Analysis of Image Super-Resolution via Reconstruction Filters for Pure Translational Motion

Fatih Kara *fatih.kara@tubitak.gov.tr*
TUBITAK-BILGEM
The Scientific and Technological Research Council of Turkey
Gebze/Kocaeli, 41470, Turkey

Cabir Vural *cabir.vural@marmara.edu.tr*
Department of Electrical-Electronics Engineering
Marmara University
Kadikoy/Istanbul, 34722, Turkey

Abstract

In this work, a special case of the image super-resolution problem where the only type of motion is global translational motion and the blurs are shift-invariant is investigated. The necessary conditions for exact reconstruction of the original image by using finite impulse-response reconstruction filters are investigated and determined. If the number of available low-resolution images is larger than a threshold and the blur functions meet a certain property, a reconstruction filter set for perfect image super-resolution can be generated even in the absence of motion. Given that the conditions are satisfied, a method for exact super-resolution is presented to validate the analysis results and it is shown that for the fully determined case, perfect reconstruction of the original image is achieved. Finally, some realistic conditions that make the super-resolution problem ill-posed are treated and their effects on exact super-resolution are discussed.

Keywords: Image Processing, Image Super-Resolution, Finite Impulse-Response Filters, Existence-Uniqueness Conditions.

1. INTRODUCTION

In most imaging applications, a high quality and high resolution (HR) image is desired, so that the level of details that the image presents to the observer will be high. However, in most cases the quality and spatial resolution of the image is degraded by several factors. For example, in space imaging, factors such as atmospheric scattering, sensor noise, non-ideal imaging optics, etc. all affect the observed image quality. Furthermore, an increase in resolution by using sensor manufacturing techniques is usually expensive and has its own problems. Thus, a signal processing technique, which is based on combining several low-resolution (LR) images of the same scene, has emerged and it is called "super-resolution image reconstruction" (or resolution enhancement). Most super-resolution methods utilize the diversity that is provided by the availability of more than one LR images. By diversity, we mean that the LR images corresponding to the same scene are different from each other in the sense of motion that they contain or the blur that they are exposed of.

Super-resolution has many application areas where multiple frames of the same scene can be obtained. For instance, in areas such as medical and space imaging, super-resolution is proven to be useful. Also, multiple frames in a video sequence can be utilized to improve the resolution for frame-freeze or zooming purposes.

Beginning from the frequency-domain approach of Huang and Tsai [1], a large number of super-resolution techniques have been proposed. Iterative back-projection [2], projection onto convex

sets (POCS) approach [3, 4], stochastic reconstruction methods such as maximum a posteriori (MAP) or maximum likelihood estimations [5, 6], hybrid MAP/POCS super-resolution algorithm [7], among others, constitute the early work on image super-resolution. Later, edge-preserving stochastic methods that perform adaptive smoothing based on the local properties of the image are studied [8]. Kang and Lee provided a least-squares solution with regularization [9]. In [10], a super-resolution algorithm that takes into account inaccurate estimates of the point spread function and the registration parameters is presented.

More recently, super-resolution methods that employ the Bayesian approach as the main framework have been studied. In [11], a Bayesian adaptive video super-resolution method is proposed in which the super-resolution process is conducted simultaneously with the motion, blur kernel and the noise level estimation. In [12], spatiotemporal dependencies are exploited and the need to estimate the subpixel motion explicitly is eliminated. Super-resolution methods using variational Bayesian analysis [13], generalized Gaussian Markov random fields [14], and symmetric alpha-stable Markov random fields [15] are also studied.

There are many other methods that are not mentioned here, the reader can refer to the tutorial papers in [16-18] for a more comprehensive reference list. Also, a number of special journal issues on super-resolution image reconstruction provide collections of studies on the topic [19, 20].

A branch of image processing that is closely related to the image super-resolution problem is multichannel image deconvolution. The purpose of multichannel image deconvolution methods is to construct an unobservable true image from several observed blurred ones, but it does not deal with the problem of increasing the resolution. It is shown that when there are more than three blurred observations and the blur functions meet a certain property, it is possible to blindly estimate the blur-free image in the absence of additive noise by using finite impulse-response (FIR) reconstruction filters [21]. Other works about the topic can be found in [22-24].

In this paper, the conditions for the existence and uniqueness of finite-impulse response restoration filters for exact reconstruction of the HR image in case of pure translational motion (or no motion) and shift-invariant blur are derived. The work presented here is based on [21] in which the same analysis is done for the multichannel image deconvolution case, but it is different in the sense that the analysis here contains the subsampling operator and (possible) motion between frames. The motion operator is eliminated by assuming only pure translational motion that allows combining the motion with the blur operator. The subsampling operator is included in the derivations and the results in [21] are modified to reflect this inclusion. In the analysis, it is found that if the number of the LR images is larger than a threshold and the blur functions meet a certain property, then a set of restoration filters can be constructed for exact HR image reconstruction even in the absence of motion for the well-posed case, i.e. no additive noise, known blur parameters, availability of adequate number of LR images, etc. Then, the factors that make the super-resolution problem ill-posed are treated and their effects on exact super-resolution are discussed.

The assumption of pure translational motion between the LR frames is somewhat limiting, but it is a valid presumption for applications where the motion is controlled and there is no local movement. For example, the scanner resolution can be increased by scanning the document more than once with slightly changed initial points. Also in some video sequences, the scene is static and image sequences are obtained by translational motion of the video camera. There are works in the literature that consider this special super-resolution case [25-27].

The paper is organized as follows. In Section 2, the observation model between the LR images and the HR image is given and the problem to be solved is defined. In Section 3, the requirements for the existence and uniqueness of perfect FIR restoration filters are derived. In Section 4, given that the conditions are met, a perfect reconstruction method is presented to verify the results of Section 3. In Section 5, some simulation results are provided in order to

validate the propositions for the well-posed case and the factors that limit exact super-resolution are discussed. Finally, some conclusions are drawn in Section 6. A preliminary version of this study has been presented at WASET 2008 conference [28].

2. OBSERVATION MODEL AND PROBLEM DEFINITION

Super-resolution is an inverse problem, where the desired unknown HR image is to be constructed from the observed LR ones. The desired and observed images are linked through linear operations such as geometric warp, blur, decimation (subsampling), and additive noise [7]. The observation model that links the desired HR image to the observed LR images is given in Fig. 1. $x(n_1,n_2)$ denotes the HR image, $y_k(n_1,n_2)$, k = 1, …, K, stand for the observed LR images and $b'_k(n_1,n_2)$, k = 1, …, K, are the blur operators. The additive noise is represented by $v_k(n_1,n_2)$, and S is the subsampling process. K is the number of LR images.

If the only type of motion between the LR images is global translational motion, then the 2-D Z-transform of the k'th warped and blurred HR image can be written as:

$$X_{k,b}(z_1,z_2) = X(z_1,z_2)\, z_1^{-H_k} z_2^{-V_k} B_k'(z_1,z_2) \tag{1}$$

where $X(z_1,z_2)$ and $X_{k,b}(z_1,z_2)$ denote for the Z-transforms of the corresponding spatial-domain images and $B'_k(z_1,z_2)$ stands for the Z-transforms of the blur functions. H_k and V_k are the horizontal and vertical shifts for the k'th observed image in terms of HR pixel units.

