14th European Signal Processing Conference (EUSIPCO 2006), Florence, Italy, September 4-8, 2006, copyright by EURASIP

A BLIND IMAGE SUPER-RESOLUTION ALGORITHM FOR PURE TRANSLATIONAL MOTION

Fatih Kara [1] *and Cabir Vural* [2]

[1] TUBITAK-UEKAE (National Research Institute of Electronics and Cryptology)
P.K. 74, 41470, Gebze, Kocaeli, Turkey
[2] Departmant of Electrical-Electronics Engineering, Sakarya University, Esentepe, Sakarya 54187, Turkey
phone: + (90) 262 6481363, fax: + (90) 262 6481100, email: fkara@uekae.tubitak.gov.tr

ABSTRACT

In almost all super-resolution methods, the blur operator is assumed to be known. However, in practical situations this operator is not available or available only within a finite extend. In this paper, a super-resolution algorithm is presented in which the assumption of availability of the blur parameters is not necessary. It is a two-dimensional and single-input multiple-output extension of the well-known constant modulus algorithm which is widely used for blind equalization in communication systems. The algorithm consists of determining a set of deconvolution filters to be applied on interpolated low-resolution and low-quality images and is suitable for pure translational motion only and shift-invariant blur. Experimental results have shown that the proposed method can satisfactorily reconstruct the high-resolution image and remove the blur especially for five or less-bit images.

1. INTRODUCTION

Super-resolution image reconstruction can be defined as the process of constructing a high-quality and high-resolution image from several shifted, degraded and undersampled ones. In areas such as medical imaging and satellite imaging, where multiple frames of the same scene can be obtained, 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.

In the literature, super-resolution is treated as an inverse problem, where the high-quality and high-resolution image to be obtained is linked to the undersampled images by a series of operators such as warping, blur, decimation and additive noise. Among recent work are projection onto convex sets (POCS) approach, iterative back-projection, maximum a posteriori (MAP) estimation, etc. Excellent tutorials about the subject can be found in [1] and [2].

In almost all above methods, in order for the high-resolution image to be reconstructed, the blur and the motion operators should be known in advance. Although the motion parameters are estimated a priori to some extend, as known to the authors, the blur operator is just assumed to be in hand. But this is hardly the case in practice. Either the blur parameters must be estimated or the high-resolution image must be constructed without the need for the blur parameters, hence the term blind image super-resolution.

In this work, a blind super-resolution image reconstruction method is developed for pure translational motion and shift-invariant blur. Generally, the blur is modelled as a 2-D finite impulse response (FIR) filter and need not be the same for all low-resolution images. The high-resolution image is estimated by superposing the degraded images after they pass through distinct adaptive FIR reconstruction filters whose coefficients are updated by using the 2-D version of the constant modulus algorithm (CMA). CMA [3, 4] is a popular tool in the area of blind equalization in communications, where the aim is to suppress the intersymbol interference (ISI). It can also be used in single-input multiple-output (SIMO) or multiple-input multiple-output (MIMO) systems as well as in single-input single-output (SISO) systems, to reduce the interuser interference (IUI) besides ISI [5].

The idea of the CMA depends on the fact that the source is of constant modulus or from a finite alphabet. In this context, image can be considered as a finite-alphabet source because each pixel is represented by a finite (usually 8) number of bits. Vural and Sethares [6] utilized this property to develop a CMA-based blind single-image blur removal algorithm. The work presented here is essentially an extention of the mentioned algorithm to the single-input multiple-output case.

The paper is organized as follows: In Section 2, the observation model that links the high-resolution image to the observed low-resolution images is presented. If the motion between the LR images consists of only pure translational motion, then the model can be simplified. Based on the simplified model, the super-resolution problem is formulated. In Section 3, the CMA-based high-resolution image reconstruction algorithm is developed. In Section 4, experimental results are presented and some conclusions are drawn in Section 5.

