The Complexity of Perception of Image Distortion: an Initial Study

Yuzhen Niu[*]

School of Computer Science
and Technology
Shandong University
Shandong, 250001, China
yuzhen@cs.wisc.edu

Feng Liu

Department of Computer Sciences
University of Wisconsin-Madison
Madison, WI 53706, USA
fliu@cs.wisc.edu

Xueqing Li

School of Computer Science
and Technology
Shandong University
Shandong, 250001, China
xqli@sdu.edu.cn

[*] This work was done during Yuzhen Niu's visit to University of Wisconsin-Madison, USA

Michael Gleicher

Department of Computer Sciences
University of Wisconsin-Madison
Madison, WI 53706, USA
gleicher@cs.wisc.edu

Abstract

Image retargeting methods adapt an image for displays with different sizes and aspect ratios. Methods are based on the assumption that some kinds of distortions are preferable to others. While a rich variety of literature exists on retargeting methods, there is little codified understanding of how these distortions are perceived. This paper shows that people's perception of image distortions is complex. We report an initial study exploring the phenomenon that shows that even in a simple form of distortion, perception depends on a myriad of factors including amount of distortion, image content, and the viewer's cultural background. These initial findings have ramifications for the design and evaluation of image retargeting, and suggest that a more thorough study is necessary.
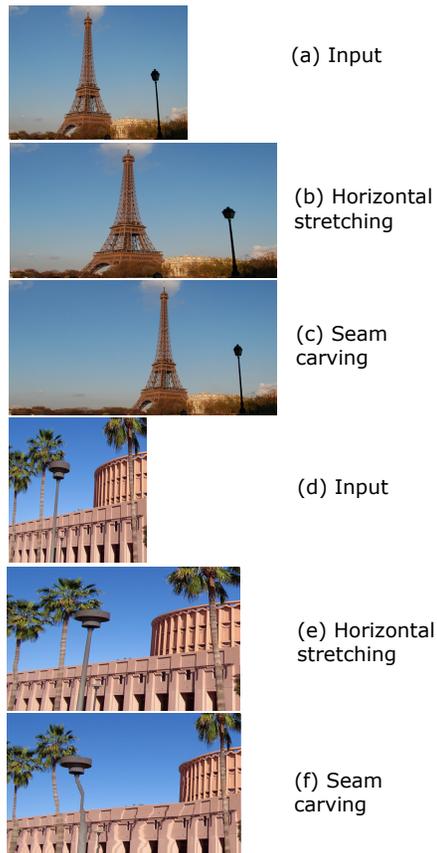
Keywords

Image stretch, image distortion, perceptual sensitivity, image retargeting, user study

ACM Classification Keywords

H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

General Terms

Design, Experimentation

(a) Input

(b) Horizontal stretching

(c) Seam carving

(d) Input

(e) Horizontal stretching

(f) Seam carving

Figure 1. Examples of image retargeting. (a) ~ (c) Photograph obtained from http://www.flickr.com/photos/panayotis/2997274618/ and reproduced under Creative Commons License http://creativecommons.org/licenses/by-sa/2.0/ (d) ~ (f) Photograph obtained from http://www.flickr.com/photos/pagedooley/2429429397/ and reproduced under Creative Commons License http://creativecommons.org/licenses/by/2.0/

Introduction

Today's world is full of a myriad of devices with displays. These displays have a diverse range of screen dimensions. Therefore, the problem of *image and video retargeting*, that is adapting an image (or video) so that it is better suited for a different display, has received considerable attention. In the past few years, there have been at least 70 technical papers (c.f. [7] ~ [10]) describing image retargeting methods, with at least 32 appearing in 2009 [1]. Commercialized versions have already appeared (for example, in Photoshop CS4).

Early methods on image retargeting either uniformly stretch and scale an input image to the target size, identify an important region of the image, and then intelligently crop off its surrounding content (c.f. [10]), or uniformly scale input image to one dimension of the target screen and add a letterbox/pillarbox to the borders of the other dimension. Recent content-aware retargeting approaches aim to adapt images by distorting them in ways that are considered to be less problematic than the above obvious adaptations. For example, seam carving methods [7] and spatially varying warp-based methods (c.f. [8]. [9]) provide solutions to image retargeting by distributing less distortion to those visually salient regions than the others. Figure 1 shows the effects of stretching method and content-aware seam carving method [7]. The distortions introduced by seam carving can be less noticeable because they are concentrated in less salient parts of the image (Figure 1(b) vs. 1(c)). But this is not true for all images. Seam carving distorts the structure of the source image, as shown in Figure 1(f), whereas people may not notice the stretch in Figure 1(e). In order to understand when these methods are

appropriate, we need a better understanding of how people perceive retargeting results.

