Video Magnification in Presence of Large Motions

Mohamed Elgharib[1], Mohamed Hefeeda[1], Frédo Durand[2], William T. Freeman[2]
[1]Qatar Computing Research Institute. [2]MIT CSAIL.

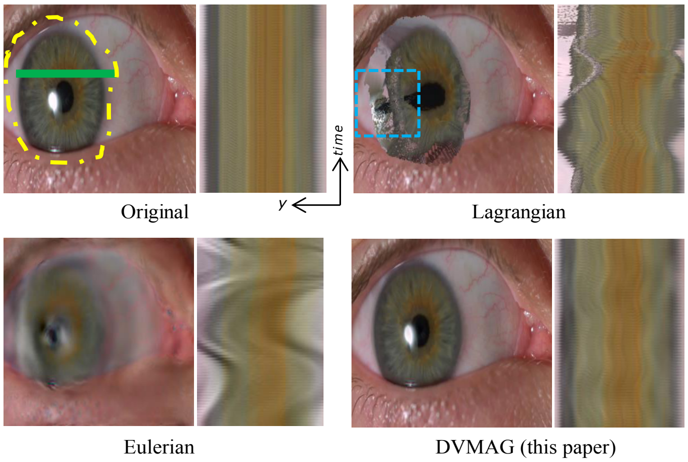

Figure 1: Comparison with state of the art. The Region of Interest is the dashed yellow region (top, left). For each magnification we show the spatio-temporal slice for the green line (top, left). For easier comparison all slices are temporally stabilized. This sequence shows an eye moving along the horizontal direction. Processing the sequence with our DVMAG technique shows that the iris wobbles as the eye moves (spatio-temporal slice). Such wobbling is too small to be observed in the original sequence (top, left). The global motion of the eye causes significant blurring artifacts when processed with the Eulerian approach [2]. The Lagrangian approach [1] sensitivity to motion errors generates noisy magnification (see dashed blue).

The world is full of small temporal variations that are hard to see with naked eyes. Variations in skin color occur as blood circulates [3], structures sway imperceptibly in the wind [2], and human heads wobble with each heart beat. While usually too small to notice, such variations can be magnified computationally to reveal a fascinating and meaningful world of small motions [1, 2, 3]. Current video magnification approaches assume that the objects of interest have very small motion. However, many interesting deformations occur within or because of larger motion. For example, our skin deforms subtly when we make large body motion. A toll gate that closes exhibits tiny vibrations in addition to the large rotational motion. And microsaccades are often combined with large-scale eye movements (Fig. 1). Furthermore, videos or objects might be handheld and may not be perfectly still, and a standard video magnification technique will amplify handshake in addition to the motion of interest. When applied to videos that contain large motions, current magnification techniques result in large artifacts such as haloes or ripples, and the small motion remains hard to see because it is overshadowed by the then magnified large motion and its artifacts (Fig. 1).

In the special case of camera motion, it might be possible to apply video stabilization as a preprocess to remove the undesirable large handshake, before magnification. However, this approach does not work for general object motion, and even in the case of camera shake, one has to be extremely careful because any small mistake in video stabilization will be amplified by the video magnification step. This problem is especially challenging at the boundary between a moving object, such as an arm, and its background, where multiple motions are present: the large motion, the subtle deformation to be amplified, and the background motion. Current video magnification such as the linear Eulerian [3] and phase-based [2] algorithms assume that there is locally a single motion. This generates a background dragging effect around object boundaries. In addition the Lagrangian approach [1] is sensitive to motion errors. This generates noisy magnifications.

This paper presents a video magnification technique capable of handling small motions within large ones. Our technique has two main components: 1. Warping to discount large motion and 2. Layer-based Magnification. Users select a region of interest (ROI) they want to magnify (Fig. 1, dashed yellow). The Warping stage removes large motion while preserving small ones, and without introducing artifact that could be magnified. For this, we use feature point tracking and optical flow, with regularized low-order parametric models for the large-scale motion. Our layer-based magnification is based on decomposing an image into a foreground, and a background through an opacity matte. We magnify each layer and generate a magnified sequence through matte inversion. We use texture synthesis to fill in image holes revealed by the magnified motion. Finally, we de-warp the magnified sequence back to the original space-time co-ordinates.

This is an extended abstract. The full paper is available at the Computer Vision Foundation webpage.

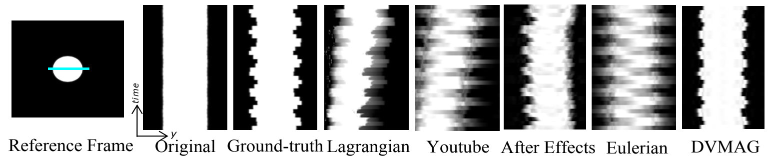

Figure 2: A frame from a synthetic sequence with mixed small and large motions (left), and its magnification using different techniques. We zoom on the blue spatio-temporal slice (see left). Our approach DVMAG best resembles ground-truth and does not generate blurring artifacts as other techniques.

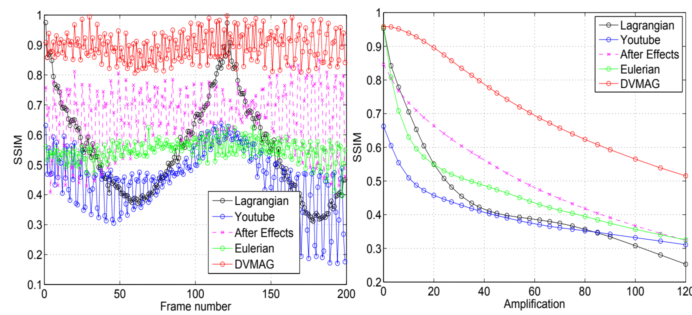

Figure 3: Left: SSIM with ground-truth for the sequence of Fig. 2. The larger the SSIM the better. DVMAG outperforms all examined techniques. Right: SSIM with ground-truth over different amplification factors. Our approach handles large amplifications with less magnification artifacts over all other techniques.

Fig. 2 shows a synthetic sequence mixed with small and large motions. A small vibrating local motion is first added to the white circle (Fig. 2, left) and then a large global motion is added to the entire sequence. We process the sequence using different magnification techniques, and our technique DVMAG best resembles ground-truth. This is also reflected numerically in In Fig. 3 (left) through SSIM. Fig. 3 (right) examines how amplification factors are handled by different magnification techniques. Results show that DVMAG handles larger amplifications with less errors over all other techniques. In addition DVMAG has the slowest rate of degradation. For instance in the range $\alpha = 0 - 40$ the slopes of Youtube, Eulerian and Lagrangian are steeper than DVMAG.

[1] Ce Liu, Antonio Torralba, William T. Freeman, Frédo Durand, and Edward H. Adelson. Motion magnification. *SIGGRAPH*, 24(3):519–526, 2005.

[2] Neal Wadhwa, Michael Rubinstein, Frédo Durand, and William T. Freeman. Phase-based video motion processing. *SIGGRAPH*, 32(4), 2013.

[3] Hao-Yu Wu, Michael Rubinstein, Eugene Shih, John Guttag, Frédo Durand, and William T. Freeman. Eulerian video magnification for revealing subtle changes in the world. *SIGGRAPH*, 31(4):65:1–65:8, 2012.