2. SYSTEM DESCRIPTION

Figure 1 shows the observation model that links the high-resolution image to the observed low-resolution ones. In this model, the LR image is obtained by successive operations such as warping, blurring and subsampling on the high-resolution image. In general, the warp operator (denoted by $W_1, W_2, \dots, W_M$, where M is the number of observed low-

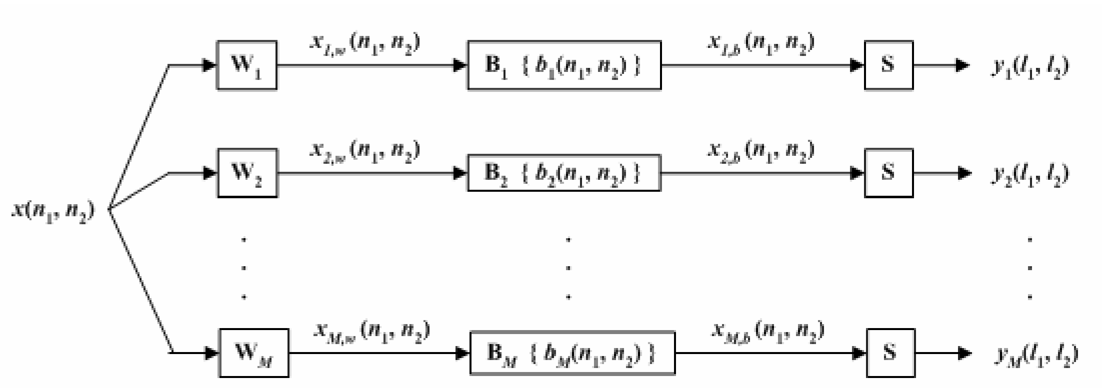

Figure 1 – Observation model relating the low-resolution images to the high-resolution image

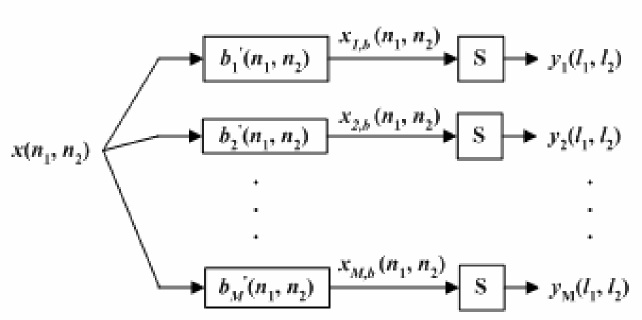

Figure 2 – Simplified observation model

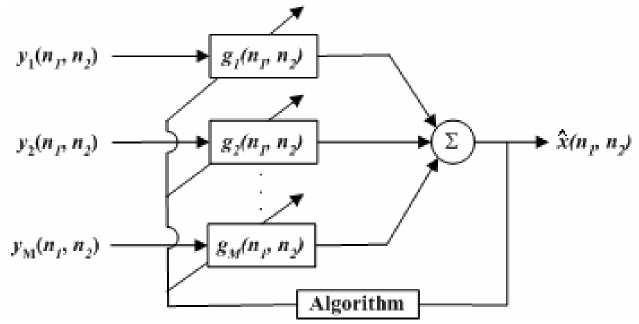

Figure 3 – Reconstruction stage

resolution images) may consist of global or local translation, rotation, etc. The motion therein must contain shifts in subpixel units in order to utilize additional information to construct the HR image. The blurring operation (B_1, …, B_M) results from factors such as relative motion between the imaging system and the scene, out of focus, point-spread function of the sensor, and so on. It is generally modeled as a linear shift-invariant finite impulse-response (FIR) two dimensional filter. The aliased low-resolution image is then generated by subsampling the warped and blurred high-resolution image. Finally, after the addition of the noise (not shown in the figure), the observed low-resolution images are generated.

If the warp operators consist only of global translational motion, then the two-dimensional z-transform of the blurred and warped high-resolution images can be written as follows:

$$X_{i,b}(z_1, z_2) = X(z_1, z_2)\, z_1^{-H_i} z_2^{-V_i} B_i(z_1, z_2). \qquad (1)$$

In Eq. (1), $X(z_1, z_2)$, $X_{i,b}(z_1, z_2)$ and $B_i(z_1, z_2)$ are the two dimensional z-transforms of the respective spatial-domain signals, and H_i and V_i are the horizontal and vertical shifts, respectively, in terms of high-resolution pixel units, for the ith low-resolution image, where $i = 1, 2, \ldots, M$.