If we define $B_k(z_1,z_2) := z_1^{-H_k} z_2^{-V_k} B'_k(z_1,z_2)$ or equivalently $b_k(n_1,n_2) := b'_k(n_1\text{-}H_k,\ n_2\text{-}V_k)$, then (1) can be rearranged as

$$X_{k,b}(z_1,z_2) = X(z_1,z_2) B_k(z_1,z_2) \tag{2}$$

or equivalently

$$x_{k,b}(n_1,n_2) = x(n_1,n_2) * b_k(n_1,n_2). \tag{3}$$

In summary, if the motion consists of only global translational motion, then the warping and blur operators can be merged as a single blur operator and the observation model given in Fig. 1 is reduced into the model seen in Fig. 2.

To obtain the HR and blur-removed image, a reconstruction filter set is applied on the LR images as seen in Fig. 3. In this paper, the existence and uniqueness conditions for this filter set will be investigated and derived. Note that in Fig. 3, the LR images must be upsampled before the filtering process. In this study, it is assumed that the subsampling operator in Fig. 2 does not reduce the size of the image, but it retains the relevant pixels while setting the value of the others to zero. Because of this assumption, the upsampling operator is not shown in Fig. 3.

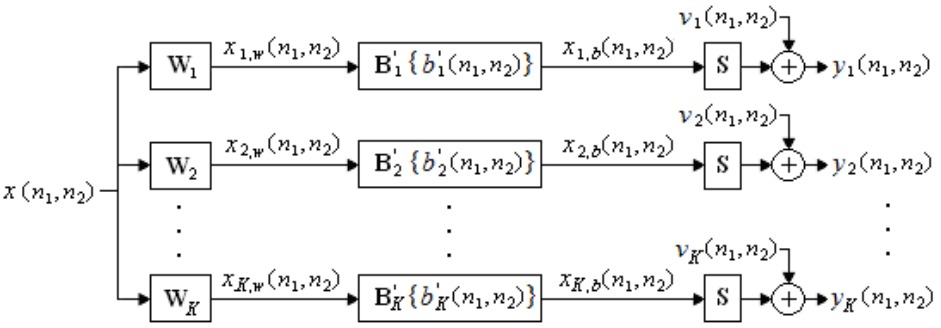

FIGURE 1: Observation Model.

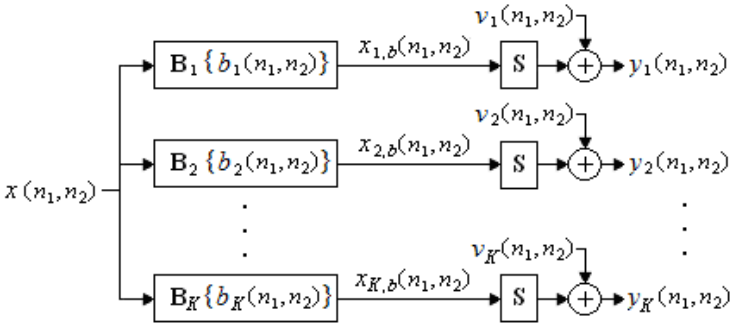

FIGURE 2: Simplified Observation Model.

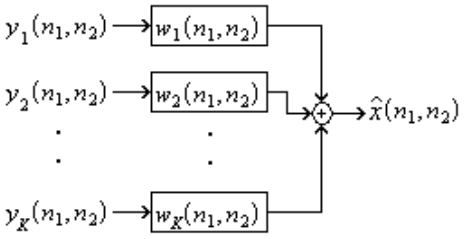

FIGURE 3: Reconstruction stage for the simplified observation model.

Throughout the analysis, the following notation and assumptions are used:

- Additive noise is ignored.
- The only type of motion is global translational motion and it is combined with the blur process as explained in (3).
- The size of each blur function is *M*x*M* and that of each reconstruction filter is *N*x*N*. The subsampling rate is D_s in both horizontal and vertical directions.
- $b^k_{ij} = b_k(i, j)$ $(0 \le i, j < M, 1 \le k \le K)$ denotes the blur coefficients.
- $w^k_{ij} = w_k(i, j)$ $(0 \le i, j < N, 1 \le k \le K)$ denotes the reconstruction filter coefficients.

It is straightforward to adapt the analysis below to more general cases (rectangular blur and reconstruction filter functions, different subsampling rates for the horizontal and vertical directions). In the analysis, concepts from the multichannel image deconvolution problem will be used. The two problems are different even if the motion is removed because of the existence of the subsampling operator in the super-resolution problem.

To determine the existence and uniqueness conditions of the reconstruction filters, the LR images must be expressed in terms of the HR image, blur operators and the subsampling operator, and the estimated HR image must be expressed in terms of the LR images and the reconstruction filters in vector-matrix notation. Let us define the input vector $\mathbf{x}(n_1,n_2)$ as the $(M+N-1)\text{x}(M+N-1)$ image segment centered at (n_1,n_2) in lexicographic notation:

$$\mathbf{x}(n_1,n_2)=\begin{bmatrix} x\left(n_1-(M+N-1)/2,n_2-(M+N-1)/2\right) \\ x\left(n_1-(M+N-1)/2,n_2-(M+N-1)/2+1\right) \\ \vdots \\ x\left(n_1+(M+N-1)/2,n_2+(M+N-1)/2\right) \end{bmatrix}^{\mathrm{T}}$$

where T denotes the transposition operator. The input vector is a row vector and it consists of $(M+N-1)^2$ elements. For (n_1,n_2) to be at the exact center of the image segment, $(M+N-1)$ must be an odd number.

The output vector, $\mathbf{y}(n_1,n_2)$, is constructed such that image segments of size $N\text{x}N$ centered at (n_1,n_2) is taken from each LR image and placed in lexicographic order, and then these segments, which are expressed by $\mathbf{y}^k(n_1,n_2)$, are arranged in a row to form the output vector $\mathbf{y}(n_1,n_2)$. That is,

$$\mathbf{y}(n_1,n_2)=[\mathbf{y}^1(n_1,n_2)\quad \mathbf{y}^2(n_1,n_2)\quad \cdots\quad \mathbf{y}^K(n_1,n_2)]$$

$$\mathbf{y}^k(n_1,n_2)=\begin{bmatrix} y_k\left(n_1-(N-1)/2,n_2-(N-1)/2\right) \\ y_k\left(n_1-(N-1)/2,n_2-(N-1)/2+1\right) \\ \vdots \\ y_k\left(n_1+(N-1)/2,n_2+(N-1)/2\right) \end{bmatrix}^{\mathrm{T}}.$$

The output vector is a row vector and its length is KN^2. Just like the case for the input vector, for (n_1,n_2) to be located at the exact center of the LR image segments, N must be an odd number. This condition, combined with the requirement that $(M+N-1)$ must be an odd number, implies that M also must be an odd number. If M, the size of the blur filters, is not an odd number, it can be done by adding zeroes to the right and bottom of the filter functions.

