

# A MAJORIZATION-MINIMIZATION APPROACH TO TOTAL VARIATION RECONSTRUCTION OF SUPER-RESOLVED IMAGES

Nikolaos Galatsanos

Department of Electrical and Computer Engineering, University of Patras, Rio 26500, Greece  
 phone: + (30) 2610969861, fax: + (30) 2610994798, email: ngalatsanos@upatras.gr

## ABSTRACT

*Super-resolution is the task of reconstructing high-resolution images from shifted, rotated, low-resolution degraded observations. It can be formulated as an inverse problem for which regularization is necessary. In this paper we adopt this formulation and use the Total-Variation criterion for regularization. Then, we employ the Majorization-Minimization (MM) methodology to reconstruct high-resolution images from low-resolution observations. Experimental results are shown, which demonstrate the advantages of the proposed algorithm compared to other methods.*

## 1. INTRODUCTION

The problem of super-resolution is defined as obtaining an image with enhanced resolution from a set of lower resolution degraded images. Many methodologies have been applied to the super-resolution problem; the interested reader is referred to the surveys [1] and [2] and the edited book [3]. An important category of them formulates this problem as an ill-posed image reconstruction problem [5] for which regularization is necessary. Next we briefly survey the literature on this problem focusing mainly on works that adopt the same problem formulation.

In [7], color images and demosaicing are considered, and regularization (image priors) based on the  $L_p$ -norm with  $p = 1$  is proposed in order to avoid the shortcomings of  $L_2$ -norm based regularization. In [6], a computationally fast method is proposed based on the  $L_1$ -norm assuming known integer pixel displacements between frames. However, in [6] and [7] the parameters that define the regularization term are chosen empirically. In [8] an Expectation-Maximization (EM) algorithm and a maximum a posteriori algorithm are presented for simultaneous registration, restoration and interpolation for super-resolution. Although in this work all parameters are chosen automatically a stationary SAR prior is used in both formulations in [8], which is equivalent to  $L_2$ -norm based regularization. In [11] a *maximum a posteriori* (MAP) formulation is proposed that uses a new non-stationary edge preserving hierarchical image prior.

In this paper we introduce the total-variation (TV) regularizer for the super-resolution problem, in order to avoid the shortcomings of the  $L_2$ -norm based regularization. TV based regularization has been very popular for inverse problems in imaging that require regularization, see for example [4], [9], and [10]. It has been shown that TV when used as a regularizer it has the ability to suppress noise in smooth areas of the image and preserve edges.

Because reconstruction with TV regularization requires the solution of a difficult non-quadratic minimization problem, even though the objective function is convex [9], we resort to the majorization-minimization (MM) approach [12]. According to this approach the non-quadratic minimization problem is converted to successive quadratic minimization problems that are easier to solve.

The task of finding the registration parameters is very important and their accurate estimation is critical to the quality of the reconstructed high-resolution image. Thus, in the herein proposed reconstruction algorithm we incorporate the two-step process (pre-registration and reconstruction) proposed in [11] in order to deal with this task.

The rest of the paper is organized as follows. In section 2 we present the imaging model. In section 3 we derive the MM algorithm for TV super-resolution and in section 4 we present the overall algorithm for simultaneous registration and super-resolution reconstruction. In section 5 we present experiments and in section 6 conclusions and thoughts for future work.

## 2. IMAGING MODEL

A linear imaging model is assumed. We denote as  $d$  the integer decimation factor. In other words, the imaging model assumes a high resolution image of size  $N_H \times 1$ , where  $N_H = Nd$ . This model also assumes as observations  $P$  low-resolution images of size  $N \times 1$  by applying the  $PN \times N_H$  degradation operator  $B$  to the high resolution image. Then, white noise is added at each observation. Let  $y$  be a  $PN \times 1$  vector, containing the  $P$  low-resolution images  $y_i$ :

$$y = [y_1^T, y_2^T, \dots, y_P^T]^T,$$

where  $y_i$  is an  $N \times 1$  vector, representing a low-resolution image. Using this notation, the observations are given by:

