A Fast Super-Resolution Reconstruction Algorithm for Pure Translation Motion and Common Space-Invariant Blur[1]

Michael Elad[2], Yacov Hel-Or[3]
HP Laboratories – Israel
HPL-98-??
December 1998

E-mail: elad@hp.technion.ac.il

super-resolution,
reconstruction,
maximum-likelihood,
translation motion

The super-resolution reconstruction problem is an inverse problem, dealing with the recovery of a single high-resolution image from a set of low quality images. In its general form, the super-resolution problem may consist of images with arbitrary geometric warp, space variant blur and colored noise. Several algorithms were already proposed for the solution of this general problem.

In this paper we concentrate on a special case of the super-resolution problem, where the warp is composed of pure translation, the blur is space invariant and constant for all the measured images, and the additive noise is a white Gaussian noise. We exploit our previous results, and develop a new highly efficient super-resolution reconstruction algorithm for this case. This algorithm separates the treatment of the blur from the fusion of the measurements, and the resulting overall algorithm is non-iterative.

The proposed algorithm is compared to known algorithms in the literature, showing that it is superior in terms of computational complexity. Simulations demonstrate the capabilities of the proposed algorithm.

[1]This paper is a short version of the report HPL-98-31 meant for external publication.
[2] HP Laboratories Israel, Technion City, Haifa 32000 Israel
[3] The Inter-Disciplinary College, Herzelia, Israel

1. Introduction

The super-resolution reconstruction problem is well known and extensively treated in the literature [1-11]. The main idea in this application is to recover a single high-resolution image from a set of low quality images of the same photographed object. Recent works [9-11] relate this problem to reconstruction theory [12,13]. As such, the problem is shown to be an inverse problem, where an unknown image is to be reconstructed, based on measurements related to it through linear operators and additive noise. This linear relation is composed of geometric warp, blur and decimation operations.

In [11] a general solution to the super-resolution reconstruction problem is given in a simple yet general algebraic form. The proposed solution can deal with a general geometric warp, space varying blur (which may even be different from one measured image to the other), spatially uniform decimation with rational resolution ratio, and colored Gaussian additive noise. The solution is based on the knowledge of the involved operators and noise characteristics. In [11], solutions based on the Maximum-Likelihood (ML), the Maximum A-posteriori Probability (MAP), and the Projection Onto Convex Sets (POCS) methods are suggested and unified.

This paper concentrates on a special super-resolution case, with the following assumptions: the blur is space invariant and the same for all the measured images; the geometric warp between the measured images is pure translation; and the additive noise is white. These assumptions are valid in cases where the images are obtained by the same camera and with slight vibrations, such as in many video scenes. Several papers already dealt with this special case [1,3,5,9], and proposed different reconstruction algorithms. In this paper, we propose a new algorithm, based on the general solution as given in [11]. The new algorithm is shown to be computationally very efficient, and with high output quality. Exploiting the properties of involved operations, it is shown that the general super-resolution reconstruction algorithm can be simplified to a large extent, resulting with a simpler algorithm. The new algorithm is shown to be superior to the existing algorithms [1-10] in terms of computational cost.

The following is an outline of the paper: Section 2 presents a definition of the super-resolution problem for the general case, and its ML based solution. Section 3 concentrates on the special case where the geometric warps are translations, and the blur is constant and LSI. The ML solution is re-developed, exploiting the specific structure of the involved operations. Section 4 reviews other existing algorithms for the treated super-resolution case, and compares them to the new method. Simulations and results are given in Section 5, and concluding remarks are drawn in Section 6.

2. General Super-Resolution

2.1 Formal Problem Definition

A photograph of a specific motionless destination is to be taken. The camera to be used is of low-quality, which means that any single image created by it consist of insufficient spatial resolution. Alternatively, the distance between the camera and the destination may be the cause for the insufficient resolution. Instead of taking a single photograph, several such images are taken, with slight different camera positions. The super-resolution application suggests a method for reconstructing the required single high quality image with a higher spatial resolution, based on the measured low-resolution images.

This reconstruction process requires first modeling of the relation between the high and the low resolution images. We denote the N measured images by $\{\underline{Y}_k\}_{k=1}^{N}$. These images are to be fused into a single improved quality image, denoted as $\underline{X}$. The images are reordered as column vectors by lexicographic ordering in order to represent arbitrary linear operators on images as matrices. Each of these images is related to the required super-resolution image through geometric warp, blur, decimation, and additive noise:

$$\underline{Y}_k = D_k H_k F_k \underline{X} + \underline{V}_k \qquad k = 1, \ldots, N \tag{1}$$

The matrix F_k stands for the geometric warp operation that exists between the images $\underline{X}$ and an interpolated version of the image $\underline{Y}_k$ (interpolation is required in order to treat the image $\underline{Y}_k$ in the higher resolution grid). The matrix H_k is the blur matrix, representing the camera's PSF. The matrix D_k stands for the decimation operation, and presents the loss of resolution in the obtained images. The vectors $\{\underline{V}_k\}_{k=1}^{N}$ represent additive measurement noise. These vectors are assumed to be Gaussian random vectors with zero mean and auto-correlation matrix $E\{\underline{V}_k \underline{V}_k^T\} = W_k$.

