Enhanced Eulerian Video Magnification

Le Liu Le Lu Jingjing Luo Jun Zhang Xiuhong Chen
School of Digital Media
Jiangnan University
Wuxi, China

Abstract— **a post-processing technique is introduced to improve the Eulerian video magnification method, which is a state-of-the-art motion magnification method to manipulate small movements in videos based on spatio-temporal filtering. The proposed method use the Eulerian video magnification as a video spatio-temporal motion analyzer to get the pixel-level motion mapping. Then the input video pixels are warped based-on this mapping to amplify the motion. This processing does not involve pixel value modifying, which makes it supports larger amplification and is significantly less influenced by the frame noise.**

Keywords- Motion magnification, Eulerian motion, Spatio-temporal analysis, Image warping, Video-based rendering

I. INTRODUCTION

Because of the limitations of human visual system on the spatio-temporal sensitivity, very small motions are difficult to be observed by naked eye. These motions yet may reveal some potential information about the world.

Liu [1] presented motion magnification technique that acts like a microscope for visual motion. It can amplify subtle motions in a video sequence, allowing for visualization of deformations that would otherwise be invisible. Just like other Lagrangian approaches to motion magnification [2-4], motion is computed explicitly and the frames of the video are warped according to the magnified velocity vectors. For example, the precise motion computations and grouping operations of the motion magnification algorithm [1] take 10 hours in a combination of C++ and Matlab code. Besides, motion de-noising [3, 5] is also by no means a trivial problem.

Recently-proposed Eulerian approach (Eulerian video magnification, EVM [6]) eliminates the need for costly flow computation, and process the video separately in space and time. As revealed by Wadhwa et al. [7], the drawback of EVM is that it supports only small magnification factors at high spatial frequencies, and can significantly amplify noise when the magnification factor is increased. To counter these issues, Wadhwa et al. [7] proposed a new Eulerian approach to motion processing, based on complex-valued steerable pyramids. Unfortunately, for sequences in which the phase signal is noisy, parts of the image in the magnified video may appear to move incoherently.

We formulate the video magnification problem in an image warping frame-work, where the goal is to recover a "smooth warping mesh" of the input video by spatio-temporal filtering to input video. The EVM is used as a motion analyzer to get a pixel-level motion mapping, which can derive a smooth warping mesh used to amplify video motion.

The main contribution of this paper is a new post-processing step of the EVM to make it support larger amplification of motions at all spatial frequencies and is significantly less influenced by the frame noise. We call this improved EVM method as the Enhanced Eulerian Video Magnification (E2VM).

II. BACKGROUND OF EULERIAN VIDEO MAGNIFICATION

The goal of Eulerian video magnification (EVM [6]) is to reveal temporal variations in videos that are difficult to see with the naked eye. This method takes a standard video sequence as input, and applies spatial decomposition, followed by temporal filtering to the frames. The resulting frames are then amplified to reveal hidden information.

The EVM approach combines spatial and temporal processing to emphasize subtle temporal changes in a video. The process is illustrated in Fig. 1. The video sequence are first decompose into different spatial frequency bands. These bands are applied a ban-pass filter and magnified differently according to the signal-to-noise ratios and spatial frequencies. This processing can amplify small motion without costly flow computation or tracking motion as in Lagrangian methods [1].

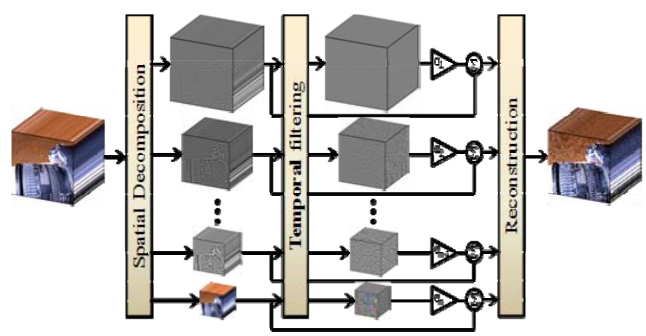


Figure 1. Overview of the EVM framework

A pseudo-code of EVM method is listed in Table I.