$$y = Bx + n, \quad (1)$$

where  $x$  the (unknown) original  $N_H \times 1$  high-resolution image to be estimated,  $B$  is a  $PN \times N_H$  imaging operator and  $n = [n_1^T, n_2^T, \dots, n_P^T]^T$  a  $PN \times 1$  vector consisting of  $P$  additive white noise vectors. We assume Gaussian statistics for the noise given by  $n_i \sim N(0, \beta_i^{-1}I)$ ,  $i = 1, 2, \dots, P$ , where  $0$  is a  $N \times 1$  vector with zeros,  $I$  the  $N \times N$  identity matrix respectively, and  $\beta_i, i = 1, \dots, P$  are the noise precisions of the observations that are assumed unknown and statistically independent with each other.  $B$  is given by:

$$B = [B_1^T, B_2^T, \dots, B_P^T]^T$$

where the individual imaging operators are assumed to be  $B_i = DH_iR(\theta_i)S(\delta_i)$  for  $i = 1, \dots, P$ . The matrix  $D$  is the

<sup>1</sup>This research project (PENED) is co-financed by E.U.-European Social Fund (75%) and the Greek Ministry of Development-GSRT (25%)

known  $N \times N_H$  decimation matrix,  $H_i$ ,  $i = 1 \dots P$ , are the shift-invariant  $N_H \times N_H$  blurring convolutional operators, and  $S(\delta_i)$ , for  $i = 1, \dots, P$ , are the  $N_H \times N_H$  shift-invariant shifting operators. Each  $\delta_i$  is a scalar which represents translation (with respect to the first image) and is assumed unknown. The shift operator,  $S(\delta_i)$ , is the Shannon interpolation operator which is shift invariant [8]. The impulse response of the shift operator is given by:

$$S_{shift}(m; \delta_i) = \frac{\sin(\pi(m - \delta_i))}{\pi(m - \delta_i)}, \quad m = 1, 2, \dots, N.$$

The shift invariant operators are assumed circulant. This is very useful for computational purposes because such matrices can be easily diagonalized in the DFT domain. One difficulty that arises in the super resolution problem is the decimation operator which is not square and thus not circulant. In this work we take advantage of the simple form of this matrix, and, despite its non-circulant nature, we obtain tractable calculations in the DFT domain.

Lastly, the  $N_H \times N_H$  matrix  $R(\theta_i)$  represents the rotation operator of each observation relative to the unknown image  $x$ . This operator is not circulant and cannot be converted efficiently in the DFT domain. The imaging model assumes that image  $i$  is a rotated (as well as a shifted) version of the first image, with angle  $\theta_i$ . Using all the above definitions, the imaging model in equation (1) can be rewritten as:

$$y_i = B_i x_i + n_i = D H_i S(\delta_i) R(\theta_i) x + n_i, \quad i = 1, \dots, P. \quad (2)$$

### 3. MM APPROACH FOR TOTAL VARIATION SUPER-RESOLUTION

Super-resolution is an ill-posed problem. In other words, direct estimation of the high-resolution image from equation (1) by minimizing

$$\|y - Bx\|_2^2$$

is not advisable, since in this estimate the high-frequency errors are enhanced [3]. For this reason, regularization is necessary. To this end, we introduce to this problem the Total-Variation (TV) criterion, which is defined as such:

$$TV(x) = \sum_{i=1}^N \sqrt{(\Delta_i^h(x))^2 + (\Delta_i^v(x))^2}, \quad (3)$$

where  $\Delta_i^h(x)$  and  $\Delta_i^v(x)$  represent the horizontal and vertical difference at the pixel location  $i$ , respectively. The TV criterion is a measure of the magnitude of the gradient across the image.

We apply the TV regularization procedure, or, in other words, we use the TV criterion as a regularizer to the super-resolution problem, across the image. TV has been used extensively as a regularizer in inverse problems. Unlike  $L_2$  regularization, it has the ability to suppress high-frequency noise in flat areas of the image and preserve edges, see for example [9] and [10]. Thus, we calculate the high-resolution image by minimizing the function

$$L(x) = \|y - Bx\|_2^2 + \lambda TV(x), \quad (4)$$

where  $\lambda$  is a constant. The addition to the data fidelity term of the TV criterion is a constraint that acts as a penalty to the roughness of the high-resolution image to be computed. The

objective when minimizing (4) is to find an  $x$  that satisfies to a degree Eq. (1) and simultaneously limit the total fluctuation of the image gradient. The reconstructed image is found by minimizing  $L(x)$ :

$$\hat{x} = \underset{x}{\operatorname{argmin}} L(x).$$

However, this minimization is not easy because of the non-quadratic form of the function  $L(x)$ , the discontinuity of the derivative of the TV term at 0, and the large dimension of  $x$ .

