Chapter 2

SUPER-RESOLUTION FROM MULTIPLE IMAGES HAVING ARBITRARY MUTUAL MOTION

Assaf Zomet and Shmuel Peleg
School of Computer Science and Engineering
The Hebrew University of Jerusalem
91904 Jerusalem, Israel

1. INTRODUCTION

Video sequences usually contain a large overlap between successive frames, and regions in the scene are sampled in several images. This multiple sampling can sometime be used to achieve images with a higher spatial resolution. The process of reconstructing a high resolution image from several images covering the same region in the world is called *Super Resolution.* Additional tasks may include the reconstruction of a high resolution video sequence [Elad and Feuer, 1996], or a high resolution 3D model of the scene [Smelyanskiy et al., 2000].

A common model for super resolution presents it in the following way: The low resolution input images are the result of projection of a high resolution image onto the image plane, followed by sampling. The goal is to find the high resolution image which fits this model. Formulating it in mathematical language:
Given K images $\{X_L^{(n)}\}_{n=1}^K$ of size $M_1 \times M_2$, find the image X_H of size $N_1 \times N_2$, which minimizes the Error function:

$$E(X_H) = \sum_{n=1}^{K} \| P_n(X_H) - X_L^{(n)} \|^2$$

where:

1. $\| \cdot \|$ - Can be any norm, usually ℓ^2.
2. $P_n(X_H)$ is the projection of X_H onto the coordinate system and sampling grid of image $X_L^{(n)}$.

When this optimization problem does not have a single solution, additional constraints may be added, expressing prior assumptions on X_H, such as smoothness.
The projection $P_n(X_H)$ is usually modeled by four stages:

1. *Geometric Transformation*
2. *Blurring*
3. *Subsampling*
4. *Additive Noise*

The major differences between most modern algorithms are in the optimization technique used for solving this set of equations, the constraints on X_H which are added to the system, and the modeling of the geometric transformation, blur and noise.

1.1 MODELING THE IMAGING PROCESS

Geometric Transformation. In order to have a unique super resolved image X_H, the coordinates system of X_H should be determined. A natural choice would be the coordinates system of one of the input images, enlarged by factor q, usually by two. The geometric transformation of X_H to the coordinates of the input images is computed by finding the motion between the input images. Motion computation and image registration are beyond the scope of this paper. It is important to mention that high accuracy of registration is crucial to the success of super resolution. This accuracy can be obtained when and assumption on the motion model holds, such as an affine or a planar-projective transformation. For highly accurate model-based methods, see [Bergen et al., 1992, Sawhney and Kumar, 1999, Zelnik-Manor and Irani, 2000].

Blur. Image blur can usually be modeled by a convolution with some low-pass kernel. This space-invariant function should approximate both the blur of the optics, and the blur caused by the sensor. The spectral characteristics of the kernel determine whether the super resolution problem is uniquely solvable: If some (high) frequencies of the kernel vanish, then there is no single solution [Papoulis, 1977]. In this case, constraints on the solution may be added [Capel and Zisserman, 2000].
The digitized result of the camera blur is called "The *PSF - Point Spread Function*". Several ways to estimate it are:

1. Use camera manufacturer information (Which is hard to get).
2. Analyze a picture of a known object [Irani and Peleg, 1991, Mann and Picard, 1994].

3. Blind estimation of the PSF from the images [Shekarforoush and Chellappa, 1999].

Some algorithms can handle space-variant blur, such as space-variant-motion blur, air turbulence, etc.

Subsampling. Subsampling is the main difference between the models of super resolution and image restoration. Sometimes the samples from different images can be ordered in a complete regular grid, for example when the motion of the imaging device is precisely controlled. In this case image restoration techniques such as inverse-filtering and De-convolution can be used to restore a high resolution image. In the general case of a moving camera, the super resolution image is reconstructed from samples which are not on a regular grid. Still, image restoration techniques inspire some of the super resolution algorithms [Elad and Feuer, 1999].

Additive Noise. In super resolution, as in similar image processing tasks, it is usually assumed that the noise is additive, normally distributed with zero-mean. Under this assumption, the maximum likelihood solution is found by minimizing the error under Mahalanobis Norm (using estimated autocorrelation matrix), or ℓ^2 norm (assuming uncorrelated "white noise"). The minimum is found by using tools developed for large optimization problems under these norms, such as approximated kalman-filter [Elad and Feuer, 1996], linear-equations solvers [Elad and Feuer, 1999, Patti et al., 1997], etc. The assumption of normal distribution of the noise is not accurate in most of the cases, as most of the noise in the imaging process is non-gaussian (quantization, camera noise, etc.), but modeling it in a more realistic way would end in a very large and complex optimization problem which is usually hard to solve.

