Visual Saliency-based Theme and Aspect Ratio Preserving Image Cropping for Small Displays

Rajarshi Pal, Pabitra Mitra, Jayanta Mukherjee
Department of Computer Science & Engineering
Indian Institute of Technology Kharagpur
Kharagpur – 721302, India
Email: (rajarshi, pabitra, jay)@cse.iitkgp.ernet.in

***Abstract*—Scaling to fit an entire image in much smaller display reduces the recognizability of the objects in the image. Therefore, salient portions of an image is cropped and then scaled to fit to a small display. To do so existing methodologies do not maintain the overall theme of the image and incorporate distortion due to difference in aspect ratios of the cropped portion of the image with that of the display. Moreover, to be executed upon resource constrained hand-held devices the method must be as simple as possible. In this paper, a method for image cropping is proposed that takes care of all these matters.**

I. Introduction

Small display size of handheld devices poses a challenge for displaying images on these devices. Images can be scaled to fit the small display but it decreases the recognizability of the objects in the image. Many attempts have been made so far to deal with this problem. Most approaches attempt to identify important regions in the image and construct a smaller image of the important region.

Most of them have used the Itti-Koch [1] visual attention model to identify the important regions. Here importance is measured with respect to low level attributes (such as intensity, color, orientation etc.). Some of the approaches have also used face detection techniques to introduce the concept of semantic importance. In [2], images are classified into three classes, i.e., landscape images, close-up images and others, and processed based on the result of the classifier. Landscape images are not being cropped. For other images the Itti-Koch visual attention model is being applied to detect the important regions. For the images those are classified into 'others' category, face detector is being applied along with the visual attention model.

Next step is cropping and displaying the important regions. In [2, 3] image is cropped in such a way that cropped portion contains all the important regions in the image. Then the cropped portion is resized to fit to target display. In [4], sizes of the important objects are exaggerated to improve their recognizability. But this exaggeration comes at the expense of realism. The relative sizes among the important and unimportant objects are not maintained and the relative distances among various objects are also distorted. In [5], several small areas that can be fit to display size are cropped surrounding each important region and displayed one by one. Here, time is compromised for space.

Moreover, aspect ratio of the cropped portion of the image must be same as the aspect ratio of the display. Otherwise, the cropped image will distort from original image while displaying. This is illustrated in figure 1. None of these approaches cared to maintain the aspect ratio of the cropped portion of the image while displaying. In figure 1, the original image is shown at the left. The cropped portion selected from the original image is shown by the white-bordered rectangular area. Now when it is displayed on an almost-square shaped display it will take the shape as shown in the right of the figure. Here, change in slope of the hills in two images is to be noticed. Beside these, to be executed upon resource-constrained hand-held devices the process of image cropping and adapting must be a simple one (unlike [2]). That means, it should consume less power and less memory.

In this paper, we propose a simple approach to crop an image covering the important portions for displaying it in a small display. The objective of our approach is not to distort the image and maintain the recognizability of the image. The specialities of our method are that we have maintained the overall theme of the original image in the cropped version and the aspect ratio of the cropped image is maintained to be as same as that of the display. In the next section (section II), a brief introduction of Itti-Koch selective visual attention model is given. In section III, we propose a methodology of maintaining the overall theme of the image while cropping. In section IV, we depict our approach to maintain the aspect ratio of the cropped image. Then in section V, we mention about the result obtained from our approach and conclude in section VI.

Fig. 1. Distortion due to mismatch of aspect ratio

II. Selective Visual Attention Model

Itti-Koch selective visual attention model [1] helps to select regions of an image that contains salient objects. Salient points are identified based on following low-level features: intensity, red/green and blue/yellow double opponent channels, and orientation at four directions (0°, 45°, 90° and 135°). For an input image these features are estimated at different scales. For each feature center-surround differences are obtained. Center-surround differences are implemented as a difference between the feature values at two scales. Then for each feature, across-scale combination of center-surround difference maps is done and finally across-feature combination of those maps is performed to get one saliency map.

III. Maintaining The Overall Theme Of The Image

Itti-Koch visual attention model [1] gives pixel wise measure of saliency. But one has to choose the set of values in the saliency map that correspond to the salient regions to be considered for cropping. After analyzing all the values in the saliency map one chooses a particular threshold value. Those pixels, for which the saliency value is greater than or equal to that chosen threshold, belong to the salient portions of the image. So a rectangular (or square) shaped image is cropped in such a way so that the rectangle (or square) contains all the salient portions. But for some cases, if we follow this general procedure, the meaning or theme of the cropped image changes significantly from the original image. This is shown in figure 2. In the left side of figure 2, the original image is given. The flowers and greeneries of a garden are shown in it. In the right side of figure 2, the cropped region is shown. It is a picture of some particular bunch of flowers at a particular location in the garden. Obviously, subjects of these two images are different.