In order to minimize  $L(x)$  we employ the majorization-minimization methodology, [12]. According to this methodology the difficult non quadratic minimization problem that has to be solved is replaced by successive quadratic minimization problems which are much easier to solve.

To accomplish this first we find a majorizer for  $TV(x)$  using the relationship

$$\sqrt{z} \leq \sqrt{z^*} + \frac{z - z^*}{2\sqrt{z^*}},$$

which holds for every positive  $z$  and  $z^*$ , and the equality holds when  $z = z^*$ . Setting

$$z_i = (\Delta_i^h(x))^2 + (\Delta_i^v(x))^2, \quad z_i^* = (\Delta_i^h(x^*))^2 + (\Delta_i^v(x^*))^2,$$

for  $i = 1, 2, \dots, N_H$ , in the above equation and ignoring the constant terms, we take the majorizer for  $L(x)$ :

$$\begin{aligned} L(x) &\leq Q(x|x^*), \\ L(x) &= Q(x^*|x^*), \end{aligned}$$

where

$$Q(x|x^*) = \|y - Bx\|_2^2 + \lambda x^T \left( \sum_{k=h,v} (\Delta^k)^T W^* \Delta^k \right) x. \quad (5)$$

$W^*$  is a diagonal matrix with components

$$W_i^* = \frac{1}{\sqrt{(\Delta_i^h(x^*))^2 + (\Delta_i^v(x^*))^2}},$$

and  $\Delta^h$  and  $\Delta^v$  are the horizontal and vertical difference operators, respectively. We have to note that some constant terms in (5) have been dropped for simplicity.

Each iteration ( $t$ ) of the majorization-minimization algorithm consist of the minimization of the majorant function, which is given by equation (5), using the solution of the previous iteration ( $t - 1$ ):

$$x^{(t)} = \underset{x}{\operatorname{argmin}} Q(x|x^{(t-1)}). \quad (6)$$

Thus, at iteration ( $t$ ) the image is estimated by solving a linear system because the function to be minimized is quadratic to  $x$ . In addition it is convex because the Hessian is positive-definite. Hence the image estimate at iteration ( $t$ ) is the single (and hence global) minimum of  $Q(x|x^{(t-1)})$  and is given by

$$x^{(t)} = \left( B^T B + \lambda \sum_{k=h,v} (\Delta^k)^T W^{(t-1)} \Delta^k \right)^{-1} B^T y. \quad (7)$$

This procedure guarantees that the algorithm converges to a minimum of  $L(x)$ . Furthermore, in the present case, it is also the global minimum of  $L(x)$ , since  $L(x)$  is convex due to the TV regularizer [9].

#### 4. REGISTRATION

Apart from reconstructing the super-resolved image, we have to deal with the registration part of the super-resolution problem. In other words, the shifts  $\delta_i$  and the rotation  $\theta_i$  in Eq. 2 have to be estimated. To do this we follow the same methodology proposed in [11]. We perform a pre-processing step in order to "correct" for rotation. During this step we also correct partially for the shifts of the observations with respect to the reference image (we assume the first image as the reference image). However, the shift is corrected only up to an integer pixel factor. The new "corrected" low-resolution images are assumed to be the observations of a new imaging model and they do not contain of rotation (the shift is still assumed present). The purpose of this pre-registration step is to remove rotations from the imaging model and thus make the calculations more tractable. More specifically, after pre-registration we use a new imaging model that does not contain rotation given by

$$z_i = DH_i S(\zeta_i)x + n_i, \quad i = 1, \dots, P, \quad (8)$$

where the new observations have been produced as:

$$z_i = R'_i S'_i y_i,$$

where  $R'_i$  and  $S'_i$  are  $N \times N$  matrices that when applied to the image  $y_i$  they transform it and bring it as closer as possible to the reference image  $y_1$  by correcting its shift and rotation. As mentioned above the shifts are corrected only up to an integer fraction of the pixel. In other words,  $\zeta_i$  are fractions smaller than one. This is intentional because low resolution images that are shifted by a fraction of a pixel are required in order to achieve super-resolution reconstruction [3].