If we define

$$B_i'(z_1, z_2) = z_1^{-H_i} z_2^{-V_i} B_i(z_1, z_2)$$

or alternatively

$$b_i'(n_1, n_2) = b_i(n_1 - H_i, n_2 - V_i)$$

then Eq. (1) can be written as

$$X_{i,b}(z_1, z_2) = X(z_1, z_2) B_i'(z_1, z_2)$$

or equivalently

$$x_{i,b}(n_1, n_2) = x(n_1, n_2) * b_i'(n_1, n_2).$$

Hence the warp and the blur operators can be merged into a single blur operator if the only motion is global translation and the observation model in Figure 1 can be reduced to what is seen in Figure 2.

For now, let us pretend that the subsample operator does not exist. In this case, if we define the system output vector $\mathbf{y}(n_1,n_2)$ and the system impulse response vector $\mathbf{B}(n_1,n_2)$ as

$$\mathbf{y}(n_1, n_2) = \begin{bmatrix} y_1(n_1, n_2) \\ \vdots \\ y_M(n_1, n_2) \end{bmatrix}, \quad \mathbf{B}(n_1, n_2) = \begin{bmatrix} b_1'(n_1, n_2) \\ \vdots \\ b_M'(n_1, n_2) \end{bmatrix}$$

then the input-output relation for this particular low-resolution image formation system can be expressed as

$$\mathbf{y}(n_1, n_2) = \mathbf{B}(n_1, n_2) * x(n_1, n_2)$$

where * denotes the 2-D convolution of the vector (or matrix) sequences. For general 2-D matrix sequences $a_{ij}(n_1,n_2)$ and $b_{ij}(n_1,n_2)$, their convolution is defined as

$$[a_{ij}(n_1, n_2)] * [b_{ij}(n_1, n_2)] = \left[\sum_k a_{ik}(n_1, n_2) * b_{kj}(n_1, n_2) \right].$$

To recover the high-resolution and distortion-free image $x(n_1,n_2)$, a set of reconstruction filters is applied to the low-resolution images $y_i(n_1,n_2)$ as in Figure 3. The impulse response sequence of the reconstruction filters, $\mathbf{G}(n_1,n_2)$, satisfies

$$\mathbf{G}(n_1,n_2)*\mathbf{B}(n_1,n_2)=\delta(n_1,n_2).$$

$\mathbf{G}(n_1, n_2)$ is defined as

$$\mathbf{G}(n_1,n_2)=[g_1(n_1,n_2) \quad g_2(n_1,n_2) \quad \cdots \quad g_M(n_1,n_2)].$$

3. ALGORITHM DEVELOPMENT

The proposed algorithm is explained in detail in this section.The constant modulus (CM) cost was introduced by Godard [3] and Treichler and Agee [4] for blind equalization of communication signals that have constant modulus or from a finite alphabet. Recently, it has been reformulated and adapted to the two-dimensional case by Vural and Sethares [14] for blind image deconvolution problem.

Throughout this paper, the high-resolution image pixel values are assumed to have odd integer values between -(L-1) and +(L-1) inclusive, where L is the number of gray levels in the original high-resolution image. It is known that most gray-scale images are 8-bit (256 grey levels), but they can be transformed by uniform or non-uniform thresholding to obtain the desired pixel values. It is also assumed that the gray levels of the true image are independent and identically distributed random variables. A suitable pre-processing for the low-resolution images such as histogram equalization may be necessary for this prerequiste. Under these assumptions, the CM cost is given by

$$J_{CM} := E\left[\left(\hat{x}^2(n_1,n_2)-\gamma\right)^2\right]$$
$$= E\left[\hat{x}^4(n_1,n_2)\right]-2\sigma_x^2\kappa_x E\left[\hat{x}^2(n_1,n_2)\right]+\sigma_x^4\kappa_x^2$$

where γ and κ_x are the dispersion constant and the normalized kurtosis of the true image, respectively. They are defined by

$$\gamma=\frac{E[x^4(n_1,n_2)]}{E[x^2(n_1,n_2)]},\quad \kappa_x=\frac{E[x^4(n_1,n_2)]}{(E[x^2(n_1,n_2)])^2}.$$

The dispersion constant and normalized kurtosis of a zero mean uniformly distributed gray scale image is given in Table 1 for several gray levels.

