Image Resizing in Saliency Histogram Domain

Hsi-Chin Hsin,[*] Chien-Kun Su,[1] and Yueh-Chi Chen

Department of Computer Science and Information Engineering, National United University,
1 Lienda, Miaoli 36003, Taiwan, R.O.C.
[1]Department of Electrical Engineering, Chung Hua University,
707, Sec. 2, WuFu Rd., Hsinchu, 30012 Taiwan, R.O.C.

The saliency map of an image provides useful information regarding the region of interest. In this paper, the normalized saliency map, called the saliency histogram, is used as a valid probability density function, based on which an efficient algorithm has been proposed for image resizing in the saliency histogram domain. Moreover, the saliency-histogram-based image resizing algorithm has been extended to video applications. Experimental results show the performance of the saliency histogram in terms of the preservation of salient objects in resized images and videos. The proposed approach is suitable for content-driven surveillance with multiresolution image sensor systems.

1. Introduction

Handheld devices such as smart phones and digital pads nowadays are in widespread use; however, there is an inevitable problem regarding the adjustment of image and video sizes.[1] Conventional methods using simple cropping and linear scaling are likely to degrade the region of interest (ROI).[2] In order to resize images/videos more effectively, it is desirable to take into account the visual content.[3] Modern resizing techniques are generally grouped into two categories: the discrete approach and the continuous approach. The content-aware seam carving (SC) algorithm is a common representative of the former.[4] As wavelet transform provides an efficient manner to represent images at multiple scales,[5] Conger *et al.* proposed the seamlet transform with wavelet filters for multiseam carving.[6] In Ref. 7, we proposed a scale-recursive algorithm for fast image resizing in the wavelet domain. The continuous approach is to warp an image by quad-shaped mesh deformation. Gal *et al.* reshaped images to preserve the proportions of prominent objects while stretching or squeezing homogeneous regions.[8] Wolf *et al.* analyzed the importance of pixels on the basis of local saliency, object detection and motion estimation for content driven video retargeting.[9] Chen *et al.* formulated image warping as a convex quadratic problem, and solved it via quadratic programming.[10] Wang *et al.* proposed the optimized scale and stretch (OSS) algorithm by allowing uniform scaling in prominent regions while hiding nonuniform stretching distortions in homogeneous backgrounds.[11] The saliency map (SM) of an image provides valuable information on the presence of prominent objects.[12] In Refs. 13–15, we proposed several algorithms based on the normalized SM, termed the saliency histogram (SH), for image resizing. In this work, we extend the potential of the SH to video resizing.

---

[*]Corresponding author: e-mail: hsin@nuu.edu.tw

The rest of the paper is organized as follows. Materials and Methods, including the proposed SH-based video resizing algorithm, are given in Sect. 2. Experimental Results are shown in Sect. 3. Conclusions can be found in Sect. 4.

2.   Materials and Methods

In contrast to the discrete approach to image resizing, the continuous approach tends to avoid the cause of annoying jags.[2] Image warping uses a mesh $M = (V, E, Q)$ composed of vertices V, edges E, and quads Q; V and E form grid lines to partition an image into Q. Wang *et al.* proposed the OSS algorithm to map the input mesh M onto the deformed mesh M' with minimum distortion energy.[11] For each quad q in Q, the OSS algorithm allows a scaled mapping $v' \in V'$ of $v \in V$:

$$v' = s_q v + t, \tag{1}$$

where s_q is the scale factor and t is the constant translation vector. The distortion energy $D(q)$ of q is defined as

$$D(q) = \sum_{(i,j)\in E(q)} \left\| (v'_i - v'_j) - s_q(v_i - v_j) \right\|^2, \tag{2}$$

where $E(q)$ is the edge of q and t is eliminated in the computation of $D(q)$. By taking the SM, $w(q)$, of an image into account, the OSS algorithm minimizes the total distortion D as

$$D = \sum_{q \in Q} w(q) \cdot D(q). \tag{3}$$

It is a quadratic problem with constraints on boundary conditions, grid line bending, and fold-over prevention.

