Table 1 Comparison results for different video saliency detection models.

Models	SR [31]	BS [32]	Proposed
AUC-ROC	0.7154	0.6897	0.7425
NSS	0.4240	0.5551	0.4309

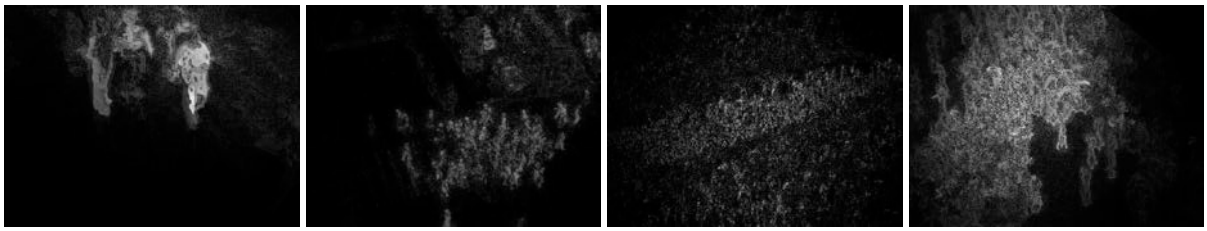

Fig.6. The saliency maps combining with additive fusion.

We also investigate the performance of another fusion strategy (i.e., additive combination) for video saliency detection. Fig.6 shows the saliency maps combining with additive fusion strategy for the same test images in Fig.5. The weight between temporal and spatial saliency is obtained by optimal training. We can see that the additive combination strategy cannot get good performance. Since the distributions of temporal and spatial saliency maps are somewhat different, the additive combination cannot suppress each other.

However, the proposed method has the following limitations: 1) the salient motion regions are not very prominent in some video sequences, because the assumption of background in low-rank matrix decomposition may be not appropriate; 2) the computation complexity of the proposed method is somewhat high, because videos should be resampled first and retransformed. These limitations should be properly considered in the future work.

5 Conclusion

In this paper, we presented a novel video saliency detection method by spatio-temporal sampling and sparse matrix decomposition. The main features of the proposed method are that: (1) we sample the input video sequence into three planes: X-T slice plane, Y-T slice plane, and X-Y slice plane; (2) we use sparse matrix decomposition to recover the background and motion features from the slices; (3) motion and static saliency maps are multiplicatively combined with central bias prior. Experimental results show the effectiveness of the proposed method.

Although the proposed scheme exhibit good performance in saliency detection, some aspects still deserve further research and improvement: 1) the connectivity of foreground and background should be further considered; 2) more low-level and high-level feature cues should be incorporated; 3) computational complexity should be an important indictor of the proposed method.

Acknowledgement:
This work was supported by the Natural Science Foundation of China (grant 61271021), and the Natural Science Foundation of Ningbo (grant 2012A610039).

References:

[1] N. J. Butko, L. Y. Zhang, G. W. Cottrell, and J. R. Movellan, “Visual saliency model for robot cameras,” in Proc. of IEEE International Conference on Robotics and Automation, pp. 2398-2403, 2008.

[2] Q. Zhang, K. N. Ngan, “Multi-view video based multiple objects segmentation using graph cut and spatiotemporal projections,” Journal of Visual Communication and Image Representation, vol. 21, no. 5, pp. 453-461, 2010.

[3] Z. Wang, Q. Li, “Information content weighting for perceptual image quality assessment,” IEEE Transactions on Image Processing, vol. 20, no. 5, pp. 1185-1198, May 2011.

[4] Z. Li, S. Qin, and L. Itti, “Visual attention guided bit allocation in video compression,” Image and Vision Computing, vol. 29, no. 1, pp. 1-14, Jan 2011.

[5] V. Setlur, T. Lechner, M. Nienhaus, and B. Gooch, “Retargeting images and video for preserving information saliency,” IEEE Computer Graphics and Applications, vol. 27, no. 5, pp. 80-88, 2007.

[6] 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, no. 11, pp. 1254-1259, 1998.

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

[8] M. Cheng, G. Zhang, N. J. Mitra, X. Huang, S. Hu, "Global contrast based salient region detection," in Proc. of IEEE International Conference on Computer Vision and Pattern Recognition, pp. 409-416, 2011.

[9] J. Harel, C. Koch, and P. Perona, "Graph-based visual saliency," in Proc. of Advances in Neural Information Processing Systems, pp. 545-552, 2006.