Figure 1 presents our model assumption in a block diagram. Going from the left to the right, we see the creation of the images $\{\underline{Y}_k\}_{k=1}^{N}$ from the required image $\underline{X}$. The super-resolution problem is a classic inverse problem - Our aim is to apply an opposite direction process of estimating $\underline{X}$ based on the known images $\{\underline{Y}_k\}_{k=1}^{N}$, and the operations they went through. In order to do that we have to know D_k, H_k, F_k, and W_k for all $k=1,\ldots,N$.

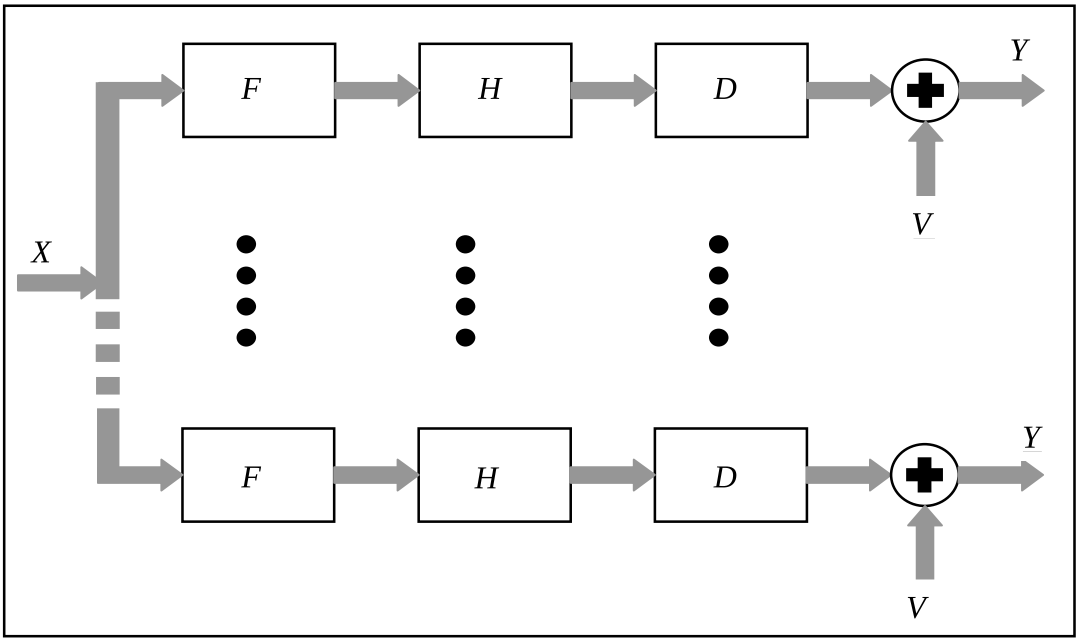


Figure 1 – Modeling the general super-resolution problem

As was said before, the recovery of the unknown image $\underline{X}$ relies on the knowledge of the involved operators. The operation F_k is obtained through motion estimation [14] between one of the images $\{\underline{Y}_k\}_{k=1}^{N}$ (chosen as a reference image), and the image $\underline{Y}_k$. In order to use the geometric displacements in terms of the finer grid, the obtained motion vectors are to be multiplied by the resolution ratio. The motion estimation process must yield sub-pixel accuracy in order to obtain good super-resolution results.

As to the decimation operation, in most applications all the decimation operations are equal - $\forall k,\ D_k = D$, and D is defined by the resolution ratio we want to obtain. Similarly, it is typically assumed that all the obtained images go through the same PSF, and therefore $\forall k,\ H_k = H$. In order to determine *H*, we can either guess the PSF, or estimate it somehow [13]. The same goes for the additive noise – in most applications the noise is assumed to be white, which means that $E\{\underline{V}_k \underline{V}_k^T\} = \sigma^2 I$. Otherwise, the noise characteristics must be estimated somehow.

2.2 Maximum Likelihood Approach

The most intuitive way to define the optimal reconstructed image $\underline{\hat{X}}$ is to choose the image that, when fed into the above system (in Figure 1), gives a set of simulated images $\{\underline{\hat{Y}}_k\}_{k=1}^N$, which are as close as possible to the original $\{\underline{Y}_k\}_{k=1}^N$. The following term presents this very idea, where the distance between the simulated and the original $\{\underline{Y}_k\}_{k=1}^N$ is given in a mean square error sense:

$$\underline{\hat{X}} = \underset{\underline{X}}{\text{ArgMin}} \left\{ \sum_{k=1}^{N} \left[\underline{Y}_k - D_k H_k F_k \underline{X}\right]^T W_k^{-1} \left[\underline{Y}_k - D_k H_k F_k \underline{X}\right] \right\} \tag{2}$$

where the term $D_k H_k F_k \underline{X}$ is the simulated image $\underline{\hat{Y}}_k$. Notice that each of the *N* terms in the above equation is weighted by the inverse of the matrix W_k, so that noisier measurements get a smaller impact on the estimated result.