The normalized SM can be used to describe the probability of a target's presence in the image domain and is therefore termed SH. For 1D shrinking, the marginal SH (MSH) is defined as

$$p_x(x) = \frac{\sum_{y=1}^{N} p(x, y)}{\sum_{x=1}^{M} \sum_{y=1}^{N} p(x, y)}; x \in [1, M], \tag{4}$$

where $p_x(x)$ denotes the MSH in the x-axis and $p(x, y)$ is the SH of an image. Figure 1(a) shows a test image, and its SM is given in Fig. 1(b). If the image is evenly partitioned into L vertical strips, it is noted from the MSH shown in Fig. 1(c) that most of the important contents are in the left and right regions. To reduce the difference in saliency between strips, a nonuniform partition is needed. On the basis of Eq. (4), the following transformation is used for adaptive nonuniform partition:

$$c_j = k : \frac{j-1}{L} < \sum_{x=1}^{k} p_x(x) \le \frac{j}{L}, \tag{5}$$

where c_j is the jth strip represented by its upper-right vertex, $j = 1, 2, \ldots, L$, and the boundary
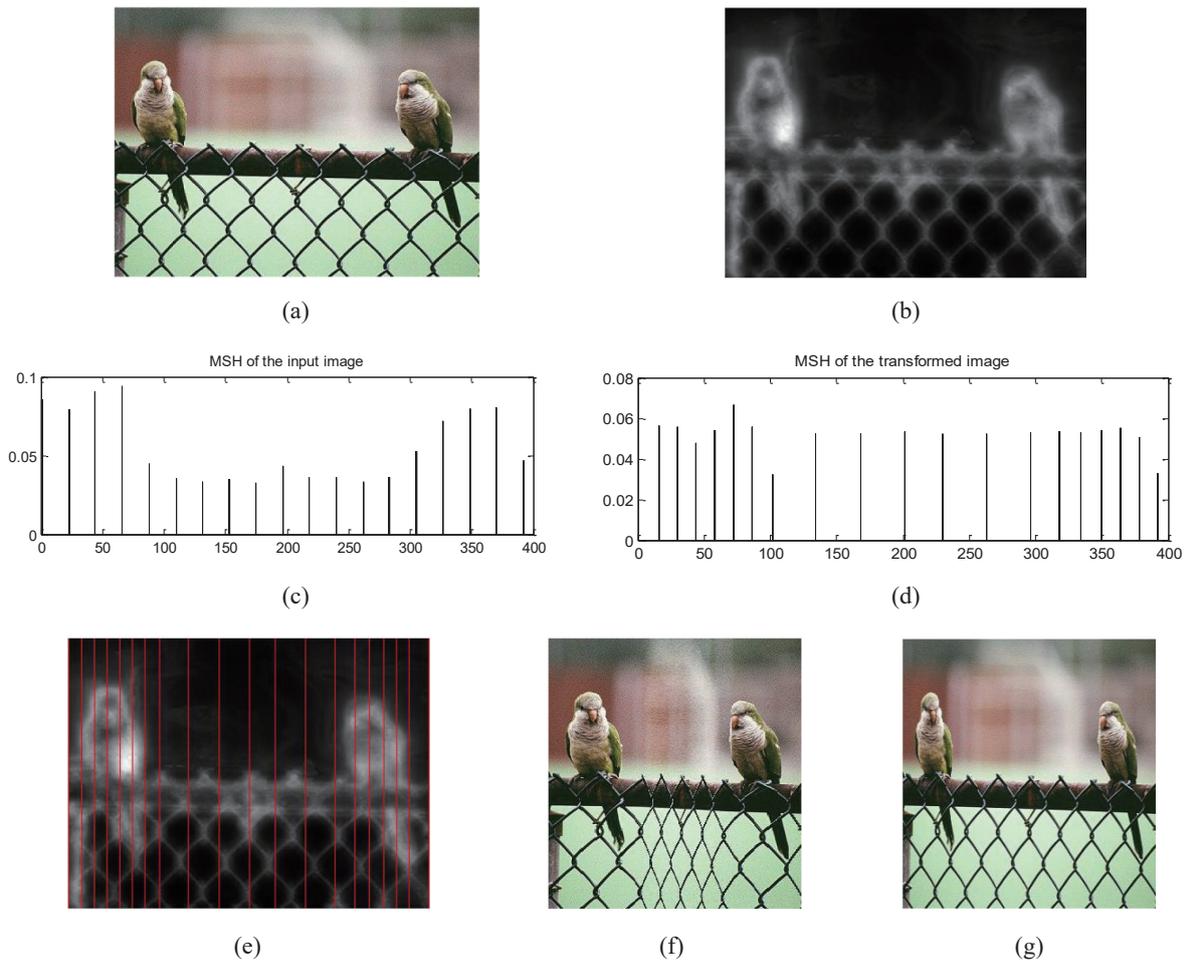
Fig. 1.   (Color online) (a) Test image.   (b) SM.   (c) MSH of (b) with uniform partition.   (d) MSH of (b) with nonuniform partition.   (e) Boundaries of the nonuniform partition overlaid on (b). (f) and (g) are the results of reducing the width of (a) using the proposed algorithm and linear scaling, respectively.

