A FULLY ROBUST FRAMEWORK FOR MAP IMAGE SUPER-RESOLUTION

Michalis Vrigkas, Christophoros Nikou and Lisimachos P. Kondi

Department of Computer Science, University of Ioannina, Greece
{mvrigkas, cnikou, lkon}@cs.uoi.gr

ABSTRACT

In this work, we propose an adaptive M-estimation scheme for robust image super-resolution. The proposed algorithm relies on a maximum *a posteriori* (MAP) framework and addresses the presence of outliers in the low resolution images. Moreover, apart from the robust estimation of the high resolution image, the contribution of the method is twofold: (i) the robust computation of the regularization parameters controlling the relative strength of the prior with respect to the data fidelity term and (ii) the robust estimation of the optimal step size in the update of the high resolution image. Experimental results demonstrate that integrating these estimations into a robust framework leads to significant improvement in the accuracy of the high resolution image.

Index Terms— Maximum *a posteriori* (MAP) image super-resolution, robust M-estimator, Tikhonov regularization.

1. INTRODUCTION

Image super-resolution (SR) is a technique for enhancing the quality and the resolution of an image. The objective is to improve the spatial resolution by using information from a set of several different low-resolution (LR) images to produce an image with more visible detail in the high spatial frequency features. The LR images may experience different degradations such as motion, point spread function blurring, subsampling and additive noise. The reconstructed high-resolution (HR) image can be successfully estimated if there exist sub-pixel shifts between the LR images. In this manner, each frame of LR sequence brings complementary information to the original HR image.

Researchers studying the direct inverse solution recognized the limitations of the interpolation, motion compensation and inverse filtering to be ill-posed due to the existence of additive noise [9, 10]. Even in cases of perfect motion registration and accurate knowledge of the point spread function of the acquisition system, a significant dependence of the estimation of the HR image on degradation conditions is observed. A large family of SR methods is based on a stochastic formulation of the problem which imposes a prior distribution on the image to be reconstructed and provides estimates in a maximum *a posteriori* (MAP) framework, where the posterior distribution of the HR image is maximized [9, 10, 12].

Violations of the assumptions of data fidelity to the assumed model are also likely to occur, because SR methods are very sensitive to inaccuracies of their parameters. However, little has been reported about suppressing the outliers artifacts. For instance, median filters have been efficiently used to treat the SR problem [6] where robustness is introduced by applying a median filter in each term of back-projected difference image.

There is a considerable interest in robust SR methods largely emphasizing in outlying data. To this end, a novel robust image SR method was reported [7]. The authors introduced a robust SR method based on the use of the L_1 norm both on the observed data and the regularization term. In the same context, a robust color image super-resolution algorithm has previously shown great potential for estimating high resolution images with crisp details [1, 2, 3]. Following a MAP scheme and adapting a robust M-estimation framework, outliers can efficiently be suppressed without the use of regularization in the objective function. Apart from using L_1 and L_2 error norms in the objective function, much research has focused on stochastic techniques in a MAP framework [4, 5]. Huber and Lorentzian error norms are used for measuring the difference between the estimation of HR image and each LR image. Moreover, a factor that affects the super-resolution quality is also the Tikhonov regularization [9, 10, 5, 7], which is used to remove artifacts from the final solutions.

In this paper, we apply a MAP scheme for robust image super-resolution where the objective function to be minimized employs a regularization term. By integrating a robust M-estimator the reconstructed image is consistently of much higher quality than in other standard approaches. Also, apart from the robust estimation of the HR image, the main contribution in this work is the robust estimation of the regularization parameters and the optimal step of the update equation. Experiments showed that the reconstructed HR image is of higher quality than in standard MAP-based methods employing robust estimation only for the HR image.