Gray levels	γ	κ_x
2	1	1
4	8.2	1.64
8	37	1.716
16	152.2	1.790
32	613	1.797
64	2456.2	1.799
128	9829	1.799
256	39320	1.8

Table 1 – Dispersion constant and normalized kurtosis for images having different gray levels

Because a closed-form solution does not exist for minimizing J_{CM} , a stochastic gradient-descent (GD) minimization method is used. A surface called the CM cost surface is generated by plotting the CM cost versus the adaptive filter parameters. The minimization algorithm tries to minimize the cost by starting at some point on the surface, then following the trajectory of the steepest descent. An instantenous estimate of J_{CM} is given by

$$J:=\frac{1}{4}\left(\hat{x}^2(n_1,n_2)-\gamma\right)^2.$$

In Fig. 3, it is shown that the degraded images $y_i(n_1,n_2)$ are applied to a set of 2-D adaptive FIR filters $g_i(n_1,n_2)$ which tries to remove the blur and generate the high-resolution image (note that the subsampling process is still ignored). An estimate of the true image is obtained as the output of the adaptive filters at the jth iteration:

$$\hat{x}_j(n_1,n_2)=\sum_{i=1}^{M}\sum_{a=-A}^{A}\sum_{b=-B}^{B} g_{i,j}(a,b)\,y_i(n_1-a,n_2-b)$$

where AxB is the support of the adaptive filters and $g_{i,j}(a,b)$ are the coefficients for the ith adaptive filter for the jth iteration. This true image estimate is used to obtain a better estimate of the adaptive filter coefficients for the next spatial location in an adaptive manner. The derivative of J with respect to the adaptive filter coefficients is needed in order to implement the GD minimization. Let $\mathbf{g}_j$ denote the lexicographically ordered vector which is composed of the coefficients of the adaptive filters at the jth iteration:

$$\mathbf{g}_j := \begin{bmatrix} \mathbf{g}_{1,j} \\ \mathbf{g}_{2,j} \\ \vdots \\ \mathbf{g}_{M,j} \end{bmatrix},\quad \mathbf{g}_{i,j} = \begin{bmatrix} g_{i,j}(-A,-B) \\ g_{i,j}(-A,-B+1) \\ g_{i,j}(-A,-B+2) \\ \vdots \\ g_{i,j}(A,B) \end{bmatrix},\quad i=1,2,\ldots,M,$$

and let $\mathbf{y}(n_1, n_2)$ be the regressor vector for the (n_1, n_2)th pixel at the jth iteration:

$$\mathbf{y}(n_1,n_2):=\begin{bmatrix} \mathbf{y}_1(n_1,n_2) \\ \mathbf{y}_2(n_1,n_2) \\ \vdots \\ \mathbf{y}_M(n_1,n_2) \end{bmatrix},$$

$$\mathbf{y}_i(n_1,n_2)=\begin{bmatrix} y_i(n_1+A,n_2+B) \\ y_i(n_1+A,n_2+B-1) \\ y_i(n_1+A,n_2+B-2) \\ \vdots \\ y_i(n_1-A,n_2-B) \end{bmatrix},\quad i=1,2,\ldots,M.$$

The estimate of the true image for the (n_1,n_2)th pixel at the jth iteration, using vectors $\mathbf{g}_j$ and $\mathbf{y}(n_1,n_2)$, can be written as

$$\hat{x}_j(n_1,n_2)=\mathbf{g}_j^T\mathbf{y}(n_1,n_2). \tag{2}$$

The derivative of J with respect to $\mathbf{g}_j$ is given by

$$\frac{dJ}{d\mathbf{g}_j}=\left(\hat{x}_j^2(n_1,n_2)-\gamma\right)\hat{x}_j(n_1,n_2)\mathbf{y}(n_1,n_2). \tag{3}$$

Hence, the adaptive filters are updated according to

$$\mathbf{g}_{j+1}=\mathbf{g}_j-\mu\frac{dJ}{d\mathbf{g}_j} \tag{4}$$

where μ is a small positive step-size that guarantees the algorithm stability.

The discussion above assumes that the subsampling operator does not exist. But this is not the case, and an interpolation method is used to obtain a scaled version of the observed low-resolution images as inputs to the adaptive filter set, i.e.

to obtain $y_i(n_1,n_2)$ from $y_i(l_1, l_2)$. The C2-continuous cubic kernel with N=6 supporting points is chosen as the interpolation method because of its superiority over other kernels [7].

