

# Detecting and Removing the Motion Blurring from Video Clips

Yang shen

Department of Computer science/Shanghai Jiaotong University, Shanghai, China  
sheny@sjtu.edu.cn

Lizhuang ma

Department of Computer science/Shanghai Jiaotong University, Shanghai, China

**Abstract**—In this paper, we give a framework to deblur the blurry frame in a video clip. Two kinds of motion blurring effects can be removed in the video, one is the blurring effect caused by hand shaking, the other is the blurring effect caused by a fast moving object. For the blurring caused by hand shaking, PSF is estimated by comparing the stable area in blurry frame and non-blurry frame, so the Richardson-Lucy algorithm can restore the blurry frame by non-blind deconvolution. We also propose a framework to deblur the motion blurring objects which move fast in the video. The background is reconstructed by the algorithm in each frame, so an accurate matte of blurry object can be extracted to deblur the moving object by alpha matting. Results show our method is effective.

**Index Terms**—Video; Motion deblurring; PSF

## I. INTRODUCTION

Camera shaking, which causes blurry frames in a video sequence, is a chronic problem for photographers. In recent years, a lot of papers discuss how to remove camera shaking from digital photography, [13] [2], but few of them discuss how to remove motion blur from video sequence. In this paper, we discuss how to remove the motion blurring from video.

In image deblurring technique, it is important to get an accurate kernel. Some corresponding methods are discussed in [13]. Li et. al uses a pair of images to estimate the kernel, one is the noise image without blur, the other is the blurry image with long exposure time. In the video, though the frames change with time, we can also find the local area that is nearly invariable in video. By finding the blurry and non-blurry pairs in the local area of different frames, we can estimate the kernel accurately.

We also propose a framework to remove the blurring effect of moving object in the video, we firstly reconstruct the background from video, then we get an accurate alpha matte of blurry object from video. Secondly, we estimate the PSF of blurring matte. Finally, motion blurring can be removed by Richardson-Lucy algorithm. Results show our method is effective.

## II. RELATED WORK

Image deblurring can be categorized into two types: blind deconvolution and non-blind deconvolution.

The former is more difficult since the kernel is unknown. Fergus et. [2] found a method to estimate kernel by natural image statistics. However, very nice results are obtained sometimes, but their method is very time-consuming. Blind deblurring algorithm is based on non-linear optimization, it is complex and hard to optimize. Agrawal et al [9] and Tai et.al [10] use the hybrid camera to deblur a video frame, they use the hybrid camera framework in which a camera simultaneously captures a high resolution image together with a sequence of low-resolution images that are temporally synchronized. In their works, optical flow is derived from the low-resolution images to compute the global motion blur of the high-resolution image, with this computed global motion kernel, deconvolution is performed to correct blur in the high-resolution image. However, their method is based on optical flow which is not accurate in some case when the noise is serious in low resolution image.

Matsushita et.al [7] deblur the blurry frame by transferring and interpolating corresponding pixels from sharper frames. However, their method could not apply to some areas which pixels is hard to find in sharper frames. Lu et .al [13] proposed a method to estimate the kernel by blurry and non-blurry pairs, the hybrid camera framework is used. Our method also uses the blurring and non-blurring pairs to solve the kernel in the video. We get the blur and non-blurry pairs from different frame, and estimate kernel efficiently, so we need not use the hybrid camera system to solve the motion invariant kernel.

Agrawal Amit et al. [1] proposed a method to deblur the blurry moving object in the video they estimate the PSF by alpha matte without hybrid camera system. In this paper, we propose a new framework to restore the blurry object from frame by background reconstruction. With the background reconstruction, we can get accurate trimap automatically, and achieve accurate alpha matte. The results show that our method is effective.

