Practical Super-Resolution from Dynamic Video Sequences

Zhongding Jiang† Tien-Tsin Wong‡ Hujun Bao†
†*State Key Lab of CAD&CG, Zhejiang University, China*
‡*The Chinese University of Hong Kong, China*
{*zdjiang,bao*}*@cad.zju.edu.cn* *ttwong@cse.cuhk.edu.hk*

Abstract

This paper introduces a practical approach for super-resolution, the process of reconstructing a high-resolution image from the low-resolution input ones. The emphasis of our work is to super-resolve frames from dynamic video sequences which may contain significant object occlusion or scene changes. As the quality of super-resolved images highly relies on the correctness of image alignment between consecutive frames, we employ the robust optical flow method to accurately estimate motion between the image pair. An efficient and reliable scheme is designed to detect and discard incorrect matchings which may degrade the output quality. We also introduce the usage of elliptical weighted average (EWA) filter to model the spatially-variant point spread function (PSF) of acquisition system in order to improve accuracy of the model. A number of complex and dynamic video sequences are tested to demonstrate the applicability and reliability of our algorithm.

1 Introduction

Super-resolution is the process of reconstructing a higher resolution image from low-resolution input ones. It received much attention [24, 16, 17, 20, 8, 9, 10, 22, 13, 2, 27, 26, 19, 11] in computer vision and image processing communities over the past decades.

Most of the proposed super-resolution algorithms belong to reconstruction-based algorithms [2] which are based on sampling theorems. However, due to the constraints on the motion models of the input video sequences, it is difficult to apply reconstruction-based algorithms. Most algorithms have either implicitly or explicitly assumed the image pairs are related by a global parametric transformations, which may not be satisfied in dynamic video.

It is challenging to design super-resolution algorithm for arbitrary video sequences. Video frames in general cannot be related through global parametric transformation due to the arbitrary individual pixel movement between image pairs. Hence local motion models, such as optical flow [15], need to be used for image alignment. As proved by Zhao *et al.* [26], an accurate alignment is the key to success of reconstruction-based super-resolution algorithms. We employ the robust optical flow algorithm [4] in our work.

Unfortunately, some phenomena in video cannot be modeled well by optical flow algorithms, such as occlusion/disocclusion, illumination mismatches and motion parallax, etc. Incorrect alignment due to the above phenomena may degrade the quality of super-resolved image. Hence it must be detected and discarded during super-resolution computation.

The point spread function (PSF) of the acquisition system is spatially-variant due to the relative sensor-object motion. In order to better approximate the spatially-variant PSF, we use elliptical weighted area (EWA) filter [12, 14, 28, 29], whose kernel can be elliptical in shape.

Based on the theoretical foundation in [26], we proposed a practical and efficient super-resolution method to tailor dynamic video sequences. The following are two main contributions we made on top of the iterative backward projection (IBP) algorithm [16, 17]:

- An efficient and reliable scheme is designed to detect and discard bad matchings in the initial alignment maps obtained by robust optical flow algorithm. Hence, the quality of super-resolved images can be improved by rejecting such errors.
- The introduction of elliptical weighted area filter (EWA) in modeling the spatially-variant point spread functions (PSF). The elliptical shape of the filter improves the accuracy of model. Therefore, we call our method EWA super-resolution.

The rest of the paper is organized as follows. We give a brief review of the existing work on super-resolution in section 2. EWA super-resolution algorithm is presented in section 3. Experimental results are described in section 4. Finally, we draw the conclusions and point out some future directions in section 5.

2 Related work

Existing super-resolution algorithms can be roughly divided into two main categories. One is reconstruction-based algorithms while the other is learning-based algorithms.