2. IMAGE FORMATION MODEL

The image degradation process [10] is modeled by motion (rotation and translation), a linear blur, and subsampling by pixel averaging along with additive Gaussian noise. We assume that p LR images, each of size $M = N_1 \times N_2$, are obtained from the acquisition process. The following observation model is assumed, where all images are ordered lexicographically

$$\mathbf{y} = \mathbf{W}\mathbf{z} + \mathbf{n}. \tag{1}$$

The set of LR frames is described as $\mathbf{y} = [\mathbf{y}_1^T, \mathbf{y}_2^T, \ldots, \mathbf{y}_p^T]^T$, where $\mathbf{y}_k$, for $k = 1, ...p$, are the p LR images. The desired HR image $\mathbf{z}$ is of size $N = l_1 N_1 \times l_2 N_2$, where l_1 and l_2 represent the up-sampling factors in the horizontal and vertical directions, respectively. The term $\mathbf{n}$ represents zero-mean additive Gaussian noise. In eq. (1), the degradation matrix $\mathbf{W} = [\mathbf{W}_1^T, \mathbf{W}_2^T, \ldots, \mathbf{W}_p^T]^T$ performs the operations of motion, blur and subsampling. Thus, matrix $\mathbf{W}_k$, for the k-th frame, may be written as

$$\mathbf{W}_k = \mathbf{D}\mathbf{B}_k\mathbf{M}(\mathbf{s}_k), \tag{2}$$

where $\mathbf{D}$ is the $N_1 N_2 \times N$ subsampling matrix, $\mathbf{B}_k$ is the $N \times N$ blurring matrix, and $\mathbf{M}(\mathbf{s}_k)$ is the $N \times N$ rigid transformation matrix with parameters (rotation angle and translation vector) denoted by $\mathbf{s}_k$ for the k-th frame. Finally, $\mathbf{n}$ is additive Gaussian noise.

3. THE PROPOSED ALGORITHM

Super-resolution reconstruction is an ill-posed inverse problem due to the existence of the additive noise. In order to stabilize the inversion process, we introduce a super-resolution algorithm that uses robust error norm in the data fidelity term of objective function. This approach is based on the class of robust M-estimators. The objective function uses a regularization term that can help the super-resolution algorithm to remove any artifacts from the final solution. We are interested in estimators whose influence function is differentiable and bounded, like the Lorentzian estimator, defined as:

$$\rho(x,\sigma) = \log\left(1+\frac{1}{2}\left(\frac{x}{\sigma}\right)^2\right), \quad \psi(x,\sigma) = \frac{2x}{x^2+2\sigma^2} \qquad (3)$$

where σ is the scale factor and ψ is the influence function, defined as the first derivative of the robust estimator ρ.

The scale factor controls a threshold beyond which all points are considered to be outliers. Violations in the mathematical model in (1) may yield large errors, which can severely influence the reconstruction process. The choice of the scale factor σ plays a crucial role in controlling the outliers. Errors falling beyond that threshold are assigned smaller weights and the corresponding outlying measures are suppressed. For small values of the scale factor, the influence function decreases faster assigning smaller weights to errors that outstrip the value of this parameter. If the value of σ is relatively small the contribution of the LR frames will be canceled leading to bad estimation of the HR image, due to the insufficient information provided by the LR frames. On the other hand, if the the value of the scale factor is chosen to be arbitrary large, outliers will significantly contribute to the estimation of the HR image.

A regularized approach using the image prior information of the HR image (Gaussian assumption) can be used to make the inverse problem well-posed [10]. Considering that each LR image may result from a different degradation process, which implies that different weighting should be given to it in the desired solution, and recasting the problem in a generalized M-estimation framework, the following channel-weighted cost function is proposed:

$$L(\mathbf{z},\mathbf{s}) = \sum_{i=1}^{M}\left[\rho\left(\mathbf{y}_k - \mathbf{W}_k(\mathbf{s}_k)\mathbf{z};\ \sigma_k\right)\right]_i + \alpha_k(\mathbf{z})||\mathbf{Qz}||^2, \quad (4)$$

where the operator $[\cdot]_i$ takes the i-th element of the vectorized matrix inside the brackets and $\mathbf{Q}$ is a matrix applying a high pass filter (in our case the Laplacian) and is used to penalize discontinuities in the final solution. The robust regularization parameters $\alpha_k(\mathbf{z})$, determine the trade-off between the fidelity of the observed data and the image prior. In (4), it is implied that the registration parameters $\mathbf{s}_k$ are collected in $\mathbf{s}$ in this type of formulation.