1.2 HISTORICAL OVERVIEW

The theoretical basis for super resolution was laid by papoulis [Papoulis, 1977], with *The Generalized Sampling Theorem* . It was shown that a continuous band-limited signal G can be reconstructed from samples of convolutions of G with different filters, assuming some properties of these filters (see 1.1-blur). This idea is simple to generalize to 2D signals (images).

A pioneering algorithm for super resolution for images was presented by Huang & Tsai [Huang and Tsai, 1984], who made explicit use of the aliasing effect, assuming the image is band limited, and the images are noise-free. Kim et. al. generalized this work to noisy and blurred images, using least square minimization [Kim et al., 1990].

Spatial domain algorithm was presented by Ur & Gross [Ur and Gross, 1992]. Assuming a known 2D translation, a fine sample grid image was created from the input images, using interpolation, and the camera blur was canceled using deblurring technique. The above methods assumed blur function which is uniform over all the images, and identical on different images. They were also restricted to global 2D translation.

A Different Approach was suggested by Irani & Peleg [Irani and Peleg, 1991, Irani and Peleg, 1993], based on previous work by Peleg et al. [Keren et al., 1988]. The basic idea, *Iterative Backward Projecting - IBP*, was adopted from computer-aided Tomography (CAT). The algorithm starts with an initial guess $X_H\{0\}$, and iteratively simulate the imaging process, reprojecting the error back to the super resolution image. This algorithm can handle general motion and non-uniform blur function, assuming they can be approximated accurately.

To reduce noise and solve singular cases, several algorithms incorporate prior knowledge into the computation by constraining the solution. Stark & Oskoui [Stark and Oskoui, 1989] and later Pati et. al. [Patti et al., 1997] base their algorithm on a set theoretic optimization tool called POCS (Projection Onto Convex Sets). It is assumed that convex constraints on the solution are known, so that their intersection is also a convex set. The implementation of the algorithm is similar to the IBP algorithm of Irani & Peleg, with a modification in the backprojection stage: The errors in the imaging are projected onto the solution image, while keeping the solution in the convex set defined by the constraints. Pati et. al. also added motion blur to the imaging process model.

Markov Random Field was also used to regularize super resolution. Shekarforoush et. al. [Berthod et al., 1994] formulated the super resolution problem in probabilistic bayesian framework, and used MRF for modeling the prior, and finding the solution. Similar formulation was presented by Schultz & Stevenson [Schultz and Stevenson, 1996], who use prior on the edges and smoothness of the image to compensate for bad motion estimation, based on Huber-Markov Random Field formulation.

There is a great similarity between super resolution and image restoration, and indeed many of the super resolution techniques are adopted from image restoration. A unifying framework for super resolution as a generalization of image restoration was presented by Elad & Feuer [Elad and Feuer, 1999]. Super resolution was formulated using matrix-vector notations, and it was shown that existing super resolution techniques are actually variations of standard quadratic minimization techniques for solving linear equations sets. Based on this analysis, they proposed other sparse matrix optimization methods for the problem.

Finally, an analytical probabilistic method was recently developed by Shekarforoush & Chellapa [Shekarforoush and Chellappa, 1999]. They proved that super resolution image can be directly constructed by a linear combination of a basis which is biorthogonal to the PSF function. The combination coefficients are the input images intensity values. They also presented an algorithm for the estimation of the camera PSF from the images.

2. EFFICIENT GRADIENT-BASED ALGORITHMS

2.1 MATHEMATICAL FORMULATION

Super resolution can be presented as a large sparse linear optimization problem, and solved using explicit iterative methods [Elad and Feuer, 1999, Capel and Zisserman, 1998, Capel and Zisserman, 2000]. In the presented framework a matrix-vector formulation is used in the analysis [Elad and Feuer, 1999], but the implementation is by standard operations on images such as convolution, warping, sampling, etc. Altering between the two formulations, a considerable speedup in the super resolution computation is achieved, by taking advantage of the two worlds: Implementing advanced gradient based optimization techniques (such as conjugate gradient), while computing the gradient in an efficient manner, using basic image operations, instead of sparse matrices multiplications.

