

Adaptive Image Cropping Based Depth of Field

Yafeng Li^{1, a}, Ying Lin^{2, b}

¹Computer School of Communication University of China, China

²Library, Beijing Normal University, China

^aliyf@cuc.edu.cn, ^bliny@lib.bnu.edu.cn

Keywords: image retargeting, image cropping, depth of field.

Abstract. The communication and sharing of visual media show a strong cross-platform feature. Massive images are shown on various devices with different resolution. Automatic image cropping, usually used to change the resolution of image, often discards important content in the image. This paper proposed a novel adaptive photo cropping method. The main idea is exploiting the characteristics of photograph aesthetics in photograph works. The algorithm infers the intention of photographer according to the depth of field in the image. The effects brought out by the focus and unfocus are utilized to exact importance information of the image. So, the method can be more satisfied with the subjective evaluation. And, it has advantage in term of computational speed. Experimentations are presented to demonstrate the validity of proposed method.

Introduction

Benefit from the development of electronic technology, images and video are easy to obtained and communicated. Moreover, the communication and sharing of visual media show a strong cross-platform feature. Anywhere and anytime, we can access the image and video on a variety of device, such as TV, pc, PDA, table pc and mobile phone, etc. These terminals for visual information have different characteristic, especially in the resolution and aspect ratio. If the resolution of device and original media is different and the aspect ratio is the same, we can just scale the media to fit the resolution of device. But in the case of that the aspect ratio of media and device isn't consistent, it is more complicated. The scaling will introduce deformations and distortions to the original media because the aspect ratio is changed. To keep the aspect ratio, there are two main methods. One is playing the media with the original aspect ratio after zooming out, and the spare part in the screen of device is filled by black, where the physical size of screen is wasted. Another standard processing is cropping, where the original media is cropped to adapt the aspect ratio of screen. In this manner, parts of the information, outer parts in common, in original media will be lost. Sometimes, the important content is discarded to cause misunderstanding. So, they are not always pleasant to users because they do not take into account the quality of the cropped region. It can be seen, in order to adapt different display devices with different resolution, the resize of visual media is a valuable research field. In the past few years, many efforts have been put on this problem. As to image, the most promising solution is content-aware image resizing, or called image retargeting, whose main idea is taking into account the visual content of images during the resizing process.

Based this scheme, different techniques for content-aware image resizing have been proposed. When resizing image, different with the traditional methods, the image retargeting don't scale image evenly. Before scaling or cropping, it analyzes the significance of content in the image firstly. Then it crops the part that consists of not salient pixels. As to scale processing, this scheme divides the image into different parts according their significance, then, warps these parts separately. Finally, these warped parts are composed to a new image. Importantly, the parts that aren't salient in visual are preferential to be wrapped, while the aspect ratio of salient parts is unchanged as far as possible. So, the image distortion is minimized. The meaningful information in image is retained.

It can be seen, estimating the significance of content in the image is essential to image retargeting. Suh et al. [1] used a low-level saliency map to make their estimates. Their decision is based the assumption that the saliency is small in a uniformly textured region, while large in a region with a complex texture, edge and/or corner. Santella et al. [2] semi-automatically determined the region of attention by estimating the gaze of a user looking at each photo. Luo [3] found the important region in image by with the structural features, e.g., the centrality and shape, and the semantic features, e.g., the face or the sky. Besides determining the salient content in the image, the wrapping and the subsequent processing of image composition are two other important problems.

The seam carving technique[4] is a typical approach for content-aware image resizing. In this method, the importance map is set up on gradient domain. The essential processing is to decrease the image width (or height) one pixel at a time by removing a seam of minimal importance. Dynamic programming can be used to do this. Its processing result is more ideal when there are enough low-importance seams to be removed, but distortions and artifacts appears when seams cut through important areas. Wang [5] also exacted salient map in gradient domain. Different with[4], they divided the image in uniform firstly. Then the content in every grid is analyzed to exact importance value. In the following processing of warping, the salient content in the grid is kept to remain unchanged as far as possible. While the parts with lower salient value get more compression.

