Hong, J.S.: Computational models and experimental investigations of effects of balance and symmetry on the aesthetics of text-overlaid images. International Journal of Human-Computer Studies 68(1), 41–56 (2010)
5. Locher, P.J., Jan Stappers, P., Overbeeke, K.: The role of balance as an organizing design principle underlying adults' compositional strategies for creating visual displays. Acta Psychologica 99(2), 141–161 (1998)

6. Locher, P., Cornelis, E., Wagemans, J., Stappers, P.: Artists' use of compositional balance for creating visual displays. Empirical Studies of the Arts 19(2), 213–228 (2001)
7. Arnheim, R.: Art and visual perception: A psychology of the creative eye. Univ of California Press (1954)
8. Liu, L., Chen, R., Wolf, L., Cohen-Or, D.: Optimizing photo composition. Computer Graphics Forum 29, 469–478 (2010)
9. Westhoff, E.: Westhoff's diagonal method
10. Otsu, N.: A threshold selection method from gray-level histograms. Automatica 11(285-296), 23–27 (1975)
11. Cerf, M., Harel, J., Einhäuser, W., Koch, C.: Predicting human gaze using low-level saliency combined with face detection. Advances in Neural Information Processing Systems 20 (2008)
12. Viola, P., Jones, M.J.: Robust real-time face detection. International Journal of Computer Vision 57(2), 137–154 (2004)
13. Itti, L., Koch, C., Niebur, E.: A model of saliency-based visual attention for rapid scene analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence 20(11), 1254–1259 (1998)
14. Cohen-Or, D., Sorkine, O., Gal, R., Leyvand, T., Xu, Y.Q.: Color harmonization. ACM Transactions on Graphics (TOG) 25, 624–630 (2006)
15. Valdez, P., Mehrabian, A.: Effects of color on emotions. Journal of Experimental Psychology: General 123(4), 394 (1994)
16. Kaya, N., Epps, H.H.: Relationship between color and emotion: a study of college students. College Student Journal 38, 396–405 (2004)