ENHANCING PRODUCT IMAGES FOR CLICK-THROUGH RATE IMPROVEMENT

Yeongnam Chae *Mitsuru Nakazawa* *Björn Stenger*

Rakuten Institute of Technology, Rakuten, Inc.

yeongnam.chae@rakuten.com

ABSTRACT

This paper proposes a statistical method to enhance image quality in order to increase the click-through rate (CTR) of product images. We build a joint probability model of global image features for photos of different product categories. The images are modified in terms of brightness, contrast, and sharpness in order to increase the expected CTR. The effectiveness of the method is evaluated using a perceptual user study, comparing it to histogram equalization methods, and by conducting an A/B test over a one-week period on the e-commerce site *Rakuten Ichiba*.

***Index Terms*—** image enhancement, click-through rate, advertisement, e-commerce

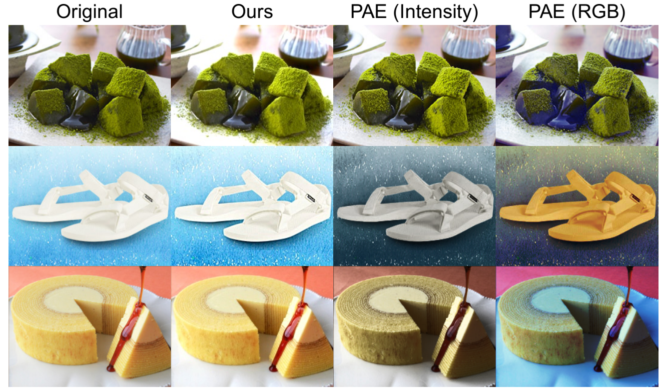

Fig. 1: Image enhancement examples. *(left column) Original input images [6][7][8], (second column) images after enhancement with the proposed statistical method, (third column) applying Piecewise Affine Equalization (PAE) [9] using intensity and (fourth column) PAE using RGB values. PAE works well on the top input image, but overly darkens the middle and bottom images.*

1 Introduction

Assessing the quality of a photograph is a matter of application context. If user preference is the quality metric, there is significant variance and dependence on the image content. Designing an automatic enhancement method that increases user preference based on adjusting image statistics such as sharpness or color distributions is not straightforward. In this paper we consider the CTR of images as a target metric. Images play a significant role in conveying information and attract potential customers to services. From a commercial perspective, it is interesting to note that advertisers spend over US$ 13 billion annually on banner ads [1], yet consumers pay little attention to them, as seen from the mean CTR value of 0.12% [2]. In order to improve the CTR of images advertisers have adopted image filtering applications for global image adjustment, *e.g.* [3, 4]. However, image filtering may not always yield the desired effect in terms of CTR increase, and advertisers rely on their experience or intuition to carefully adjust image parameters to generate a more attractive image. Here we propose an image processing method that helps automating the image enhancement process, thereby empowering merchants who are less experienced in image processing. Instead of defining an appearance-based quality measure of a given image, we directly aim to increase its CTR value. We first collect real-scale product images and their CTR from a large e-commerce website [5]. We then construct a statistical model of visual features and the corresponding CTR. Finally, we enhance an image by maximizing an objective function based on the image features.

The contribution of this paper is twofold. First, we propose an image enhancement method to increase the expected CTR. This is the first study that enhances the image based on CTR. The method can be applied to images of various categories, as shown Fig. 1. Second, we study the effectiveness of the method, using both a preference survey, comparing it with standard histogram equalization methods, and an A/B test on an e-commerce website [5], where the CTRs of original and modified images are measured over a one-week period.

2 Prior work

There has been significant prior work on enhancing the CTR for advertising images. The visual factors that have been investigated include:

Color. According to psychological studies, the dominant image color has effects on the viewer[10]. For example, blue tones convey trust, while red tones convey excitement[11]. To confirm how colors affect the CTR, North and Ficorilli [12] conducted A/B tests, and reported that on the website of an insurance company dominantly blue colors showed significantly higher CTR than red ones. In tests on both B2B and B2C websites, Lohtia *et al.* [13] found that moderate colors achieve higher CTR than dark or bright colors.

Image size. While large images tend to increase CTR [14], there are clearly limitations given the screen real-estate. Berke [15] compared the CTR for three different image sizes, reporting the highest CTR for medium sizes of 300×250, a finding also supported in [12].