Cropping is a kind of basic way in the field of photo editing, which is wildly used for removing the unwanted subjects in a photo to change its aspect ratio or to improve its overall composition. This work is often manual operation by the users according to their intention. But, for a large amount of photos, this iterative user operation is too time-consuming to be feasible. In the case, automatic photo cropping is very necessary.

Gianluigi [6]proposed a novel automatic photo cropping method with a classifier, which classified the original photos into two types of close-up and landscape. Then different type is processed with different way. The key point of this scheme is classifier, whose robustness seriously affects the results.

Despite the technological advances in the context of content-aware, image resizing algorithms are still far to be satisfied to the users because of its weakness in taking into consideration the agreeability that the cropped regions.

In fact, we know, form the knowledge of photography aesthetics, the photographer will make his decision on the theme before shooting up. And, he presents his intention in the photo using the characteristic of human visual system. For example, the subject can be described by means of brightness contrast and color variety. Among these skills, the depth of field plays an important role. The photographer applies the effect brought out by the unfocused and focused of objects to show the main object in the scene. Based this phenomena, the proposed paper develops a kind of adaptive and automatic cropping method for resizing image.

Cropping Based the Depth of Filed

Unfocused and Point Spread Function. General optical instruments such as camera can be described as the model of lens. Given a object point p and a lens with aperture A , the point p get a image p' in the focus plane through the lens. The object distance u , image distance v and focal length f yield the formula $1/f = 1/u + v/l$. If the image sensor plane is in the distance s from the lens, the object is not in focus. The image of p is a defocused called blur circle. The diameter of the blur circle is a function of the depth and parameter of lens:

$$d(x, y) = sA\left(\frac{1}{f} - \frac{1}{u(x, y)} - \frac{1}{s}\right) \quad (\text{Eq. 1})$$

It can be seen form the Fig 1, for a certain lens system, the size of the blur circle only depends on the distance of the object to the lens. Taking into account the diffraction and other non-ideal situation, the two-dimensional Gauss function is used to describe the unfocus approximately.

$$h(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (\text{Eq. 2})$$

Where, k is constant.

This formula is called Point Spread Function(PSF) or blur function. It is a basic parameter for evaluating the imaging quality of optical instruments. Usually, the blur function is only calibrated by the blur diameter, and the blur diameter is determined by the distance of object to lens and parameter of lens [7].

According to the description of optical knowledge, when the lens in a camera focuses on a object at a distance, all subjects in the plane at the same distance and being perpendicular to the lens axis of the lens center position are sharply focused, and will get a clear image on the image sensor or film. Subjects that are not at the same distance are out of focus and theoretically are not sharp. The point gets a image of a blur circle. We call it unfocused. However, since human eyes cannot distinguish very small degree of unsharpness, some subjects that are in front of and behind the sharply focused subjects can still appear sharp. The zone of acceptable sharpness is referred to as the depth of field. Namely, depth of field (DOF) is the distance between the nearest and farthest objects in a scene that appear acceptably sharp in an image. Because the blur degree in objects varies with the distance, we can utilize it to matting objects or computing distance [8].

Because of the visual effects brought about by DOF, the photographers often utilize it to highlight the theme, particularly as it relates to close-up and portrait. Obviously, the photographers have shown the different importance of the objects by DOF. The focused object is the subject of the photograph works, while the importance of unfocused objects is relatively lower. To some extent, this can be seen as a kind of salient map emphasized in the most image retargeting technologies.

Proposed Algorithm. General optical instruments such as camera can be described as the model of lens. Given a object point p and a lens with aperture A , the point p get a image p' in the focus plane through the lens. The object distance u , image distance v and focal length yield the formula $1/f = 1/u + v/1$. If the image sensor plane is in the distance s from the lens, the object is not in focus. The image of p is a defocused called blur circle. The diameter of the blur circle is a function of the depth and parameter of lens:

The adaptive image cropping can be described as: transforming an image with resolution of $W \times H$, to a new image of $W' \times H'$. The resolution of new image is often lower than the old one's. The restriction is that new image is similar with the old one from the view of user. It means the important objects in the photo must be retained or be distorted slightly. For cropping operation, its aim is to find image block $I_{x,y}, W', H'$ in the original image, where W', H' is the width and height of the cropped image, and (x, y) is the upper-right corner coordinates of the rectangular area in the original image. We hope that the main objects are included in the image as much as possible.