The solution defined above can be shown to be the Maximum Likelihood estimator of the given problem [11], since we have assumed that the additive noise vectors $\{\underline{V}_k\}_{k=1}^N$ are Gaussian. The solution for the above quadratic minimization problem is given by the equation:

$$\sum_{k=1}^{N} \left[D_k H_k F_k\right]^T W_k^{-1} \left[\underline{Y}_k - D_k H_k F_k \underline{X}\right] = 0 \ \Rightarrow \mathbf{R}\underline{\hat{X}} = \underline{\mathbf{P}}$$
$$\text{where} \quad \mathbf{R} = \sum_{k=1}^{N} F_k^T H_k^T D_k^T W_k^{-T} D_k H_k F_k \qquad \underline{\mathbf{P}} = \sum_{k=1}^{N} F_k^T H_k^T D_k^T W_k^{-1} \underline{Y}_k \tag{3}$$

Solving the above equation directly is practically impossible due to its dimensions. If, for example, the size of image $\underline{\hat{X}}$ is $1000 \cdot 1000$ pixels, the matrix **R** is a $10^6 \cdot 10^6$ matrix – which is very difficult to invert. Inversion of such a huge matrix can be obtained using iterative algorithms. Note that the actual inverse of **R**, namely $\mathbf{R}^{-1}$, is not required, but rather the solution of the linear equation $\mathbf{R}\underline{\hat{X}} = \underline{\mathbf{P}}$. Such iterative methods are very common and very efficient [11].

In special cases, the matrix **R** may have a specific simplified structure, which can be exploited in order to apply the inversion directly. As we shall see in the next Section, this is exactly the case we treat in this paper.

In this section we present the iterative approach, utilizing one of the simplest possible algorithms – the Steepest Descent (SD) algorithm. The obtained equations will be used later for constructing the solution for the special case this paper deals with. The SD algorithm suggests the following iterative equation for the solution of $\mathbf{R}\underline{\hat{X}} = \underline{\mathbf{P}}$:

$$\underline{\hat{X}}_{j+1} = \underline{\hat{X}}_j + \mu\left[\underline{\mathbf{P}} - \mathbf{R}\underline{\hat{X}}_j\right] \tag{4}$$

where $\underline{\hat{X}}_0$, the initialization vector, can be any vector. The above algorithm is guaranteed to converge to the unique solution of $\mathbf{R}\underline{\hat{X}} = \underline{\mathbf{P}}$, provided that $\mu > 0$ is small enough [15-16]. Putting the terms for **R** and $\underline{\mathbf{P}}$ from Equation (9) into the above equation we get:

$$\underline{\hat{X}}_{j+1} = \underline{\hat{X}}_j + \mu \sum_{k=1}^{N} F_k^T H_k^T D_k^T W_k^{-T} \left[\underline{Y}_k - D_k H_k F_k \underline{\hat{X}}_j\right] \tag{5}$$

3. Super-resolution - The Special Case

Let us first repeat the special case properties we intend to exploit:

(i) All the decimation operations are the same, i.e. $\forall k,\ D_k = D$.

(ii) All the blur operations are the same, i.e. $\forall k, H_k = H$. Moreover, the matrix H is assumed to be block circulant, representing a linear and space invariant blur [12,17].

(iii) All the warp operations correspond to pure translations. Thus, the matrices F_k are all block-circulant as well [12,17]. Moreover, we assume that F_k is represented through the nearest neighbor displacement paradigm [12], which means that the displacement in the finer grid is rounded and F_k applies only integer translations. This assumption simplifies the analysis and the obtained results. Its implications on the output quality are negligible, since the rounding is done in the finer resolution grid.

(iv) The additive noise is white and the same for all the measurements, i.e. $\forall k, W_k = \sigma^2 I$.

Putting these assumptions into equation (5), we get that the iterative equation becomes:

$$\underline{\hat{X}}_{j+1} = \underline{\hat{X}}_j + \mu \sum_{k=1}^{N} F_k^T H^T D^T \left[\underline{Y}_k - DHF_k \underline{\hat{X}}_j\right] \tag{6}$$

Exploiting the fact that block circulant matrices commute [12,17], we get that $F_k^T H^T = H^T F_k^T$ and $HF_k = F_k H$. Thus:

$$\underline{\hat{X}}_{j+1} = \underline{\hat{X}}_j + \mu H^T \sum_{k=1}^{N} F_k^T D^T \left[\underline{Y}_k - DF_k H \underline{\hat{X}}_j\right] \tag{7}$$

Let us define the blurred super-resolution image by $\underline{\hat{Z}}_j = H \underline{\hat{X}}_j$. Multiplying both sides of equation (7) with H we get:

$$\begin{aligned} & H\underline{\hat{X}}_{j+1} = H\underline{\hat{X}}_j + \mu HH^T \sum_{k=1}^{N} F_k^T D^T \left[\underline{Y}_k - DF_k H \underline{\hat{X}}_j\right] \\ \Rightarrow \quad & \underline{\hat{Z}}_{j+1} = \underline{\hat{Z}}_j + \mu HH^T \sum_{k=1}^{N} F_k^T D^T \left[\underline{Y}_k - DF_k \underline{\hat{Z}}_j\right] = \\ & \quad = \underline{\hat{Z}}_j + \mu HH^T \left[\sum_{k=1}^{N} F_k^T D^T \underline{Y}_k - \sum_{k=1}^{N} F_k^T D^T DF_k \underline{\hat{Z}}_j\right] = \underline{\hat{Z}}_j + \mu HH^T \left[\underline{\tilde{\mathbf{P}}} - \tilde{\mathbf{R}} \underline{\hat{Z}}_j\right] \end{aligned} \tag{8}$$

where we have used the notations:

$$\underline{\tilde{\mathbf{P}}} = \sum_{k=1}^{N} F_k^T D^T \underline{Y}_k \quad \text{and} \quad \tilde{\mathbf{R}} = \sum_{k=1}^{N} F_k^T D^T DF_k \ . \tag{9}$$

Since the matrix HH^T is a positive semi-definite matrix [17], the above iterative equation stands as a general SD algorithm with a weight matrix HH^T [15-16]. It is known [15-16] that such an iterative equation converges to the same final solution as the one without the weight matrix, as long as this matrix is positive semi-definite. Therefore, the steady-state solution of the difference equation (8) is given $\underline{\hat{Z}}_\infty = \tilde{\mathbf{R}}^{-1}\underline{\tilde{\mathbf{P}}}$. In order to be precise, since the matrix *H* might be singular, the steady-state solution consist of two parts: the first is the part of the initialization vector $\underline{\hat{Z}}_0$ which is in the null-space of HH^T, and a second is the part of the solution $\tilde{\mathbf{R}}^{-1}\underline{\tilde{\mathbf{P}}}$ which is vertical to this null-space.