## III. FIND THE BLURRY FRAME AND STABLE AREA

We find the blurry frame from video and estimate the motion invariant kernel in blurry frame. At first, we use the inverse of the sum of squared gradient measure to evaluate the relative blurriness. The blurriness measure is defined by following formula: [6]



Figure 1: the left figure is the blurry frame, the right figure is the non-blurry frame.



Figure 2: the left figure is the stable area in blurry frame, the center figure is the stable area in non-blurry frame. the right figure is the projective translation of the center figure

$$b_t = \frac{1}{\sum_{pt} \{(f_x * I_t)^2 + (f_y * I_t)^2\}} \quad (1)$$

$f_x$  and  $f_y$  are the derivative filters along the x and y directions respectively, This blurriness measure give a relative evaluation of image blurriness. Therefore, this measure is used in a limited number of neighboring frames where significant change is not observed.

Relatively blurry frames are determined by examining  $bt / bt'; t' \in Nt$ , when  $bt / bt'$  is larger than 1, frame  $I_t$  is considered to be sharper than frame  $I_{t'}$ .

By the above method, we find the blurry image from the video clip, and deblur the blurry frame. However, blind deconvolution algorithm is time-consuming and hard to get perfect result because of nonlinear optimization problem. In this paper, we use non-blind deconvolution algorithm, we estimate the kernel of the blurry frame in video clip by following step:

At first, we use SIFT algorithm and optical flow to estimate the stable area from video clip; secondly, we get optimized blurry and non-blurry pairs from stable area in different frame. Finally, we estimate the kernel by blurry and non-blurry pairs in video clip.

In the non-blurry frames, we use the optical flow to estimate the movement of object in the video. The area which is rigid is called the stable area. We assume that the stable areas in different frames which maintain the projective transformation, so we use RANSAC algorithm to find the projective transformation of area between neighbor frames.

Though the optical flow technique can find the projective transformation between neighbor frames, it could not be applied to blurry frames. However, ee find that the SIFT algorithm is robust and effective in matching the feature points between non-blurry frame and blurry frame, so we use SIFT algorithm to match the feature points between each frame. By locating the SIFT feature points of stable area between blurry frame and non-blurry frame, we can align the non-blurry frame and blurry frame and get the blurry and non-blurry pairs in the video. By the blurry and non-blurry pairs[8], we estimate the spatial invariant kernel f accurately by follow formula:

$$E(f) = (\sum \|\delta^* L \otimes f - \delta^* I\|_2^2) + \|f\|_1 \quad (2)$$

The convolution is a linear operator, it can be denoted as multiplication of matrix [13].

$$E(f) = (Af - B\|_2^2) + \|f\|_1 \quad (3)$$



Figure 3: The left are two blurry frames, the center is the kernel solved by Fergus et. al [2]. the right figure is the kernel solved by our method.

$B$  is the vector form of blurry frame,  $f$  is the vector form of kernel,  $A$  is a matrix form of non-blurry image  $L$  which is decided by  $L$  and the size of kernel. Equation 3 is a standard format of the problem defined in [5], and computes the optimal solution with an enhanced interior point method.

We also use the method of Yuan et al.[13] to estimate the kernel, they estimate the kernel by the formula similar with formula 3. Different from formula 3, Yuan use 2- norm of  $f$  instead of 1-norm of  $f$ . However, we find that the method in [13] is not stable, Yuan's method could not converge to a real kernel in many cases.

In kernel estimation stage, only the salient edges have influences on optimization of the kernel because convolution of zero gradients is always zero. To speedup the algorithm, we use the high contrast area in the stable area of blurry frame and un-blurry frame.

In Fig 2, we show the stable area of blurry frame and non-blurry frame. By the alignment of two images, we can solve the kernel of motion blurring.

As shown in Fig 3, we calculate the kernel by method in [2] and our method. Results in Fig 4 show that our kernel is more accurate. Blind deconvolution algorithm is time-consuming. It is obvious that our method is more efficient. We use blind convolution in [2] to deblur the picture, and spend 20 minutes on deblurring

a  $600 * 480$  image with  $20 * 20$  kernel. Our method is based on non-blind deconvolution algorithm. We spend 1 second on solving the kernel, and spend 2 seconds on deblurring the image.