In the analysis part images are represented as column vectors. (with any arbitrary order of the pixels). Basic image operations such as convolution, subsampling, upsampling and warping are linear, and thus can be represented as matrices operating on these vector images.

The image formation process can be formulated in the following way [Elad and Feuer, 1999]:

$$\underline{x}_L^{(n)} = DH_nW_n\underline{x}_H + \underline{e}_n$$

where:

- $\underline{x}_H$ is the high resolution image X_H of size $[N_1 \times N_2]$, reordered in a vector.
- $\underline{x}_L^{(n)}$ is the n-th image of size $[M_1 \times M_2]$, reordered in a vector.
- $\underline{e}_n$ is the normally distributed additive noise in the n-th image, reordered in a vector.
- W_n is the geometric warp matrix, of size $[N_1N_2 \times N_1N_2]$

- H_n is the blurring matrix, of size $[N_1N_2 \times N_1N_2]$
- D is the decimation matrix, of size $[M_1M_2 \times N_1N_2]$

Stacking the vector equations from the different images into a single matrix-vector:

$$\begin{bmatrix} \underline{x}_L^{(1)} \\ \vdots \\ \underline{x}_L^{(K)} \end{bmatrix} = \begin{bmatrix} DH_1W_1 \\ \vdots \\ DH_KW_K \end{bmatrix} \underline{x}_H + \begin{bmatrix} \underline{e}_1 \\ \vdots \\ \underline{e}_k \end{bmatrix} \Longleftrightarrow \underline{x}_L = A\underline{x}_H + \underline{e}$$

For practical reasons it is assumed the noise is uncorrelated and has uniform variance. In this case, the maximum likelihood solution is found by minimizing the functional:

$$E(\underline{x}_H) = \frac{1}{2} \parallel \underline{x}_L - A\underline{x}_H \parallel^2$$

taking the derive of E with respect to $\underline{x}_H$, and setting the gradient to zero:

$$\nabla E = 0 \Longrightarrow A^T(A\underline{x}_H - \underline{x}_L) = 0 \Longleftrightarrow \sum_{n=1}^{K} W_n^T H_n^T D^T (DH_nW_n\underline{x}_H - \underline{x}_L^{(n)}) = 0$$

Gradient-based iterative methods can be used without explicit construction of these large matrices. Instead, the multiplication with A and A^TA is implemented using only image operations such as warp, blur and sampling.

The matrix A^TA operates on vectors $\underline{x}_H$, corresponding to an image of the size of the super resolution solution X_H,

$$A^TA\underline{x}_H = \sum_{n=1}^{K} W_n^T H_n^T D^T D W_n T_n \underline{x}_H$$

The matrix A^T operates on vectors $\underline{x}_L$, stacking of the input images $X_L^{(1)} \ldots X_L^{(K)}$ reordered in column vectors $\underline{x}_L^{(1)} \ldots \underline{x}_L^{(K)}$

$$A^T\underline{x}_L = \sum_{n=1}^{K} W_n^T H_n^T D^T \underline{x}_L^{(n)}$$

The matrices W_n, H_n, D model the image formation process, and their implementation is simply the image warping, blurring and subsampling respectively. The implementation of the transpose matrices is also very simple:

- D^T is implemented by upsampling the image without interpolation, i.e. by zero padding.

- H_n^T - For a convolution blur, this operation is implemented by convolution with the flipped kernel, i.e.if $h(i,j)$ is the imaging blur kernel, then the flipped kernel $\hat{h}$ satisfies $\forall i,j, \hat{h}(i,j) = h(-i,-j)$. For space-variant blur H_n^T is implemented by forward projection of the intensity values, using the weights of the original blur filter.

- W_n^T - If W_n is implemented by backward warping, then W_n^T should be the forward warping of the inverse motion.

The simplest implementation of this framework is using Richardson iterations [Kelley, 1995], a from of steepest-descent with iteration step:

$$\underline{x}_H\{m+1\} = \underline{x}_H\{m\} + \sum_{n=1}^{K} W_n^T H_n^T D^T (\underline{x}_L^{(n)} - DH_n W_n \underline{x}_H\{m\})$$

This is a version of the Iterated Back Projection [Irani and Peleg, 1991], using a specific blur kernel and forward warping in the back projection stage.