Estimation of the registration parameters $\mathbf{s}$ and the HR image $\mathbf{z}$ may be obtained by an alternating optimization scheme [9, 10]. At a first step, the registration parameters may be computed by a variety of methods involving block matching schemes [9, 10] or algorithms combining feature extraction and mutual information [12]. Having fixed the registration parameters, we may use a gradient descent method with a properly calculated step size to minimize (4) with respect to the HR image. Therefore, the HR image may be obtained by the following minimization problem:

$$\mathbf{z}^* = \arg\min_{\mathbf{z}}\left\{\sum_{i=1}^{M}\left[\rho\left(\mathbf{W}_k\mathbf{z} - \mathbf{y}_k;\sigma_k\right)\right]_i + \alpha_k(\mathbf{z})||\mathbf{Qz}||^2\right\} \quad (5)$$

where we have omitted the dependence of matrix $\mathbf{W}_k$ from the registration parameters $\mathbf{s}_k$ to simplify the notation. Notice that, in (5), different outlier thresholds are assigned to different LR frames.

In our previous work [10], it was shown that the regularization parameters $\alpha_k(\mathbf{z})$ may be obtained in closed form from the images. In the robust framework proposed herein, the related expression is:

$$\alpha_k(\mathbf{z}) = \frac{\sum_{i=1}^{M}\left[\rho\left(\mathbf{y}_k - \mathbf{W}_k\mathbf{z};\ \sigma_k\right)\right]_i^2}{2||\mathbf{y}_k||^2 - ||\mathbf{Qz}||^2}. \qquad (6)$$

which provides a per frame robust regularization parameter.

To obtain a robust solution of (5), a gradient descent method may be employed. By computing the gradient of (5) with respect to the HR image $\mathbf{z}$, we obtain the following update:

$$\hat{\mathbf{z}}^{n+1} = \hat{\mathbf{z}}^n - \sum_{k=1}^{p}\varepsilon_k^n\mathbf{W}_k^T\psi\left(\mathbf{W}_k\hat{\mathbf{z}}^n - \mathbf{y}_k;\ \sigma_k\right) + \alpha_k(\mathbf{z})||\mathbf{Qz}||^2 \quad (7)$$

where the influence function ψ of the robust estimation is now involved.

It must be noted that the choice of the step-size parameter ε_k^n plays an important role in the behavior of the gradient descent method. This parameter must be small enough to prevent divergence and large enough to provide convergence. A constant step-size could be the easiest solution but this is an inappropriate approach for the most of the robust image super-resolution problems. After some manipulation and following the spirit in [9], a robust closed form solution of the optimal step size is obtained, which is given in (8) in the next page for space purposes. Notice that both the robust estimator ρ and its influence function ψ appear in (8).

This optimal step size (8) is calculated for every single LR image. Having an adaptive step size, provides a better convergence and also keeps off the algorithm from trapping into "bad" solutions.

In robust image super-resolution reconstruction, it is necessary to define a process for automatically computing the value of the outlier threshold parameter. In statistics, the *Median Absolute Deviation* (*MAD*) criterion [8] is considered to be one of the most accurate robust measure of the variability of a univariate sample of quantitative data. For the k-th LR image:

$$\begin{aligned} MAD_k^n = &\operatorname*{median}_i\left\{|r_{k,i}^n([\mathbf{W}_k\mathbf{z}^{n-1};\mathbf{y}_k]_i) - \right. \\ &\left.\operatorname*{median}_j(r_{m,j}^n([\mathbf{W}_m\mathbf{z}^{n-1};\mathbf{y}_m)]_j)|\right\} \end{aligned} \qquad (9)$$