To maintain the overall subject of the original image in the cropped one, we propose the following approach. The saliency map will be partitioned into four quadrants as figure 3. Then for each individual quadrant a threshold is selected considering the values only from that quadrant of the saliency map. Thus the selected threshold of a particular quadrant is independent of the saliency values from other quadrants. Now the pixels of a quadrant belong to the salient portions if the saliency value for that particular pixel is greater than or equal to the chosen threshold for that quadrant. The threshold is selected so that only 2% of the pixels of a particular quadrant belong to salient portion. The fact that 2% of the pixels of a particular quadrant will belong to the salient portion is empirically determined. Now we crop the image with the minimum rectangle that contains all the identified salient portions from four quadrants.

Fig. 2. Theme of the image has changed

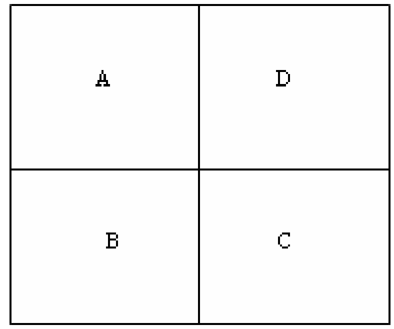


Fig. 3. Partitions of the saliency map

To explain the reason behind partitioning the saliency map into four quadrants we can think of two extreme cases. At the one extreme, the saliency map will be treated as an unpartitioned one. The disadvantage of that case is stated above (figure 2). At the other extreme, the saliency map will be partitioned into as many partitions as possible, i.e., one partition corresponding to one pixel in the image. Then selecting salient points from each partition means selecting each point in the map. This corresponds to selecting the entire image for cropping. That means the image is not being cropped at all. We have to keep number of partitions as less as possible but to make sure that the overall theme of the image will be maintained.

IV. Maintaining The Aspect Ratio Of The Cropped Portion

To avoid distortion of the image the aspect ratio of the cropped portion of the image must be preserved while displaying. But while cropping with a minimum rectangle which contains all the identified important regions, the aspect ratio of the cropped rectangular area, in most cases, does not match the aspect ratio of the display. Because the approach described in previous section does not keep the aspect ratio of display in head.

Let h and w be the height and width of the display and h′ and w′ be the height and width of the rectangular cropped portion from the original image. It is important to mention that heights (h and h′) and widths (w and w′) are measured in terms of number of pixels. Therefore, aspect ratio of the display (a) is w/h and aspect ratio of the cropped portion (a′) is a w′/h′. If a′ equals a then no adjustment regarding aspect ratio is needed. If a′ is less than a then a′ has to be increased.

To increase a', w' has to be increased. Decreasing the value of h' also increases a', but we cannot decrease h' because decreasing h' means some important regions will be omitted. So we have no other option but to increase w', i.e., width of the rectangular cropped region, by adding one column at a time. But the question is in which direction we will increase the area, i.e., left or right. The choice of direction will depend upon the summations of the saliency values corresponding to immediate left column and immediate right column of the rectangle. In whichever direction sum is greater, the column in that direction will be added. This will continue until either the aspect ratio of the region equals to aspect ratio of the display or we have added the entire width of the original image.

On the other hand if a' is greater than a then a' has to be decreased. To decrease a' the value of w' cannot be decreased because decreasing w' means some important regions will be omitted. So in this regard the only option is to increase h', i.e., height of the rectangular cropped region, by adding one row at a time. Again the question is in which direction we will increase the region, i.e., up or down. The choice of direction will depend upon the summations of the saliency values corresponding to immediate top row and immediate bottom row of the rectangle. In whichever direction sum is greater, the row in that direction will be added. Similar to the case stated in above paragraph, this will continue until either the aspect ratio of the region equals to aspect ratio of the display or we have added the entire height of the original image.