The ability to compute the gradient of the super resolution functional by image operations opens possibilities to efficiently use advanced gradient-based optimization techniques:

- Fast non-constrained optimization. An example is the Conjugate-Gradient method, which is elaborated in the following section.

- Constrained minimization, bounding the solution to a specific set. An example for this is the POCS solution [Patti et al., 1997, Stark and Oskoui, 1989].

- Constrained minimization, using linear regularization term. The regularization operator should be also easily implemented using image operations. An example is given in the following section.

2.2 SUPER RESOLUTION BY THE CG METHOD

The Conjugate Gradient method. To demonstrate the practical benefit of computing the gradient by image operations, a super resolution algorithm using the conjugate-gradient (CG) method was implemented. The CG method is an efficient method to solve linear systems defined by symmetric positive definite matrices.

Definition 1 *Let Q be a symmetric and positive definite matrix. A vector set* $\{V_i\}_{i=1}^n$ *is Q-conjugate if*

$$\underline{V}_i^T Q \underline{V}_j = 0, \forall i \neq j$$

.

A Q-conjugate set of vectors is linearly independent, and thus form a basis. The solution to the linear equation is therefore a linear combination of these vectors. The coefficients α_i are very easily found:

$$Q\underline{X} = \underline{Y} \Longrightarrow Q(\sum_{i=k}^{n} \alpha_k \underline{V}_k) = \underline{Y} \Longrightarrow \alpha_k = \frac{\underline{V}_k^T \underline{Y}}{\underline{V}_k^T Q \underline{V}_k}$$

The CG algorithm iteratively creates a conjugate basis, by converting the gradients computed in each iteration to vectors which are Q-conjugate to the previous ones (e.g. Graham-Shmidt procedure). Its convergence to the solution in n steps is guaranteed, but this is irrelevant to the super resolution problem, since n, the matrix size in super resolution, is huge. Still, the convergence rate of the CG is superior to steepest descent methods. Below is an implementation of CG.

CG($X, Y, Q, \epsilon, mMax$)
solves $QX = Y$, ϵ and $mMax$ limit the number of iterations

1. $r = Y - QX, \rho_0 = \| r \|^2, m = 1$
2. Do While $\sqrt{\rho_{m-1}} > \epsilon \| Y \|^2$ and $m \leq kmax$
   (a) if m = 1 then $p = r$
   else $\beta = \frac{\rho_{m-1}}{\rho_{m-2}}$ and $p = r + \beta p$
   (b) $w = Qp$
   (c) $\alpha = \rho_{m-1}/p^T w$
   (d) $x = x + \alpha p, r = r - \alpha w, \rho_m = \| r \|^2, m = m + 1$

CG super resolution. In the CG implementation to super resolution, the input includes the low resolution images, $\epsilon, mMax$ and the estimated blur function. First the motion between the input images is computed, and an initial estimate to the solution $X_H\{0\}$ is set, for example the average of the bilinearly upsampled and aligned input images.

Then, in order to use the CG code, two functions are implemented, project and backProject. The simple vector operations, such as inner product and multiplication with a scalar are easily translated to operations on images (correlation and multiplication by scalar). The matrix operations are handled in the following way:

- Step 1 - In the super resolution case $b = A^T Y$ and $Q = A^T A$. The code to compute the residual r is therefore:
  r = 0
  for n=1 to K do $r = r$ + backProject($X_L^{(n)}$-project($X_H\{0\}, n), n$)

- steps 2-b is replaced by the following code:
  w=0
  for n=1 to K do $w = w$ + backProject(project(p, n), n)

and The functions backProject,project are simply:

- I_3=project(p, n)
  - I_1 = blur(p, n) $\Longrightarrow$ blur image p by the blur operator H_n (e.g. convolution filter $h(i, j)$)
  - I_2 = backwardWrp(I_1, n) $\Longrightarrow$ Warp I_1 using backward warping, i.e. for each pixel in I_2, find its sub-pixel location in I_1, based on the motion to the n-th image, and use interpolation to set its value.
  - return subsample(I_2) $\Longrightarrow$ decimate the image, to get an image of the size of the input image $X_L^{(n)}$.
- I_1=backProject(p, n)
  - I_2 = upsample(p) enlarge p to the size of the super resolved image, by zero padding.
  - I_1 = forwrdWrp(I_2, n) $\Longrightarrow$ Warp I_1 using forward warping, i.e. for each pixel in I_2, find its sub-pixel location in I_1, based on the motion to the n-th image. Spread the intensity value of the pixel on the pixels of I_2, proportionally to the interpolation coefficients.
  - return blur(I_1, n) $\Longrightarrow$ blur image p by the transpose of the blur operator H_n (e.g. in the case H is defined by a convolution filter $h(i, j)$, its transpose is implemented by convolution with the flipped filter $\hat{h}(i, j) = h(-i, -j)$)