Reconstruction-Based Super-Resolution The base of reconstruction-based super-resolution is uniform/non-uniform sampling theories. It assumes the original high-resolution signal (image) can be well predicted from the low-resolution input samples (images). Most super-resolution algorithms fall into this category. A detail review

Proceedings of the 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'03)
1063-6919/03 $17.00 © 2003 IEEE

can be found in [5]. Among them, the frequency-domain method is a pioneering approach [24]. Iterative backward projection (IBP) algorithm [16, 17] are proposed later. Recently, a unifying framework for super-resolution has been presented. It formulates the problem using matrix-vector notation [8]. To reduce noise and solve singular cases, projection-onto-convex-sets algorithm (POCS) [20] can incorporate prior knowledge into the computation by applying constraints. The super-resolved images can also be regarded as the MAP solution of stochastic optimization, and the prior smoothness assumptions are used to reduce the effects of inconsistent measurements [22, 13, 7]. In most cases, the enforced smoothness constraint suppresses high-frequency components and hence the results are usually blurred. Regularization method can be used when the scene is strongly rigid, such as the case of a binary text image [7]. In [9], super-resolution can be solved by recursively applying Kalman filter. Super-resolution can also be performed simultaneously in time and in space [23]. By applying perturbation theory to the linear system assembled by reconstruction-based algorithms, the explicit limits for a major class of them are provided[18].

Several refinements have been proposed to address the robustness issue of super-resolution algorithms. One approach handles the case of moving object by motion segmentation [10]. An accurate motion segmentation is hence crucial. Unfortunately, accurate segmentation is hard to obtain in the presence of aliasing and noise. Recently, a robust median estimator is used in an iterative super-resolution algorithm [27]. Since it uses the median of the enlarged input images as initial guess, only the resolution of background rather than the whole target image is enhanced.

Learning-Based Super-Resolution This kind of algorithms *create* high-frequency image details by using the learned generative model from a set of training images. Several algorithms have been proposed for specific types of scene, such as faces and text [2, 19]. Recently, Freeman *et al.* [11] proposed a novel approach for interpolating high-frequency details from a training set. Learning-based super-resolution algorithms are awkward to handle the dynamic real-world video sequences.

3 EWA Super-Resolution

The input to our algorithm includes: 1) multiple low-resolution video frames, (including the target frame and its neighboring frames), 2) the desired magnification factor and 3) two elliptical Gaussian filters chosen from the given filter banks. The filters are chosen according to the video quality and the magnification factor. The output is a high-resolution image reconstructed at the target frame.

Our practical super-resolution algorithm is based on IBP. There are two major reasons we use it as our base. Firstly, it is a simple and effective algorithm operating in image domain. Image processing techniques, such as convolution, warping, etc, can be used to accelerate the computation. In contrast, most of other reconstruction-based super-resolution algorithms are formulated as the matrix solving with/without regularization optimization. Although the matrix formulation is beauty in theory, it may not be very practical for implementation. For example, if the image is high in resolution, the matrices involved are usually sparse and large in dimension. As our goal is to develop a practical system, we prefer a stable and less demanding IBP approach. Secondly, accurate motion segmentation, which is still an open problem, is not required in IBP approach.

Our algorithm first attempts to improve the accuracy of motion estimation and reduce the artifacts due to the presence of poor matchings. We employ the robust optical flow technique [4] to establish the correspondence maps between the low-resolution frames and the high-resolution image. An efficient and reliable scheme is designed to measure the quality of alignment between image and its warped counterpart. An alignment quality map is hence recorded. A pixel in low-resolution frame is then classified as valid if its value in this map exceeds a threshold. Only the residual values at the valid pixels are used (backward projected) in the iterative reconstruction of high-resolution image. This rejection scheme allows us to handle local model inconsistencies, such as highlights, large occlusion/disocclusion, and motion parallax.

Secondly, we attempt to approximate the spatially-variant PSF and the backward projection filter defined at each valid pixel using elliptical weighted area (EWA) filtering techniques. We predefine a list of typical Gaussian filters to approximate the PSFs of acquisition systems. According to the input video quality and the desired magnification factor, the user selects one of them as input to our algorithm. The spatially-variant PSF for each valid pixel is computed by warping the selected filter kernel according to its local affine approximation of the mapping from the low-resolution frame to the desired high-resolution image. The spatially-variant backward projection filter is handled in the same manner.