Assuming that we somehow found $\underline{\hat{Z}}_\infty$, the above analysis implies that an image restoration process must be applied in order to remove the effect of the blur matrix *H*. This way, based on $\underline{\hat{Z}}_\infty = H\underline{\hat{X}}_\infty$, we recover the required image $\underline{\hat{X}}_\infty$. Moreover, if we neglect the fact that *H* might be singular, we can say that this process of first finding a blurred version of the super-resolution image, and later restoring the image itself, is as optimal as the direct approach. The fact that the treatment of the blur can be separated from the fusion of the measurements part was already proposed in other works [1,2,5].

We return now to the recovery of the image $\underline{\hat{Z}}_\infty$. As it turns out, computing $\underline{\hat{Z}}_\infty$ is very easy because of the following result:

Theorem: Based on the assumptions, mentioned at the beginning of this section, the matrix $\tilde{\mathbf{R}} = \sum_{k=1}^{N} F_k^T D^T D F_k$ is a diagonal matrix.

Proof: We first note that the matrix $D^T D$ is diagonal. This is easily verified by noticing that the operation $D^T D$ stands for decimation followed by interpolation. Thus, if *D* decimate by factor of *r*, applying $D^T D$ causes all the positions $[1+mr, 1+nr]$ for integer [*m*,*n*] to stay unchanged, whereas the remaining pixels are replaced by zeros. Thus, $D^T D$ stands for a masking operation, which is represented by diagonal matrix.

Let us look at the expression $\tilde{\mathbf{R}}_k = F_k^T D^T D F_k$ for some *k*. We use the notation $f_k(j)$ to denote the j^{th} column of the matrix F_k. If the displacement vector represented by this matrix is $[d_x, d_y]$, $f_k(j)$ will have '1' value at the position[4] $p\{j, d_x, d_y\}$, and zeros elsewhere. In the general case, the [*m*,*n*] entry of the matrix $\tilde{\mathbf{R}}_k$ is given by $f_k^T(m) D^T D f_k(n)$. Since $D^T D$ is diagonal, if $m \neq n$, we get that $p\{m, d_x, d_y\} \neq p\{n, d_x, d_y\}$, and thus, $\tilde{\mathbf{R}}_k[m,n] = 0$. If $m = n$, we get that $\tilde{\mathbf{R}}_k[m,m] = 1$. Since $\tilde{\mathbf{R}} = \sum_k \tilde{\mathbf{R}}_k$, we get that $\tilde{\mathbf{R}}$ is diagonal as well, and the claim is proved. ■

Since the matrix $\tilde{\mathbf{R}}$ is diagonal, obtaining $\underline{\hat{Z}}_\infty = \tilde{\mathbf{R}}^{-1}\underline{\tilde{\mathbf{P}}}$ is easy to achieve. The super-resolution reconstruction process thus consist of the following stages:

(i) Compute the pair $\underline{\tilde{\mathbf{P}}}$ and $\tilde{\mathbf{R}}$ based on equation (9). The matrix $\tilde{\mathbf{R}}$ can be simply represented as a mask image of the same size as the image $\underline{X}$. The same goes to the vector $\underline{\tilde{\mathbf{P}}}$ - it can be represented as an image of the same size as $\underline{X}$. Note that these operations require only additions as both the entries of *D* and F_k are zeros and ones.

(ii) Compute $\underline{\hat{Z}}_\infty = \tilde{\mathbf{R}}^{-1}\underline{\tilde{\mathbf{P}}}$. This operation requires only one division per each pixel. One possible problem with this stage is the possibility that some of the entries at the main diagonal of $\tilde{\mathbf{R}}$ may be zeros. In these positions, the division $\underline{\hat{Z}}_\infty = \tilde{\mathbf{R}}^{-1}\underline{\tilde{\mathbf{P}}}$ becomes singular. As it turns out, however, it is easy to verify that whenever $\tilde{\mathbf{R}}(m,m) = 0$, it is guaranteed that $\underline{\tilde{\mathbf{P}}}(m) = 0$, and thus, we get that $\underline{\hat{Z}}_\infty(m) = 0/0$. In such cases, a simple interpolation can be used to fill these positions.

(iii) Restore $\underline{\hat{X}}_\infty$ from $\underline{\hat{Z}}_\infty$, which can be done in various ways [12-13]. This part of the process is the computationally demanding part. If this part of the process is done directly (e.g. by Wienner filtering), the overall reconstruction algorithm is non-iterative.

[4] This mapping is one-to-one, i.e., given the motion vector and *p*, one can compute the value *j*.