Adding regularization. In many cases the super resolution does not have a unique solution, and the matrix $A^T A$ is not invertible. This can be solved by introducing constraints on the solution, e.g. smoothness. If the constraints f are differentiable, and their derivative can be approximated from the images, then they can be easily combined with our proposed framework, by minimizing:

$$E(\underline{x}_H) = \frac{1}{2}(\| \underline{x}_L - A\underline{x}_H \|^2 + \lambda f(\underline{x}_H))$$

Where λ is the regularization coefficient. Taking the derivative of E with respect to $\underline{x}_H$, results in a set of equations:

$$\nabla E = 0 \Longrightarrow A^T(A\underline{x}_L - \underline{x}_H) + \lambda \nabla f(\underline{x}_H) = 0$$

In each iteration the image corresponding to $\lambda\nabla f(\underline{x}_H\{m\})$ is added to the image corresponding to $A^T A\underline{x}_H\{m\})$, For example, when f can be expressed by a linear operator M:

$$\nabla f(\underline{x}_H) = M^T M\underline{x}_H$$

This image can be computed from the image corresponding to $\underline{x}_H$, by applying the operator M and its transpose. (The implementation of M^T in the image domain is derived similarly to the transpose of the blur operators). The selection of the optimal f and λ is beyond the scope of this paper [Capel and Zisserman, 2000].

3. COMPUTATIONAL ANALYSIS AND RESULTS

To demonstrate the computational benefit of the proposed framework, the running time of the CG super resolution is compared to another image-based non-constrained algorithm, the IBP of Irani & Peleg. Images of a planar scene were captured by a hand held camera, and the projective-planar motion between them was computed. Then both methods of super resolution were applied, and the computation time and results were compared.

The graph in Figure 2.1 presents the projection error E as a function of the running time. The first iteration in the CG method is slower, since it requires additional multiplication with the matrix $A^T A$. The next iterations of both of the methods require a single multiplication with A and A^T, so the running time is similar (with small advantage to the CG method). This means that the comparison of the running time of these algorithm depends mainly on the convergence rate. It is notable in the graph that the convergence of the CG method in the first crucial iterations is much faster, yielding better results in a very short time. This can be further accelerated by using efficient image operations in the computation of the gradient.

The results of the super resolution algorithm are presented in Fig. 2.2. A set of images were captured by a hand-held camera. First the motion between the images was computed [Bergen et al., 1992]. Then, the proposed super resolution algorithm was applied (Fig. 2.2:E). For comparison, the images were enlarged and warped to a common coordinate system, and their median was computed (Fig. 2.2:C). Both the median and the SR improved the readability dramatically. The SR result is sharper than the median. After applying high-pass filter to the median results (Fig. 2.2:D), the readability is improved, but the result is not as good as the SR result.

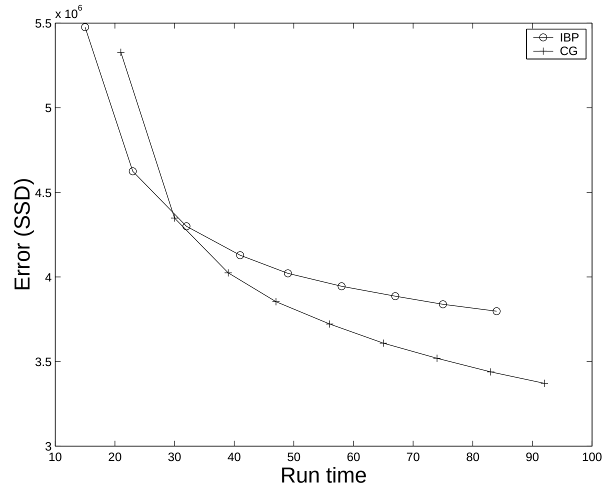

Figure 2.1 The sum of squared error in the images as a function of the running time (measured in seconds). The circles mark the iteration times of the CG algorithm, and crosses mark the iteration times of the IBP algorithm.