Static and dynamic images. In the early days of the web many animated images were used for advertising. The study by Yoo *et al.* [16] reported that dynamic ads produced higher CTRs than static images, however, a more recent study of the effect of static and dynamic images on four company websites [17] reported the opposite effect. The effect of animations is therefore not conclusive.

3 Statistical Image Enhancement

The target metric adopted in this paper is the expected CTR of an image, which is the ratio of the number of clicks and the number of times an image was displayed on the web page. The CTR of an image is affected by various factors, such as the image content, visual quality, relative placement, and context on the page it is displayed. Here we are interested in how the image quality affects the CTR, and consider all other factors as nuisance variables. We formulate the task as finding a suitable enhancement function f_θ, with parameters θ which, when applied to an input image I leads to a high expected value of the CTR, r:

$$\theta^* = \text{argmax}_\theta \; p(r|f_\theta(I)). \tag{1}$$

We consider functions f_θ of the form

$$f_\theta(I_L) = (\alpha I_L + \beta) - \mathcal{G}_\sigma * I_L, \tag{2}$$

where I_L is the luminance channel of the image in Lab representation, and the parameters $\theta = (\alpha, \beta, \sigma)$ include the scaling parameter α, the offset β and the σ to determine the variance in the Gaussian kernel $\mathcal{G}_\sigma$. We take a learning approach to determine the parameters of the enhancement function. In order to estimate the perceptual effect of the image enhancement, we first extract global image features $\phi(I)$ and learn a mapping from these features to the expected CTR value

$$E(r|\phi(I)), \tag{3}$$

where $\phi(I)$ the 4-dimensional feature vector as

$$\phi(I) = (\varphi_b, \varphi_c, \varphi_s, \varphi_{sal})^T, \tag{4}$$

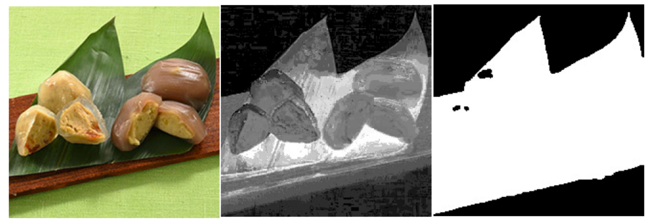

Fig. 2: Saliency coverage feature. *From the input image [20] (left), we compute a saliency map (middle) using a method that takes objectness into account [18]. The binarized saliency map (right) is then used to compute the area ratio.*

where φ_B is the mean brightness, φ_C is the image contrast, defined as the scaled standard deviation of the luminance channel:

$$\varphi_c = \zeta_\psi \left(|L|^{-1} \sum_{i \in I_L} l - \mu_{I_L} \right)^{\frac{1}{2}}, \tag{5}$$

with scale factor ζ_ψ. The sharpness, φ_S is defined as the maximum of the Laplacian:

$$\varphi_s = \zeta_\lambda \max_{i \in I_L} \left(\frac{\delta i}{\delta x} + \frac{\delta i}{\delta y} \right), \tag{6}$$

with scale factor ζ_λ. To compute the saliency value, φ_{sal}, we compute a saliency map using the method in [18], and binarize it using Otsu's method [19]. The saliency feature is the ratio of the salient and non-salient areas, see Fig. 2. This feature ensures that when adjusting brightness, contrast, or sharpness, the main object remains visually salient.

Category-based models. The images contain products from different categories. Fig. 3 shows marginal distributions of visual features and CTR values for two different product categories. Since these distributions differ for different products we create separate models for each product category. The feature distributions $\phi(I_k)$ for training images $\{I_K\}_{k=1}^K$ from a particular category are modeled using histograms, and we compute the expected CTR for each bin (Fig. 4(a)). Histogram bin-weighting is applied to down-weight bins containing few data points, thus reducing the effect of outliers. The resulting model with bin weighting is shown in Fig. 4 (b).