The overall super-resolution algorithm works as follows. After pre-registration in Eq. (8), the high-resolution image is reconstructed and the shift parameters are found using a stationary SAR model and the EM algorithm [8]. The resulting high-resolution image  $x_{STAT}$  is then used to initialize the MM algorithm  $x^{(0)} = x_{STAT}$ . The registration parameters are also initialized using the results of the EM algorithm.

To minimize  $L(x)$  with respect to the image  $x$  and the shift parameters for the super-resolution problem:

$$[\hat{x}, \hat{\zeta}_1, \dots, \hat{\zeta}_P] = \arg \min_{[x, \zeta_1, \dots, \zeta_P]} L(x, \zeta_1, \dots, \zeta_P),$$

where

$$L(x, \zeta_1, \dots, \zeta_P) = \sum_{i=1}^P \|y_i - B_i(\zeta_i)x\|_2^2 + \lambda TV(x).$$

Thus, we alternate between the following two equations Eq. (7) and

$$\zeta_i^{(t)} = \operatorname{argmin}_{\zeta_i} \|y_i - B_i(\zeta_i)x^{(t)}\|_2^2, \quad i = 1, \dots, P, \quad (9)$$

where

$$B = [B_1^T, B_2^T, \dots, B_P^T], \quad B_i = DH_i S(\zeta_i), \quad i = 1, 2, \dots, P.$$

Equation (7) is solved using the conjugate gradient algorithm. The minimization in Eq. (9) is performed in the discrete Fourier domain (DFT) using a Newton-Raphson algorithm. This is possible because the new imaging model in Eq.



Figure 1: Low resolution degraded observation,  $SNR = 20$ , uniform blur  $7 \times 7$

(8) does not include rotations and the remaining operators are either circulant or have a tractable structure in the DFT domain [11]. This also allows calculation of the first and second order derivatives required by the Newton-Raphson algorithm in closed form. Thus, the optimization in Eq. (9) is very fast and efficient.

We have to note that although one dimensional notation is used for the registration parameters for brevity, in our experiments we use a 2-D vector for each image  $\zeta = [\zeta_i(1) \zeta_i(2)]^T$ . This makes the calculations a bit more complicated, because solution of equation 9 involves the inversion of a Hessian matrix of dimension  $2 \times 2$ .

#### 5. EXPERIMENTS

In order to test the proposed methodology, we used both artificially generated and real data. We compared the new TV-based super-resolution algorithm with the EM super-resolution algorithm in [11] that uses a stationary prior. We also compared our super-resolution algorithm with one that uses total variation (TV) regularization [8]. For this comparison a gradient descent algorithm was used given by

$$x^{(t+1)} = x^{(t)} - \alpha \left( B^T (Bx^{(t)} - z) + \lambda' \sum_{k=h,v} (\Delta^k)^T v_k^{(t)} \right) \quad (10)$$

where the superscript  $(t)$  denotes the iteration number,  $v_k^{(t)} = \operatorname{sign}(\Delta^k x^{(t)})$ , with  $\Delta^h x^{(t)}$  and  $\Delta^v x^{(t)}$  the first order horizontal and vertical differences of the image at the  $(t)$ th iteration,  $\lambda'$  the regularization parameter and  $\alpha$  the step of the algorithm. Lastly, we compared the proposed algorithm with the MAP algorithm that uses a non-stationary prior proposed in [11]. In the following experiments the parameters  $\lambda$ ,  $\lambda'$  and parameters were selected by trial and error to provide the best possible results. However, as a general rule as  $\lambda$  and  $\lambda'$  increase the reconstructed image becomes blurrier. For all methods we used the herein proposed registration methodology.