For the adaptive image cropping, the essential procedure is to determine the importance of each part of the contents in the image. Most researches exact a data structure named salient map from the image, considering about the information of color, lightness, skin color and silhouette etc. Based the above analysis on the DOF, the blurring brought out by the DOF can be used to determine the importance of image content. The measure of blurring of a image is a basic problem in the field of Computer Vision. Various methods in space domain and frequency domain have been proposed. Among them, the approach of measure by the edge sharpness is simple and efficient, but is weak in robustness. Taking into account the characteristics of the DOF, we proposed an improved algorithm. From the Eq.2, we can see that the impact of the point spread function will be in attenuation with the increasing of distance from the center point increase. For a block of size $(2B+1) \times (2B+1)$ in the image, the value of the centre pixel (x, y) is denoted as $g(i, j)$. Then, the difference of the centre pixel and other pixels in the block is calculated as,

$$d(k, l) = g(i, j) - g(i+k, i+l) \quad k, l \in \{-B, \dots, B\} \quad (\text{Eq. 3})$$

Finally, the weighted sums of these differences are got. Considering the characteristics of the point spread function, Gaussian function can be used as the weights:

$$s(i, j) = \sum [d(k, l) * \frac{1}{2\pi} e^{-\frac{(i-k)^2 + (j-l)^2}{2}}] \quad (\text{Eq. 4})$$

The higher value of a pixel reveals that the adjacent filed gets better focus. So, these values indicate the importance of each pixel in the image.

After obtaining the importance value, the next step is to find the focus centre. This can be treated as a problem of center-of-gravity position of object. Defining the resolution of image as $M \times N$, the coordination of focus centre is:

$$(\sum_{i=1, j=1}^{i=M, j=N} (i, j) * (s(i, j)) / \sum_{k=1, l=1}^{k=M, l=N} s(k, l)) \quad (\text{Eq. 5})$$

This is just the coordination of cropping operation.

Experimental Results

To verify the effectiveness of the proposed method, we used a data set composed of 300 images. These images were downloaded from Internet, including the type of portrait, close-up and landscape. The Fig. 1 shows a part of the experimental results. The algorithm can be evaluated subjectively.