We also use the method of Shan et al.[8] to deblur the image, we use the blind deconvolution algorithm and non-blind deconvolution algorithm of [8]to deblur the frame. We find that their method is hard to adjust the parameter, and we could not get an ideal result in (c) and (d) of Fig 4.

Our method can be used to solve motion variant kernel. When the kernel is spatial variant, we can use the hybrid imaging system to get the kernel. When a blurring frame are recorded by high resolution sensor with long exposure time, a series of non-blurry image are recorded in the low resolution sensor with low exposure time, so we can get the motion path of camera in each discrete time by following equation:

$$I_n(x) = H_n I_0(x) \tag{4}$$

$H_n$  is the homography that project the stable area in frame  $f_0$  to stable area in frame  $f_n$ ,  $I_n(x)$  is the  $n_{th}$  stable area in low resolution sensor, so the spatial variant kernel can be represented by a series of projective transformation:

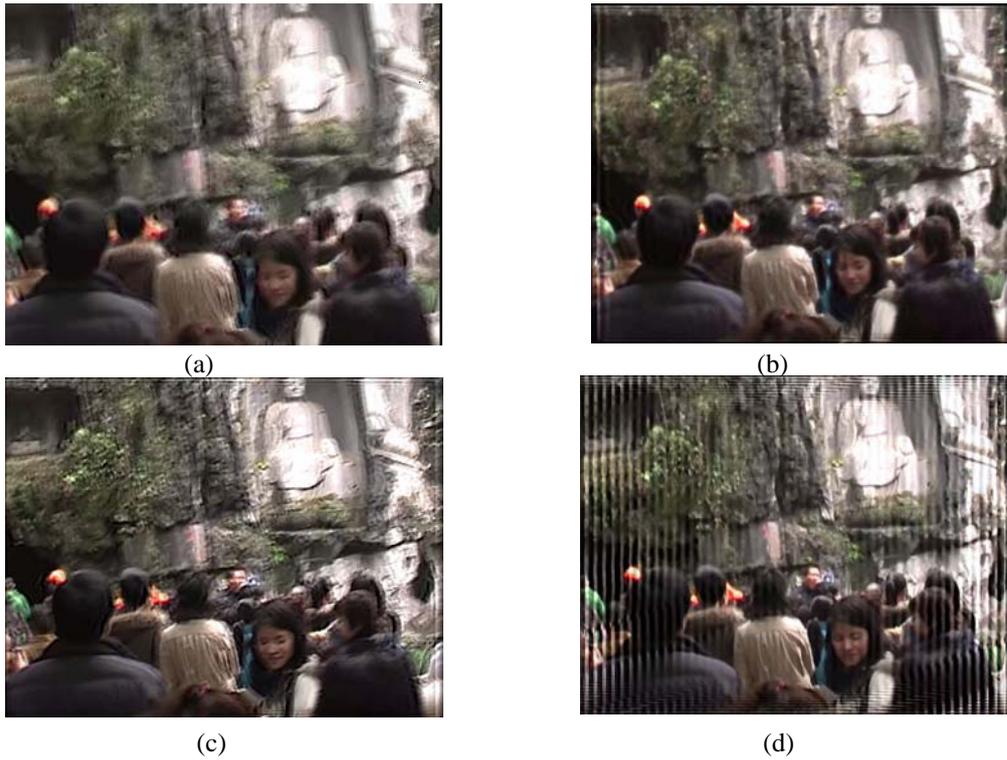


Figure 4: (a) The result of Fergus et al. [2], (b) the result of our method,(c) the result of blind deblurring algorithm in Shan et al.[8] ,(d) the result of non-blind deblurring algorithm in Shan et al.[8]

$$I(x) = \int_0^T \Delta I(x,t) dt = \sum_{i=1}^N \Delta I(x,t_i)$$

$$\Delta I(x,t_i) = \Delta I(\prod_{j=1}^i h_j x, t_0) = \frac{1}{N} I_0(H_i x) \quad (5)$$

$$I(x) = \frac{1}{N} \sum_{i=1}^N I_0(H_i x) \quad (6)$$

$I(x)$  is the blurry stable area in high resolution sensor.