The blur matrix, $\mathbf{B}$, is constructed from the coefficients of the blur functions:

$$\mathbf{B}=\begin{bmatrix}\mathbf{B}_1 & \mathbf{B}_2 & \cdots & \mathbf{B}_K\end{bmatrix},\quad \mathbf{B}_k=\begin{bmatrix} \mathbf{B}_1^k & \underline{\mathbf{0}} & \cdots & \underline{\mathbf{0}} \\ \vdots & \mathbf{B}_1^k & & \vdots \\ \mathbf{B}_M^k & \vdots & \ddots & \underline{\mathbf{0}} \\ \underline{\mathbf{0}} & \mathbf{B}_M^k & & \mathbf{B}_1^k \\ \vdots & \vdots & \ddots & \vdots \\ \underline{\mathbf{0}} & \underline{\mathbf{0}} & \cdots & \mathbf{B}_M^k \end{bmatrix},\quad \mathbf{B}_i^k=\begin{bmatrix} b_{i1}^k & 0 & \cdots & 0 \\ \vdots & b_{i1}^k & & \vdots \\ b_{iM}^k & \vdots & \ddots & 0 \\ 0 & b_{iM}^k & & b_{i1}^k \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & b_{iM}^k \end{bmatrix} \tag{4}$$

The size of $\mathbf{B}_i^k$ is $(M+N-1)\text{x}N$, the size of $\mathbf{B}_k$ is $(M+N-1)^2\text{x}N^2$, and $\mathbf{B}$ is a $(M+N-1)^2$ x KN^2 matrix (i=1,…,M, k=1,…,K). $\underline{\mathbf{0}}$ denotes for the zero matrix of appropriate size.

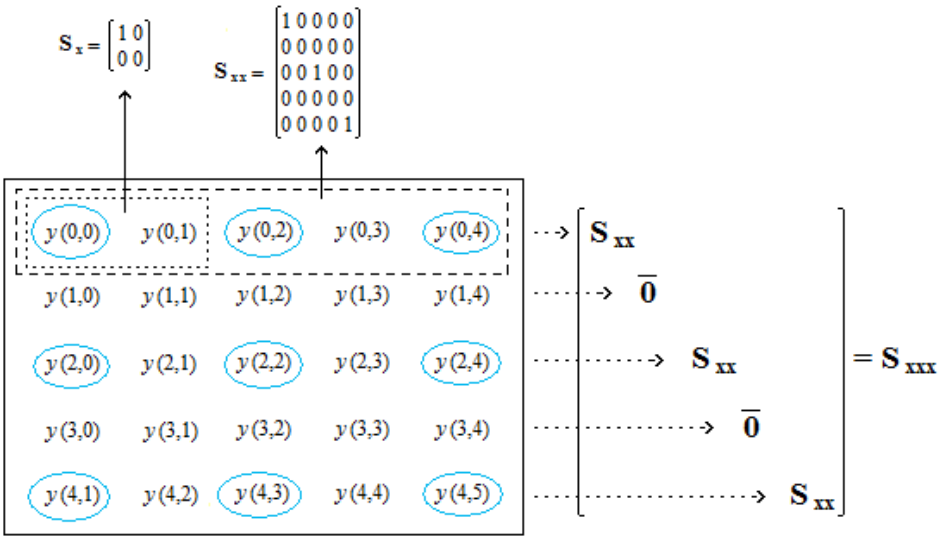

FIGURE 4: The formation of the subsampling matrix for the case of one LR image.

The subsampling matrix **S** can be defined as:

$$\mathbf{S}_x = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}, \quad \mathbf{S}_{xx} = \begin{bmatrix} \mathbf{S}_x & \underline{\mathbf{0}} & \cdots & \underline{\mathbf{0}} \\ \underline{\mathbf{0}} & \mathbf{S}_x & \cdots & \vdots \\ \vdots & \vdots & \ddots & \underline{\mathbf{0}} \\ \underline{\mathbf{0}} & \cdots & \underline{\mathbf{0}} & \end{bmatrix}, \quad \mathbf{S}_{xxx} = \begin{bmatrix} \mathbf{S}_{xx} & \underline{\mathbf{0}} & \cdots & \underline{\mathbf{0}} \\ \underline{\mathbf{0}} & \underline{\mathbf{0}} & \cdots & \underline{\mathbf{0}} \\ \vdots & \vdots & \mathbf{S}_{xx} & \vdots \\ \underline{\mathbf{0}} & \underline{\mathbf{0}} & \cdots & \ddots \end{bmatrix},$$

$$\mathbf{S} = \begin{bmatrix} \mathbf{S}_{xxx} & \underline{\mathbf{0}} & \cdots & \underline{\mathbf{0}} \\ \underline{\mathbf{0}} & \mathbf{S}_{xxx} & \cdots & \vdots \\ \vdots & \vdots & \ddots & \underline{\mathbf{0}} \\ \underline{\mathbf{0}} & \cdots & \underline{\mathbf{0}} & \mathbf{S}_{xxx} \end{bmatrix}. \tag{5}$$

The size of $\mathbf{S_x}$ is D_sxD_s, the size of $\mathbf{S_{xx}}$ is NxN, the size of $\mathbf{S_{xxx}}$ is N^2xN^2, and **S** is a KN^2xKN^2 square matrix. The repetition rate of $\mathbf{S_{xx}}$ on the main diagonal of $\mathbf{S_{xxx}}$ is $1/D_s$. As explained before, the subsampling matrix defined here does not reduce the size of the image; it retains the desired pixel values while zeroing the others.

A visual example is given in Fig. 4 to have an insight on how the subsampling matrix defined above performs the subsampling process on the blurred input vector. In this example, the following values are used:

- The number of LR images: K = 1,
- Subsampling rate: D_s = 2,
- The size of the reconstruction filters: N = 5,
- Location of the pixel to be reconstructed: (n_1,n_2) = (2, 2).

The pixels in circles are the pixels we want to retain. The example is given for K=1. For K >1, the subsampling matrix **S** is formed by repetition of $\mathbf{S_{xxx}}$ on the main diagonal.

Based on the definitions above, the input-output relation of the observation model given in Fig. 2 is given as:

$$\mathbf{y}(n_1,n_2) = \mathbf{x}(n_1,n_2)\cdot\mathbf{B}\cdot\mathbf{S}\,. \tag{6}$$

The problem we wish to solve is the following: Given $\mathbf{y}(n_1,n_2)$, $\mathbf{B}$ and $\mathbf{S}$, determine the existence and uniqueness conditions for the reconstructing filter set shown in Fig. 3 such that $x(n_1,n_2)$ is obtained perfectly for all (n_1,n_2).

