2012 International Conference on Software and Computer Applications (ICSCA 2012)
IPCSIT vol. 41 (2012) © (2012) IACSIT Press, Singapore

Cooperative Resizing Technique for Stereo Image Pairs

Huei-Yung Lin [1], Chin-Chen Chang [2 +] and Chia-Hao Hsieh [1]

[1] Department of Electrical Engineering
National Chung Cheng University, Minhsiung, Chiayi 621, Taiwan
[2] Department of Computer Science and Information Engineering
National United University, Miaoli 360, Taiwan

Abstract. In this paper, a cooperative stereo image resizing technique is presented. The proposed approach first extracts three feature maps: disparity map, gradient map and saliency map. After that, the proposed approach constructs an importance map which combines the three feature maps by the weighted sum. Finally, the proposed approach constructs the target image using the seam carving method based on the importance map. The experimental results show that the proposed approach performs well in terms of the resized quality.

Keywords: Image resizing, Image retargeting, Stereo image, Feature map, Seam carving

1. Introduction

Non-standard screen aspect ratios will be applied more extensively because of cellular phones, portable multimedia players and so on. In such cases, different image sizes are required to adapt to the display devices. Scaling and cropping are two standard methods for resizing images. Scaling resizes the image uniformly over an entire image. However, when the display screen is too small, the image loses some of its detail in adjusting to the limitations of the display screen. Cropping resizes the image by discarding boundary regions and preserving important regions. This method provides a close up of a particular image section, but prevents users from viewing the rest of the image.

Recently, several retargeting techniques [1, 5, 6, 7] for resizing image based on image contents has been proposed. These methods require a certain understanding of image content and do not adjust the size of the image as a whole. Retargeting preserves important regions and discards less important regions, to achieve a target image size. Since the creation of stereo images for a 3D display from the 2D images is important, developing techniques for stereo image retargeting is essential.

In this paper, a stereo image resizing technique is proposed. The proposed approach first extracts three feature maps: disparity map, gradient map and saliency map. After that, the proposed approach constructs an importance map which combines the three feature maps by the weighted sum. Based on the importance map, the important regions are preserved and less important regions are discarded. Finally, the proposed approach constructs the target image using the seam carving method [1] based on the importance map. The experimental results show that the proposed approach resizes stereo image pairs effectively.

2. Related Works

Avidan and Shamir [1] proposed a method for adjusting image size based on image content. They

[+] Corresponding author. Tel.: + 886-37-381893; fax: +886-37-354326.
E-mail address: ccchang@nuu.edu.tw.

analyzed the relationships of energy distribution in the image and compared methods of image resizing. The proportion of residual energy after image resizing indicated the quality of the resizing. By iteratively adding or removing seams, their approach can alter the size of images.

Kim et al. [6] used the adaptive scaling function, utilizing the importance map of the image to calculate the adaptive scaling function, which indicated the reduction level for each row of the original image. Kim et al. [7] used Fourier analysis for image resizing. After constructing the gradient map, they divided the image into strips of various lengths, and then used Fourier transform to determine the spectrum of each strip. The spectrums were then used as a low-pass filter to obtain an effect similar to smoothing. The level of horizontal reduction for each strip was then determined according to the influence of the filter.

Hwang and Chien [5] used a neural network method to determine the subject of images. They also used face recognition techniques to ensure the human faces within images. For ratios that could not be compressed using the seam carving method, they used proportional ratio methods to compress the subject of images. Rubinstein et al. [10] proposed a method of improvement for the procedure of seam carving. This method utilized techniques of forward energy and backward energy to reduce discontinuity in images. Wang et al. [14] proposed a method that simultaneously utilized techniques of stereo imaging and inpainting. This method had the capacity to remove image objects that caused occlusion, restoring original background image and depth information.

3. The Proposed Approach

First, the proposed approach extracts three feature maps, namely, disparity map, gradient map and saliency map, from input stereo images. After that, the proposed approach integrates all the feature maps to an importance map by the weighted sum. Finally, the proposed approach constructs the target image using the seam carving method [1].

3. 1 Importance Map

The importance map is defined as a weighted sum of the three feature maps: disparity map, gradient map and saliency map. The idea is to combine the visually important features gradient and saliency, and physically prominent feature depth to remove the less important regions in the image.

3.1.1 Disparity Map

Depth information is one of the most important features for an image. In many cases, the importance of the object of interest in an image is inversely proportional to its depth to the camera. Therefore, the proposed approach extracts a disparity map from the stereo images to represent depth information.