TABLE I. THE PSEUDO-CODE OF EVM

Step	*processing*
Input	Video $I(\mathbf{x},t)$, $t=0$
<1>	Read current frame I
<2>	Computing Gaussian and Laplacian pyramid: $\{G_l(I_k)\}_{l=0}^{n}\{L_l(I_k)\}_{l=0}^{n}$
<3>	Applying band-pass filter: $y_L(\mathbf{x},t)=\omega_L L_l(\mathbf{x},t)+(1-\omega_L)L_l(\mathbf{x},t-1)$ $y_H(\mathbf{x},t)=\omega_H L_l(\mathbf{x},t)+(1-\omega_H)L_l(\mathbf{x},t-1)$ $B_l(\mathbf{x},t)=y_H(\mathbf{x},t)-y_L(\mathbf{x},t)$
<4>	Space-time video processing: $\tilde{L}_l(\mathbf{x},t)=L_l(\mathbf{x},t)+(1+\alpha_l)B_l(\mathbf{x},t)$
<5>	Reconstruction: $\tilde{L}_l(I)\leftarrow\tilde{L}_l(I)+upSample(\tilde{L}_{l+1}(I))$ $\tilde{I}\leftarrow\tilde{L}_0(I)$
<6>	Rewrite current frame: $I\leftarrow\tilde{I}$
<7>	If $t<t_{\max}$, $t++$ goto <1>

The drawback of EVM is that it supports only small magnification factors at high spatial frequencies, and can significantly amplify noise when the magnification factor is increased.

In section 3 and section 4, we will show how to use EVM to extract video motion and magnified small motion by image warping without any noise amplification.

III. MOTION ANALYSIS BY EVM

In this section, we show that EVM can be used as a pixel-level motion analyzer by only the spatio-temporal filtering. Just like Wu et al. [6], we consider the simple case of a 1D signal undergoing translational motion. This analysis generalizes directly to locally-translational motion in 2D.

Let $I(x,t)$denote the image intensity at position *x* and time *t*. In Eulerian video magnification, the expressed the observed intensities with respect to a displacement function $\delta(t)$, such that $I(x,t)=f(x+\delta(t))$ and $I(x,0)=f(x)$. The motion magnification result in Eulerian video method is a synthesized the pixel value

$$\hat{I}(x,t)=f(x+(1+\alpha)\delta(t)) \tag{1}$$

for a fixed amplification factor α.

Assuming the image can be approximated by a first-order Taylor series expansion, one can write the image$f(x+\delta(t))$ in a first-order Taylor expansion about *x*. This leads to the following formulation.

$$f(x+(1+\alpha)\delta(t))\approx f(x)+(1+\alpha)\delta(t)\frac{\partial f(x)}{\partial x} \tag{2}$$

Let $B(x,t)$ be the result of applying a broadband temporal band-pass filter to $I(x,t)$ at every position *x*. In Eulerian video magnification, the magnified motion is just the $B(x,t)$, then we have

$$B(x,t)=(1+\alpha)\delta(t)\frac{\partial f(x)}{\partial x} \tag{3}$$

Combining Eqs. 1, 2 and 3, we have

$$(1+\alpha)\delta(t)\approx\frac{f(x+(1+\alpha)\delta(t))-f(x)}{\frac{\partial f(x)}{\partial x}}\approx\frac{\hat{I}(x,t)-I(x,0)}{\frac{\partial I(x,0)}{\partial x}} \tag{4}$$

As a result, we have got a 2D motion mapping $[U(x,y),V(x,y)]$ at point (x,y) by calculate the difference between the input video and EVM processing result:

$$U(x,y)=\frac{\hat{I}(x,y)-I(x,y)}{\frac{\partial I(x,y)}{\partial x}} \qquad V(x,y)=\frac{\hat{I}(x,y)-I(x,y)}{\frac{\partial I(x,y)}{\partial y}} \tag{5}$$

IV. MOTION MAGNIFICATION BY IMAGE WARPING

After completing the calculation of the input video's 2D motion mapping, we can use it as the basis of an image warping grid to amplify the video motion. In this section, this image warping framework (shown in Fig. 2) will be presented.

We formulate the motion magnification problem in an image warping framework, where the goal is to amplify the video motion by warping some frame pixels along the motion mapping direction. In order to improve the video processing speed, the motion mapping is sub-sampling to get a spares grid.

The warped frame of the output video can be calculated by the following formula：

$$\hat{I}(x,y)=I(x+\beta(x,y)U(x,y),y+\beta(x,y)V(x,y)) \tag{6}$$

where $\beta(x,y)$ is a smooth scale matrix to further adjust the magnification of the input video. Meanwhile, we set the elements of $\beta(x,y)$ in the image boundary close to 0, which can remove subtle frame shake caused by the motion noise from the EVM.

A pseudo-code of EVM method is listed in Table II.