#### IV. DEBLUR THE MOVING OBJECT IN VIDEO FRAME

In this section, we discuss how to remove motion blurring produced by fast movement in video; we only care about the blurring object with spatial-invariant kernel in the frame.

To deblur the blurry object in the video, we firstly extract the blurry object from frame by alpha matting. Alpha matte contains a lot of information which can be used to restore the latent image. For example, accurate alpha matte can help us to get a accurate kernel, Jiaya Jia ya et al. [4] proposed a single image deblurring method, they estimated the PSF by alpha matte.

In some frames of video with fast moving objects, the pixels of structured background and the pixels of blurry object mix together in large field. It is hard to

extract the alpha matte accurately because the background pixel is unknown. However, if the background is known in those areas, we can get a more accurate alpha matte.

Comparing with foreground object which moves fast in the video, background objects move slowly in the video, so we can reconstruct the background object by neighbor frames in video. Different from image completion method, we need to reconstruct the real background in the unknown region of trimap, but not the other patches in video frames.

When background is still, we can reconstruct the background by background reconstruction method. In our paper, we use the method in [3] to reconstruct the background.

Let  $(I_1, I_2, \dots, I_n)$  represent a image sequence,  $f_i(x, y)$  represents the intensity value of the point  $f_i(x, y)$  of the  $i_{th}$  frame where  $i = 0, 1, 2, \dots, N$ . At each point  $(x, y)$  of video, a indicator function is defined by following formula:

$$a_i(x, y) = \begin{cases} 1 & \text{if } |f_i(x, y) - f_{i-1}(x, y)| > \varepsilon \\ 0 & \text{if } |f_i(x, y) - f_{i-1}(x, y)| \leq \varepsilon \end{cases} \quad (7)$$

At each point  $(x, y)$  of video, we find many stable intervals. In each stable intervals  $Interval_{ab}$ , the

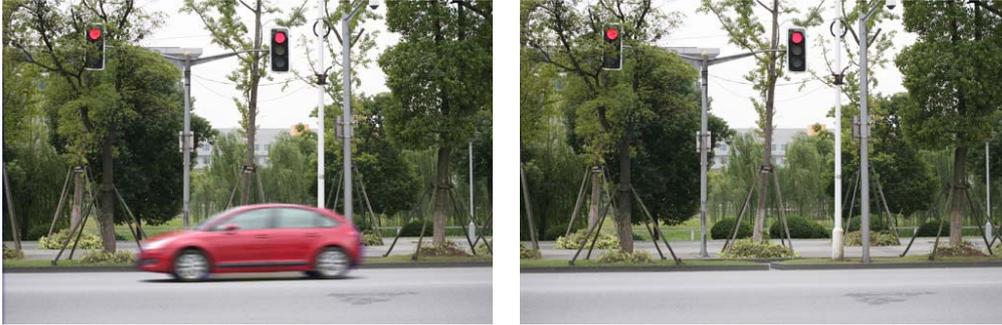


Figure 5: the left figure is the car with motion blurring, the right figure is the background.