3. EXISTENCE-UNIQUENESS OF THE RECONSTRUCTION FILTERS

To reconstruct the original image, an FIR filter set is applied on the degraded images as shown in Fig. 3. Let us define the reconstruction vector $\mathbf{w}$ as the concatenation of the coefficients of each reconstruction filter:

$$\mathbf{w} = \left[\left(\mathbf{w}^1\right)^{\mathrm{T}} \quad \left(\mathbf{w}^2\right)^{\mathrm{T}} \quad \cdots \quad \left(\mathbf{w}^K\right)^{\mathrm{T}} \right]^{\mathrm{T}}, \quad \mathbf{w}^k = \left[w_{11}^k \quad w_{12}^k \quad \cdots \quad w_{(N-1)N}^k \quad w_{NN}^k \right]^{\mathrm{T}}.$$

$\mathbf{w}^k$ is a column vector of length N^2 and $\mathbf{w}$ is a column vector of length KN^2. The estimated HR image can be expressed as:

$$\hat{x}(n_1,n_2) = \mathbf{y}(n_1,n_2)\cdot\mathbf{w} \tag{7}$$

For exact reconstruction, the estimated HR image must be equal to a shifted version of the original image for all (n_1,n_2) (possibly excluding the borders):

$$\hat{x}(n_1,n_2) = x(n_1-\alpha, n_2-\beta) \tag{8}$$

The shift is expressed by (α, β). Combining (6) and (7), the requirement for exact reconstruction can be written as

$$\mathbf{B}\cdot\mathbf{S}\cdot\mathbf{w} = \mathbf{e}_{\alpha,\beta} \tag{9}$$

where $\mathbf{e}_{\alpha,\beta}$ is a column vector of size $(M+N-1)^2$ in which the location of the only nonzero element is determined by the shift (α, β). When $\mathbf{B}$, $\mathbf{S}$ and (α, β) are given, the reconstruction vector $\mathbf{w}$ that satisfies (9) becomes the appropriate column of the (pseudo) inverse of $\mathbf{B}\cdot\mathbf{S}$. For (9) to have at least one solution (i.e. the existence of the filter set that satisfies (9)), $\mathbf{B}\cdot\mathbf{S}$ must have full row rank, and the uniqueness of the solution is guaranteed when either $\mathbf{B}\cdot\mathbf{S}$ is a non-singular square matrix or it has full column rank. Consequently, before deriving the existence and uniqueness conditions, the rank properties of $\mathbf{B}\cdot\mathbf{S}$ will be investigated first.

Definition. A set of 2-D FIR transfer functions, $B_1(z_1,z_2)$, $B_2(z_1,z_2)$, ..., $B_k(z_1,z_2)$, are said to be strongly co-prime if there does not exist a zero (ζ_1,ζ_2) common to all transfer functions [21], i.e., there does not exist (ζ_1,ζ_2): $B_k(\zeta_1,\zeta_2) = 0$, $\forall k = 1, ..., K$.

Theorem 1. Let $\mathbf{B}$ and $\mathbf{S}$ be defined as (4) and (5), respectively. Then, the rank of $\mathbf{B}\cdot\mathbf{S}$ becomes $(M+N-1)^2$ if the following two conditions are satisfied:

1. $K \cdot \left\lfloor N / D_s^2 \right\rfloor^2 \geq (M+N-1)^2$
2. The 2-D Z-transforms of the blur filters, $b_1(n_1,n_2)$, $b_2(n_1,n_2)$, …, $b_K(n_1,n_2)$, are strongly co-prime.

Proof. In the proof of Theorem 1, we will utilize the Sylvester's Inequality that gives an upper bound for the rank of product of two matrices [29]. Sylvester's Inequality is given by:

$$\text{rank}(\mathbf{B}\cdot\mathbf{S}) \le \min\{\text{rank}(\mathbf{B}),\text{rank}(\mathbf{S})\}. \quad (10)$$

Eq. (10) means that the rank of **B**·**S** can not be larger than the rank of **B** or the rank of **S**. Recall that the sizes of **B** and **S** are $(M+N-1)^2 \times KN^2$ and $KN^2 \times KN^2$, respectively. Therefore, **B**·**S** is full row rank if its rank is equal to $(M+N-1)^2$. If the rank of **B**·**S** is $(M+N-1)^2$, then the rank of **B** is also $(M+N-1)^2$ according to the Sylvester's Inequality, that is, **B** must have full row rank. **B** having full row rank means that when the subsampling operator does not exist (when **S** is the unity matrix), i.e. in the multichannel image deconvolution problem, the reconstruction filters that satisfy (9) do exist. The analysis of this situation is done in [21] and the necessary conditions are derived as: (i) **B** must have more columns than rows $\{KN^2 \ge (M+N-1)^2\}$, (ii) The 2-D Z-transforms of the blur filters must be strongly co-prime. The proof of this proposition is given in [21].

According to the Sylvester's Inequality, the rank of **S** must be at least $(M+N-1)^2$ for **B**·**S** to have full row rank. If we consider that **S** is a diagonal matrix whose nonzero elements appear only on the main diagonal, then the rank of **S** becomes the number of 1's that it contains:

$$\begin{aligned}\text{rank}(\mathbf{S}) &= (\#\text{ of 1's in } \mathbf{S_x}) \times (\#\text{ of } \mathbf{S_x}\text{'s in } \mathbf{S_{xx}}) \times (\#\text{ of } \mathbf{S_{xx}}\text{'s in } \mathbf{S_{xxx}}) \times (\#\text{ of } \mathbf{S_{xxx}}\text{'s in } \mathbf{S}) \\ &= \lfloor N/D_s^2 \rfloor \cdot \lfloor N/D_s^2 \rfloor \cdot K \end{aligned} \quad (11)$$

where $\lfloor x \rfloor$ means the smallest integer larger than x. As a result, the following inequality arises for the rank of **S**:

$$K \cdot \lfloor N/D_s^2 \rfloor^2 \ge (M+N-1)^2 \quad (12)$$

Note that if the requirement in (12) is satisfied, then the requirement about the rank of **B** which is given by $KN^2 \ge (M+N-1)^2$ is also satisfied. Let us rewrite the Sylvester's Inequality when the conditions about the ranks of **B** and **S** are satisfied:

$$\text{rank}(\mathbf{B}\cdot\mathbf{S}) \le (M+N-1)^2 \quad (13)$$

To prove Theorem 1, we must show that the inequality in (13) becomes equality, i.e. the rank of **B**·**S** is $(M+N-1)^2$ if the requirements about the ranks of **B** and **S** are satisfied. Three matrix properties will be utilized to get the result:

Let **C** be an arbitrary matrix of size $m \times n$ and **D** of size $n \times p$.

Property 1. Let $\mathbf{C}^c$ be constructed by applying column permutations on **C**, and let $\mathbf{D}^r$ be constructed by applying the corresponding row permutations on **D**. Then the following property holds:

$$\mathbf{C}\cdot\mathbf{D} = \mathbf{E} \quad \Rightarrow \quad \mathbf{C}^c \cdot \mathbf{B}^r = \mathbf{E}.$$

Property 2. If **C** is of the form [$\mathbf{C}_1$ **0**] and if we express **D** as $\mathbf{D} = \begin{bmatrix}\mathbf{D}_1 \\ \mathbf{D}_2\end{bmatrix}$ where the number of rows in $\mathbf{D}_2$ is equal to the number of columns in **0** matrix, the following equality holds:

$$\mathbf{C}\cdot\mathbf{D} = \mathbf{C}_1 \cdot \mathbf{D}_1.$$

Property 3. rank(**C**) = $n \Rightarrow$ rank(**C**·**D**) = min{ n, rank(**D**) } [29].