4. Relation To Other Methods

Several papers addressed the general super-resolution problem and suggested practical reconstruction algorithms for solving it [6-11]. Such are the IBP method [6-7], the POCS based solution [8,9,11], and the MAP based algorithms [10-11]. These algorithms typically tend to be complex, as they attempt to treat any kind of geometric warp, complicated kinds and blur (such as space variant ones), possible colored noise, and more complicated decimation patterns. Thus, these general algorithms fail to compete with methods designed for the special case treated in this paper.

When facing the special simplified super-resolution problem treated in this paper, one can use one of the following three options: (i) frequency domain methods [1-3]; (ii) generalized sampling theorems [3,4]; and (iii) the method described in the previous section. In all these three approaches, the ability to separate the treatment of the blur from the fusion of the images can be (and actually is) exploited.

The frequency approach was proposed initially by Kim, Bose and valenzuela [1,2]. Their approach suggests applying a 2D-DFT per each of the input images, combining the images in the frequency domain exploiting aliasing relationships, and then applying an inverse 2D-DFT. As in our case, blur treatment is done at the end of the recovery algorithm. One of the benefits of the frequency algorithm is its ability to be recursive, i.e. the ability to add more measurements as they come. Actually, similar behavior can be identified in our algorithm, since both $\underline{\tilde{\mathbf{P}}}$ and $\tilde{\mathbf{R}}$ are computed as a direct some of terms, which correspond to different measurements.

As to the computational complexity of the frequency domain algorithm: In the non-recursive approach, the frequency domain algorithm requires the accumulation of a complex matrix of size $[N \times r]$ per each pixel[5], and the inversion of it. The recursive approach requires an inversion of a $[r \times r]$ matrix, and more multiplications in order to apply the RLS algorithm. We have to remember that above all these comutations, the DFT operations must be taken into account. Thus, the frequency approach is far more complicated, compared to our way of fusing the measurements.

The generalized sampling theorems by Yen [3] and later by Papulis [4] where used as the basis for the method proposed by Ur and Gross [5]. Their method also separates the treatment of

[5] r is the resolution ratio and N is the number of given images

the blur from the fusion process. The proposed algorithm is highly sensitive to measurements with close spatial positions, and to over measurements. This algorithm totally disregards the possibility of additive noise. The given samples are considered as the ground truth, and the reconstruction result is merely an interpolation between them.

As to this algorithm computational complexity, the recovered signal is computed by summing sufficiently many interpolation functions, which are based on a generalization of the Sinc function. In [5] it is claimed that $O\left\{r^2\right\}$ multiplications per one output pixel are required for the fusion process. Our method, on the other hand, requires only one division per an output pixel.

5. Results

We start with a simple synthetic example. We have taken an image of size 720×884 pixels, and created from it 9 different 240×294 images by 3:1 decimation at each axis and starting at the 9 possible different locations. Each of these image is shifted by integer multiplication of $1/3$ at each axis, and these displacements are exactly known. Furthermore, by simply interlacing these images together, we get the original image, which stands for 3:1 resolution improvement result.

We have applied the reconstruction process on these 9 images. The displacements were estimated using an algorithm described in [14], and were found to be (after the rounding for the nearest neighbor) the exact ones. Thus, the vector $\underline{\tilde{\mathbf{P}}}$ consists of exactly the required image, and the main diagonal of $\tilde{\mathbf{R}}$ is actually constant and equals to 1. Thus $\underline{\hat{Z}}_\infty = \tilde{\mathbf{R}}^{-1}\underline{\tilde{\mathbf{P}}}$ is the exact super-resolution image. We have assumed a PSF kernel is [0.25, 0.5, 0.25], and applied an appropriate restoration. The restoration was applied by a Wienner filter of size 15 by 15, manually searching for the best parameter in this filter.

Figure 2 presents the results for the synthetic example. We have chosen to show two informative blocks from the input, the output and the original images. The input blocks are taken from the reference image (we could take it from any of the images – the quality is the same), and scaled up by a factor 3 using a NN interpolation. The output and the original blocks correspond to the same portions of the image, and the improvement is self-evident.

