Enhanced Depth Features of Human Bodies for Image Retargeting

Jen-Jung Cheng and Chin-Chen Chang
Department of Computer Science and Information Engineering, National United University, Miaoli 360, Taiwan
Email: {M0124004, ccchang}@nuu.edu.tw

Huei-Yung Lin
Department of Electrical Engineering and Advanced Institute of Manufacturing with High-tech Innovations, National Chung Cheng University, Chiayi 621, Taiwan
Email: hylin@ccu.edu.tw

***Abstract*—In this paper, we propose enhanced depth features of human bodies for image retargeting. First, the depth map of an input image is generated. Second, with the depth map we divide the depths into several depth layers to separate human bodies and the background. Next, the proposed approach performs the morphological dilation operation on the human bodies and enhances the edges of the human bodies to generate an enhanced depth map. Moreover, we assign different weights for the human bodies and their edges according to their importance. Finally, we generate the target image based on the enhanced depth map. The proposed approach can obtain pleasing resized images for a wide range of images.**

***Index Terms*—image retargeting, feature, depth map, seam carving**

I. INTRODUCTION

Numerous devices for displaying digital images with different aspect ratios and resolutions have been developed. To meet various demands, changing display content becomes crucial. Image or video retargeting techniques have been studied extensively. Common methods for image resizing include scaling and cropping. However, these two approaches often fail to identify and protect important regions.

Several content-based image resizing approaches have been proposed [1]-[12]. Resizing images based on image content is an approach that removes insignificant regions and maintains the important regions of original images. Seam carving [1] is an effective image resizing approach based on image content. However, it cannot protect important objects in images if the energy of the main objects is low as regards to their surroundings. Also, it often fails to recognize edges and thus removes borders of the main objects for some images.

In this paper, we present an image retargeting approach based on enhanced depth features of human bodies. Fig. 1 depicts a flowchart of the proposed approach. We first obtain the depth map of an input image. After that, the depths are divided into several depth layers based on the depth map to separate human bodies and the background. Then, we construct an enhanced depth map by performing the morphological dilation operation on the human bodies to enhance the borders of the human bodies. Also, different weights for the human bodies and their edges are assigned according to their importance. Finally, the proposed approach constructs the target image based on the enhanced depth map.

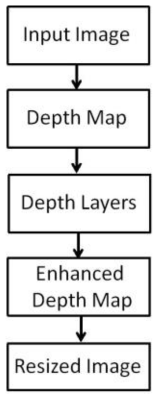


Figure 1. The flowchart of the proposed approach.

The rest of this paper is organized as follows. In Section 2, we briefly discuss related works. Section 3 is the proposed approach. Section 4 presents results. Finally, Section 5 is conclusions.

II. RELATED WORKS

Avidan and Shamir [1] presented an approach for resizing images based on image content. They proposed a simple technique for image processing using seams. Seams are lines with eight connecting pixels that vertically or horizontally cross images. By iteratively adding or removing seams, it can change the aspect ratio of an image. However, because the content of images is often complex, how to determine the correct positions of subjects according to image characteristics is a goal for further studies.

Kim *et al.* [8] used the importance map of the image to calculate the adaptive scaling function. This function

Manuscript received March 10, 2015; revised June 17, 2015.

doi: 10.18178/joig.3.1.52-55

means the reduction level for each row of the original image. Kim *et al.* [9] used Fourier analysis for image resizing. They divided the image into bars to determine the spectrum of each bar. They used the spectrums as a low-pass filter to obtain a smoothing effect. The horizontal reduction level for each bar was then determined based on the influence of the filter.

Visual significant area detection is a part of object detection. The traditional approach for determining the salient objects in an image is to set many parameters and then use the training approach to determine objects [2]-[10]. However, the human eye can quickly locate common objects [13]-[15]. Several approaches have been proposed to simulate the functions of the human eye such as Saliency ToolBox [14] and Saliency Residual (SR) [5]. The Saliency ToolBox needs a lot of computation. By comparison, SR is faster. It transforms the image into Fourier space and then determines the difference between the log spectrum and averaged spectrum of the image. The region showing the difference is the possible region of visual saliency.