In order to conduct experiments where the ground truth is known, we used synthetic data. Six sets each of eight  $128 \times 128$  low-resolution images were generated by performing translation and rotation to the well-known "Cameraman" image of size  $256 \times 256$ , before blurring and down-sampling by a factor of 2, and lastly contaminated with additive white Gaussian (AWG) noise. Two types of blur and three of noise



Figure 2: Sample of low-resolution observations

were used (resulting in six image sets): uniform point spread function (PSF) of size  $7 \times 7$  and Gaussian with shape parameter equal to 2; AWG noise corresponding to signal to noise ratio (SNR)  $SNR = 40, 20$ , and  $10dB$ . This metric is defined as

$$SNR = 10 \log_{10} \frac{\|z_i\|_2^2}{N\sigma^2} dB,$$

where  $\sigma^2$  is the variance of the additive noise and  $N_H$  is the size of the image  $z_i$ . The Mean Square Error metric (MSE) between the restored image and the original was used to evaluate the performance of the algorithm. The MSE is defined as

$$MSE = \frac{\|x - \hat{x}\|_2^2}{N_H},$$

where  $x$  and  $\hat{x}$  are the original and estimated images, respectively.

In Figure 1 the low-resolution image of the experiment with uniform blur  $7 \times 7$  and  $SNR = 20$  is shown. In Figure 3 we show the super-resolved images and the corresponding MSE for this experiment, (a) using the stationary prior in [8], (b) using regularization as performed in equation (10), (c) the MAP algorithm with the non-stationary prior in [11] and (d) the herein proposed TV regularization via the MM algorithm. In this experiment the MSE using the TV regularizer is lower and the visual quality of the images has improved. In Table 1 we provide the *MSE* results for the six previously described experiments. From these results it is clear that the proposed TV based algorithm provides superior results in all cases as compared to the methods in [8] and [6]. As far as the method in [11] is concerned it seems that the herein proposed method is superior in the cases with the lower *SNR*.

We also used the proposed super-resolution algorithm on a real data set that includes four low-resolution degraded images that contain both translations and rotations. One of them is shown in Figure 2. Each low-resolution image is of size  $128 \times 128$ . The  $4x$  super-resolved images (a) using the stationary prior in [8], (b) using equation (10), (c) the MAP algorithm in [11] and (d) the herein proposed TV regularization via the MM algorithm are shown in Figure 4.

In estimating the shape of the blur for the real data set, a Gaussian-shaped blur was assumed. This choice was motivated by the observation that Gaussian shaped functions are smooth and have good approximation properties. The width of each blur was experimentally estimated using trial and error experiments. The width is captured by the variance of Gaussian PSF. The values of the variances of the Gaussian shaped PSFs was set equal to 4.



(a)



(b)



(c)

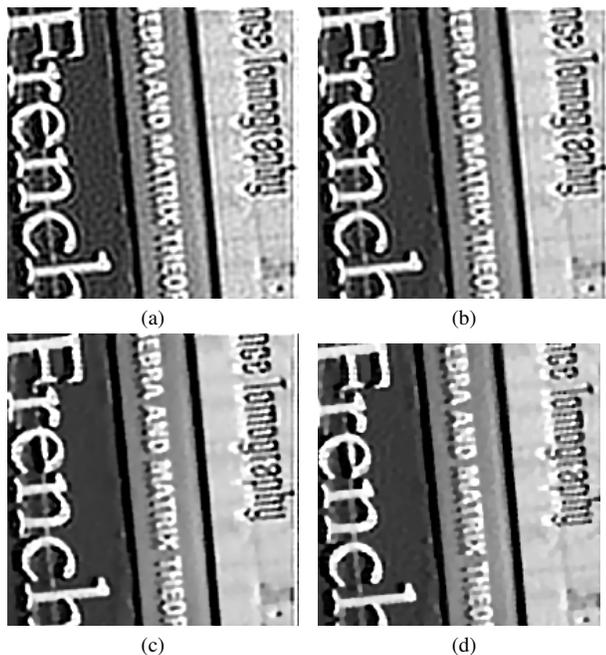


(d)

Figure 3: 2x super-resolved images: (a) stationary [8] ( $MSE = 273$ ), (b)  $L_1$  ( $MSE = 227$ ) (c) non-stationary prior [11] ( $MSE = 202$ ), and (d) TV based regularization ( $MSE = 195$ )

Table 1: MSE's for the experiments using synthetic data

Method	Uniform blur $7 \times 7$			Gaussian blur $\sigma^2 = 2$		
	SNR = 40	SNR = 20	SNR = 10	SNR = 40	SNR = 20	SNR = 10
[8]	208	273	364	187	207	273
[6]	185	227	269	173	185	206
[11]	<b>167</b>	205	260	<b>161</b>	169	203
TV	169	<b>195</b>	<b>256</b>	164	<b>165</b>	<b>200</b>

Figure 4:  $4x$  Super-resolved images: (a) stationary [8], (b)  $L_1$  and (c) non-stationary [11], and (d) TV based regularization.