Local neighborhood weighting. We use weighting of the expected CTR value function in a local neighborhood of the input image features $\phi(I)$ to obtain the following objective function for an image I_j:

$$\mathcal{S}(\phi(I_j)) = w(\phi(I), \phi(I_j)) \, E(r(\phi(I_j))), \tag{7}$$

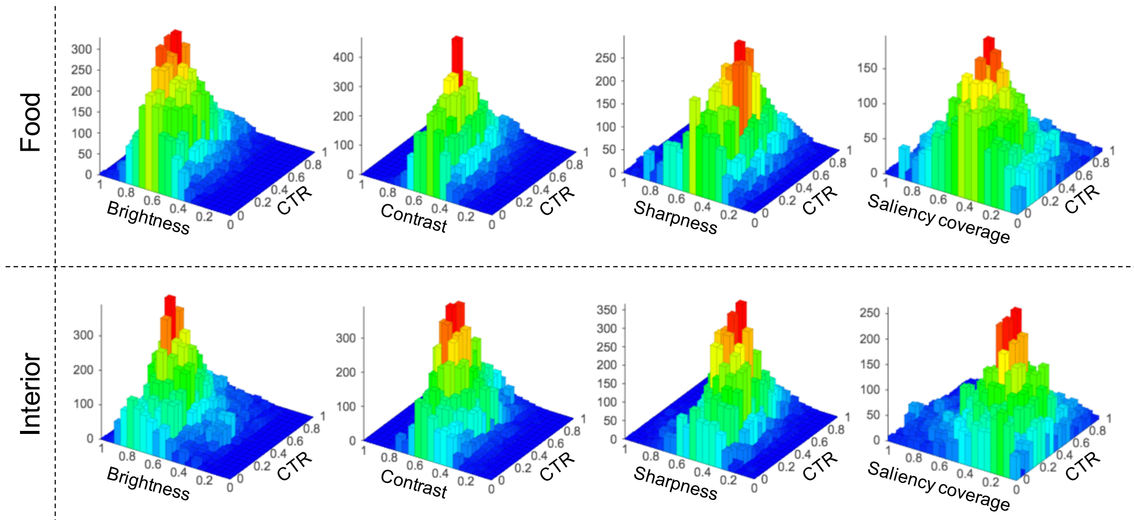

Fig. 3: Histograms of visual features and relative CTR. *(top) Food category, and (bottom) interior product category. The distributions change dependent on the image content, and are modeled separately for each category. Marginal histograms are shown for visualization.*

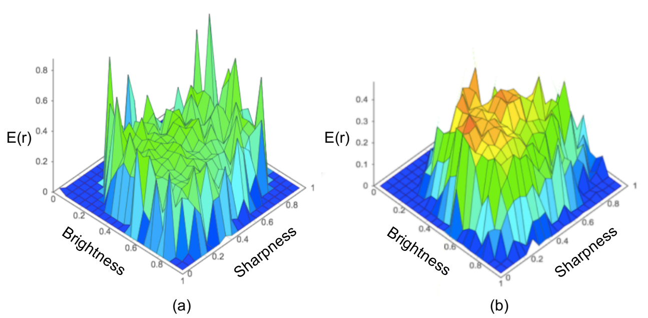

Fig. 4: Reducing the effect of bins with few entries. *The effect of outliers due to histogram bins with few entries in (a) is reduced by re-weighting (b). Only two of the visual features are shown here for visualization purposes.*

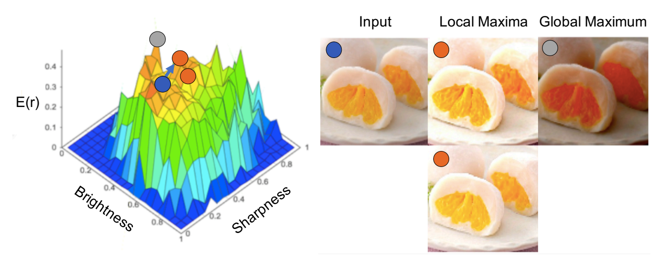

Fig. 5: Nearby local maxima vs. global maximum. *Local weighting of the expected CTR ensures that the enhancement parameters to not significantly distort the image values. (left) Input image [21], (right) Enhanced images using locally weighted maxima vs. the global maximum.*

by applying Gaussian weighting with zero mean and standard deviation $\sigma = 0.01$

$$w(\phi(I), \phi(I_j)) = \mathcal{G}_{\mu,\sigma}(||\phi(I) - \phi(I_j)||))/\zeta_w, \qquad (8)$$

where ζ_w is a scaling factor. The global maximum averages over all image parameters of a certain object category, and can therefore lead to undesired visual changes. The importance of local weighting is demonstrated in Fig. 5.