4. SUMMARY

Super resolution can be presented as a large sparse linear system. The presented framework allows for solving this system efficiently using image-domain operations. With the rapid advance in computing speed, applying super resolution algorithms on large video sequences becomes more and more practical. There is still work to be done in improving the noise model and the noise sensitivity of super resolution algorithms.

References

[Bergen et al., 1992] Bergen, J., Anandan, P., Hanna, K., and Hingorani, R. (1992). Hierarchical model-based motion estimation. In *European Conf. on Computer Vision*, pages 237–252.

[Berthod et al., 1994] Berthod, M., Shekarforoush, H., Werman, M., and Zerubia, J. (1994). Reconstruction of high resolution 3D visual information using sub-pixel camera displacements. In *IEEE Conf. on Computer Vision and Pattern Recognition*, pages 654–657.

[Capel and Zisserman, 1998] Capel, D. and Zisserman, A. (1998). Automated mosaicing with super-resolution zoom. In *IEEE Conf. on Computer Vision and Pattern Recognition*, pages 885–891.

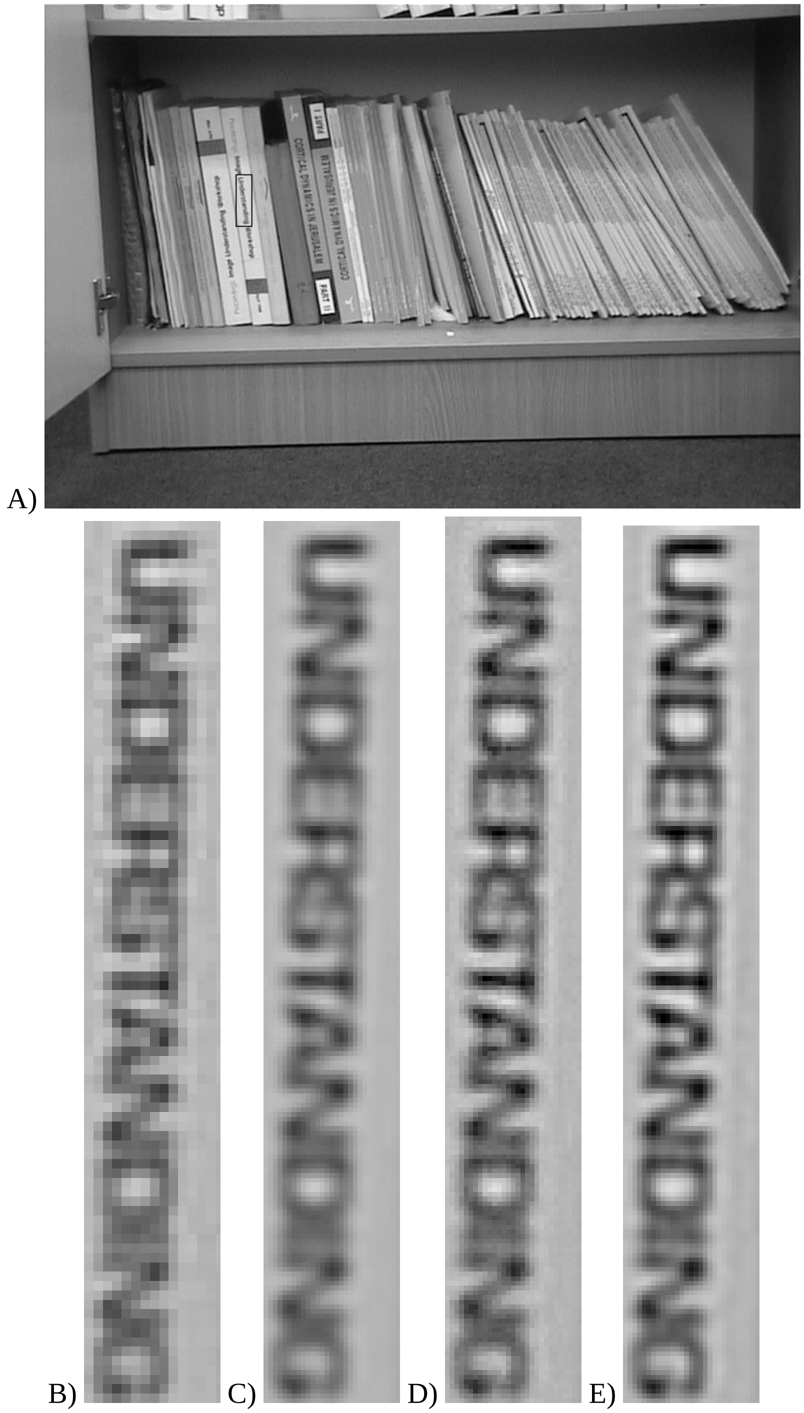

Figure 2.2 Super resolution results.
A) An original frame.
B) A small region from the original frame. The region is marked in Fig. A.
C) The median of the aligned+enlarged images
D) The results of applying high-pass filter to Fig. C.
E) Super resolution.

[Capel and Zisserman, 2000] Capel, D. and Zisserman, A. (2000). Super-resolution enhancement of text image sequences. In *Int. Conf. Pattern Recognition*, pages Vol I: 600–605.

[Elad and Feuer, 1996] Elad, M. and Feuer, A. (1996). Super-resolution reconstruction of continuous image sequence. *IEEE Trans. on Pattern Analysis and Machine Intelligence*.

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

[Huang and Tsai, 1984] Huang, T. and Tsai, R. (1984). Multi-frame image restoration and registration. In Huang, T., editor, *Advances in Computer Vision and Image Processing*, volume 1, pages 317–339. JAI Press Inc.

[Irani and Peleg, 1991] Irani, M. and Peleg, S. (1991). Improving resolution by image registration. *Graphical Models and Image Processing*, 53:231–239.

[Irani and Peleg, 1993] Irani, M. and Peleg, S. (1993). Motion analysis for image enhancement: Resolution, occlusion, and transparency. *Journal of Visual Communication and Image Representation*, 4:324–335.