## 6. CONCLUSIONS AND FUTURE WORK

The resolution of the super-resolved images shown in Figure 3 and 4 has greatly improved. Also, from the MSE in Table 1 with the experiments on the artificial data, for low SNR the proposed algorithm seems to be superior to all other tested algorithms.

In future we plan to formulate the problem in a Bayesian framework using a Student-t based prior in product form which preserves edges in different directions and also allows Bayesian inference [13]. This will allow systematic evaluation of the regularization parameters.

## ACKNOWLEDGEMENT

The real low-resolution image data used herein were provided by Sina Farsiu and Peyman Milanfar.

## REFERENCES

- [1] S. Park, M. Park, and M. Kang, "Super-resolution image reconstruction: a technical overview," *IEEE Signal Processing Magazine*, vol. 20, no. 3, pp. 21-36, May 2003.
- [2] S. Farsiu, D. Robinson, M. Elad, and P. Milanfar, "Advances and challenges in super-resolution," *International Journal of Imaging Systems and Technology*, vol. 14, no. 2, pp. 47-57, August 2004.
- [3] S. Chaudhuri, Editor, Super-Resolution Imaging, Chapter 4, "Reconstruction of a high resolution image from multiple low resolution images," by B. Tom, N. Galatsanos and A. Katsaggelos, pp. 71-105, Kluwer, 2001.
- [4] T. F. Chan, S. Esedoglu, F. Park, and M. H. Yip, Recent developments in total variation image restoration, in *Handbook of Mathematical Models in Computer Vision*, Springer Verlag, 2005.
- [5] G. Demoment, "Image restoration and reconstruction: overview of common estimation structures and problems," *IEEE Transactions on Image Processing*, vol. 37, no. 12, pp. 2024-2036, December 1989.
- [6] S. Farsiu, M. D. Robinson, M. Elad, P. Milanfar, "Fast and robust multi-frame super-resolution," *IEEE Transactions on Image Processing*, vol. 13, no. 10, pp. 1327-1344, October 2004.
- [7] S. Farsiu, M. Elad, P. Milanfar, "Multiframe demosaicing and super-resolution of color images," *IEEE Transactions on Image Processing*, vol. 15, no. 1, pp. 141-159, January 2006.
- [8] N. A. Woods, N.P. Galatsanos and A.K. Katsaggelos, "Stochastic methods for joint restoration, interpolation and registration of multiple under sampled images," *IEEE Transactions on Image Processing*, vol. 15, no. 1, pp. 201-213, January 2006.
- [9] J. M. Bioucas-Dias, M. A. T. Figueiredo and J. P. Oliveira, "Total-variation based image deconvolution: a majorization-minimization approach," *Proceedings of EUSIPCO 2006*, Florence, Italy, September 2006.
- [10] S. D. Babacan, R. Molina and A. K. Katsaggelos, "Total Variation Image Restoration and Parameter Estimation using Variational Posterior Distribution Approximation," *IEEE International Conference on Image Processing*, vol. I, pp. 97-100, San Antonio, Texas (USA), September 2007.
- [11] G. Chantas, N. P. Galatsanos, and N. Woods, "Super-resolution based on fast registration and maximum a posteriori reconstruction," *IEEE Transactions on Image Processing*, vol. 16, no. 7, pp. 1821-1830, July 2007.
- [12] K. Lange, Optimization, *Springer Texts in Statistics*, Springer-Verlag, 2004.
- [13] G. Chantas, N. Galatsanos and A. Likas, "Bayesian image restoration based on variational inference and a product of experts image prior," *IEEE Int. Workshop in Machine Learning for Signal Processing*, Thessaloniki, Greece, August, 2007.