Note that because the visual features $\phi(I)$ are interdependent, it is difficult to find a mapping from the processed image $f_\theta(I)$ to CTR maxima directly. In order to estimate the image transformation parameters θ, we generate samples from a uniform distribution of the parameters $\{\theta_m\}_{m=1}^{M}$ within a local region, apply each f_{θ_m} to the image and compute the resulting image features. We select the one with the largest value of the objective function $\mathcal{S}$.

4 Results

We collected 127,121 images of 26 different product categories, collected from an e-commerce site [5]. The feature histogram was built using 20 bins for each dimension. Scale factors for the image features were chosen according to their theoretical range. The number of samples is chosen as $M = 300$.

We carried out two experiments to assess the performance of the method, a user preference study, and an A/B test. For

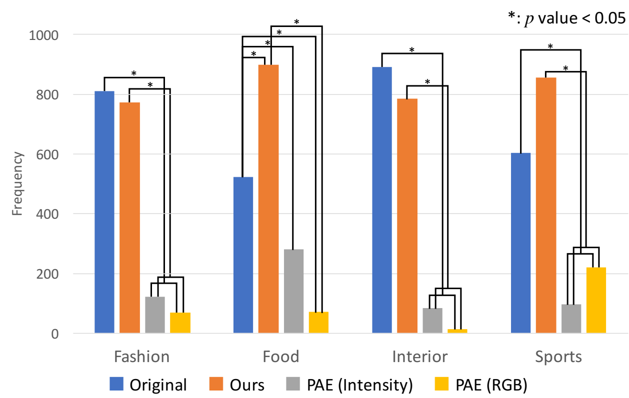

Fig. 6: User preference study. *The proposed method performs well on the food and sports goods categories, while performing slightly worse on fashion and interior categories. Statistically significant differences ($p < 0.05$) are shown with black connecting lines. Both PAE methods perform poorly.*

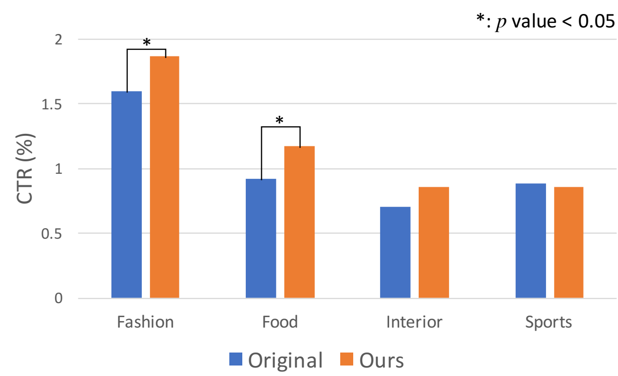

Fig. 7: Result of A/B test. *Change of CTR, when showing the original images and the enhanced images using the proposed method, for different product categories. Improvements are observed for fashion, food, and interior product photos. Only results on fashion and food are statistically significant.*

the two experiments, we used 48 images of 270×200 resolution, selected from four product categories: fashion, food, interior products and sports goods.

4.1 Preference survey

For each input image we computed the enhanced image using the proposed method, and two enhancement methods based on histogram equalization, proposed in [9]. These four images were displayed in randomized order to 148 users together with the question: *Which of these images makes you want to purchase the product most*?

The result is shown in Fig. 6. To measure the statistical significance, we calculated the p values. Both PAE methods led to low preference, even compared to the original image. PAE performed poorly on images containing large white areas as can be seen in the middle and bottom row of Fig. 1. The proposed method showed significantly higher preference than PAE. Compared to the input images, the it performed well on the food and sports goods categories, while performing slightly worse on fashion and interior categories. However, these differences, are statistically not significant ($p > 0.05$, see Fig. 6).

4.2 A/B test

In the second experiment we carried out an A/B test on an e-commerce website. The original images and images enhanced with the proposed method were shown at the same location of the page and all other page content was identical. The number of impressions and the number of click-throughs were measured over the period of one week, yielding a total of 100,720 impressions and 1,119 clicks.

The overall result is shown in Fig. 7. On average, the CTR was improved by 15.7%. By category, the enhanced images show significant improvement of 27.4% for food images, 16.9% on fashion images ($p < 0.05$). Although there is no significant difference for other categories, the interior and the sports category show 21.6% better and 3.1% worse performance, respectively.

4.3 Discussion.

We have carried out both a preference survey and an A/B test, which are in fact measuring two different metrics. While the survey measures conscious preference, this does not necessarily imply a higher CTR. We use the assumption that the survey serves as a useful proxy for an A/B test in a real setting, which carries more financial risk. Further, we also found cases in which the original image was already of high quality and the enhanced version did not differ significantly, leading to similar results of the two images.