3.1 Notations

Before we continue, let's first discuss some useful definitions and notations. Figure 1 is referred throughout the problem formulation.

Definition 1 *A low-resolution pixel y is influenced by a high-resolution pixel x, if x is in the footprint [25] of the PSF defined at y.*

Definition 2 *A low-resolution image g is influenced by a high-resolution pixel x, if g contains a pixel y influenced by x.*

The following notations are used with the following meanings throughout this paper:

- x denotes a high-resolution pixel
- y denotes a low-resolution pixel influenced by x
- f denotes the desired high-resolution image

- $f^{(n)}$ is the approximation of f obtained after n-th iterations
- g_k denotes the k-th low-resolution image
- $g_k^{(n)}$ denotes the low-resolution image obtained by projecting $f^{(n)}$ onto low-resolution lattice
- $h_{k,y}^{PSF}$ is the spatially-variant PSF defined at $g_k(y)$
- $h_{k,y}^{BP}$ is the spatially-variant backward projection filter defined at $g_k(y)$
- m_k denotes the mapping from the low-resolution image g_k to the high-resolution image f
- Q_k denotes the alignment quality map between g_k and its warped counterpart

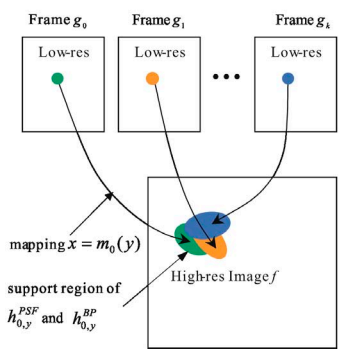

Figure 1: EWA super-resolution. *In each iteration, residual values at valid pixels are backward projected to the high-resolution image. The warped Gaussian filters are computed in the forward and backward projections.*

3.2 Iterative Reconstruction

Figure 1 illustrates the basic idea of our algorithm. It starts with an initial estimation $f^{(0)}$ for the high-resolution image f. Then a forward projection process (from the current high-resolution image to low-resolution images) is carried out to obtain a simulated low-resolution image. A simulated pixel value $g_k^{(0)}(y)$ can be obtained by this forward projection for each valid low-resolution pixel $g_k(y)$. If $f^{(0)}$ is the true high-resolution image, the simulated pixel value $g_k^{(0)}(y)$ should be identical to the input one $g_k(y)$. The residual value $r_k^{(0)}(y)$ between $g_k(y)$ and $g_k^{(0)}(y)$ is then computed. It is backward projected onto the high-resolution image to improve the approximation in the next iteration $f^{(1)}$. Instead of a naïve isotropic filter kernel, we use an anisotropic EWA filter $h_{k,y}^{BP}$ defined at $g_k(y)$. This process is repeated to minimize the error function $e^{(n)}$,

$$e^{(n)} = \sqrt{\sum_k (g_k(y) - g_k^n(y))^2}$$

The forward projection process can be expressed by

$$g_k^{(n)}(y) = \sum_x f^{(n)}(x) h_{k,y}^{PSF}(m_k(y) - x)$$

The residual after n iterations is computed by

$$r_k^{(n)}(y) = g_k(y) - g_k^{(n)}(y) \tag{1}$$

The iterative update to the high-resolution image f is expressed by

$$f^{(n+1)}(x) = f^{(n)}(x) + \lambda \delta^{(n)}(x) \tag{2}$$

where

$$\delta^{(n)}(x) = \frac{\sum_{k,y} W_k Q_k(y) r_k^{(n)}(y) h_{k,y}^{2BP}(m_k(y) - x)}{\sum_{k,y} W_k Q_k(y) h_{k,y}^{BP}(m_k(y) - x)}$$

In Equation (2), the value of $f^{(n)}(x)$ at each high-resolution pixel x is updated with the contribution of all valid low-resolution pixels y influenced by x. The contribution of the valid pixel $g_k(y)$ is weighted by the frame weight W_k, its alignment quality measurement $Q_k(y)$, and backward projection kernel $h_{k,y}^{BP}$. The frame weight W_k can be defined to be inversely proportional to the temporal distance between the k-th frame and the target frame. The scaling factor λ is to control the increment step size.

