

# Content-Aware Video Seam Carving Based on Visual Cubes

Duan-Yu Chen\* and Yi-Shiou Luo

Department of Electrical Engineering

Yuan-Ze University

Chung-Li 32003, Taiwan, ROC

dychen@saturn.yzu.edu.tw; s984629@email.yzu.edu.tw

*Received 18 April 2010; Revised 19 May 2010; Accepted 20 June 2010*

**Abstract.** Seam carving for still images has attracted lots of attention in recent years. Approaches that can work well in this domain may not sufficiently robust enough to be applied to consecutive video frames due to the nature of visual dynamics in videos. Carving in consecutive frames with different criteria would usually result in discontinuity of visual perception. Therefore, how to preserve the visual continuity in video frames is the most critical issue in the field of video seam carving. In this paper, we propose a novel approach for modeling dynamic visual attention based on spatiotemporal analysis in order to detect the focus of interest automatically. The continuously varied co-sited blocks in a video cube are first detected and their variations are characterized as a bag of visual cubes, which are further employed to determine a proper extent of salient regions in video frames. Once the proper extent through video cubes is determined, the carving process then can be conducted to find the global optimum. Our experiment shows that the proposed content-aware video seam carving based on spatiotemporal bag of visual cubes can effectively generate resized videos while keeping their isotropic manipulation and the continuous dynamics of visual perception.

**Keywords:** Seam carving, visual cubes, spatiotemporal analysis

## References

- [1] S. Avidan and A. Shamir, "Seam Carving for Content-Aware Image Resizing," *ACM Transactions on Graphics*, Vol. 26, No. 3, 2007.
- [2] R. Achanta and S. Susstrunk, "Saliency Detection For Content-Aware Image Resizing," *Proceedings of the 16th IEEE International Conference on Image Processing*, pp. 1005-1008, 2009.
- [3] W. James, *The Principles of Psychology*, Harvard Univ. Press, Cambridge, Massachusetts, 1980/1981.
- [4] J. Yuan, Z. Liu, Y. Wu, "Discriminative Subvolume Search for Efficient Action Detection," *Proceedings of the 2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp.2442-2449, 2009.
- [5] Y. Ke, R. Sukthankar, M. Hebert. "Event Detection in Crowded Videos," *Proceedings of IEEE International Conference on Computer Vision*, pp. 1-8, 2007.
- [6] Y. Zhai and M. Shah, "Visual Attention Detection in Video Sequences Using Spatiotemporal Cues," *Proceedings of ACM International Conference on Multimedia*, pp.815-824, 2006.
- [7] L. Laptev and T. Lindeberg, "Space-Time Interest Points," *Proceedings of IEEE International Conference on Computer Vision*, pp.432-439, 2003.
- [8] G. Hua, C. Zhang, Z. Liu, Z. Zhang, Y. Shan, "Efficient Scale-Space Spatiotemporal Saliency Tracking for Distortion-Free Video Retargeting," *Proceedings of the 9th Asian Conference on Computer Vision*, pp. 182-192, 2009.
- [9] F. Liu and M. Gleicher, "Video Retargeting: Automating Pan and Scan," *Proceedings of the 14th Annual ACM International Conference on Multimedia*, pp. 241-250, 2006.

---

\* Correspondence author

- [10] S. Li and M.C. Lee, "An Efficient Spatiotemporal Attention Model and its Application to Shot Matching," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol.17, No.10, pp.1383-1387, 2007.
- [11] Y.F. Ma, L. Lu, H.J. Zhang, M. Li, "A User Attention Model for Video Summarization," *Proceedings of ACM Multimedia*, pp.533-541, 2002.
- [12] L. Itti, C. Koch, E. Niebur, "A Model of Saliency-Based Visual Attention for Rapid Scene Analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.20, No.11, pp.1254-1259, 1998.
- [13] V. Navalpakkam and L. Itti, "An Integrated Model of Top-Down and Bottom-Up Attention for Optimizing Detection Speed," *Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Vol.2, pp. 2049-2056, 2006.
- [14] L. Itti, and C. Koch, "Computational Modeling of Visual Attention," *Nature Reviews Neuroscience*, Vol.2, No. 3, pp. 194-203, 2001.
- [15] N. Otsu, "A Threshold Selection Method from Gray-Level Histograms," *IEEE Transactions on In Systems, Man and Cybernetics*, Vol. 9, No. 1, pp. 62-66, 1979.
- [16] N. Dalal and B. Triggs, "Histograms of Oriented Gradients for Human Detection," *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Vol. 1, pp. 886-893, 2005.
- [17] L. Wolf, M. Guttman, D. Cohen-Or, "Non-homogeneous Content-driven Video-retargeting," *Proceedings of the 11th IEEE International Conference on Computer Vision*, pp. 1-6, 2007.