condition is $c_L = N$. As noted in Fig. 1(d), the MSH of the nonuniformly partitioned strips is much smoother than that in Fig. 1(c). Figure 1(e) shows the boundaries of the nonuniformly partitioned strips overlaid on the SM. For horizontal shrinking, the output image is evenly partitioned into L vertical strips with an equalized MSH, and the mapping of the output strips can be realized from the corresponding nonuniform input strips by simple interpolation. Figures 1(f) and 1(g) show the results of reducing the width of the image in Fig. 1(a) using the MSH-based algorithm and linear scaling. As expected, the mapping shows the potential of the SH. 2D shrinking can be realized by using the tensor product of two 1D shrinking operations, i.e., width reduction followed by height reduction, or vice versa. Details of the saliency histogram equalization (SHE) and hybrid SHE-SC algorithms can be found in Refs. 13–15.

Visual saliency provides useful information on the likelihood of ROI, which is beneficial to the development of content-driven image/video resizing systems. In general, the nonuniform mesh constructed by SHE is composed of different quad sizes, and high-saliency quads are smaller than low-saliency quads. Both of the SHE algorithm and the hybrid SHE-SC algorithm can be

applied to the first video frame to construct the initial mesh. For computational simplicity, a fast smoothing scheme is proposed to stabilize the construction of successive meshes with temporal coherence for the rest of a video. Specifically, let M_{k-1} and M_k be the meshes obtained from the SH of two consecutive video frames, F_{k-1} and F_k, at times $k-1$ and k, respectively. Each vertex, $v_{i,j,k}$, at coordinates (i, j) in the mesh M_k is linearly smoothed as

$$v_{i,j,k} = \alpha \cdot v_{i,j,k-1} + (1 - \alpha) \cdot v_{i,j,k}, \tag{6}$$

where $v_{i,j,k-1}$ is the corresponding vertex in the previous mesh M_{k-1}, and $0 < \alpha < 1$ implies the degree of temporal coherence between the frames F_{k-1} and F_k.

As a video is usually composed of different scenes, the proposed video resizing algorithm first divides the input video into sequences of scenes using the scene change detection algorithm,[16] and then resizes each sequence of frames independently since the temporal coherence between different scenes is not necessary. Figure 2 exemplifies the detection result with sharp peaks indicating the occurrences of scene changes for a segment of the big buck bunny sequence.

3. Experimental Results

The SH-based image and extended video resizing algorithms were implemented in Matlab. The discriminative regional feature integration (DRFI) SM[17] was used to produce the SH of an image. Figure 3 shows the potential of the SH-based image resizing algorithm compared with those of other state-of-the-art algorithms.[9,10] The width of the original image was reduced. This shows that the proposed algorithm is preferable in terms of the preservation of salient objects.

The extended SH-based video resizing algorithm was compared with other existing algorithms.[18,19] The discrete cosine transform (DCT)-based pairwise comparison method[16] was used for scene change detection. Different scenes were resized separately as temporal coherence is not necessary for disjointed scenes. Figure 4 shows a visual comparison.[18] As can be seen, there exist curving artifacts at the right border of the resized frame in the warp method.[18] They are absent in the image resized using the proposed algorithm. Figure 5 presents visual comparison with the improved SC.[19] Note that most of the ROI in the middle area can be preserved using the
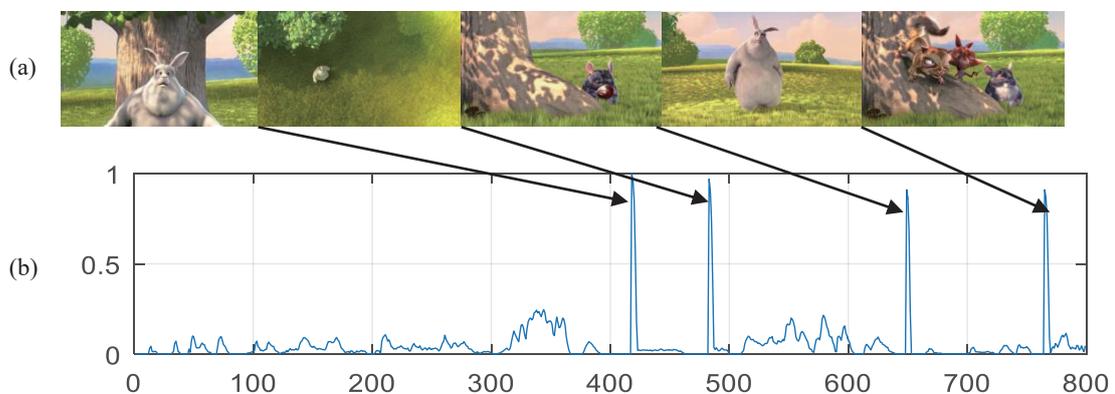