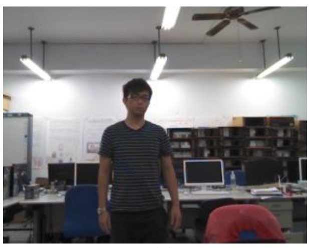
Hwang and Chien [6] determined the main objects of images using a neural network technique. Also, they used face recognition techniques to protect the completeness of human faces. They used proportional ratio techniques to resize images for ratios that could not be resized using the seam carving algorithm. Rubinstein *et al.* [12] proposed a method for improving the seam carving. This method utilized techniques of forward energy and backward energy to reduce discontinuity in images for image resizing.

Mishiba and Yoshitome [11] proposed an approach based on a warping technique for image retargeting. Their approach finds an optimal transformation by solving an energy minimization problem. It can maintain the relative arrangement and protect important regions.

III. The Proposed Approach

A. Enhanced Depth Map

In our approach, the Kinect [16] camera is used to obtain a color image and its depth information. The camera uses a 3D scanner system using near-infrared light to determine the depth and illuminate the objects of the image. Fig. 2 shows (a) an input image and (b) the depth map.


(a)


(b)

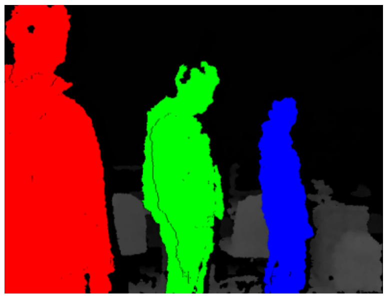
Figure 2. (a) An input image and (b) the depth map.

After obtaining the depth map, we divide the depths into several depth layers to separate of human bodies and the background. We focus the depths of the human bodies and strengthen the importance of the human bodies and their borders using the morphological dilation operation. Hence, we can enhance the edges of human bodies. Thus, the edges of the human bodies can be effectively reserved. Moreover, we set different weights for the human bodies and their edges based on their layers. The enhanced depth map is $E_{EnhancedDepthMap}$ defined by

$$\begin{aligned} &E_{EnhancedDepthMap}(x,y) \\ &= \sum_{i=1}^{n} HBE_i(x,y) * wHBE_i + \\ &\sum_{i=1}^{n} HB_i(x,y) * wHB_i + \\ &Background(x,y) * wB \end{aligned} \quad (1)$$

where n is the number of human bodies, $HBE_i(x,y)$ is the edges of human body i, $HB_i(x,y)$ is human body i, and $Background(x,y)$ is the background. Moreover, $wHBE_i$, wHB_i and wB are the corresponding weights. In our approach, the enhanced depth map is used as the importance map for seam carving to resize images. Fig. 3 shows (a) an input image, (b) the depth map with several depth layers, and (c) the enhanced depth map.


(a)


(b)

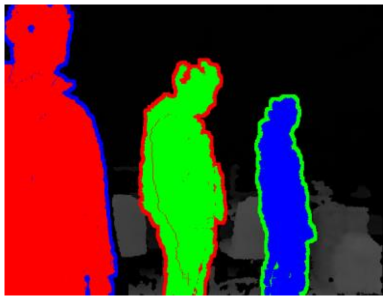

(c)

Figure 3. (a) An input image, (b) the depth map with several depth layers, and (c) the enhanced depth map.

B. *Image Retargeting*

We apply the seam carving [1] for image retargeting. Let I be an $n \times m$ image. A vertical seam is an 8-connected line. Each row only contains a single pixel. A vertical seam is defined by

$$\begin{aligned} & s^x = \{s_i^x\}_{i=1}^n = \{(x(i), i)\}_{i=1}^n, \\ & \text{s.t. } \forall i, \ |x(i) - x(i-1)| \le 1 \end{aligned} \tag{2}$$

where x is a mapping $x : [1, \ldots, n] \to [1, \ldots, m]$.