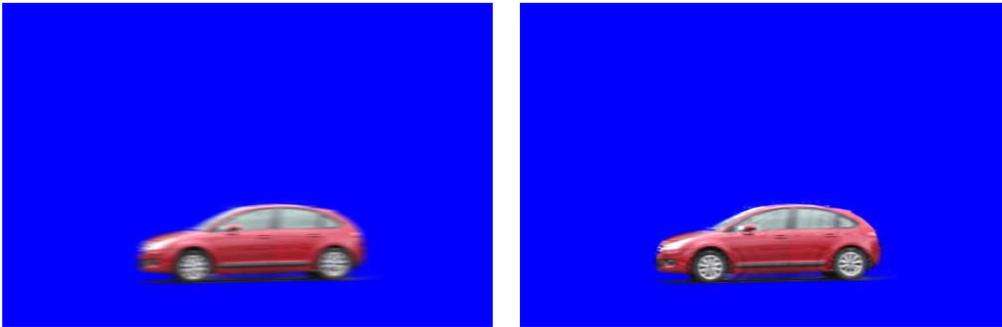


Figure 6: the left figure is the alpha matting by our method, the right figure is debluring result by the RL algorithm

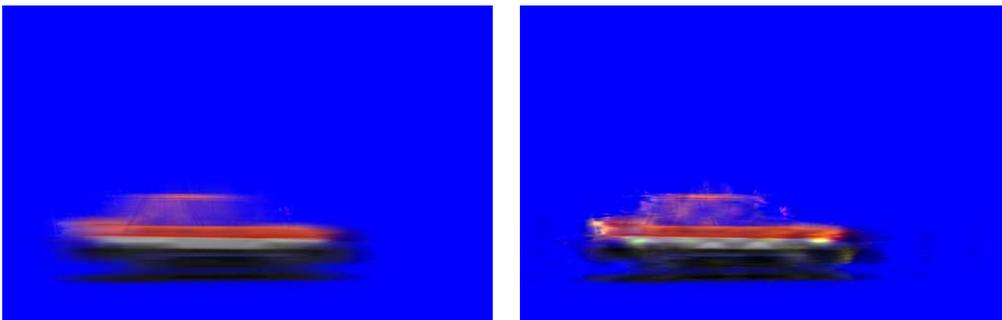


Figure 7: the left figure is the alpha matting of blurry car with large kernel. The right figure is debluring result by the RL algorithm.

intensity of pixels is similar, so  $a_i(x, y) = 0, i \in [a, b]$ . We combine the stable intervals in the video with similar intensity by following formula:

$$\begin{aligned} \text{Interval}_{\text{combine}} &= \sum_{[a_i, b_i] \in \Omega} [a_i, b_i], \\ g(a, b) &= \text{average}(\text{Interval}_{ab}) \\ |g_{ab} - g_{cd}| &< \varepsilon, (ab, cd \in \Omega) \end{aligned} \quad (8)$$

We think that the background is in the interval combined with largest length, so we can estimate the intensity of background by finding the interval combined with the largest length.

When the background is reconstructed, we use robust matting with known background to extract the foreground matte [12]. In this case, robust matting can be simplified. Because the background color is known, the distance ratio  $R_d(F_j, B_j)$  and weight  $w(F_j, B_j)$  in [12] can be simplified by following formula:

$$\begin{aligned} \hat{\alpha} &= \frac{(I - B)(F_j - B)}{\|F_j - B\|^2} \\ R_d(F_j) &= \frac{\|I - (\hat{\alpha}F_j + (1 - \hat{\alpha})B)\|}{F_j - B} \\ w(F_j) &= \exp(-\|F_j - I\|^2 / D_F^2) \end{aligned} \quad (9)$$

$D_B$  is the minimum distance between background sample and the current pixel. The confidence value  $f(B^j)$  for a background pixel  $B^j$  can be represented by following equation

$$f(B^j) = \exp - \frac{R_d(F^j)^2 \cdot w(F^j)}{\sigma^2} \quad (10)$$

The alpha value can be solved by minimizing following formula:

$$J(\alpha) = \alpha^T L \alpha + (\alpha - \hat{\alpha})^T \hat{\Gamma} (\alpha - \hat{\alpha}) \quad (11)$$

$\alpha^T L \alpha$  is the smoothness term defined in [6],  $(\alpha - \hat{\alpha})^T \hat{\Gamma} (\alpha - \hat{\alpha})$  is the sampling term.  $\hat{\alpha}$  is the sampling alpha value with highest confidence  $\hat{f}$ .  $\hat{\Gamma}$  is the diagonal matrix which defines the relative weighting between sampling term and smoothness term. The diagonal element can be  $r * \hat{f}$ . The algorithm relies more on propagation in low confidence regions and relies more on sampling in high confidence regions.