3.3 Quality Motion Estimation

The correspondence maps $\{m_k\}$ between the low-resolution frames $\{g_k\}$ and the high-resolution image f are established before the iteration in the following manner. Firstly, the correspondence maps between low-resolution neighboring frames and the target frame (low-resolution as well) are determined. Then they are projected onto the desired high-resolution image using a Gaussian pyramid. From now on, 'inspection image' denotes one of the low-resolution neighboring frames while 'reference image' denotes the target low-resolution frame. Correspondence map between the inspection and reference frames is established using optical flow [15], since this allows both camera and independent object motions. One of the sophisticated optical flow techniques is robust optical flow proposed by Black *et al.* [4]. It employs statistics to avoid large errors caused by outliers and to allow for discontinuities in the flow field. Using multi-resolution approach, it can also handle large motion. Hence we adopt it to determine the correspondence maps in a robust manner.

Even a robust optical flow technique is used, the resultant correspondence maps in general may still contain flaws due to the violation of basic assumptions (such as Lambertian surfaces). Therefore, a quality measure of correspondence is needed. We compute an alignment quality map [21] for a pair of images. Based on this map, a low-resolution pixel is classified as valid if the corresponding value exceeds the threshold. Only the residual values (Equation 1) of valid pixels are backward projected to improve the high-resolution image during the iterative reconstruction.

The alignment quality map is computed between the original image and the warped counterpart. We adopt a quality measure based on correlation coefficient of the original image with the warped image [6]. The correlation coefficient assigns a value in [-1, 1] to each pixel. Two variations of correlation coefficient are incorporated. We first compute the variances within the correlation window of each pixel. The image variance and mean are used to compute a

mean normalized variance. Then, the quality measure *Cval* is computed as follows,

if $((\sigma_1^2 \leq T_1$ and $\sigma_2^2 \leq T_1)$ or
$(\sigma_{N1}^2 \leq T_2$ and $\sigma_{N2}^2 \leq T_2))$ then
{
if $(|\mu_1 - \mu_2| \leq k\mu_1)$ then
$Cval = 1.0$;
else
$Cval = 0.0$;
}
else
$Cval = \frac{\sum_p (I_1(p)-\mu_1)(I_2(p)-\mu_2)}{N\sigma_1\sigma_2}$

where $I_i(p)$ returns the pixel value at p in image I_i; σ_1^2 and σ_2^2 are the respective image variances within the correlation window; μ_1 and μ_2 are the respective means; $\sigma_N = \sigma^2/(\mu^2 + c)$ is the mean normalized variance with μ as the mean and c as a stabilizing constant to avoid singularity due to zero mean; T_1, T_2 and k are parameters, and N is the number of pixels in the correlation window.

3.4 Elliptical Weighted Area Filtering

In this subsection we discuss how to compute $h_{k,y}^{PSF}$ and $h_{k,y}^{BP}$ which play important roles in our super-resolution algorithm. IBP algorithm assumes the PSF and backward projection filter are fixed in their kernel shapes and sizes. PSF is estimated by studying a picture of a known object [16]. However, this approach may not be feasible in general, because the original acquisition system must be known. Furthermore, the PSF is spatially-variant due to the relative sensor-objection motion in real-world video sequences. Estimation of PSF from the input images is hard due to the complexity of arbitrary real-world motion.