To utilize Property 3, **B**·**S** multiplication matrix must have a special structure. For this purpose, we must operate on **B**·**S** by using Properties 1 and 2. These operations are as follows:

i. $(\mathbf{B}\cdot\mathbf{S})^T = \mathbf{S}^T\cdot\mathbf{B}^T = \mathbf{S}\cdot\mathbf{B}^T$. If $\mathbf{B}\cdot\mathbf{S}$ is full row rank, then $(\mathbf{B}\cdot\mathbf{S})^T$ is full column rank. $\mathbf{B}^T$ is a matrix of size KN^2 x $(M+N-1)^2$ and has full column rank as explained above.

ii. By applying column exchange operations on matrix **S**, all the columns that are comprised of zeroes are forced to be on the righthandside of the matrix **S**. The same operations are done on the rows of $\mathbf{B}^T$. Let us denote the resulting matrices as $\mathbf{S}^c$ and $\mathbf{B}^{T,r}$. According to Property 1, the multiplication $\mathbf{S}\cdot\mathbf{B}^T$ will be equal to $\mathbf{S}^c\cdot\mathbf{B}^{T,r}$.

iii. $\mathbf{S}^c$ matrix is of the form $[\mathbf{S}_1^c \ \underline{\mathbf{0}}]$. Let us write the matrix $\mathbf{B}^{T,r}$ as $\mathbf{B}^{T,r} = \begin{bmatrix}\mathbf{B}_1^{T,r}\\ \mathbf{B}_2^{T,r}\end{bmatrix}$. By using Property 2, we have $\mathbf{S}\cdot\mathbf{B}^T = \mathbf{S}^c\cdot\mathbf{B}^{T,r} = \mathbf{S}_1^c\cdot\mathbf{B}_1^{T,r}$.

iv. The size of $\mathbf{S}_1^c$ is KN^2 x $K\cdot\lfloor N/D_s^2\rfloor^2$ and its rank is $K\cdot\lfloor N/D_s^2\rfloor^2$. The size of the reduced transposed blur matrix, $\mathbf{B}_1^{T,r}$, is $K\cdot\lfloor N/D_s^2\rfloor^2$ x $(M+N-1)^2$. If we think that the requirement given in (12) is valid, then the number of rows of $\mathbf{B}_1^{T,r}$ is still larger than its number of columns. Consequently, $\mathbf{B}_1^{T,r}$ is still of full column rank, i.e. its rank is $(M+N-1)^2$. Now we can compute the rank of $\mathbf{S}\cdot\mathbf{B}^T$ by using Property 3:

$$\text{rank}(\mathbf{S}\cdot\mathbf{B}^T) = \text{rank}(\mathbf{S}_1^c\cdot\mathbf{B}_1^{T,r}) = \min\{K\cdot\lfloor N/D_s^2\rfloor^2, (M+N-1)^2\} = (M+N-1)^2. \quad (14)$$

v. The rank of a matrix is equal to the rank of its transpose, i.e. $\text{rank}(\mathbf{B}\cdot\mathbf{S}) = \text{rank}(\mathbf{S}\cdot\mathbf{B}^T) = (M+N-1)^2$. Thus, we can conclude that when (12) is satisfied, the multiplication $\mathbf{B}\cdot\mathbf{S}$ has full row rank and the proof is complete. ■

Assuming that **B.S** has full row rank, the existence and uniqueness conditions for (9) is given in Theorem 2.

Theorem 2. Let us assume that $K\cdot\lfloor N/D_s^2\rfloor^2 \geq (M+N-1)^2$ and $\mathbf{B}\cdot\mathbf{S}$ has full row rank, i.e. its rank is $(M+N-1)^2$. Then for a given shift (α, β), the reconstruction filter set **w** that satisfies (9), exist. The uniqueness of the solution is guaranteed if the inequality becomes equality (in this case, $\mathbf{B}\cdot\mathbf{S}$ becomes a square matrix) or the minimum norm solution is obtained.

Proof. For the system given in (9) to have at least one solution, the full row rank property of $\mathbf{B}\cdot\mathbf{S}$ is adequate [30]. The full row rank property of $\mathbf{B}\cdot\mathbf{S}$ requires that $\mathbf{B}\cdot\mathbf{S}$ must have more columns than rows. The inequality $K\cdot\lfloor N/D_s^2\rfloor^2 \geq (M+N-1)^2$ means that $\mathbf{B}\cdot\mathbf{S}$ has more columns than rows. Consequently, when the inequality is valid and $\mathbf{B}\cdot\mathbf{S}$ has full row rank, the system has at least one solution. If (12) is satisfied as an equality, then $\mathbf{B}\cdot\mathbf{S}$ is a square non-singular matrix, and the solution is unique. If (12) is satisfied as an inequality, then the minimum norm solution which is known to be unique is selected. ■

To sum up, the requirement in (12) and the co-primeness condition of the 2-D Z-transforms of the blur functions are the necessary and sufficient conditions for the existence of the perfect reconstruction filters. Some direct results of (12) can be stated as follows:

- K must be at least D_s^2+1.
- Super-resolution without motion is possible as long as the requirement in (12) and the co-primeness condition is satisfied.

4. A PERFECT SUPER-RESOLUTION ALGORITHM

Using (9), the reconstruction filter vector can be found as:

$$\hat{\mathbf{w}} = \text{pinv}(\mathbf{B} \cdot \mathbf{S}) \cdot \mathbf{e}_{\alpha,\beta} \tag{15}$$

where pinv(**A**) denotes for the pseudo-inverse or generalized inverse of a matrix **A**. Let us consider a set of linear equations expressed as **y** = **Ax**. Generally the solution **x** cannot be found or it is not unique. However, a vector **x** can always be found that minimizes $||\mathbf{Ax}-\mathbf{y}||^2$ where ||.|| denotes the Euclidean norm. When formulated in this manner, the solution is unique, and it is given by **x** = pinv(**A**).**y**. Singular value decomposition is used to compute the pseudo-inverse of a matrix. If **A** has full row rank, then the pseudo-inverse becomes pinv(**A**) = $\mathbf{A}^T(\mathbf{A}\mathbf{A}^T)^{-1}$. If **A** does not have full row rank, a method like the one in [22] can be utilized to compute pinv(**A**). In the simulations, the pinv(.) function of MATLAB is used.