The path of a vertical seam is given by

$$I_s = \{I(S_i)\}_{i=1}^n = \{I(x(i), i)\}_{i=1}^n \tag{3}$$

All pixels will move leftward or upward to fill the gaps of deleted pixels.

Horizontal reduction can be equated with deleting the vertical seam; the enhance depth map is used to select seams. Given an energy function e, the energy of a seam is determined by the energy occupied by the positions of all pixels as

$$E(s) = E(I_s) = \sum_{i=1}^{n} e(I(s_i)) \tag{4}$$

When cutting an image horizontally, deleting the seam with the lowest energy by

$$s^* = \min_s E(s) = \min_s \sum_{i=1}^{n} e(I(s_i)) \tag{5}$$

Dynamic programming is employed to calculate s^*. The smallest accumulated energy M is calculated with each possible point on the seam (i, j) from the second to the last row of the image as

$$\begin{aligned} & M(i, j) = e(i, j) + \\ & \min(M(i-1, j-1), M(i, j-1), M(i+1, j-1)) \end{aligned} \tag{6}$$

Finally, the backtracking technique is adopted iteratively to delete the seams by gradually searching upward for the seams with a minimum energy sum from the pixel with the lowest energy in the last row.

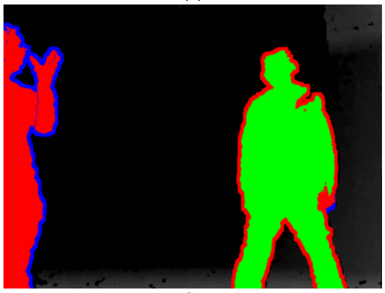
IV. RESULTS

Our approach was implemented and applied to images to evaluate the performance for adjusting the size of an image while preserving its visual quality. The platform is a PC with an Intel Core2 i5-3750 CPU @ 3.40GHz and 4.0GB of memory and the Windows 7. Our approach was compared with the seam carving [1]. The size of the input image is 580×440 and that of the resized images is 348×440.

Fig. 4 shows (a) an input image, (b) the enhanced depth map, (c) the resized image by the seam carving, and (d) the resized image by our approach. From the results, our approach can preserve objects more satisfactorily. Furthermore, with the advantage of the enhanced edges, the distortion is minimized. Some distortion occurred in the results of the seam carving method.

Fig. 5 shows (a) an input image, (b) the enhanced depth map, (c) the resized image by the seam carving, and (d) the resized image by our approach. The results show that our approach can obtain smooth resizing results. The previous method produces some discontinuities, particularly in areas containing the objects, such as the boxes and the black chair. Therefore, the proposed approach can improve the results of seam carving.

(a)



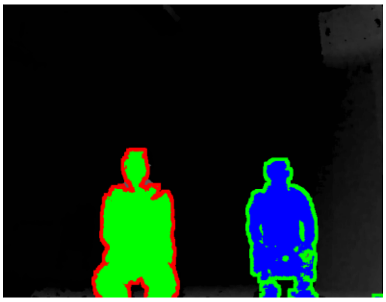
(b)

(c)

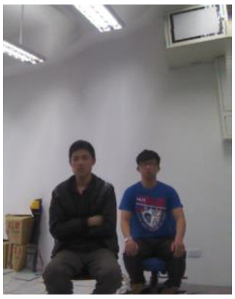
(d)

Figure 4. (a) An input image, (b) the enhanced depth map, (c) the resized image by the seam carving, and (d) the resized image by our approach.

(a)



(b)



(c)

(d)

Figure 5. (a) An input image, (b) the enhanced depth map, (c) the resized image by the seam carving, and (d) the resized image by our approach.

V. Conclusions

We have proposed an image retargeting approach based on enhanced depth features of human bodies. According to the depth map, human bodies and the background are separated and different importance is assigned for the human bodies and their edges. The proposed approach can produce target images with fewer distortions. In the future research, further studies are needed to develop a technique for video retargeting.