where $n = 0, 1, 2, \ldots$ refers to the n-th iteration of the algorithm and $r_{k,i}^n(\mathbf{W}_k\mathbf{z}^{n-1};\mathbf{y}_k) = [\mathbf{W}_k\mathbf{z}^{n-1} - \mathbf{y}_k]_i$ is the residual error of the i-th datum between the estimation of the degraded HR image and the k-th LR frame. The MAD is a measure of statistical dispersion. It is a robust statistic, being more resilient to outliers in a data set. In order to use MAD criterion as a consistent estimator for the estimation of the scale factor we consider $\sigma_k^n = K \cdot MAD_k^n$, where K is a constant which depends on the distribution. For normally distributed data with standard deviation 1, K is taken to be $1/\Phi^{-1}(3/4) \approx 1.4826$, where Φ^{-1} is the inverse of the cumulative distribution function for the standard normal distribution. In that case, for the k-th LR frame the scale factor σ_k is computed as follows:

$$\sigma_k^n = 1.4826 \cdot MAD_k^n, \quad k = 1, 2, \cdots, p. \qquad (10)$$

$$\varepsilon_k^n = \frac{\sum_{m=1}^{pM} [\mathbf{Wg}]_m \left[\psi\left(\mathbf{W}_k\mathbf{z} - \mathbf{y}_k;\ \sigma_k\right)\right]_m + \sum_{k=1}^{p} \alpha_k(\mathbf{z}) \sum_{i=1}^{N} [\mathbf{Qg}(\mathbf{z})]_i [\mathbf{Qz}]_i}{\sum_{m=1}^{pM} [\mathbf{Wg}]_m^2 \left[\rho\left(\mathbf{W}_k\mathbf{z} - \mathbf{y}_k;\ \sigma_k\right)\right]_m + \sum_{k=1}^{p} \alpha_k(\mathbf{z}) \sum_{i=1}^{N} [\mathbf{Qg}(\mathbf{z})]_i} \tag{8}$$

4. EXPERIMENTAL RESULTS

In order to evaluate the proposed methodology, experiments were conducted on synthetic data sets. Sequences of low resolution images were created by blurring, down-sampling and degrading by noise an original image. At first, the images were downsampled by a factor of 2 (4 pixels to 1). Then, a point spread function of 5×5 Gaussian kernel with standard deviation of 1 was applied and the resulting images were degraded by white Gaussian noise in order to obtain a signal to noise ratio of 30 dB. Finally, 50% of the LR frames were degraded by salt and pepper noise (two cases were examined: corruption of 5% and 10% of the pixels in the respective frame).

To highlight the importance of the proposed fully robust super-resolution scheme, we compared it to the standard approach that employs a robust estimator only in the HR image update and integrates a heuristic scheme for the step size. We also compared several robust estimators in that framework: the truncated least squares (TLS), the Geman-McClure and the Lorentzian error norms [8]. We used 20 frames of the *Susie* image sequence in our experiments (Fig. 1).

In all of the experiments, in order to have a first estimate of the HR image, a LR image was chosen at random and it was upscaled by bicubic interpolation. Convergence of the super-resolution algorithm was achieved when $\|\hat{\mathbf{z}}^{n+1} - \hat{\mathbf{z}}^n\|/\|\hat{\mathbf{z}}^n\| < 10^{-5}$. A quantitative evaluation of the obtained HR images is given by the peak signal to noise ratio (PSNR).

(a) (b) (c)

Fig. 1. Representative frames of low-resolution images for *Susie* sequence: (a) original sequence and sequence degraded by salt & pepper noise at (b) 5% and (c) 10%.

Table 1 presents the statistics of the PSNR for the compared algorithms for 10 realizations of the experiment in each case. In this table, the term “no robust” refers to the employment of a robust estimator only for the computation of the HR image using (7) but not for the parameters $\alpha_k(\mathbf{z})$ and ε_k^n. The term term “robust” indicates that a robust estimator was also employed for the computation of $\alpha_k(\mathbf{z})$ and ε_k^n, which were computed by (6) and (8) respectively. As it may be observed, our fully robust method outperforms the algorithm that employs a robust estimator only for the SR image update. The improvement in PSNR is significant, particularly when using the TLS estimator, which is the one that provides the better results for the *Susie* sequence. Representative SR images of the compared methods are shown in figures 2 and 3 for *Susie* test sequence. It is clearly depicted that the robust estimation of the step size ε_k^n and the regularization parameters $\alpha_k(\mathbf{z})$ improves significantly the quality of the super-resolved images for all the employed robust estimators. On the other hand, it may be easily observed that the robust estimation of the intensities of the HR image is not sufficient if it is not combined with the appropriate robust values for ε_k^n and $\alpha_k(\mathbf{z})$.