Let us express $\hat{\mathbf{w}}$ in (15) as $\mathbf{w}_{\alpha,\beta}$ to indicate that it is obtained for a shifted version of the original image. In this case, the estimated image can be written as:

$$\hat{x}_{\alpha,\beta}(n_1,n_2) = \sum_{k=1}^{K} y_k(n_1,n_2) * w^k_{\alpha,\beta}(n_1,n_2) \tag{16}$$

where * stands for the 2-D convolution operator and $w^k_{\alpha,\beta}(n_1,n_2)$ are the reconstruction filters for a given shift (α, β). Note that the size of the original image is D_sN_1xD_sN_2, where N_1xN_2 is the size of the LR images. The estimation of the image in (16) does not suffice to represent all of the original image, it is equal to the original image for pixel locations (n_1, n_2) = ($i.D_s$, $j.D_s$), i = 0, 1, …, N_1, j = 0, 1, …, N_2 and zero for others (if the existence requirements are satisfied). The reason for this phenomena is that there is not a single subsampling matrix. One can construct D_s^2 distinct subsampling matrices. As a result, there will be D_s^2 reconstruction filter sets. For each N_1xN_2 sized portion of the original image, a different reconstruction filter set will have to be be used.

Instead of creating D_s^2 different subsampling matrices, all of the original image (possibly excluding the borders) can be reconstructed by utilizing the information about the shift (α, β). This method is given as follows:

- Estimate N using (12) when M and K are given.
- Construct **B** and **S**.
- For all $0 \le \alpha, \beta < D_s$-1, find $\mathbf{w}_{\alpha,\beta}$ using (15).
- For all α and β, find $\hat{x}_{\alpha,\beta}(n_1,n_2)$ using (16).
- For all α and β, shift $\hat{x}_{\alpha,\beta}(n_1,n_2)$ by (-α, -β).
- Reconstruct $\hat{x}(n_1,n_2) = x(n_1,n_2)$ by using $\hat{x}(n_1,n_2) = \sum_{\alpha=0}^{Ds-1}\sum_{\beta=0}^{Ds-1}\hat{x}_{\alpha,\beta}(n_1,n_2)$.

The purpose of the method presented above is to validate the propositions given in Section 3 by simulation. We do not claim that it can be used in real situations. In reality, the problem is not fully determined like assumed here. It is ill-posed because of reasons such as additive noise, non-availability of the blur operators, errors in estimation of the motion vectors, insufficient number of available LR images, etc. In the next section, the effects of these factors in estimating the HR image will be investigated.

5. SIMULATION RESULTS

To validate the analysis above, some computer simulations are performed on a Lena image of size 200x200. The image is first blurred, then subsampled (motion is considered only in the last part of the simulations, as long as the conditions for existence are satisfied, there is no need for

motion). Blurring and subsampling are applied numerous times on the original image to obtain many LR images. The parameters of the blur functions are chosen randomly, the only constraint about the blur functions is that they do not change the energy of the original image. A number of simulations are performed for different blur sizes, different subsampling rates and different number of available LR images. For each case, the HR image is estimated using the method explained in the previous section.

In Table 1, the minium size of the reconstruction filters for perfect reconstruction of the original image is given for different subsampling rates, number of LR images and blur filter sizes. The size of the reconstruction filters is estimated by using (12). D_S is the subsampling rate, K is the number of available LR images, and M is the size of the blur filters. N is the size of the FIR reconstruction filters for exact super-resolution under given conditions. It is assumed that all the blur filters are of the same size, and the reconstruction filters are also of the same size. In the table, it can be seen that as long as the number of LR images is more than the square of the subsampling rate, a reconstruction filter set can be defined for error-free reconstruction of the original image. When the number of available LR images increases, the size of reconstruction filters for exact super-resolution decreases. There is a relationship between the blur and reconstruction filter sizes, when the former increases, the latter also increases as expected.

Now, the effect of various factors on exact super-resolution will be investigated. These factors are: (i) the number of available LR images, (ii) the reconstruction filter size, (iii) the case of linearly dependent blur functions, (iv) additive noise, and (v) existence of motion. Before the investigation, exact super-resolution is performed for the ideal case. The parameters are chosen as D_S=3, K=11, and M=5, and the reconstruction filter size is chosen as 19 for error-free super-resolution. The mean-square error (MSE) value between the original and the reconstructed images defined as

$$\text{MSE} = \frac{1}{D_S^2 N_1 N_2} \sum_{i=0}^{D_S N_1 - 1} \sum_{j=0}^{D_S N_2 - 1} \left[x(i,j) - \hat{x}(i,j) \right]^2 \tag{17}$$

is used as a measure to assess the estimation performance. When the method is run with the defined parameters, the original image can be perfectly reconstructed with zero MSE. Now, the factors that affect exact super-resolution will be investigated one by one.

(i) The effect of the number of LR images on exact super-resolution:

While holding the other factors same (D_S=3, M=5, N=19), two simulations are performed for K=10 and K=9. The results are shown in Fig. 5. For K=10, the reconstructed image does not look very different from the original one although the MSE is a little bit high. For K=9, visual artifacts are visible and the MSE is too high.

In Fig. 6, the change of the signal to residual noise ratio (SRNR) with respect to the number of LR images is given. By residual noise, we mean the noise that is still present (that could not be removed) after the reconstruction of the HR image. SRNR is given by

D_S	K	M	N	D_S	K	M	N
2	≤ 4	> 1	∞	3	≤ 9	> 1	∞
	5	3	9		10	3	5
		5	25			5	37
		7	43			7	73
	6	3	5		11	3	5
		5	13			5	19
		7	23			7	37

TABLE 1: Minimum reconstruction filter size for exact super-resolution for different cases.

(a) (b) (c)

Figure 5: The effect of the number of LR images. (a) Original image. (b) The result for *K*=10 (MSE=9.89). (c) The result for *K*=9 (MSE=155.98).

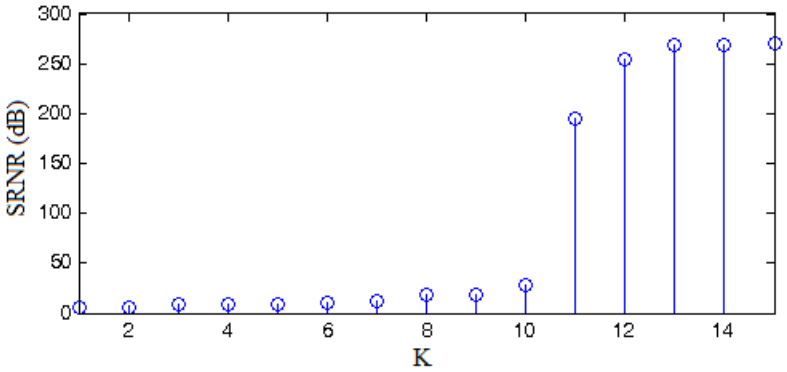

FİGURE 6: The change of signal to residual noise ratio with respect to the number of LR images.

$$\mathrm{SRNR} = 10\log_{10}\left(\frac{\sum_{i=0}^{D_sN_1-1}\sum_{j=0}^{D_sN_2-1}[x(i,j)]^2}{\sum_{i=0}^{D_sN_1-1}\sum_{j=0}^{D_sN_2-1}[x(i,j)-\hat{x}(i,j)]^2}\right) \quad (18)$$

For SRNR levels larger than 70 dB, perfect reconstruction is said to be achieved, because when the estimated image is quantized, it is basically the same as the original image for such SRNR levels. If the SRNR is smaller than 70 dB but larger than 30 dB, perfect reconstruction is not achieved, but the estimated image does not contain visually disturbing artifacts. In Fig. 6, it can be seen that perfect reconstruction of the original image is achieved for *K*=11 or larger. For *K*=10, we do not have perfect reconstruction, but the estimated image is visually acceptable. For *K*=9 or less, the SRNR is less than 30 dB, and the estimated image has visual artifacts. Although the visual results are not shown here, increasing the size of the reconstruction filters improves the visual quality when there are insufficient number of LR images.