While many retargeting methods have been evaluated for their relative performance or desirability, there has been little examination of the fundamental principle of how distortions are perceived. Without this grounding in perceptual foundations, retargeting methods are necessarily *ad hoc:* we have few good principles to know when and where and what kinds of distortions should be applied. A better understanding of how people perceive distortion will lead to better retargeting methods and better metrics for evaluating approaches.

In this paper, we show that the perception of image distortion is a complex phenomenon, depending on many factors. We report on an initial study that considers a very simple case: horizontal stretching. Our findings show that even in this simple case, people's ability to notice distortions depends on several factors including the amount of distortion, the content of the image, and the viewer's cultural background. These findings have direct implications for the design and evaluation of retargeting methods, as aspect ratio change is a focus of method development. They also suggest the need for and challenges in developing a more wide-scale understanding of image distortion.

Initial Study

Our hypothesis is that the perception of image stretch depends on many factors. To investigate this, we conducted a web-based study to measure how well people can notice stretching. The experiment was designed to examine the interactions of a number of
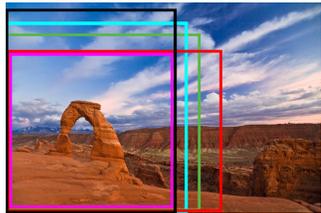
(a) AOI (pink) and cropping windows (in other colors)

(b) 0%

(c) 20%

(d) 40%

(e) 60%

Figure 2. An input image and its four cropped and stretched results. Photograph obtained from http://www.flickr.com/photos/4gott/2336126105/ and reproduced under Creative Commons License http://creativecommons.org/licenses/by/2.0/

effects including stretch scale, image content, and the viewer's cultural background.

Materials
Our first challenge was to select an appropriate set of images for the study. Our design (below) required 32 images that with sufficient quality and diversity of content. We selected 32 images from *Flickr* based on three criteria. First, this image set should cover a wide range of image categories. In our study, the image set covered all of the 14 categories used in [6] and covered 19 out of 20 categories given by [5]. We omitted textures and patterns because most are of artificial design. Second, all of the images in the set should be unstretched. To verify this, each image was checked to see if its aspect ratio matches the camera with which it was taken. Lastly, all selected images should be of high stretching of each image won't be noticeable to the participants. All the selected images were at least 800 Next, we needed to create 4 versions of each image with varying levels of stretch. In order to keep the experiment duration reasonable in this initial study, we focused on horizontal stretch, and chose uniformly increased stretch scales, that are 0%, 20%, 40%, and 60%. These variants were created in a way that minimized possible clues of stretching, such as size and quality of composition. To do this, we created 4 variants of the same size (400 X 300) by cropping appropriately sized regions and scaling them to this aspect ratio as shown in Figure 2. To perform the cropping, we first manually identified an area of appropriately sized cropping regions that included the AOI. Cropped regions contained approximately the same number of pixels to avoid having different amounts of image information and similar resampling artifacts. From these candidate croppings, we manually

selected one that we considered to be the best in terms of its compositional quality, preferring ones that avoid cutting objects and that approximate compositional rules (such as centering or the rule of thirds) [3].

The result of this preparation was a set of 32 diverse, high quality images. For each, we had 4 variants that were created using different amounts of horizontal scale, but were otherwise similar in size, composition, and resampling artifacts.

Design
In our study, the stretch scale is a within subject factor. To counterbalance stretch scales across all participants and images, a 4X4 Latin square design is used [2]. For every 4 participants, we assign each of them to one of the four participant groups. We also randomly assign 32 images to 4 image groups such that each image group gets 8 images. Each participant is shown all the image groups, where each group of images is stretched at a different scale. And the order in which each image is shown to each participant was randomized.

Participants
We conducted this web-based user study with two different subject pools. In U.S., 62 participants (51 males and 11 females, from 21 to 57 years old, average age of 26.5), all of whom were graduate students, voluntarily participated in our study. In China, 84 participants (54 males and 30 females, from 20 to 36 years old, average age of 26.4), all of whom were either undergraduate or graduate students, voluntarily participated in our study. After the study, we excluded the data from 8 U.S. participants and 15 China participants because they either did not complete all of the images or gave unreliable answers.

Category	20%	40%	60%	post-hoc
Image set	0.83	1.35	1.89	**|**|**
Human	1.57	2.45	3.01	**|**|**
Animal	1.14	2.39	2.89	o |**|**
Plant	0.68	1.43	2.07	o |**|**
Man-made object	0.88	1.26	1.98	**|**|**
Architecture	0.71	1.31	1.60	**|**|**
Ocean	0.38	0.36	0.94	o|o|*
Mountain	0.21	0.24	1.08	o|o|**

Table 1. Perceptual sensitivities for the whole image set and 7 categories. ?% column is the sensitivity at ?% stretch scale. The last column shows the post-hoc tests following ANOVA, on all U.S. participants' behavior at 0% and 20% scale, 0% and 40% scale, and 0% and 60% scale. Three results are separated by |, and symbol o, *, and ** mean $p \geq 0.05$, $p < 0.05 \leq 0.01$, and $p < 0.01$ respectively. So * and ** stand for people can detect the stretch, while o stands for people can't detect the stretch.