As shown in Fig 5 and Fig 6, compared with robust matting, our matting algorithm achieves more effective result. There are two methods to solve the spatial invariant PSF in video, one method is to solve the PSF by optical flow. PSF can be found by multiplying the image-space object velocity  $v(\text{pix}/\text{ms})$  with the exposure time(ms) [1]. The other method is to solve the PSF by

alpha matte directly [4]. Different from the PSF produced by camera shake, PSF of fast moving object in video is always 1D, 1D kernel  $f$  can be solved by following formula by alpha matte [4]:

$$\begin{aligned} f &= [f_{-n}, \dots, f_0, \dots, f_n]^T \\ f_j &= \alpha_{j+n} - \alpha_{j+n-1} \quad (-n < j < n) \\ &= \alpha_0 \quad (j = -n) \end{aligned} \quad (12)$$

In formula 12,  $f$  is the kernel of motion blurring. In 1-D translational motion,  $f$  is denoted as a  $(2n+1)$  by 1 vector along motion direction,  $\alpha_m$  is the nonzero alpha values along motion direction

When PSF is estimated, we use the Richardson-Lucy algorithm to restore the blurry object in frames which is shown in Fig 6.

We use the Richardson-Lucy algorithm to restore the blurry image. If the kernel is spatial invariant, many enhanced Richardson-Lucy algorithm can be used to solve the deconvolution problem such as [14]. For the spatial variant kernel, we can deblur the blurry frame by Projective Richardson-Lucy algorithm [11].

However, our method could not be used in blurry frame with large kernel. As shown in Fig [7], the deblurring result is shown in right figure. Two reasons lead to bad result, one is that the alpha matting algorithm could not get accurate alpha matte of serious blurring object, the other is that the noise is serious in large blurry object, so we are hard to restore the latent object accurately.

## V. CONCLUSION

In this paper, we propose a new method to deblur the video. We estimate the PSF by stable areas between blurry frame and non-blurry frames, so we can deblur the video frame by non-blind deconvolution method. We also propose a new framework to deblur the blurry object in frames. In the future, we will study how to estimate the motion variant kernel in the image and apply the method to [15][16][17].

## ACKNOWLEDGMENT

This work is supported by National Basic Research Program of China (Program 973) under Grant No. 2006CB-303105, Natural Science Foundation of China (No. 60873136) and Fund for National High Technology Research and Development Program of China (Program 863 No. 2009AA01Z334).

## REFERENCES

- [1] Amit Agrawal, Yi Xu, and Ramesh Raskar. Invertible motion blur in video. *ACM Trans. Graph.*, 28(3):1–8, 2009.
- [2] Rob Fergus, Barun Singh, Aaron Hertzmann, Sam T. Roweis, and William T. Freeman. Removing camera shake from a single photograph. In *SIGGRAPH '06: ACM*