(ii) The effect of the size of the reconstruction filters on exact super-resolution:

While holding the other factors unchanged (D_s = 3, M = 5, K = 11), two simulations are performed for N = 15 and N = 11. The simulation results are shown in Fig. 7. For N = 15, the MSE is low, and the visual difference between the original and estimated images can be seen only with a careful examination. For N = 11, the MSE value is high and the estimated image has visual degradations.

(a) (b)

FİGURE 7: The effect of the size of the reconstruction filters. (a) The result for *N*=15 (MSE=11.05). (b) The result for *N*=11 (MSE=35.34).

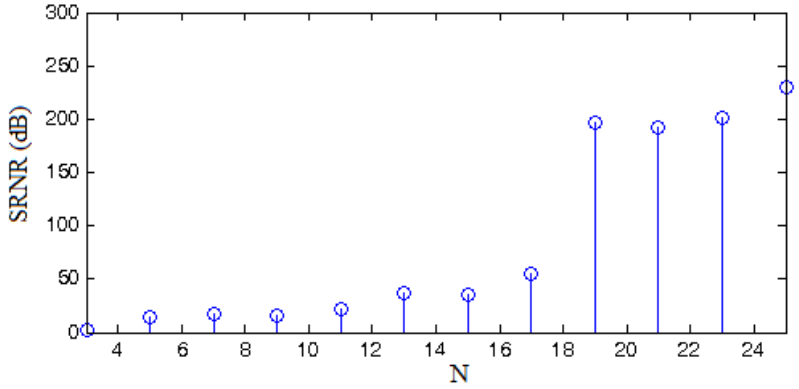

FİGURE 8: The change of SRNR with respect to the size of the reconstruction filters.

(a) (b)

FİGURE 9: The effect of linearly dependent blur functions. (a) The result for one linearly dependent blur function (MSE=42.84). (b) The result for three linearly dependent blur functions (MSE=77.35).

In Fig. 8, the change of SRNR with respect to the reconstruction filter size is given. In accordance with the theory, error-free super-resolution is accomplished when *N*=19 or larger. For *N*=17, 15 and 13, exact super-resolution is not attained, but the results are visually satisfactory. For *N*=11 or below, the SRNR is too low and the reconstructed image has visual artifacts. Hence, it can be said that the threshold value for the reconstruction filter size specified by (12) can be tolerated to some extend if there is a good reason to do that, i.e. to save some processing time.

(iii) The effect of linear dependence of the blur functions on exact super-resolution:

While the other factors are same, two simulations are done for the cases of one or more blur functions that are dependent on the others. In the first simulation, one of the blur functions is linearly dependent on the others, while in the second simulation, three of them are dependent on the others. The visual results are given in Fig. 9. The visual results show that degradation in the estimated image starts even if the number of linearly independent blur functions is just one below the necessary value. The images that have dependent blur functions do not provide the diversity that is necessary for super-resolution. Having dependent blur functions is essentially the same as having insufficient number of LR images.

(iv) The effect of additive noise on exact super-resolution:

The additive noise was assumed to be absent in all above analysis and simulations. Now, while the other factors are the same, two simulations are performed for the case of additive noise. The noise is added on the LR images such that the signal to noise ratio (SNR) is 60 dB in the first experiment and 40 dB in the second experiment. The visual results are given in Fig. 10. The result is visually good for 60 dB case while it is very disturbing for 40 dB additive noise. In Fig. 11, the change of MSE versus additive noise SNR is given. After a point between 40-50 dB, a significant increase in MSE is observed. This is expected because noise is not considered in the analysis of Section 3. In [21], it is stated that increasing the filter size or the number of available degraded images provides more robustness against noise. The result when the size of the reconstruction filters is almost doubled and K is set to 13 is shown in Fig. 10(c). The MSE is significantly dropped and the image is visually more pleasing. The extension of the analysis in Section 3 to the noisy case is left as an open topic.

(a) (b) (c)

FİGURE 10: The effect of additive noise. (a) The result for SNR=60 dB (MSE=21.57). (b) The result for SNR=40 dB (MSE=1474.13). (c) The result for SNR=40 dB, K=13, N=39 (MSE=47.12).

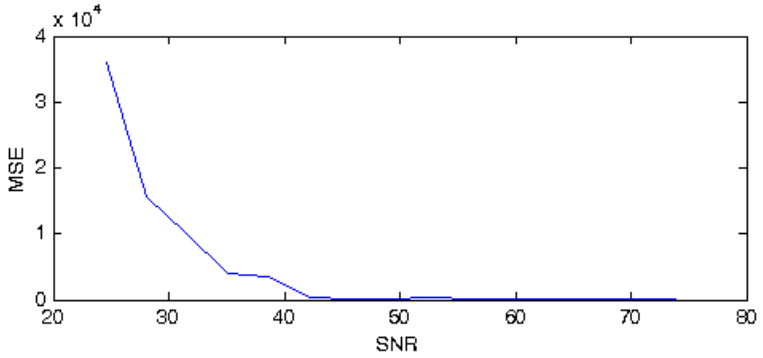

FİGURE 11: The change of MSE versus SNR.

(v) The effect of motion on exact super-resolution:

In this simulation, the effect of motion on super-resolution image reconstruction is investigated. In the first simulation, four blur functions out of eleven is made the same. Normally, this will cause trouble as shown in previous simulations. But the four LR images that are blurred by the same four blur functions are shifted by different amounts before being fed to the algorithm. In this case, the algorithm produced exact results with zero MSE. When the images are shifted in different amounts, the combined motion-blur functions become linearly independent although the blur operators are the same. In other words, the diversity which is necessary for the perfect reconstruction can be achieved by the motion as well as by the linear independence of the blur operators. Then the same simulation is performed for the case of nine similar blur functions and corresponding shifted images, and the same result is observed.

In brief, the simulation results presented in this section practically prove the the analysis of the existence-uniqueness conditions of the reconstruction filters for exact super-resolution. Besides, they aim to add insight to the case when exact super-resolution is not possible. For instance, increasing the reconstruction filter size improves the HR image quality when the number of LR images is low or there is noise in the LR images.