5 Conclusion

In this paper we have presented a statistical image enhancement method, adjusting image brightness, contrast, and sharpness, with the objective to increase the click-through rate of product images. We have shown in a user preference study that our method outperforms standard enhancement methods using histogram equalization. In an A/B test over one week we have measured the change of CTR over the period of a week and found an average increase of 15.7%, demonstrating the effectiveness of the method.

6 References

[1] eMarketer, "US digital display ad spending to surpass search ad spending in 2016," https://www.emarketer.com/Article/US-Digital-Display-Ad-Spending-Surpass-Search-Ad-Spending-2016/1013442 (accessed on Feb 12, 2018).

[2] eMarketer, "Consumers get engaged with rich media," https://www.emarketer.com/Article/Consumers-Engaged-with-Rich-Media/1011282 (accessed on Feb 12, 2018).

[3] LateNiteSoft, "Camera+," http://camera.plus/ (accessed on Feb 8, 2018).

[4] Everimaging, "Fotor," https://www.fotor.com/ (accessed on Feb 8, 2018).

[5] Rakuten Inc., "Rakuten ichiba," https://global.rakuten.com/en/ (accessed on Feb. 8, 2018).

[6] "Source image," `http://image.ias.rakuten.co.jp/acp/ec/11461/12216177/1-1-1/e32f86f5ca46849881050ecd00a040df.jpg` (accessed on May 23, 2018).

[7] "Source image," `http://image.ias.rakuten.co.jp/acp/ec/41771/12216179/1-1-1/b40c18e2108918be208a2f7417349510.jpg` (accessed on May 23, 2018).

[8] "Source image," `http://image.ias.rakuten.co.jp/acp/ec/175992/12216186/1-1-1/82a1928a280677b5b87d5bc616d9c477.jpg` (accessed on May 23, 2018).

[9] J. Lisani, A. Petro, and C. Sbert, "Color and contrast enhancement by controlled piecewise affine histogram equalization," *Journal of Image Processing On Line*, , no. 2, pp. 49–60, 2012.

[10] G. Ciotti, "The psychology of color in marketing and branding," https://www.helpscout.net/blog/psychology-of-color/ (accessed on Feb 11, 2018).

[11] L. Widrich, "Why facebook is blue: The science of colors in marketing," https://blog.bufferapp.com/the-science-of-colors-in-marketing-why-is-facebook-blue (accessed on Feb 11, 2018).

[12] M. North and M. Ficorilli, "Click me: an examination of the impact size, color, and design has on banner advertisements generating clicks," *Journal of Financial Services Marketing*, vol. 22, no. 3, pp. 99–108, 2017.

[13] R. Lohtia, N. Donthu, and E. Hershberger, "The impact of content and design elements on banner advertising click-through rates," *Journal of Advertising Research*, vol. 43, no. 4, pp. 410–418, 2004.

[14] C. Cho, "How advertising works on the www: Modified elaboration likelihood model," *Journal of Current Issues & Research in Advertising*, vol. 21, no. 1, pp. 34–50, 1999.

[15] A. Berke, "What is the best ad size?," https://blog.adroll.com/best-practices/best-ad-size (accessed on Feb 8, 2018).

[16] C. Y. Yoo, K. Kim, and P. A. Stout, "Assessing the effects of animation in online banner advertising," *Journal of Interactive Advertising*, vol. 4, no. 2, pp. 49–60, 2004.

[17] J. O'Toole, "Which works better: Static vs. animated ads," https://www.webtrends.com/blog/2013/12/which-works-better-static-vs-animated-ads/ (accessed on Feb 8, 2018).

[18] G. Yildirim and S. Süsstrunk, "FASA: fast, accurate, and size-aware salient object detection," in *ACCV*, 2014.

[19] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Transactions on Systems, Man and Cybernetics.*, vol. 9, no. 1, pp. 62–66, 1979.

[20] "Source image," `http://thumbnail.image.rakuten.co.jp/@0_mall/otoshuclub/cabinet/item6/8300182_r.jpg` (accessed on May 23, 2018).

[21] "Source image," `https://event.rakuten.co.jp/mother/tieup/otoshu/img/20160415/item04_thumb.jpg` (accessed on May 23, 2018).