5. CONCLUSIONS

A robust MAP image super-resolution algorithm was proposed in this paper. The robust estimation scheme was integrated into the estimation of the high resolution image, the regularization parameters and the step size in the update of the high resolution image. To the best of our knowledge, this is the first time that these issues are addressed in a robust formulation. Several M-estimators were employed in the comparison of the proposed method with a method employing a robust estimation only in the HR image update. The obtained improvement is on average 5 dB in PSNR.

6. REFERENCES

[1] N. A. El-Yamany and P. E. Papamichalis, “Robust color image superresolution: An adaptive M-estimation framework,” *EURASIP Journal on Image and Video Processing*, 2008, ID=763254.

[2] N. A. El-Yamany and P. E. Papamichalis, “An adaptive M-estimation framework for robust image superresolution without regularization,” *in Visual Communications and Image Processing*, vol. 6822, pp. 1–12, January 2008.

[3] N. A. El-Yamany and P. E. Papamichalis, “Using bounded-influence M-estimators in multi-frame super-resolution reconstruction: A comparative study,” *Image Processing, 2008. ICIP 2008. 15th IEEE International Conference on*, pp. 337–340, 2008.

[4] V. Patanavijit, S. Tae-O-Sot, and S. Jitapunkul, “A robust iterative superresolution reconstruction of image sequences using a Lorentzian Bayesian approach with fast affine block-based registration,” *Proceedings of IEEE International Conference on Image Processing (ICIP ’07)*, vol. 5, pp. 393–396, 2007.

[5] V. Patanavijit, S. Tae-O-Sot, and S. Jitapunkul, “A lorentzian stohastic estimation for a robust iterative multiframe suprresolution reconstruction with lorentzian-tikhonov regularization,” *EURASIP Journal on Advatages in Signal Processing*, 2007, ID=34821.

[6] A. Zomet, A. Rav-Acha, and S. Peleg, “Robust super-resolution,” *in Proc. of the IEEE Workshop on Applications of Computer Vision*, vol. 1, pp. 645–650, 2001.

[7] S. Farsiu, M. D. Robinson, M. Elad, and P. Milanfar, “Fast and robust multiframe super resolution,” *IEEE Transactions on Image Processing*, vol. 13, no. 10, pp. 1327–1344, October 2004.

[8] P. J. Rousseeuw and A. M. Leory, *Robust Regression and Outlier Detection*, John Wiley & Sons,, 1987.

[9] R. C. Hardie, K. J. Barnard, and E. E. Armstrong, “Joint MAP image registration and high-resolution image estimation using