TABLE II. THE PSEUDO-CODE OF E2VM

Step	*processing*
Input	Video $I(\mathbf{x},t)$, $t=0$
<1>	Read current frame I
<2>	Computing Gaussian and Laplacian pyramid: $\{G_l(I_k)\}_{l=0}^{n}\{L_l(I_k)\}_{l=0}^{n}$
<3>	Applying band-pass filter: $y_L(\mathbf{x},t)=\omega_L L_l(\mathbf{x},t)+(1-\omega_L)L_l(\mathbf{x},t-1)$ $y_H(\mathbf{x},t)=\omega_H L_l(\mathbf{x},t)+(1-\omega_H)L_l(\mathbf{x},t-1)$ $B_l(\mathbf{x},t)=y_H(\mathbf{x},t)-y_L(\mathbf{x},t)$

Step	*processing*
<4>	Space-time video processing: $\tilde{L}_l(\mathbf{x},t)=L_l(\mathbf{x},t)+(1+\alpha_l)B_l(\mathbf{x},t)$
<5>	Reconstruction: $\tilde{L}_l(I)\leftarrow\tilde{L}_l(I)+upSample(\tilde{L}_{l+1}(I))$ $\tilde{I}\leftarrow\tilde{L}_0(I)$
<6>	Motion analysis: $U(x,y)=\frac{\hat{I}(x,y)-I(x,y)}{\frac{\partial I(x,y)}{\partial x}}$ $V(x,y)=\frac{\hat{I}(x,y)-I(x,y)}{\frac{\partial I(x,y)}{\partial y}}$
<7>	Warping frame: $\hat{I}(x,y)=I\big(x+\beta(x,y)U(x,y),y+\beta(x,y)V(x,y)\big)$
<8>	Rewrite current frame: $I\leftarrow\tilde{I}$
<9>	If $t<t_{\max}$, $t++$ goto <1>

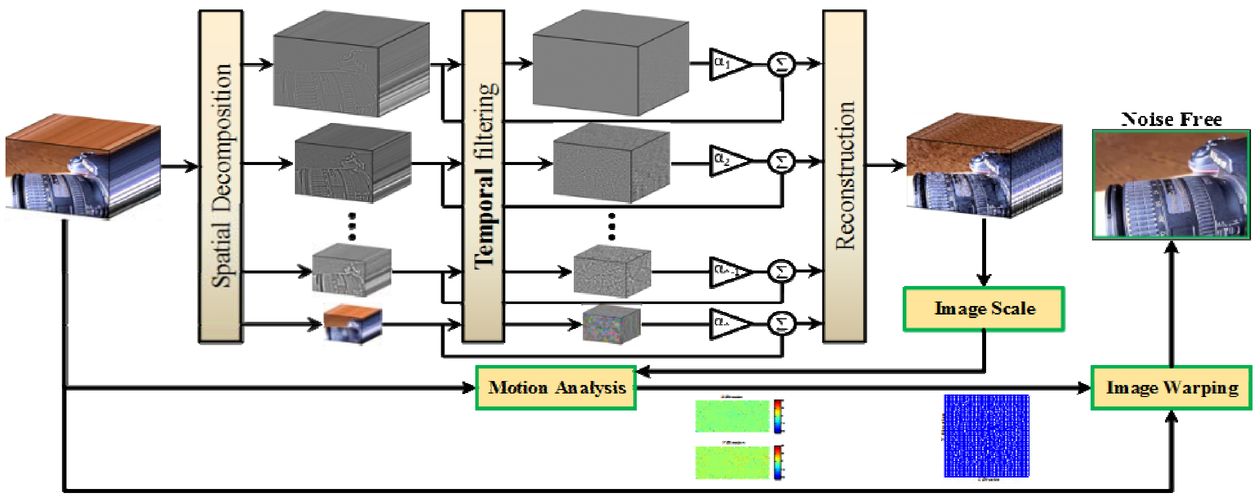


Figure 2. Overview of the E2VM framework

V. Results

In this section, we present some experimental results to evaluate our E2VM video magnification method. These experiments indicate that our method is effective. We process a one-megapixel frame in about 0.2 seconds using an un-optimized Matlab program on a normal PC with 3.1 GHz Intel i5 CPU.

A. The Motion Mapping

A pixel-level motion mapping is shown in Fig. 3, which was calculated from the difference between the input video and EVM processing result.