Initialization of the adaptive filters is worth mentioning at this point. For each adaptive filter, using a 2-D spike characterized by a non-zero coefficient whose location is determined by the motion vector of the corresponding filter input, i. e.

$$g_{i,1} = \delta(n_1 + H_i, n_2 + V_i), \quad i = 1,2,...,M \tag{5}$$

is found to be useful because it eliminates the effect of the motion at the beginning by initially shifting the interpolated low-resolution images to their original motion-free spaces.

Based on the above discussion, the proposed blind image super-resolution algorithm is summarized below:

(i) Interpolate the observed low-resolution images $y_i(l_1, l_2)$ to obtain $y_i(n_1,n_2)$.

(ii) Initialize the adaptive filters as described in Equation (5).

(iii) Perform Equations (2) to (4) to obtain an estimation of the true image and the adaptive filter coefficients.

(iv) Repeat Step (iii) until a predefined number of iterations is reached or the difference between consecutive image (or adaptive filter) estimations falls below a threshold value.

4. SIMULATION RESULTS

To demonstrate the usefulness of the proposed method, two computer simulation results are provided in this section. A computer-generated one-bit image and Lena images of one, two, three, four and five-bits are used in the simulations. Some pre-processing is applied on the images to fulfill the assumptions made at the beginning of Section 3. Then for each image, four low-resolution images are obtained by simulating the image formation model which is given in Figure 1. As the blur operator, a 3x3 random blur whose coefficients add up to one is chosen to be applied on each shifted image. Note that for each shifted image, a different random blur is used. Then the shifted and blurred images are subsampled by a factor of two in each dimension. Finally, zero-mean Gaussian noise is added to the low-resolution images to complete the image formation model such that the Blurred Signal-to-Noise Ratio (BSNR) of each noisy low-resolution image is approximately 60 dB.

In Figure 4, the original one-bit computer generated image, the image obtained by bilinear interpolation of one of the four low-resolution images, and the image obtained by the proposed method are shown. It is seen that the method has generated the high-resolution image and removed the blur to some degree without the knowledge of the blur parameters. Figure 5 shows the increase as iteration number increases in terms of the frequently-used metric Improvement in Signal-to-Noise Ratio (ISNR).

The results obtained from one and two-bit Lena figures are shown in Figures 6 and 7, respectively. As the number of the bits used to represent the true image pixel values increases,

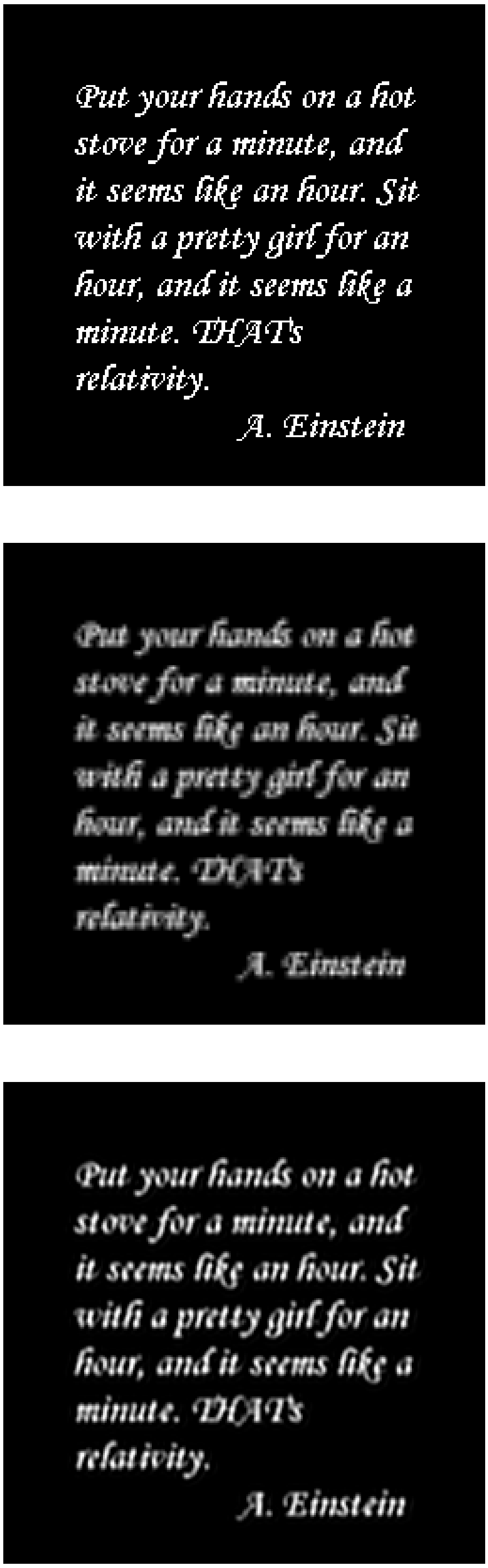

Figure 4 – (Top) computer generated one-bit image, (middle) the image obtained by bilinear interpolation of one of the low-resolution images, (bottom) the image obtained by the proposed method