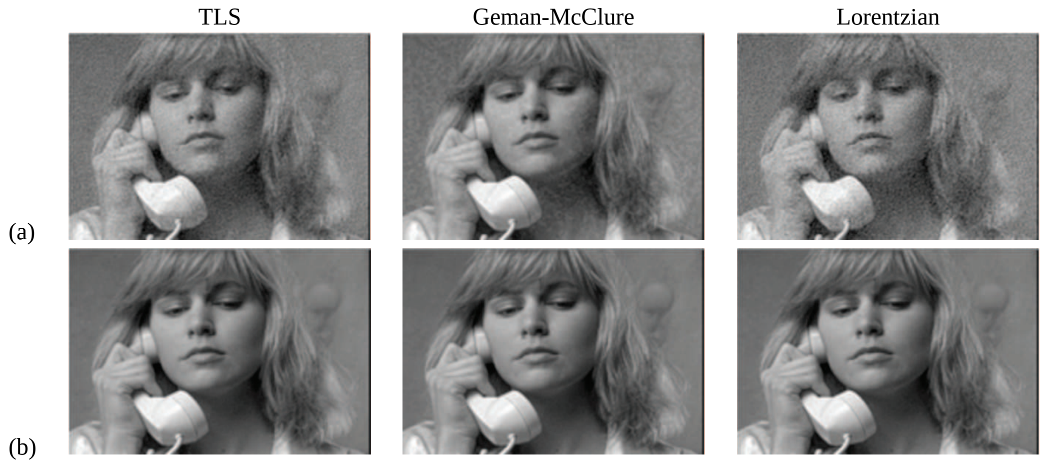

Fig. 2. Reconstructed high-resolution images for 20 frames of the *Susie* sequence, with salt & pepper noise at 5%, using the compared methods. (a) No robust parameters and (b) robust parameters ε_k and $\alpha_k(\mathbf{z})$.

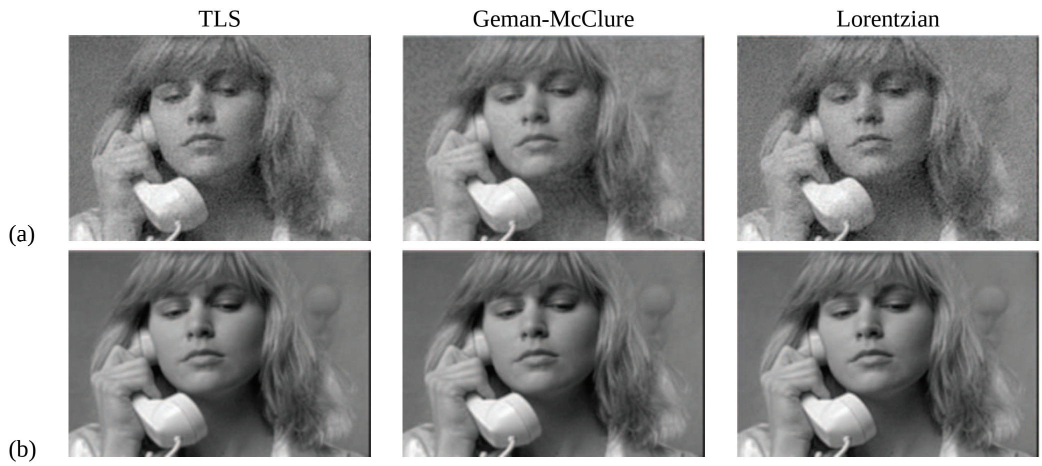

Fig. 3. Reconstructed high-resolution images for 20 frames of the *Susie* sequence, with salt & pepper noise at 10%, using the compared methods. (a) No robust parameters and (b) robust parameters ε_k and $\alpha_k(\mathbf{z})$.

Table 1. Performance evaluation (PSNR in dB) for the sequence *Susie*.

Robust Estimator	Salt & Pepper at 5%						Salt & Pepper at 10%					
	no robust			robust			no robust			robust		
	mean	std	median	mean	std	median	mean	std	median	mean	std	median
TLS	18.1	0.6	17.7	**25.8**	0.4	**25.7**	15.3	0.9	15.9	**23.6**	0.3	**23.7**
Geman-McClure	22.1	0.6	22.1	**24.4**	0.4	**24.3**	19.5	0.4	19.5	**21.5**	0.5	**21.5**
Lorentzian	20.8	0.3	20.8	**25.1**	0.7	**25.0**	19.0	0.3	19.1	**22.9**	0.5	**22.7**

a sequence of undersampled images," *IEEE Transactions on Image Processing*, vol. 6, no. 12, pp. 1621–1633, December 1997.

[10] H. He and L. P. Kondi, "An image super-resolution algorithm for different error levels per frame," *IEEE Transactions on Image Processing*, vol. 15, no. 3, pp. 592–603, March 2006.

[11] H. He and L. P. Kondi, "Resolution enhancement of video sequences with simultaneous estimation of the regularization parameter," *SPIE Journal of Electronic Imaging*, vol. 13, no. 3, pp. 586–596, 2004.

[12] M. Vrigkas, C. Nikou, and L. P. Kondi, "On the improvement of image registration for high accuracy super-resolution," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'11)*, Prague, Czech Republic, 2011, pp. 981–984.