The proposed approach adopts the method of Konolige [8], which applies the sum of absolute difference (SAD) to find out the corresponding blocks of the left-right images. The corresponding points are the centers of the corresponding blocks with the strongest SAD value. This method can be applied to any pair of stereo images [4, 5] that have been rectified. As regard to the rectified pair of stereo images, the stereo matching method for block matching is divided into three parts:

1. Preprocessing: balancing the image intensity and intensifying the texture features.
2. Using SAD window to find the corresponding points of the left-right images along the horizontal coaxial.
3. Post processing: remove the false corresponding points.

The balance of image intensity can reduce errors in SAD approximation, intensify texture features in order to increase the accuracy of SAD in high texture regions. After SAD finds the corresponding points, the proposed approach applies left-right consistency to remove all the wrong corresponding points, which results in the disparity map as $E_{disparity}$.

3.1.2 Gradient Map

The human visual system is sensitive to edge information in an image. Therefore, the proposed approach extracts a gradient map from the source image to represent edge information.

The Sobel calculation on original image I results in the gradient map. Considering the direction of image resizing, the proposed approach offers weight individually to the Sobel horizontal and vertical directions as

$$E_{gradient} = \sqrt{w_h \cdot \left|\frac{\partial I}{\partial x}\right|^2 + (1 - w_h) \cdot \left|\frac{\partial I}{\partial y}\right|^2} \ ,$$

where w_h is the weight for the Sobel horizontal direction.

When the image reduces or extends horizontally, the increase of weight of Sobel horizontal direction can reduce the energy dispersal caused by the vertical gradient.

3.1.3 Saliency Map

Visual saliency is an important factor for human visual system. Therefore, the proposed approach extracts a saliency map from the source image. The computer vision [2, 9] tries to imitate the possible visual perception of the human eye, from object detection, object classification to object recognition.

The most-studied feature of natural images is the invariant of extension and reduction, which is also called $1/f$ law [11, 12]. After adding several images together and proceeding the fast Fourier transform, the amplitude $A(f)$ of the averaged Fourier spectrum are observed as

$E\{A(f)\} \propto 1/f$.

Based on the residual image in frequency space [3, 4, 13], Hou and Zhang [4] proposed a log spectrum representation to find out the relation of polylines and visual features in log spaces. The method can rapidly detect conspicuous objects without extra references. Given an image I , the intensity and the phase spectrum of the image can be obtained by transferring the image to Fourier space as

$A(f) = \Re(F[I])$, and $P(f) = \Im(F[I])$,

where F denote the Fourier Transform. $L(f)$ is the intensity information obtained after the reduction of the original image to one of $size_r \times size_r$ by Fourier transform as follows:

$L(f) = \log(A(f))$.

Therefore, the spectral residual $R(f)$ can be obtained by

$R(f) = L(f) - h_n(f) * L(f)$,

where $h_n(f)$ is an $n \times n$ matrix defined by

$$h_n(f) = \frac{1}{n^2} = \begin{pmatrix} 1 & 1 & \dots & 1 \\ 1 & 1 & \dots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \dots & 1 \end{pmatrix},$$

and $h_n(f) * L(f)$ is the averaged spectrum approximated by convoluting the input image.

Using inverse Fourier transform, the saliency map in spatial domain can be constructed as

$E_{saliency} = g(x) * F^{-1}[\exp(R(f) + P(f))]^2$,

where F^{-1} denote the inverse Fourier transform and $g(x)$ is a Gaussian filter smoothing the saliency map.

3.2 Image resizing

The proposed approach applied the method proposed by Avidan and Shamir [1] for image retargeting. Let I be an $n \times m$ image and the vertical seam is defined as

$s^x = \{s_i^x\}_{i=1}^n = \{(x(i), i)\}_{i=1}^n$, s.t. $\forall i, |x(i) - x(i-1)| \le 1$,

where x is a mapping $x : [1, \dots, n] \rightarrow [1, \dots, m]$.

A vertical seam is an 8-connected line. Every row only contains a single pixel. Carving the seam interactively is considered an advantage because it can prevent horizontal displacement during the deleting process. Horizontal displacement appears if the number of deleted pixels in each row is different, resulting in changes in the shape of the object. Therefore, the route of the vertical seam is indicated as $I_s = \{I(S_i)\}_{i=1}^{n} = \{I(x(i),i)\}_{i=1}^{n}$. 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 energy map is used to select seams. Given an energy function *e*, the energy $E(s) = E(I_s) = \sum_{i=1}^{n} e(I(s_i))$ of a seam is determined by the energy occupied by the positions of each pixel. When cutting a particular image horizontally, deleting the seam with the lowest energy $s^* = \min_s E(s) = \min_s \sum_{i=1}^{n} e(I(s_i))$ first is essential.