- SIGGRAPH 2006 Papers, pages 787–794, New York, NY, USA, 2006. ACM.
- [3] Zhiqiang Hou. A background reconstruction algorithm based on pixel intensity classification in remote video surveillance system.
- [4] Jiaya Jia. Single image motion deblurring using transparency. In CVPR, 2007.
- [5] Seung-Jean Kim, Kwangmoo Koh, Michael Lustig, and Stephen Boyd. An efficient method for compressed sensing. In ICIP (3), pages 117–120. IEEE, 2007.
- A. Levin, D. Lischinski, and Y. Weiss. A closed form solution to natural image matting. In IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, June 2006.
- [6] Yasuyuki Matsushita, Eyal Ofek, Xiaoou Tang, and Heung yeung Shum. Full-frame video stabilization. In In Proc. Computer Vision and Pattern Recognition, pages 50–57, 2005.
- [7] Qi Shan, Jiaya Jia, and Aseem Agarwala. High-quality motion deblurring from a single image. ACM Transactions on Graphics (SIGGRAPH), 2008.
- [8] A. Agrawal and R. Raskar. Resolving objects at higher resolution from a single motion-blurred image. In CVPR, 2007.
- [9] Yu-Wing Tai, Hao Du, Michael S. Brown, and Stephen Lin. Image/video deblurring using a hybrid camera. In CVPR, 2008.
- [10] Yu-Wing Tai, Hao Du, Michael S. Brown, and Stephen Lin. Correction of spatially varying image and video motion blur using a hybrid camera. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2009.
- [11] Jue Wang and M. F. Cohen. Optimized color sampling for robust matting. In Computer Vision and Pattern Recognition, 2007. CVPR '07. IEEE Conference on, pages 1–8, 2007.
- [12] Lu Yuan, Jian Sun, Long Quan, and Heung-Yeung Shum. Image deblurring with blurred/noisy image pairs. In SIGGRAPH '07: ACM SIGGRAPH 2007 papers, page 1, New York, NY, USA, 2007. ACM.
- [13] Lu Yuan, Jian Sun, Long Quan, and Heung-Yeung Shum. Progressive inter-scale and intra-scale non-blind image deconvolution. ACM Trans. Graph., 27(3):1–10, 2008.
- [14] Kun Xu, Jiaping Wang, Xin Tong, Shi-Min Hu, Baining Guo. Edit Propagation on Bidirectional Texture Functions. Computer Graphics Forum, Special issue of Pacific Graphics 2009, Vol. 28, No. 7, 1871-1877.
- [15] Kun Xu, Yong Li, Tao Ju, Shi-Min Hu, Tian-Qiang Liu. Efficient Affinity-based Edit Propagation using K-D Tree. ACM Transactions on Graphics, Vol. 28, No. 5, Article No. 118.
- [16] Ming-Ming Cheng, Fang-Lue Zhang, Niloy J. Mitra, Xiaolei Huang, Shi-Min Hu. RepFinder: Finding Approximately Repeated Scene Elements for Image Editing. 2010, Vol. 29, No. 3, ACM SIGGRAPH 2010



**Yang Shen** was born in Anhui Province, P.R. China, in 1975. He received the master degree in the field of computer science at Hefei University of Technology

He is currently working toward the Ph.D. degree in the Department of Computer Science and

Technology, Shanghai Jiao Tong University. His research interests include computer graphics and image processing.



**Lizhuang Ma** received his B.Sc. and Ph.D. degrees at Zhejiang University, China in 1985 and 1991, respectively. He was a Post-doctor at the Department of Computer Science of Zhejiang University from 1991 to 1993.

Dr. Ma was promoted as an Associate Professor and Professor in 1993 and 1995, respectively.

He is also the Chairman of the Center of Information Science and Technology for Traditional Chinese Medicine at Shanghai Traditional Chinese Medicine University. His research interests include computer-aided geometric design, computer graphics, scientific data visualization, computer animation, digital media technology, and theory and applications for computer graphics, CAD/CAM.

Dr. Ma was promoted as an Associate Professor and Professor in 1993 and 1995, respectively. He stayed at Zhejiang University from 1991 to 2002. He is now a Full Professor, Ph.D. Tutor, and the Head of Digital Media Technology and Data Reconstruction Lab at the Department of Computer Science and Engineering, Shanghai Jiao Tong University, China since 2002. Dr. Ma has published more than 100 academic research papers in both domestic and international journals.