We design the PSF and the backward projection filter to be Gaussian and their warped versions are EWA filters [12, 14]. By adopting an anisotropic and spatially varying filter, EWA, we are able to better model the spatially-variant PSF and backward projection filter. EWA changes its supporting region according to the local motion. Additionally, EWA is less computational intensive than the ideal low-pass $sinc$ function. It gives similar result as the $sinc$ function. EWA can be implemented using only a 1D lookup table that can be hardware-accelerated. These nice properties of EWA are the underlying reasons of our method outperforming IBP.

EWA filter was originally introduced to address the aliasing problem in texture mapping [12, 14]. Recently, it has been extended to represent and render texture functions on irregularly point-sampled surfaces [28]. EWA filter has also been used in volume rendering to reduce aliasing artifacts of splatting [29]. In our algorithm, it is used in the image warping between the low-resolution images $\{g_k\}$ and the high-resolution image f. Valid pixel $g_k(y)$ is regarded as a circular region, and its support region in f is hence an ellipse whose orientation is determined by the mapping $m_k(y)$ as illustrated in Figure 2.

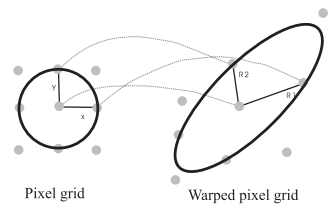

Figure 2: EWA filter for super-resolution.

An elliptical Gaussian $G_V(y)$ with variance matrix V is defined as:

$$G_V(y) = \frac{1}{2\pi|V|^{\frac{1}{2}}} e^{-\frac{1}{2}yV^{-1}y^T}$$

where $|V|$ is the determinant of V. We denote the variance matrices of input PSF $h^{PSF}(y)$ and backward projection filter $h^{BP}(y)$ as V^{PSF} and V^{BP}, respectively. If we apply a warp $u = yM$, where M is a 2×2 matrix, the warped PSF and warped backward projection filter can then be computed as follows:

$$h_w^{PSF}(u) = G_{V^{PSF}}(uM^{-1}) = |M| G_{M^T V^{PSF} M}(u) \quad (3)$$

$$h_w^{BP}(u) = G_{V^{BP}}(uM^{-1}) = |M| G_{M^T V^{BP} M}(u) \quad (4)$$

Thus $h_{k,y}^{PSF}(u) = h_w^{PSF}(u)|_{g_k(y)}$ and $h_{k,y}^{BP}(u) = h_w^{BP}(u)|_{g_k(y)}$ are determined by the matrix M defined at the low resolution pixel $g_k(y)$.

Matrix M can be computed in the following manner. For each low-resolution pixel $g_k(y) = y_0$, we find its corresponding point u_0 in the reference image. Then we establish the local coordinate systems in the inspection and reference images centered at y_0 and u_0, respectively. If we define $u = yM$, then M is the Jacobian of the mapping from local y space to local u space, and hence $M = \begin{pmatrix} u_x & v_x \\ u_y & v_y \end{pmatrix}$.

4 Results

To verify our algorithm, we tested it with five video clips, namely *Beach* (720×480, 30 fps, Figure 3), *Canyon* (720×480, 30 fps, Figure 4), *Carphone* (176×144, 30 fps, Figure 8(a)), *Glasgow* (176×144, 30 fps, Figure 7(a)), and *Salesman* (176×144, 30 fps, Figure 7(d)). All experiments and timing statistics are carried out and recorded on executing the unoptimized code on Dell Inspiron 8100 with Pentium III 1.0 GHz CPU and 512 MB memory.