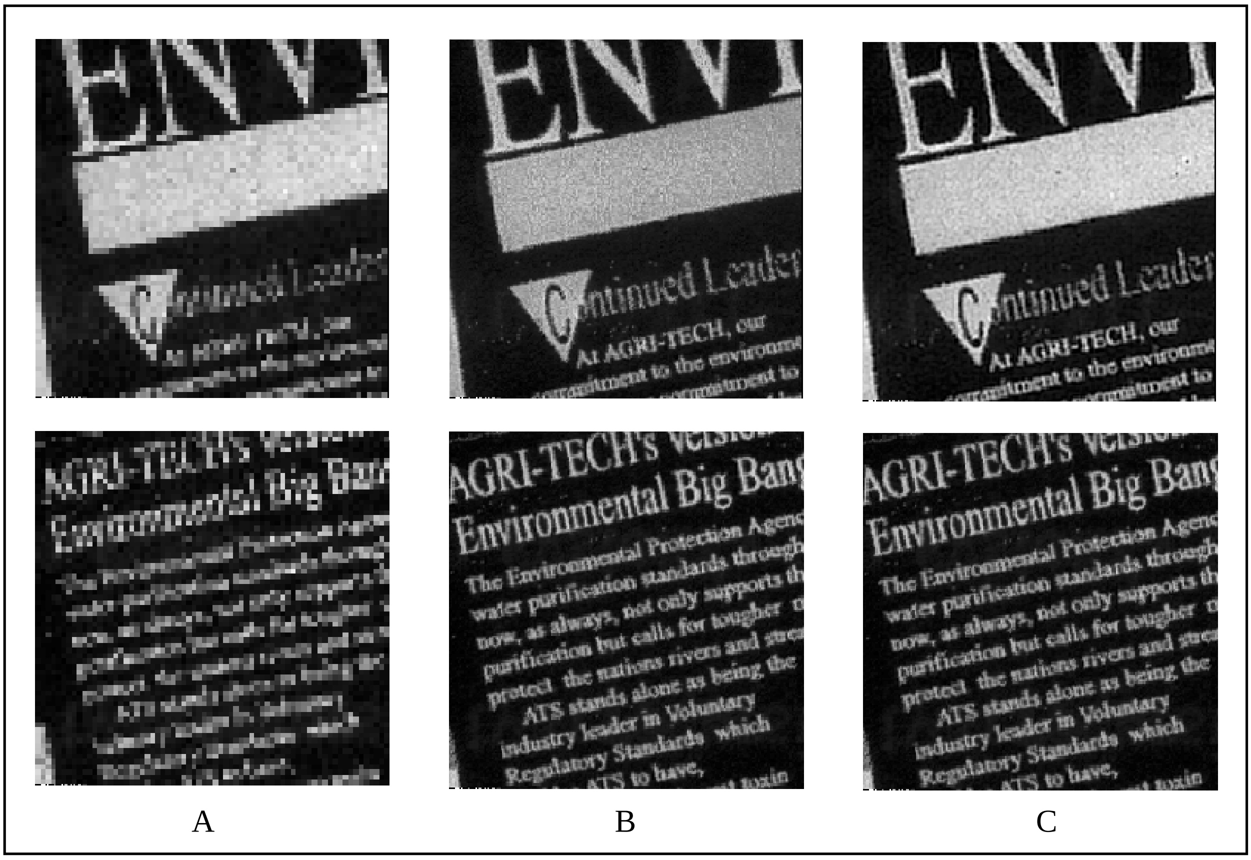


Figure 2 – Results of the synthetic test: A – Reference image, B – Original image, C – Reconstruction Results.

Figure 3 presents the results for a sequence of 12 images of size 318×411. The resolution is increased by a factor 2. Again, two blocks are shown for comparison, and the improvement is evident – the result is sharper and with more details. The original is a dither image, which explains the textured regions, both in the photographed image and the super-resolution result. The assumed PSF in this and the next cases were Gaussian blur operations with manually found variance.

Figures 4 and 5 correspond to the third and last test. This sequence contains 16 images of size 300×303 Pixels. The resolution is increased by a factor 2 in each axis. Figure 4 shows the entire images (one of the measured images, and the result), and figure 5 presents magnified small portions of them for better comparison.