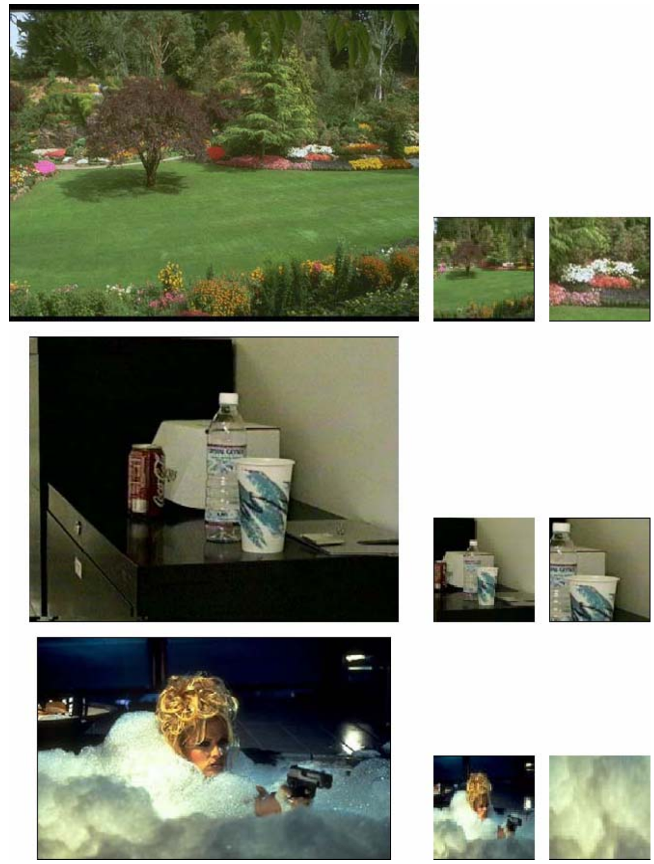
V. Experimental Results

The output of the proposed algorithm was tested by 14 volunteers. Each of them was shown 285 images. Those images were of various classes, e.g., personal indoor or outdoor images, news pictures and scenery photos. Volunteers were asked two questions.

1. Do the cropped images maintain the theme of the original images?
2. Have the objects been distorted in the cropped version while displaying in small display?

According to response from the volunteers, of the 285 test cases 276 maintain the theme of the original images in cropped versions and in all of these 276 cases objects shown are not distorted. This means our approach gives 96.84% correct result. Some of the results are shown in figure 4 (b). Original images are shown in figure 4 (a). Figure 5 shows some more results according to our proposed approach.

For the same set of images we have also conducted an experiment where a single threshold is selected over the entire saliency map and image is cropped based on that. Some of the results are shown in figure 4 (c). Volunteers are asked afore mentioned two questions. According to them, of the 285 test cases 160 maintain the theme of the original images and in 144 of these 160 cases images are not distorted. This means this approach gives correct result in 50.53% cases where as our proposed approach gives correct results for 96.84% cases.



(a) (b) (c)

Fig. 4. Results

VI. Conclusion

In this paper, an approach is proposed to crop an image for displaying it in small displays. Itti-Koch model of selective visual attention is applied to identify salient portions of the image. But unlike existing approaches effort has been made to maintain the overall theme of the image and to make the aspect ratio of the cropped portion as same as the aspect ratio of the display where it will be displayed so that the image is not distorted.

Another important point is that we have not used any image classifier to categorize the image. All the images are treated with same manner. We have not used face detector to detect the presence of human faces in the image. In spite of these, we have got somewhat good results. It indicates that in resource-constrained mobile devices this proposed method is more suitable than any other method that needs image classifier or face detector along with Itti-Koch selective visual attention model.

Fig. 5. More results

REFERENCES

[1] L. Itti, C. Koch, E. Niebur, “A model of saliency based visual attention for rapid scene analysis”, IEEE Trans. on PAMI , Vol. 20, Issue 11, pp. 1254-1259, Nov. 1998.

[2] G. Ciocca, C. Cusano, F. Gasparini, R. Schettini, “Self-adaptive image cropping for small displays”, Proc. of Inter. Conf. on Consumer Electronics ICCE ’07, pp. 1-2, Jan. 2007.

[3] B. Suh, H. Ling, B. B. Bederson, D. W. Jacobs, “Automatic thumbnail cropping and its effectiveness”, 16th Annual ACM Symposium on User Interface Software and Technology UIST ’03, pp. 95-104 Nov. 2003.

[4] V. Setlur, S. Takagi, R. Raskar, M. Gleicher, B. Gooch, “Automatic image retargeting”, 4th Inter. Conf. on Mobile and Ubiquitous Multimedia MUM ’05, pp. 59-68, Dec. 2005.

[5] X. Fan, X. Xie, W. Y. Ma, H. J. Zhang, H. Q. Zhou, “Visual attention based image browsing on mobile devices”, Inter. Conf. on Multimedia and Expo ICME ’03, vol. 1, pp. 53-56, July 2003.