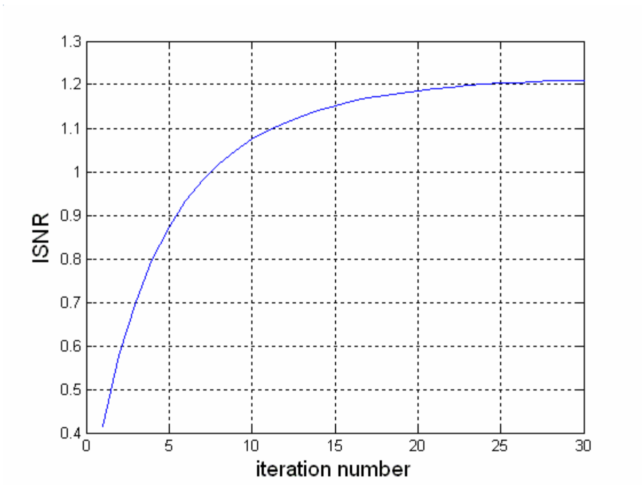

Figure 5 – ISNR vs. iteration number

the improvement in the quality of the image becomes less apparent. This is because, (i) the cost surface flattens as the image kurtosis increases, hence the number of iterations needed to achieve improvement becomes excessively large, and (ii) the step size must be chosen much smaller in order to maintain algorithm stability [8].

5. CONCLUSION

A new blind image super-resolution method in which the blur parameters are not assumed to be known is presented in this paper. It is a 2-D and single-input, multiple-output extension of the CMA which is widely used in the area of blind equalization. An important property of the method is that the blur need not be the same for all low-resolution images, but the only motion allowed is pure translational motion. Simulation results show that the method can recover the high-resolution image and remove the blur for especially five or less-bit images. As the number of gray levels used to represent the true image decreases, the performance of the method increases.

REFERENCES

[1] S. C. Park, M. K. Park, and M. G. Kang, “Super-resolution image reconstruction - a technical overview”, *IEEE Signal Processing Magazine*, vol. 20, pp. 21-36, May 2003.

[2] S. Borman and R. L. Stevenson, “Super-resolution from image sequences – a review”, in *Proc. 1998 Midwest Symp. Circuits and Systems*, 1999, pp. 374-378.

[3] D. Godard, “Self-recovering equalization and carrier tracking in two dimensional data communication systems”, *IEEE Trans. Commun.*, vol. 28, no. 11, pp. 1867-1875, Nov. 1980.

[4] J. R. Treichler and B. G. Agee, “A new approach to multipath correction of constant modulus signals”, *IEEE Trans. Commun.*, vol. 31, no. 2, pp. 459-473, Apr. 1983.

[5] Y. Li and K. J. Ray Liu, “Adaptive blind source separation and equalization for multiple-input/multiple-output systems”, *IEEE Trans. On Information Theory*, vol. 44, no. 7, pp. 2864-2876, Nov. 1998.

[6] C. Vural and W. A. Sethares, “Blind image deconvolution via dispersion minimization”, *Digital Signal Processing* 16 (2006) pp. 137-148.

[7] T. M. Lehmann, C. Gönner, and K. Spitzer, “Survey: interpolation methods in medical image processing”, *IEEE Trans. On Medical Imaging*, vol. 18, no. 11, pp. 1049-1074, Nov. 1999.

[8] J. R. Johnson, et al., “Blind equalization using the constant modulus criterion: A review”, *Proc. IEEE* 86 (10) 1998, pp. 1927-1950.

Figure 6 – (Left) original one-bit Lena image, (middle) the image obtained by bilinear interpolation of one of the low-resolution images, (right) the image obtained by the proposed method

Figure 7 – (Left) original two-bit Lena image, (middle) the image obtained by bilinear interpolation of one of the low-resolution images, (right) the image obtained by the proposed method