6. CONCLUSION AND FUTURE WORK

In this work, the existence and uniqueness conditions for finite impulse-response restoration filters for exact image super-resolution in the case of pure translational motion (or no motion) and shift-invariant blur are investigated and derived. A method for exact reconstruction is proposed to validate the analysis results. Experimental results demonstrate that as long as the conditions are met, exact super-resolution is possible. If the necessary diversity is provided by the co-primeness of the blur functions, motion is not necessary. When the exact reconstruction conditions are not met, the quality of the estimated HR image is assessed and the conditions for acceptable super-resolution are experimentally derived. When exact or acceptable recovery is jeopardized by factors such as additive noise and insufficient number of LR images, increasing the size of the reconstruction filters is found to be useful in increasing the quality of the result. Analyzing the conditions and choosing the optimum parameters values in case of imperfections is left as a future work.

The reconstruction method proposed is not intended to perform comparably with the well-known super-resolution methods; its sole purpose is to provide a tool to verify the analysis results. It may have practical importance when the conditions are near ideal: known blur functions, little additive noise, sufficient number of LR images, etc. Mostly this is not the case in practice. Based on the analysis about the reconstruction filters, study must be carried out on a more practical super-resolution method that estimates the HR image directly from the LR images and works on more noisy conditions.

7. ACKNOWLEDGMENTS

This work was supported by Scientific and Technological Research Council of Turkey (TUBITAK) under Project Number 107E193.

8. REFERENCES

[1] R.Y. Tsai and T.S. Huang, “Multi-frame image restoration and registration”, Adv. Comput. Vis. Image Process., vol. 1, pp. 317–339, 1984.

[2] M. Irani and S. Peleg, “Improving resolution by image registration”, CVGIP: Graph., Models, Image Process., vol. 53, pp. 231-239, 1991.

[3] H. Stark and P. Oskoui, “High-resolution image recovery from image-plane arrays, using convex projections”, J. Opt. Soc. Amer., vol. 6, pp. 1715-1726, 1989.

[4] A.J. Patti, M.I. Sezan, and A.M. Tekalp, “Super-resolution video reconstruction with arbitrary sampling lattices and nonzero aperture time”, IEEE Trans. Image Process., vol. 6, pp. 1064-1076, 1997.

[5] R.R. Schultz and R.L. Stevenson, “Extraction of high-resolution frames from video sequences”, IEEE Trans. Image Process., vol. 5, pp. 996-1011, 1996.

[6] B.C. Tom and A.K. Katsaggelos, “Reconstruction of a high-resolution image by simultaneous registration, restoration, and interpolation of low-resolution images”, IEEE Int. Conf. Image Process., pp. 539-542, Washington DC, 1995.

[7] M. Elad and A. Feuer, “Restoration of a single super-resolution image from several blurred, noisy, and undersampled measured images”, IEEE Trans. Image Process., vol. 6, pp. 1646-1658, 1997.

[8] S. Farsiu, D. Robinson, M. Elad, and P. Milanfar, “Fast and robust multi-frame super-resolution”, IEEE Trans. Image Process., vol. 13, pp. 1327-1344, 2004.

[9] M. Kang and E. Lee, “Regularized adaptive high-resolution image reconstruction considering inaccurate subpixel registration”, IEEE Trans. Image Process., vol. 12, pp. 826-837, 2003.

[10] H. He and L.P. Kondi, “An image super-resolution algorithm for different error levels per frame”, IEEE Trans. Image Process., vol. 15, pp. 592-603, 2006.

[11] C. Liu and D. Sun, “On Bayesian adaptive video super-resolution”, IEEE Trans. Pattern Anal. Mach. Intell., vol. 36, pp. 346-360, 2014.

[12] H. Takeda, P. Milanfar, M. Protter, and M. Elad, “Super-resolution without explicit subpixel motion estimation”, IEEE Trans. Image Process., vol. 18, pp. 1958-1975, 2009.

[13] S. Babacan, R. Molina, and A. Katsaggelos, “Variational Bayesian super-resolution”, IEEE Trans. Image Process., vol. 18, pp. 1958-1975, 2009.

[14] J. Chen, J. Nunez-Yanez, and A. Achim, “Video super-resolution using generalized Gaussian Markov random fields,” IEEE Signal Process. Lett., vol. 19, pp. 63–66, 2012.

[15] J. Chen, J. Nunez-Yanez, and A. Achim, “Bayesian video super-resolution with heavy-tailed prior models”, IEEE Trans. Circuits Syst. Video Technol., vol. 24, pp. 905-914, 2014.

[16] S.C. Park, M.K. Park and M.G. Kang, “Super-resolution image reconstruction - a technical overview”, IEEE Signal Process. Mag., vol. 20, pp. 21-36, 2003.

[17] S. Farsiu, D. Robinson, M. Elad and P. Milanfar, “Advances and challenges in super-resolution”, Int. J. Imaging Syst. Technol., vol. 14, pp. 47-57, 2004.

[18] J. Tian and K.K. Ma, “A survey on super-resolution imaging”, Signal, Image, Video Process., vol. 5, pp. 329-342, 2011.

[19] Special issue on high resolution image reconstruction, Int. J. Imaging Syst. Technol., vol.14, 2004.

[20] Special issue on super-resolution imaging: analysis, algorithms, and applications, EURASIP J. App. Signal Process., 2006.

[21] G.B. Giannakis and R.W. Heath, Jr., “Blind identification of multichannel FIR blurs and perfect image restoration”, IEEE Trans. Image Process., vol. 9, pp. 1877-1896, 2000.

[22] G. Harikumar and Y. Bresler, “Perfect blind restoration of images blurred by multiple filters: theory and efficient algorithms”, IEEE Trans. Image Process., vol. 8, pp. 202-219, 1999.

[23] H.T. Pai and A.C. Bovik, “On eigenstructure-based direct multichannel blind image restoration”, IEEE Trans. Image Process., vol. 10, pp. 1434-1446, 2001.

[24] F. Sroubek and J. Flusser, “Multichannel blind iterative image restoration”, IEEE Trans. Image Process., vol. 12, pp. 1094-1106, 2003.

[25] M. Elad and Y. Hel-Or, “A fast super-resolution reconstruction algorithm for pure translational motion and common space-invariant blur”, IEEE Trans. Image Process., vol. 10, pp. 1187-1193, 2001.

[26] F. Kara and C. Vural, “Blind image resolution enhancement based on a 2D constant modulus algorithm”, Inverse Probl., vol. 24, article 015010, 2008.

[27] F. Sroubek, G. Cristobal, and J. Flusser, “Simultaneous super-resolution and blind deconvolution,” J. Physics: Conf. Series, vol. 124, article 012048, 2008.

[28] F. Kara and C. Vural, “Exact image super-resolution for pure translational motion and shift-invariant blur”, World Academy of Science, Engineering and Technology, pp. 23-26, Prague, Czech Republic, 2008.

[29] F.R. Gantmakher, The Theory of Matrices, American Mathematical Society, AMS Bookstore, Providence, RI, 2000.

[30] C.D. Meyer, “Generalized inverses and ranks of block matrices”, SIAM J. Appl. Math., vol. 25, pp. 597-602, 1973.