Procedure

The study begins with an introductory page that informs participants the task motivation and description. The first two images in the experiment are examples (not used in analysis), one that is clearly stretched, and another that is not. These examples provide a check that the participant understands the instructions. Following the examples, the 32 images are shown one by one. For each image, each participant was asked whether the image was stretched or not. At the end of the study, each participant was shown how well she/he did as well as the images for which she/he gave incorrect answers. On average, it took participants about five minutes to complete this study.

Results

Our study shows that people's perception of image stretch depends on a variety of factors, such as stretch scale, image content, and subjects. We performed a mixed-model analysis of variance (ANOVA) with stretch scale (0%, 20%, 40%, and 60%) and image category (human, animal, architecture, man-made object, ocean, mountain and plant) as within-participants factors and country of origin (U.S. and China) as a between-participants factor. The results are based on all participants' answers to the images shown to them. We found two significant main effects of stretch scale (F(3,315)=338.99, p<0.001) and image category (F(6,690)=85.60, p<0.001) and a marginal main effect of country of origin (F(1,85)= 3.09, p=0.08). We also identified a significant interaction between stretch scale and image category (F(18,1267)=1.62, p<0.001). No other main effects or interaction effects were found. Post-hoc tests were conducted for stretch scales over the whole image set and each image category, comparing 20%, 40%, 60% stretch scales to 0%

stretch scale, and revealed that the stretch scale had a different effect on different images. Full details of post-hoc tests can be seen in Table 1.
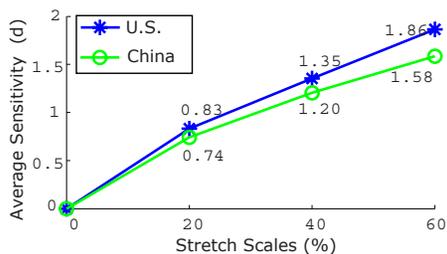
Stretch Scale

Our data show that people's sensitivity to image stretch increases with the stretch scale. We calculated people's sensitivity to image stretch based on Signal Detection Theory [4]. Sensitivity measures the discrepancy between the hit rate (H), the frequency that stretched images are correctly detected; and the false-alarm rate (F), the frequency that unstretched images are incorrectly judged as stretched ones. Perfect sensitivity means a hit rate of 1 and a false-alarm rate of 0. According to the Signal Detection Theory, sensitivity, d, is calculated in terms of z that is the inverse of the normal distribution function: $d = z(H) - z(F)$.

Figure 3 shows people's average sensitivity over 20%, 40%, and 60% stretch scales plus the origin (zero sensitivity at 0 stretch scale). There we can see that people's sensitivity to image stretch increases with the stretch scale, as supported by a chi-square analysis (U.S.: c^2 (1, N=3) = 70.9, $p < 0.001$; and China: c^2 (1, N=3) = 63.5, $p < 0.001$).

Image Content

We analyzed people's perceptual sensitivity to different image categories in this section. We found that people are more sensitive to stretch in images with humans, animals and man-made objects than those with natural scenes. For clarity, here we only report the common observations we had from this study over the subject pools in U.S. and China. We take the U.S. subject pool as an example here. The difference between these two pools is reported and analyzed in the following subsection.

Figure 3. People's average sensitivity over four stretch scales.
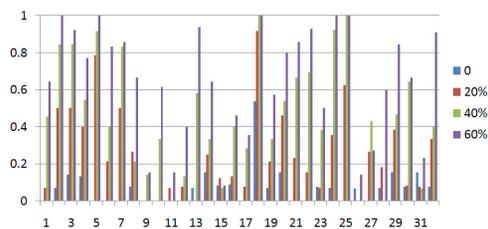
Figure 4. Perceived stretch rates of 32 images. Image number ranges from 1 to 32.

(a) Unstretched train (b) Stretched train

Figure 5. An example where Chinese participants is less sensitive to stretching than participants in U.S.. Photograph obtained from http://www.flickr.com/photos/eldelinux/3525659943/ and reproduced under Creative Commons License http://creativecommons.org/licenses/by/2.0/

Figure 4 shows each image's perceived stretch rate, the fraction of times that people thought it was stretched. Here we can see that it is easier to detect stretch in some images than in the others. We sorted the 32 images according to the sum of people's sensitivity to them over all stretch scales. We find that all of the 5 images that people are most sensitive to when stretched are images with humans, animals and man-made objects. And all of the 5 images that people are least sensitive to when stretched are images with natural scenes, like oceans and mountains.