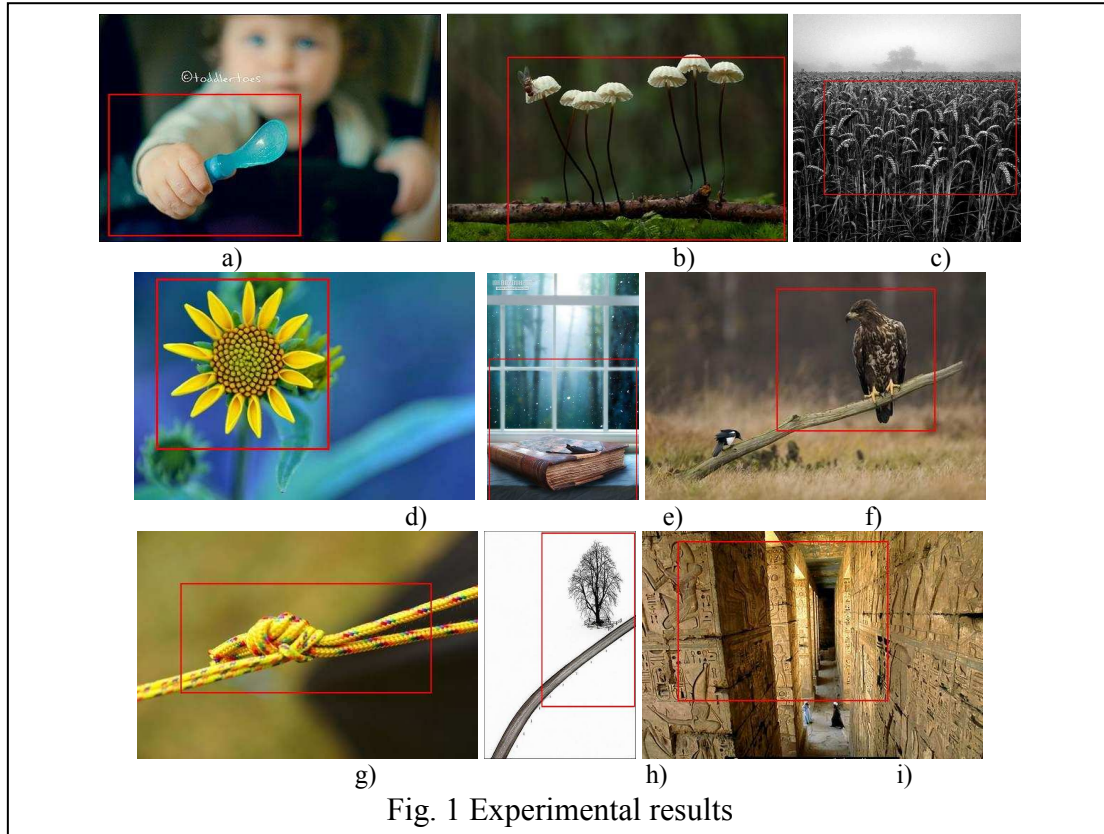


Fig. 1 Experimental results

It can be seen from the results that this method satisfy user more for the images with clear focus. As to the images that have obviously difference between subject and background, especially the background with single texture, it also works well. While the photograph works are taken at a small aperture without apparent difference in focus and unfocus zone, a mistake often happens. This weakness can be overcome by image classifier adopted by the paper[6]. After being classified, different type of image can be treated with appropriate scheme.

Summary

For the adaptive cropping in the field of image retargeting, utilizing the knowledge of photograph aesthetics, a simple and efficient scheme is proposed. The proposed cropping method can reflect the intension of the user. Besides, it has advantage in term of computational speed. So, it is particularly suitable for real-time massive image processing. And, this method is more practical for the handheld devices such as mobile phones which are relatively weak in computing performance. But the proposed method is restricted in the situations of application. It can get more satisfied results for the photos of close-up and portrait, but is often lost in the photos such as landscape and architecture. As part of our future work, we intend to increase satisfaction of the proposed method by its combining with other image retargeting methods.

References

- [1] B. Suh, H. Ling, B. Bederson, and D. Jacobs, Automatic thumbnail cropping and its effectiveness, in Proc. ACM UIST, 95 – 104, 2003.
- [2] A. Santella, M. Agrawala, D. DeCarlo, D. Salesin, and M. Cohen, Gaze-based interaction for semi-automatic photo cropping, in Proc. SIGCHI Conf . Human Factors in Computing Systems, 771 – 780, 2006.
- [3] J. Luo, Subject content-based intelligent cropping of digital photos. in Proc. ICME, 2218 – 2221, 2007.
- [4] S. Avidan and A. Shamir, Seam carving for content-aware image resizing, ACM Trans. Graph, 26(3), 2007.
- [5] Y.S. Wang, C.L.Tai, O.sorkine,T.Y.Lee, Optimized scale-and-stretch for image resizing, ACM Trans.Graph,27,1-8, 2008.
- [6] Gianluigi Ciocca, Self-Adaptive Image Cropping for Small Displays, IEEE Tran. on Consumer Electronics, 53(4) 1622-1627.
- [7] A Pentland, A.Paul, A new sense for depth of field , IEEE trans. PAMI, 9(4) 523-531,1987.
- [8] Y. Luo, X.O. Tang, Photo and Video Quality Evaluation: Focusing on the subject, ECCV '08 Proceedings of the 10th European Conference on Computer Vision, 386-399, Springer-Verlag Berlin, Heidelberg ,2008.