Fig. 2. (Color online) (a) Segment of the big buck bunny sequence and (b) percentage of changed DCT blocks[16] with sharp peaks indicating scene changes.

(a)   (b)   (c)   (d)

Fig. 3.   (Color online) (a) Test images. Resized images using (b) Wolf,[9] (c) OSS,[11] and (d) proposed algorithm.

Fig. 4.   (Color online) Visual comparison: (from left to right) source video frame and results of SH-based video resizing algorithm.[18]

Fig. 5.   (Color online) Visual comparison: (from left to right) source video frame and results of SH-based video resizing algorithm.[19]

(a)

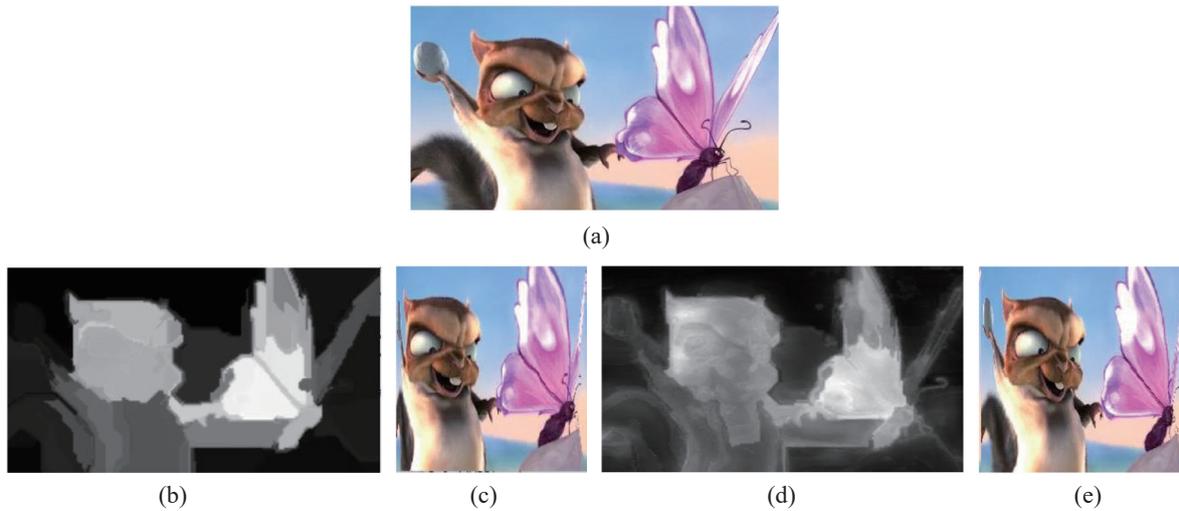(b)   (c)   (d)   (e)

Fig. 6.   (Color online) (a) Source video frame. (b) SH of (a) obtained by DRFI[17] with low saliency near left border. (c) Resized frame using proposed algorithm with (b). (d) Enhanced SH with additional features proposed by Judd *et al.*[20] (e) Improved result with the aid of (d).

improved SC at the cost of border backgrounds. In contrast, the SH-based video resizing algorithm is capable of preserving border regions with acceptable deformation.

In most cases, the SH-based algorithm is superior.[18,19] As the warping approach is mainly dependent on the estimation of SH, there is a possibility of overwarping in regions with low saliency. One example is given in Fig. 6. As observed from the estimated SH [Fig. 6(b)] obtained by DRFI[17] of a video frame [Fig. 6(a)], the saliency values near the left border are small, and consequently, there is overwarping at the left border of the resized frame. To solve this problem, one of the promising approaches is to supplement the SH with additional features, e.g., those proposed by Judd *et al.*[20] Figure 6(e) shows the improved result with the enhanced SH shown in Fig. 6(d).