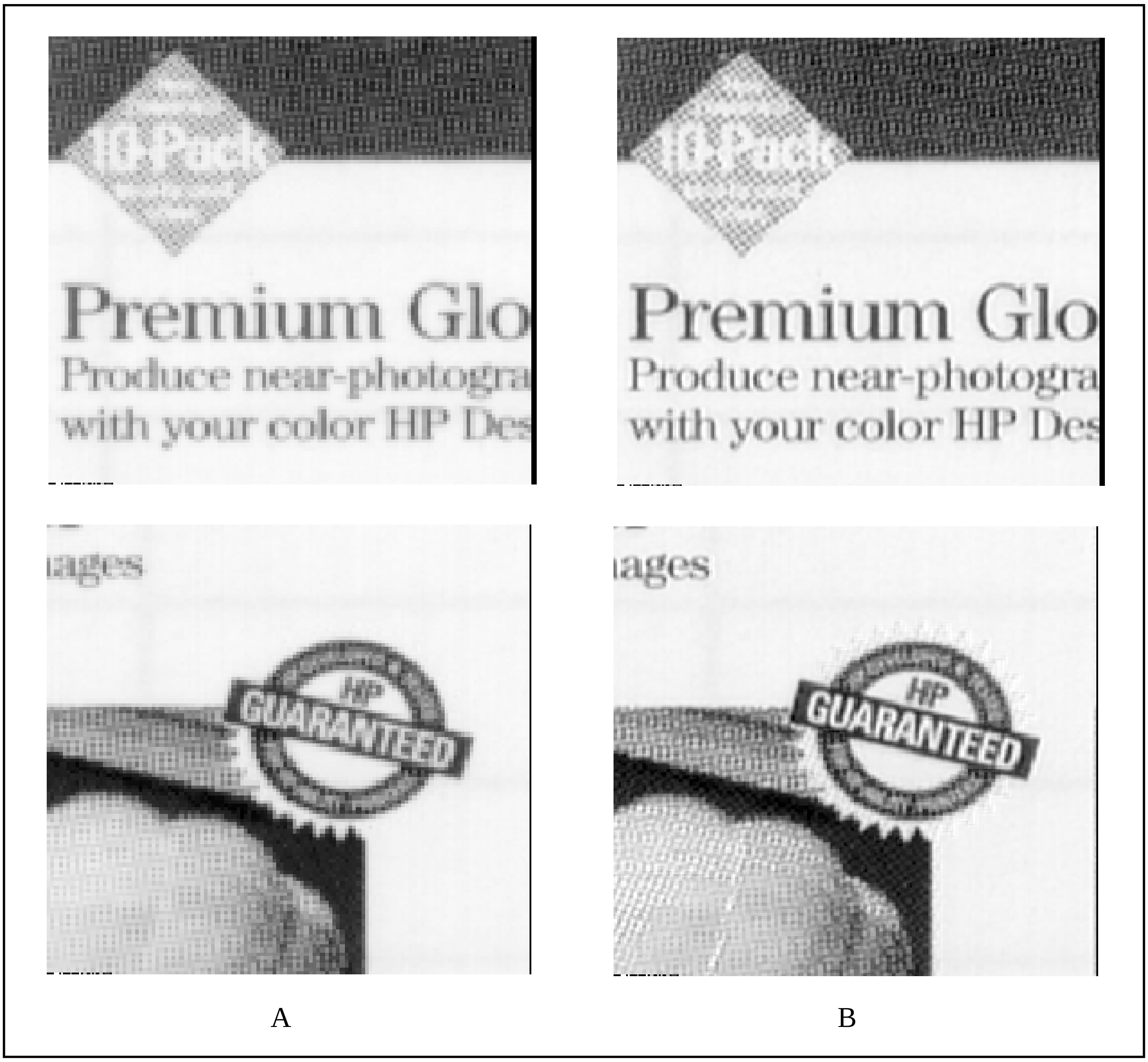


Figure 3 – Results of the second sequence: A – The original measured images, B – The result

6. Conclusions

In this paper we have presented a new algorithm for super-resolution reconstruction, for the special case were the geometric warp between the given images consist of pure translation, the blur is the same for all the measurements, and is space invariant, and the additive noise is white. The proposed algorithm is shown to be very efficient in terms of computational cost, compared to other algorithms. Simulation results demonstrate its capabilities in terms of output quality.

A

B

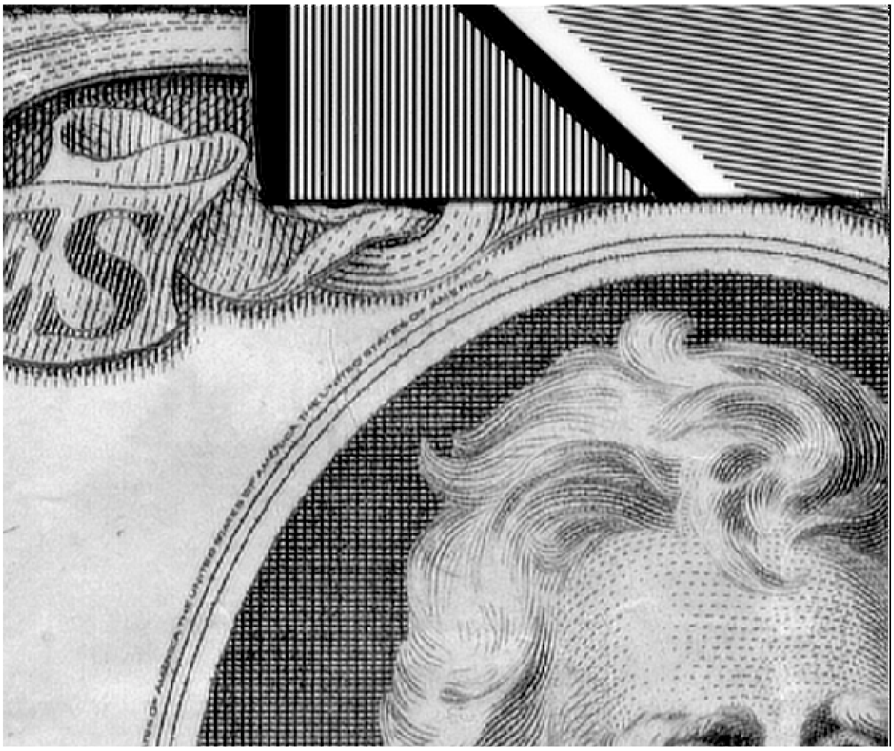
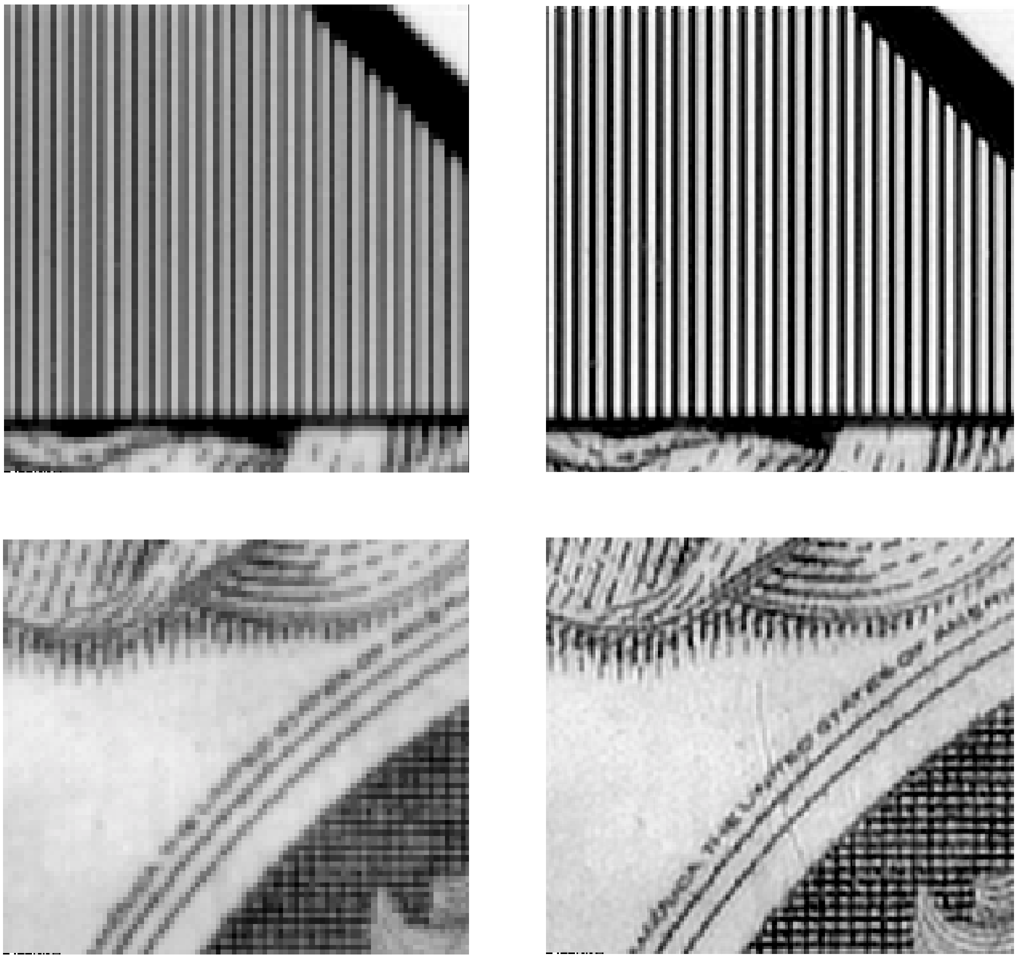