[Kelley, 1995] Kelley, C. T. (1995). *Iterative Methods for Linear and Nonlinear Equations*. SIAM, Philadelphia, PA.

[Keren et al., 1988] Keren, D., Peleg, S., and Brada, R. (1988). Image sequence enhancement using sub-pixel displacements. In *IEEE Conf. on Computer Vision and Pattern Recognition*, pages 742–746.

[Kim et al., 1990] Kim, S. P., Bose, N. K., and Valenzuela, H. M. (1990). Recursive reconstraction of high-resolution image from noisy undersampled frames. *IEEE Trans. Acoust., Speech and Sign. Proc.*, 38:1013–1027.

[Mann and Picard, 1994] Mann, S. and Picard, R. (1994). Virtual bellows: Constructing high quality stills from video. In *Int. Conf. on Image Processing*.

[Papoulis, 1977] Papoulis, A. (1977). Generalized sampling expansion. *IEEE Trans. Circuits Systems*, CAS-24:652–654.

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

[Sawhney and Kumar, 1999] Sawhney, H. and Kumar, R. (1999). True multi-image alignment and its application to mosaicing and lens distortion correction. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 21(3):245–243.

[Schultz and Stevenson, 1996] Schultz, R. and Stevenson, R. (1996). Extraction of high-resolution frames from video sequences. *IEEE Trans. Image Processing*, 5(6):996–1011.

[Shekarforoush and Chellappa, 1999] Shekarforoush, H. and Chellappa, R. (1999). Data-driven multichannel superresolution with application to video sequences. *Journal of the Optical Society of America*, 16(3):481–492.

[Smelyanskiy et al., 2000] Smelyanskiy, V., Cheeseman, P., Maluf, D., and Morris, R. (2000). Bayesian super-resolved surface reconstruction from images. In *IEEE Conf. on Computer Vision and Pattern Recognition*, pages I:375–382.

[Stark and Oskoui, 1989] Stark, H. and Oskoui, P. (1989). High resolution image recovery from image-plane arrays, using convex projections. *J. Opt. Soc. Am. A*, 6(11):1715–1726.

[Ur and Gross, 1992] Ur, H. and Gross, D. (1992). Improved resolution from subpixel shifted pictures. *Graphical Models and Image Processing*, 54(2):181–186.

[Zelnik-Manor and Irani, 2000] Zelnik-Manor, L. and Irani, M. (2000). Multi-frame estimation of planar motion. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 22(10):1105–1116.