We first compare our method with naïve bicubic interpolation in Figure 5. The *Beach* example (Figure 3) shows the target frame in the video clip with panning motion. To generate our result, two neighboring low-resolution frames plus the target frame are used. The displacement between the consecutive frames is almost 8 pixels in some cases. The super-resolved image is magnified two times, *i.e.* 1440 × 960 in resolution. The program takes 84 sec. to generate

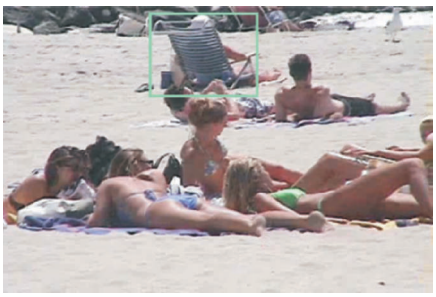

Figure 3: One target frame of *Beach*.

Figure 4: One target frame of *Canyon*.

(a) Bicubic interpolation (b) Our algorithm (c) Sharpened (a)

Figure 5: Comparison with bicubic interpolation. In this experiment, we magnify the target frame of *Beach* two times in both dimensions using bicubic interpolation (a) and our algorithm (b). The result by bicubic interpolation is blocky and blurry. With sharpening (c), errors are exaggerated. Our algorithm generates image with better quality.

(a) Bicubic interpolation (b) Our algorithm (c) Ground truth image

Figure 6: Comparison with the ground truth image. The peak signal-to-noise ratio (PSNR) values for bicubic interpolation and our algorithm are 37.78 dB and 38.81 dB, respectively.

the result. We blow up part of the image to highlight the difference between image by our method (Figure 5(b)) and that of bicubic interpolation (Figure 5(a)). Result from bicubic interpolation exhibits blocky artifact (see the back of the chair) when comparing with our result. One may argue that by sharpening the bicubic interpolated image, a better image may be resulted. Figure 5(c) shows the result of sharpening Figure 5(a). Although the image looks sharper, the quality is not improved because error is also exaggerated.

In the *Canyon* example, we compare our method with the ground truth image. The low-resolution input frames are simulated by down-sampling the original frames (original resolution: 720×480) to 360×240. One target frame of the down-sampled version is shown in Figure 4. We blow up part of the image to highlight the difference between image by our method (Figure 6(b)), that of bicubic interpolation (Figure 6(a)) and the ground truth image (Figure 6(c)). Experimental result shows that image generated by our algorithm outperforms that of bicubic interpolation by 1.03 dB in terms of peak signal-to-noise ratio (PSNR).

We also compare our method with the original iterative backward projection algorithm (IBP). Two examples, *Glasgow* (Figure 7(a)) and *Salesman* (Figure 7(d)), are tested as there is fast local motion in both videos. Again, we blow up the region containing quick motion to highlight the difference between our result and that from IBP. In both examples (Figure 7(b) & (e)), IBP returns erroneous results (see the left person in Figure 7(b) and the moving hand in Figure 7(e)) due to its inability to address the local motion. By accounting local motion with optical flow, our method correctly aligns the pixels and generates better super-resolved images (Figure 7(c) & (f)). It takes 25 seconds to reconstruct the super-resolved images in both examples.

Finally, we compare our algorithm with robust super-resolution [27] in the *Carphone* example (Figure 8(a)). There is fast local motion and occlusion around the finger. We magnify it two times in both dimensions. Four neighboring frames and small filter kernels are used to compute the super-resolved image (Figure 8(b)). The corresponding neighboring frames are not shown due to the page limit. We carefully choose program parameters to explicitly rejecting bad matchings. Our algorithm generates better result than that of robust super-resolution (Figure 8(c)).

Our algorithm can treat mild motion blur and spatially varying blur in real-world video clips. The input video clip in Figure 3 presents much noise and temporal aliasing. We use large filter kernels to generate the result in Figure 5(b). However, severe blurring needs more efforts[3].