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

$$M(i,j) = e(i,j) + \min(M(i-1,j-1), M(i,j-1), M(i+1,j-1)).$$

Then, the backtracking method was adopted to iteratively delete the seams with relatively weak energy by gradually searching upward for the seams with a minimum energy sum from the point with the weakest energy in the last row.

4. Results

The camera used in this experiment was a Microsoft HD-5000. Two web cameras were connected to a digital camera binocular. The digital camera binocular is installed onto a laptop to create a dolly shot when the scene requires the camera to shift.

The proposed algorithm based on the disparity map was compared with the seam carving [1] incorporating the standard gradient map. Figure 1 shows the original left-right images (the top row), the disparity map and resized results by the proposed approach (the middle row), and the standard gradient map and resized results by the seam carving approach (the bottom row). The results significantly reveal the features of the disparity map, which is not determined by the textural information of the image. Therefore, the disparity map is not affected by the complex background of the image, and can protect the subject from being destroyed by the seam carving algorithm based on the standard gradient map.

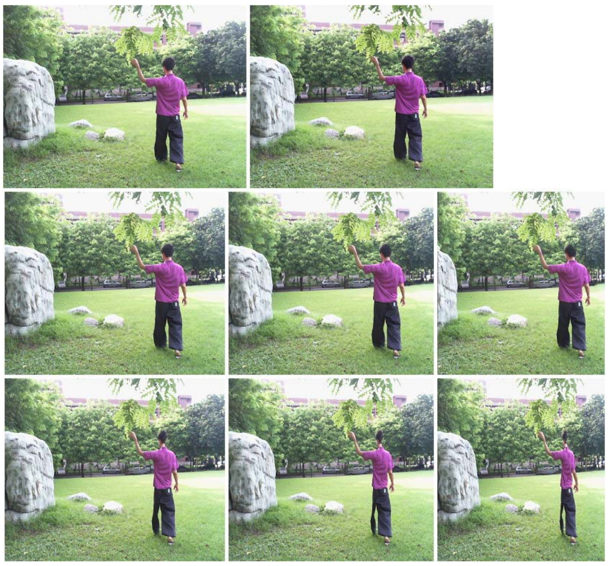

Fig. 1: Original left-right images (the top row), resized results by the proposed approach (the middle row), and the resized results by the seam carving approach (the bottom row).

5. Conclusions

This paper proposes a cooperative resizing technique for stereo image pairs. Several analyses were conducted, including the energy of disparity, gradient and visual saliency. Moreover, different types of energy were integrated as importance maps for stereo image resizing. Therefore, a perfect protection of the subject was achieved. In the future research, further studies are needed to develop an improved technique for video retargeting using the seam carving.

6. Acknowledgements

The support of this work in part by the National Science Council of Taiwan, R.O.C. under Grant NSC-96-2221-E-194-016-MY2 is gratefully acknowledged.

7. References

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

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

[3] C. Guo, Q. Ma, and L. Zhang, "Spatio-temporal saliency detection using phase spectrum of quaternion fourier transform," *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1-8, 2008.

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

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

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

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

[8] K. Konolige, "Small vision systems: Hardware and implementation," *Proceedings of Eighth International Symposium on robotics Research*, vol.8, pp.111-116, 1997.

[9] 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*, 2010.

[10] M. Rubinstein, A. Shamir, and S. Avidan, "Improved seam carving for video retargeting," *ACM Transactions on Graphics*, pp.1-9, 2008.

[11] D. Ruderman, "The statistics of natural images," *Network: Computation in Neural Systems*, vol. 5, no. 4, pp. 517-548, 1994.

[12] A. Srivastava, A. Lee, E. Simoncelli, and S. Zhu, "On advances in statistical modeling of natural images," *Journal of Mathematical Imaging and Vision*, vol. 18, no. 1, pp. 17-33, 2003.

[13] A. vander Schaaf and J. VanHateren, "Modelling the power spectra of natural images: statistics and information," *Vision Research*, vol. 36, pp. 2759-2770, 1996.

[14] L. Wang, H. Jin, R. Yang, and M. Gong, "Stereoscopic inpainting: Joint color and depth completion from stereo images," *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1-8, 2008.