Acknowledgment

The authors would like to thank the Ministry of Science and Technology of Taiwan for financially supporting this research under Contract No. MOST 103-2221-E-239-030-.

References

[1] S. Avidan and A. Shamir, "Seam carving for content-aware image resizing," *ACM Transactions on Graphics*, vol. 26, no. 3, pp. 10, 2007.

[2] R. Fergus, P. Perona, and A. Zisserman, "Object class recognition by unsupervised scale-invariant learning," in *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2003, vol. 2, pp. II-264-II-271.

[3] D. Gao and N. Vasconcelos, "Integrated learning of saliency, complex features, and object detectors from cluttered scenes," in *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2005, vol. 2, pp. 282-287.

[4] M. Gao, Z. He, Z. Li, G. Ma, X. Ye, and H. Yu, "A new improved seam carving content aware image resizing method," in *Proc. IEEE Conference on Industrial Electronics and Applications (ICIEA)*, 2013, pp. 738-741.

[5] X. Hou and L. Zhang, "Saliency detection: A spectral residual approach," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, 2007, pp. 1-8.

[6] D. S. Hwang and S. Y. Chien, "Content-Aware image resizing using perceptual seam carving with human attention model," in *Proc. 2008 IEEE International Conference on Multimedia and Expo*, 2008, pp. 1029-1032.

[7] L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, pp. 1254-1259, 1998.

[8] J. H. Kim, J. S. Kim, and C. S. Kim, "Image and video retargeting using adaptive scaling function," in *Proc. 17th European Signal Processing Conference (EUSIPCO 2009)*, 2009.

[9] J. S. Kim, J. H. Kim, and C. S. Kim, "Adaptive image and video retargeting technique based on Fourier analysis," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, 2009, pp. 1730-1737.

[10] T. Liu, Z. Yuan, J. Sun, J. Wang, N. Zheng, X. Tang, and H. Shum, "Learning to detect a salient object," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 2, pp. 353-367, 2010.

[11] K. Mishiba and T. Yoshitome, "Image retargeting with protection of object arrangement," in *Proc. 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2014, pp. 5814-5818.

[12] M. Rubinstein, A. Shamir, and S. Avidan, "Improved seam carving for video retargeting," *ACM Transactions on Graphics*, vol. 27, no. 3, pp. 1-9, 2008.

[13] A. Treisman and G. Gelade, "A feature-integration theory of attention," *Cognitive Psychology*, vol. 12, pp. 97-136, 1980.

[14] D. Walther and C. Koch, "Modeling attention to salient proto-objects," *Neural Networks*, vol. 19, pp. 1395-1407, 2006.

[15] D. Wang, A. Kristjansson, and K. Nakayama, "Efficient visual search without top-down or bottom-up guidance," *Perception & Psychophysics*, vol. 67, pp. 239-253, 2005.

[16] Kinect. [Online]. Available: http://www.microsoft.com/en-us/kinectforwindows/

Jen-Jung Cheng received his M.S. degree in Computer Science and Information Engineering from National United University in Miaoli, Taiwan. His research interests include image processing, computer graphics, and interactive multimedia.

Chin-Chen Chang received his Ph.D. degree in Computer Science from National Chiao Tung University in Hsinchu, Taiwan, in 1998. At present, he is a professor in the Department of Computer Science and Information Engineering, National United University in Miaoli, Taiwan. His research interests include computer graphics, image processing, and digital multimedia.

Huei-Yung Lin received his BS degree in Applied Mathematics from National Chiao Tung University, Taiwan, and his MS and PhD degrees in electrical and computer engineering from the State University of New York at Stony Brook. In 2002 he joined the Department of Electrical Engineering, National Chung Cheng University, Taiwan, and currently is a full professor. His current research interests include computer vision, robotics, pattern recognition, and image processing. He serves as an organizing committee member of several robotics related IEEE international conference. He is also a founding member and deputy secretary general of Taiwan Society of Robotics.