4.   Conclusion

Saliency information plays an important role in content-aware image/video resizing. The normalized SM, called the SH, can be used as a valid probability density function for image warping with nonuniform meshes. In this study, images/video frames are partitioned into nonuniform quad meshes in the SH domain. Note that the content of salient objects can be well preserved in the resized images and video frames after mapping from the input quads onto the corresponding output quads. In the future, the proposed content-driven approach will be suitable for intelligent surveillance with multiresolution image sensors.[21] More specifically, the salient objects can be well sampled in a high-resolution mode with more pixels by sampling the backgrounds with fewer pixels.

Acknowledgments

References

1 X. Yang, X. Qian, and Y. Xue: IEEE Trans. Image Process. **24** (2015) 1709.
2 M. Rubinstein, D. Gutierrez, O. Sorkine, and A. Shamir: ACM Trans. Graphics **29** (2010) 9.
3 D. Vaquero, M. Turk, K. Pulli, M. Tico, and N. Gelfand: Proc. SPIE **7798** (2010) p. 15.
4 S. Avidan and A. Shamir: ACM Trans. Graphics **26** (2007) 1.
5 J. C. Goswami and A. K. Chan: Fundamentals of Wavelets: Theory, Algorithm, and Applications (John Wiley & Sons, Inc., Hoboken, New Jersey, 2011).
6 D. D. Conger, M. Kumar, and H. Radha: Proc. SPIE **7870** (2011) p. 7.
7 H.-C. Hsin, Y.-C. Chan, and C.-K. Su: Proc. APCEAS (2015) pp. 146–152.
8 R. Gal, O. Sorkine, and D. Cohen-Or: Proc. Eurographics Symp. Rendering (2006) pp. 297–303.
9 L. Wolf, M. Guttmann, and D. Cohen-Or: Proc. IEEE ICCV (2007) pp. 1–6.
10 R. Chen, D. Freedman, Z. Karni, C. Gotsman, and L. Liu: Proc. CVPR (2010) pp. 13–18.
11 Y. S. Wang, C. L. Tai, O. Sorkine, and T. Y. Lee: ACM Trans. Graphics **27** (2008) 118.
12 A. Borji, M.-M. Cheng, H. Jiang, and J. Li: IEEE Trans. Image Process. **24** (2015) 5706.
13 H.-C. Hsin: IET Image Process. **10** (2016) 787.
14 H.-C. Hsin: IET JoE **3** (2017) 7.
15 H.-C. Hsin, C.-K. Su, and Y.-C. Chen: Proc. ICSS (2017) p. 6.
16 J. Zhang, S. Li, and C.-C. J. Kuo: IEEE Trans. Image Process. **23** (2014) 797.
17 H. Jiang, J. Wang, Z. Yuan, Y. Wu, N. Zheng, and S. Li: Proc. IEEE CVPR (2013) pp. 2083–2090.
18 P. Krähenbühl, M. Lang, A. Hornung, and M. Gross: ACM Trans. Graphics **28** (2009) 10.
19 M. Rubinstein, A. Shamir, and S. Avidan: ACM Trans. Graphics **27** (2008) 1.
20 T. Judd, K. Ehinger, F. Durand, and A. Torralba: Proc. IEEE ICCV (2009) pp. 2106–2113.
21 D. Kim, M. Song, B. Choe, and S. Y. Kim: Sensors **17** (2017) 1.

About the Authors

Hsi-Chin Hsin received his M.S. and Ph.D. degrees in electrical engineering from University of Pittsburgh, PA, USA, in 1992 and 1995, respectively. From 1995 to 1996, he was a researcher in the Optoelectronics and Systems Laboratories at Industrial Technology Research Institute, Taiwan, and a consultant from 1996 to 1997. From 1996 to 2005, he was an Associate Professor in the Department of Electrical Engineering, Chung Hua University, Taiwan, and served as Chairman in 2004. From 2005 to 2007, he was a Professor in the Department of Computer Science and Information Engineering, National Formosa University, Taiwan. He was a Vice President at National United University, Taiwan from 2009 to 2012. Currently, he is a Professor in the Department of Computer Science and Information Engineering at National United University. His research interests include wavelet transform, multimedia techniques, and communication systems.

Chien-Kun Su received his Ph.D. degree in electrical and computer engineering from National Chiao Tung University, Taiwan in 2010. Currently, he is an Associate Professor in the Department of Electrical Engineering, Chung Hua University, Taiwan. His research interests include wavelet transform, multiresolution processing, and image processing.

Yueh-Chi Chen received his B.S. degree in computer science and information engineering from National United University, Taiwan in 2016. He is a graduate student in the Department of Computer Science and Information Engineering, National United University. His research interests include wavelet transform and image processing.