5 Conclusions

In this paper, we present a practical super-resolution algorithm that is capable of reconstructing high-resolution images from complex and dynamic video sequences, which may contain large object occlusion and scene changes. By integrating the robust optical flow and EWA filtering techniques into the iterative reconstruction process, the super-resolved images are generated in a short period of time. Our EWA super-resolution algorithm is a practical solution to the super-resolution problem due to its robustness, simplicity and efficiency. A number of complex and dynamic video

(a) Target frame of *Glasgow* (b) IBP algorithm (c) Our algorithm

(d) Target frame of *Salesman* (e) IBP algorithm (f) Our algorithm

Figure 7: Comparison with IBP algorithm. Image (a) is the target frame from *Glasgow* video and image (d) is from *Salesman* sequence. Both clips are in QCIF resolution (176×144). We magnify the images two times using IBP ((b)&(e)) and our algorithm ((c)&(f)). Parts of the super-resolved images are blown up for clarity. Images by IBP exhibit apparent errors at the left-most person in (b) and at the hand in (e).

(a) Target frame (b) Our algorithm (c) Robust super-resolution

Figure 8: Comparison with robust super-resolution. Image (a) is one target frame from *Carphone* sequence which contains fast local motion and occlusion around the finger. Our super-resolved image (b) is generated by carefully choosing program parameters to explicitly rejecting bad matchings. Image (c) is the result produced by robust super-resolution. Our image is visually better than that of robust super-resolution. However, you can still find few artifacts around the finger in (b) when examining it carefully.

sequences are tested to demonstrate the applicability of our algorithm. The performance of our algorithm depends on the accuracy of parameter estimation. Estimating those parameters is difficult in real-world cases. Nevertheless, our algorithm can still produce better results than those of bicubic interpolation with/without sharpening.

To further improve the speed performance, we are now investigating less computational intensive optical flow algorithm, such as hierarchical Lucas-Kanade optical flow with multi-window techniques. Another direction to speed up is to accelerate the evaluation of EWA by using programmable graphics hardware. This will be a cost-effective direction to achieve real-time performance. Quantization information in the compressed video bitstream provides additional information[1]. We are currently working on incorporating it into our algorithm to further improve the image quality.

Acknowledgments

This project is supported in part by the Research Grants Council of the Hong Kong Special Administrative Region, under RGC Earmarked Grants(Project No. CUHK4186/00E), the Natural Science Foundation for Innovative Research Groups of China(Project No.60021201) and the 973 program of China (Project No.2002CB312104).

References

[1] Y. Altunbasak, A. Patti and R. Mersereau. Super-Resolution Still and Video Reconstruction From MPEG-Coded Video. *IEEE Trans. on Circuits and Systems for Video Technology*, **12**(4):217–226, 2002.

[2] S. Baker and T. Kanade. Limits on Super-resolution and How to Break Them. *IEEE Trans. Pattern Anal. and Machine Intell*, **24**(9):1167–1183, 2002.

[3] B. Bascle, A. Blake and A. Zisserman. Motion Deblurring and Super-resolution from an Image Sequence. *Proc. ECCV'96*, pp.573–582, 1996.

[4] M. Black and P. Anandan. The robust estimation of multiple motions: Parametric and piecewise-smooth flow fields. *Journal of Computer Vision and Image Understanding*, **63**(1):75–104, 1996.

[5] S. Borman and R. Stevenson. Spatial Resolution Enhancement of Low-Resolution Image Sequences: A Comprehensive Review with Directions for Future Research. *Technical Report, University of Notre Dame*, 1998.

[6] L. Brown. A survey of image registration techniques. *ACM Computing Surveys*, **24**(4):325–376, 1992.

[7] D. Capel and A. Zisserman. Super-Resolution enhancement of text image sequence. *Proc. ICPR'2000*, pp.600–605, 2000.

[8] M. Elad and A. Feuer. Restoration of Single Super-resolution Image from Several Blurred, Noisy and Down-sampled Measured Images. *IEEE Trans. on Image Processing*, **6**(12):1646–1658, 1997.

[9] M. Elad and A. Feuer. Super-resolution reconstruction of image sequences. *IEEE Trans. Pattern Anal. and Machine Intell*, **21**(9):817–834, 1999.