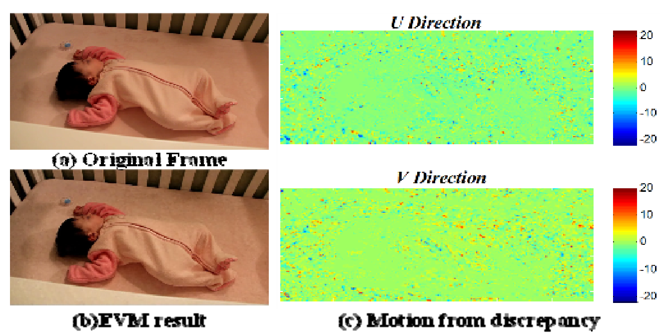


Figure 3. Motion from discrepancy between original frame and EVM result

In the motion mapping of Fig. 3, Red positions indicate forward movement, blue positions indicate reverse movement, and green positions indicate no movement.

B. The Warping Grid

Using the motion mapping $[U(x,y),V(x,y)]$ in Fig. 3, we can get a warping grid from the expression $[x+\beta(x,y)U(x,y),y+\beta(x,y)V(x,y)]$(shown in Fig. 4).

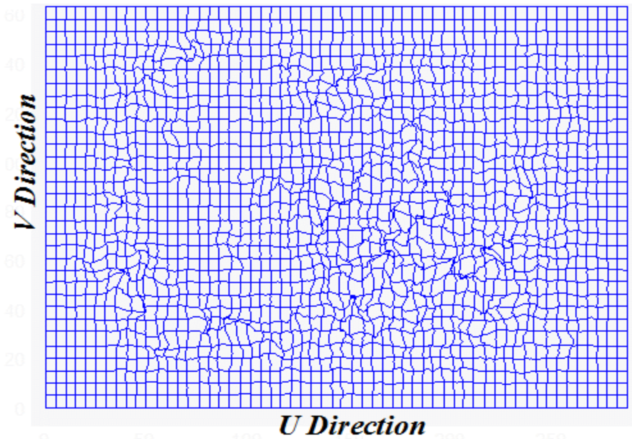


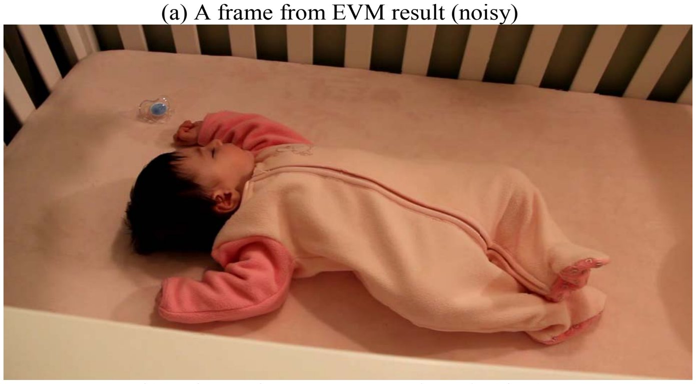
Figure 4. Image warping mesh from the motion mapping

To ensure that the video frame is unaffected by the motion noise from EVM, the boundary of the warping grid was smoothed by the mean processing.

C. *The Output Video of E2VM*

Using image warping equation (Eq. 6), we can amplify the video motion by warping some frame pixels along the motion mapping direction. From the output frame shown in Fig. 5, we demonstrate that E2VM provides a better post-processing way of EVM at handling noise.

(a) A frame from EVM result (noisy)



(b) A frame from E2VM result (noise free)

Figure 5. Results of EVM (a) and E2VM (b)

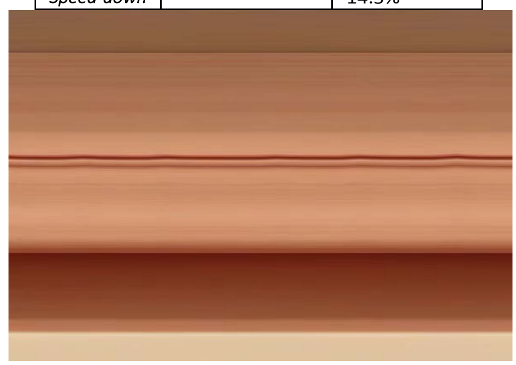
In comparison shown in Fig. 6, our new warping-based E2VM supports larger magnification factors with significantly fewer artifacts and less noise.