Figure 4 – Results of the third sequence: A – The measured image, B – The result



A B

Figure 5 – Results of the third sequence – Sections taken from the images in Figure 4:

A - The measured image, and B – the result

References

[1] S. P. Kim, N. K. Bose, and H. M. Valenzuela, "Recursive Reconstruction of High-resolution Image from Noisy Undersampled Multiframes", *IEEE Trans. Acoustics, Speech and Signal Processing*, Vol. 38, pp. 1013-1027, June 1990.

[2] N. K. Bose, H. Kim, and B. Zhou, "Performance Analysis of the TLS algorithm for Image Reconstruction from a Sequence of Undersampled Noisy and Blurred Frames", *in Proc. ICIP, Austin – Texas*, pp. 571-575, 1994.

[3] L. J. Yen, "On Non-Uniform Sampling of Bandwidth Limited Signals", *IRE Trans. Circuits Theory*, Vol. 3, pp. 251-257, April 1956.

[4] A. Papulis, “Generalized sampling Theorem”, *IEEE Trans. on Circuits and Systems*, Vol. 24, pp. 652-654, November, 1977.

[5] H. Ur and D. Gross, “Improved resolution From Sub-Pixel Shifted Pictures”, *CVGIP: Graph. Models and Image Proc.*, Vol. 54, pp. 181-186, March, 1992.

[6] A. M. Tekalp, M. K. Ozkan, and M. I. Sezan, “High Resolution Image Reconstruction From Low-Resolution Image Sequences, and Space Varying Image Restoration”, IEEE ICASSP, San Francisco, CA., March 1992, Vol. III, pp. 169-172.

[7] A. J. Patti, M. I. Sezan, and A. M. Tekalp, “Super-Resolution Video Reconstruction With Arbitrary Sampling Lattices and Nonzero Aperture Time”, *IEEE Trans. on Image Processing*, Vol. 6, pp. 1064-1076, August 1997.

[8] M. Irani and S. Peleg, “Improving Resolution by Image Registration”, *CVGIP: Graphical Models and Image processing*, Vol. 53, pp.231-239, May 1991.

[9] M. Irani and S. Peleg, “Motion Analysis for Image Enhancement: Resolution, Occlusion, and Transparency”, *Journal of Visual Communication and Image Representation*, Vol. 4, pp.324-335, April, 1993.

[10] R. R. Schultz and R. L Stevenson, “Extraction of High-Resolution Frames From Video Sequences”, *IEEE Trans. on Image Processing*, Vol. 5, pp. 996-1011, June 1996.

[11] M. Elad and A. Feuer, “Restoration of a Single Super-Resolution Image from Several Blurred, Noisy and Undersampled Measured Images”, *IEEE Trans. Image Processing*, Vol. 6, pp. 1646-1658, December 1997.

[12] A. K. Jain, **FUNDAMENTALS OF DIGITAL IMAGE PROCESSING**, Prentice Hall, Englewood Cliffs, NJ 07632, 1989.

[13] R. L. Lagendijk and J. Biemond, **ITERATIVE IDENTIFICATION AND RESTORATION OF IMAGES**, Kluwer Academic, Boston MA. 1991.

[14] J. L. Barron, D. J. Fleet, and S. Beauchemin, “Performance of Optical Flow Techniques”, *International Journal of Computer Vision*, Vol. 12, pp. 43-77, 1994.

[15] D. Betrsekas, **NONLINEAR PROGRAMMING**, Athena Scientific, Belmont Massachusetts, 1995.

[16] D. M. Young, **ITERATIVE SOLUTION OF LARGE LINEAR SYSTEMS**, Academic, New-York, 1971.

[17] R. A. Horn and C. J. Johnson, **MATRIX ANALYSIS**, Cambridge University Press, Cambridge MA. 1985.