[10] P. Eren, M. Sezan and A. Tekalp. Robust, object based high resolution image reconstruction from low-resolution video. *IEEE Trans. on Image Processing*, **6**(10):1446–1451, 1997.

[11] W. Freeman, T. Jones, and E. Pasztor. Example-Based Super-Resolution. *IEEE Computer Graphics and Applications*, **22**(2):56–65, 2002.

[12] N. Greene and P. Heckbert. Creating raster ominimax images from multiple perspective views using the elliptical weighted average filter. *IEEE Computer Graphics and Applications*, **6**(6):21–27, 1986.

[13] R. Hardie, K. Barnard and E. Armstrong. Joint MAP Registration and High-resolution Image Estimation Using a Sequence of Undersampled Images. *IEEE Trans. on Image Processing*, **6**(12):1621–1633, 1997.

[14] P. Heckbert. Fundamentals of Texture Mapping and Image Warping. *Master's thesis*, University of California at Berkeley, Department of Electrical Engineering and Computer Science, June 17, 1989.

[15] B. Horn. *Robot Vision*. MIT Press, 1986.

[16] M. Irani and S. Peleg. Improving Resolution by Image Registration. *Journal of Computer Vision, Graphics, and Image Processing*, **53**(3):231–239, 1991.

[17] M. Irani and S. Peleg. Motion Analysis for Image Enhancement: Resolution, Occlusion, and Transparency. *Journal of Visual Communication and Image Representation*, **4**(4):324–335, 1993.

[18] Z. Lin and H. Shum. On the Fundamental Limits of Reconstruction-Based Super-resolution Algorithms. *Proc. CVPR'2001*, pp.1171–1176, 2001.

[19] C. Liu, H. Shum and C. Zhang. A Two-Step Approach to Hallucinating Faces: Global Parametric Model and Local Non-Parametric Model. *Proc. CVPR'2001*, pp.192–198, 2001.

[20] A. Patti, M. Sezan and A. Tekalp. Superresolution video reconstruction with arbitrary sampling lattices and nonzero aperture time. *IEEE Trans. on Image Processing*, **6**(8):1064–1076, 1997.

[21] H. Sawhney, Y. Guo, K. Hanna, R. Kumar, S. Adkins and S. Zhou. Hybrid Stereo Camera: An IBR Approach for Synthesis of Very High Resolution Stereoscopic Image Sequences. *Proc. SIGGRAPH'2001*, pp.451–460, 2001.

[22] R. Schultz and R. Stevenson. Extraction of High-resolution Frames from Video Sequences. *IEEE Trans. on Image Processing*, **5**(6):996–1011, 1996.

[23] E. Shechtman, Y. Caspi and M. Irani. Increasing Space-time Resolution in Video. *Proc. ECCV'2002*, pp.753–768, 2002.

[24] R. Tsai and T. Huang. Multi-frame Image Restoration and Registration. *Advances in Computer Vision and Image Processing*, vol.1, pp. 317–339, 1984.

[25] L. Westover. Footprint Evaluation for Volume Rendering. *Computer Graphics(SIGGRAPH'90)*, **24**(4):367–376, 1990.

[26] W. Zhao, H. Sawhney. Is Super-Resolution with Optical Flow Feasible? *Proc. ECCV'2002*, vol.1, pp.599–613, 2002.

[27] A. Zomet, A. Rav-Acha and S. Peleg. Robust Super-resolution. *Proc. CVPR'2001*, pp.645–650, 2001.

[28] M. Zwicker, H. Pfister, J. Baar and M. Gross. Surface Splatting. *Proc. SIGGRAPH'2001*, pp.371–378, 2001.

[29] M. Zwicker, H. Pfister, J. Baar and M. Gross. EWA Volume Splatting. *Proc. IEEE Visualization'2001*, pp.29–36, 2001.