To better understand people's perceptual sensitivity to different image categories, we show people's average sensitivity to the whole image set and 7 categories and the post-hoc tests following ANOVA results which show whether participants can detect 20%, 40%, and 60% stretches in Table 1. We sorted the 7 categories according to the sum of their sensitivities over four stretch scales and then arranged all categories in Table 1 by the sensitivity rank.

Viewer's Cultural Background
While our results do not show a significant difference in the overall sensitivity between the U.S. and China subject pools, we do find significant differences in some cases. For example, on images of man-made objects, the sensitivity is significantly different between pools. At the 20% scale, U.S. participants were more likely to notice the stretch (sensitivity: 0.88 vs. 0.07). An image that exemplifies the issue is shown in Figure 5. The stretched object appeared more consistent with Chinese participants' experience based on the post-test questionnaire comments.

Discussion and Future Work

In this paper, we presented an initial study on people's perception of image stretch. We find that people's perception of image stretch is a complex phenomenon. It depends on a variety of factors. We believe that even more factors can affect the perception of distortion. For example, the perspective shown in an image can affect how distortion is perceived. For example, in our study people frequently mistake an unstretched image as a stretched one; the false-alarm rate of it is 53.9% for U.S. participants and 41.2% for Chinese participants. We speculate that the cause is the unusual perspective foreshortening.

Our results have implications for designing image retargeting methods. For example, our findings suggest that using saliency maps as the importance measurement is insufficient because salience does not seem to be a predictor of sensitivity. Many of the images for which stretch is difficult to notice contain salient objects. Salient objects exist in images of all categories, yet sensitivity varies across the categories. Higher level information, such as content category, may serve as a useful importance metric. Our study also shows the challenges in evaluating image retargeting methods. As the perception of distortion depends on many factors, the acceptance of retargeting results will have similar dependencies. Appropriately diverse image sets, aspect ratio change amounts and participants seem essential to get reliable evaluation.

We plan to extend the current study as follows.

1. Our current study only provides data on a simple case of distortion. While horizontal stretch is an important special case, studying other types of

distortion is important to build a more general understanding. We will first combine vertical stretch with horizontal stretch, and then extend to spatially-varying distortions and seam carving.

2. We will design new studies to gain a deeper and more robust understanding of how the three factors found in this initial study affect people's perception to general distortions and the interactions among these factors. We also plan to seek thresholds for when image distortions are noticeable for various specific. To achieve these goals, we will incorporate more images to cover more image categories. This means we need to recruit more participants. We plan to deploy our study on Amazon Mechanical Turk which allows us to control the location of our participants, and provides potential for broad geographic diversity. Based on people's perception for different image categories (Table 1), we also plan to design different stretch scales for different image categories.

3. We believe there are other factors that affect people's perception of image distortion. Our future experiments may incorporate factors including perspective angle, time, and image/display size.

4. Understanding questions beyond noticeability will be important in developing a general understanding of the perception of image distortion. For example, in some cases it may be important to know what distortions are objectionable, and to examine distortions in context.

Acknowledgements

References

[1] Niu, Y.-Z. and Liu F. and Gleicher M. Multimedia retargeting bibliography. http://sites.google.com/site/mediaretargeting.

[2] Box, G. and Hunter, J. S. and Hunter.W. *Statistics for experimenters: design, innovation, and discovery*. Wiley-Interscience, 2005.

[3] Krages, B. P. *The Art of Composition*. Allworth Communications Inc., 2005.

[4] Macmillan, N. and Creelman, D. *Detection theory: a user's guide*. Cambridge University Press, 1991.

[5] Mojsilović , A. and Rogowitz, B. Semantic metric for image library exploration. *IEEE Trans. on Multimedia*, 6(6):828-838, 2004.

[6] Torralba, A. and Oliva, A. Statistics of natural image categories. *In Network: Computation in Neural Systems*, pages 391–412, 2003.

[7] Avidan, S. and Shamir, A. Seam carving for content-aware image resizing. *ACM Trans. Graph.*, 26(3):267–276, 2007.

[8] Wang, Y.-S and Tai, C.-L. and Sorkine, O. and Lee, T.-Y. Optimized scale-and-stretch for image resizing. *ACM Trans. Graph.*, 27(5), 2008.

[9] Wolf, L. and Guttmann, M. and Cohen-Or, D. Non-homogeneous content-driven video-retargeting. In *ICCV'07*, 2007.

[10] Santella, A. and Agrawala, M. and DeCarlo, D. and Salesin, D. and Cohen, M. Gaze-based interaction for semi-automatic photo cropping. In *CHI '06*.