[10] R. Achanta, S. Hemami, F. Estrada, and S. Susstrunk, "Frequency tuned salient region detection," in Proc. IEEE International Conference on Computer Vision and Pattern Recognition, pp. 1597-1604, 2009.

[11] E. Erdem and A. Erdem, "Visual saliency estimation by nonlinearly integrating features using region covariance," Journal of Vision, vol.13, no. 4, article 11, 2013.

[12] R. Margolin, A. Tal, and L. Manor, "What makes a patch distinct?" in Proc. of IEEE International Conference on Computer Vision and Pattern Recognition, pp. 1139-1146, 2013.

[13] W. Zhang, Q. Xiong, and S. Chen, "Data-driven saliency region detection based on undirected graph ranking," WSEAS Transactions on Computers, vol. 13, pp. 310-319, 2014.

[14] X. Lv, D. Zou, L. Zhang, and S. Jia, "Feature coding for image classification based on saliency detection and fuzzy reasoning and its application in elevator videos," WSEAS Transactions on Computers, vol. 13, pp. 266-276, 2014.

[15] Y. Zhai and M. Shah, "Visual attention detection in video sequences using spatiotemporal cues," in Proc. of ACM Multimedia, pp. 815-824, 2006.

[16] E. Rahtu, J. Kannala, M. Salo, and J. Heikkilä, "Segmenting salient objects from images and videos," in Proc. of European Conference on Computer Vision, pp. 366-379, 2010.

[17] L. Xu, J. Jia, and Y. Matsushita, "Motion detail preserving optical flow estimation," in Proc. of IEEE Conference on Computer Vision and Pattern Recognition, pp. 1293-1300, 2010.

[18] S. Mathe, C. Sminchisescu, "Dynamic eye movement datasets and learnt saliency models for visual action recognition," in Proc. of European Conference on Computer Vision, pp. 842-856, 2012.

[19] Y. Fang, W. Lin, Z. Chen, C. Tsai, C. Lin, "A video saliency detection model in compressed domain," IEEE Transactions on Circuits and Systems for Video Technology, vol. 24, no. 1, pp. 27-38, 2014.

[20] K. Muthuswamy, and D. Rajan, "Salient motion detection in compressed domain," IEEE Signal Processing Letters, vol. 20, no. 10, pp. 996-999, 2013.

[21] X. Cui, Q. Liu, and D. Metaxas, "Temporal spectral residual: fast motion saliency detection," in Proc. of ACM Multimedia, pp. 617-620, 2009.

[22] Z. Zivkovic, "Improved adaptive Gaussian mixture model for background subtraction," in Proc. of International Conference on Pattern Recognition, pp. 28-31, 2004.

[23] V. Mahadevan and N. Vasconcelos, "Spatiotemporal saliency in dynamic scenes," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 32, no. 1, pp. 171-177, 2010.

[24] Y. Luo, J. Yuan, "Salient object detection in videos by optimal spatio-temporal path discovery," in Proc. of ACM Multimedia, pp. 509-512, 2013.

[25] Y. Xue, X. Guo, X. Cao, "Motion saliency detection using low-rank and sparse decomposition," in Proc. of IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 1485-1488, 2012.

[26] W.-T. Li, H.-S. Chang, K.-C. Lien, H.-T. Chang, Y. F. Wang, "Exploring visual and motion saliency for automatic video object extraction," IEEE Transactions on Image Processing, vol. 22, no. 7, pp. 2600-2610, 2013.

[27] J. Tünnermann, B. Mertsching, "Region-based artificial visual attention in space and time," Cognitive Computation, vol. 6, no. 1, pp 125-143, 2014.

[28] E. J. Candès, X. Li, Y. Ma, J. Wright, "Robust principal component analysis," Journal of the ACM, vol. 58, no. 3, Article No. 11, 2011.

[29] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Susstrunk, "SLIC superpixels compared to state-of-the-art superpixel methods," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 34, no. 11, pp. 2274-2282, 2012.

[30] N. Riche, M. Mancas, D. Culibrk, V. Crnojevic, B. Gosselin, and T. Dutoit, "Dynamic saliency models and human

attention: a comparative study on videos," in Proc. of Asian Conference on Computer Vision, pp. 586-598, 2013.

[31] L. Itti, P. Baldi, "A principled approach to detecting surprising events in video," in Proc. of IEEE Conference on Computer Vision and Pattern Recognition, pp. 631-637, 2005.

[32] H. J. Seo, P. Milanfar, "Static and space-time visual saliency detection by self-resemblance," Journal of Vision, vol. 9, no. 12, pp. 1-17, 2009.