D. *Running-times*

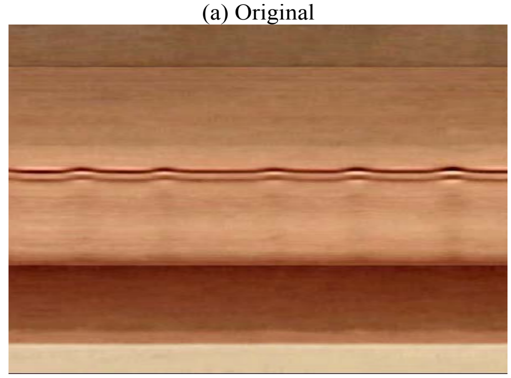
The proposed E2VM method involves only a small resolution image difference calculation and sparse grid image warping. Compared with the original EVM method, E2VM does not increase too much additional computation times and still has a comparative advantage. The proof is shown in Table III.

TABLE III. RUNNING-TIMES OF EVM AND E2VM (UNIT: SECONDS)

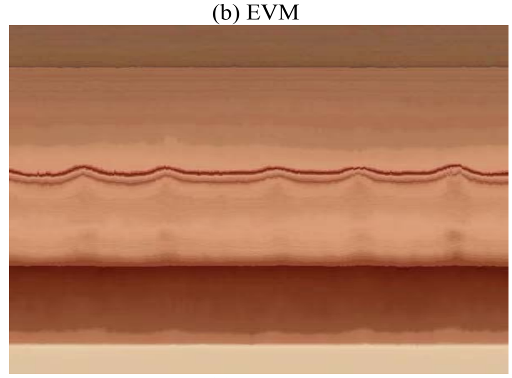
Method	*Frame resolution*	*Running-times*
EVM	960×554×300	107.1
E2VM		122.6
Speed-down	---	14.5%



(a) Original



(b) EVM



(c) E2VM

Figure 6. A spatiotemporal X-T slice of the video along the profile at the middle line.

VI. Conclusion

We have presented an efficient and noise free motion magnification method by spatio-temporal filtering and image warping. To amplify motion, our method uses Eulerian video magnification method as a spatio-temporal motion analyzer to get the pixel-level motion mapping. Our enhanced Eulerian video magnification method then magnifies the temporal video motion by image warping based-on the previous motion mapping. We demonstrated that this image warping-based technique improves the state-of-the-art in Eulerian motion processing and provides a better post-processing way to handle the frame noise.

Using a negative motion mapping, our image warping-based method may also be used to remove small motion of video for a variety of applications, such as de-animating [8], motion de-noising [3] and video stabilization [9-10]. It will be focused in later studies.

Acknowledgements

This paper is supported by the national undergraduate training programs for innovation and entrepreneurship (201310295054), fundamental research funds for the central universities (JUSRP1046) and natural science foundation of Jiangsu province of China (BK20130158).

References

[1] Liu C, Torralba A, Freeman W T, et al. Motion Magnification. ACM Transactions on Graphics. 2005, 24(Jul): 519-526.

[2] Gautama T, Van Hulle M M. A Phase-Based Approach to the Estimation of the Optical Flow Field Using Spatial Filtering. IEEE Transactions on Neural Networks. 2002, 13(5): 1127-1136.

[3] Rubinstein M, Liu C, Sand P, et al. Motion De-noising with Application to Time-lapse Photography. In: IEEE Computer Vision and Pattern Recognition. Colorado Springs, USA: IEEE Computer Society, 2011. 313-320.

[4] Freeman W T, Adelson E H, Heeger D J. Motion Without Movement. Computer Graphics. 1991, 25(4): 27-30.

[5] Fuchs M, Tongbochen, Wangc O, et al. Real-time temporal shaping of high-speed video streams. Computer and Graphics. 2010, 34(5): 575-584.

[6] Wu H, Rubinstein M, Shih E, et al. Eulerian Video Magnification for Revealing Subtle Changes in the World. ACM Transactions on Graphics. 2012, 31(4): 65.

[7] Wadhwa N, Rubinstein M, Durand F, et al. Phase-Based Video Motion Processing. ACM Transactions on Graphics. 2013, 32(4).

[8] Bai J, Agarwala A, Agrawala M, et al. Selectively De-Animating Video. ACM Transactions on Graphics. 2012, 31(4): 66.

[9] Liu F, Niu Y, Jin H. Joint Subspace Stabilization for Stereoscopic Video. In: IEEE International Conference on Computer Vision. Sydney, Australia: IEEE Computer Society, 2013. 73-80.

[10] Liu F, Gleicher M, Wang J, et al. Subspace Video Stabilization. ACM Transactions on Graphics. 2011, 30(1): 4.