A ROBUST IMAGE SUPER-RESOLUTION SCHEME BASED ON REDESCENDING *M*-ESTIMATORS AND INFORMATION-THEORETIC DIVERGENCE

Noha A. El-Yamany [(1)], Panos E. Papamichalis [(1)] and William R. Schucany [(2)]

[(1)] Department of Electrical Engineering, Southern Methodist University, Dallas, TX 75725-0338
[(2)] Department of Statistical Science, Southern Methodist University, Dallas, TX 75275-0332
E-mail: {nelyaman,panos}@engr.smu.edu, schucany@smu.edu, FAX: 214-768-3573

ABSTRACT

This paper proposes a novel image super-resolution (SR) algorithm in a robust estimation framework. SR estimation is formulated as an optimization (minimization) problem whose objective function is based on robust *M*-estimators and its solution yields the SR output. The novelty of the proposed scheme lies in the selection of this class of estimators and the incorporation of information-theoretic similarity measures. Such a choice helps in dealing with violations (outliers) of the assumed mathematical model that generated the low-resolution images from the "unknown" high-resolution one. The proposed approach results in high-resolution images with no estimation artifacts. Experimental results demonstrate its superior performance in comparison to both L_1 and L_2 estimation in terms of robustness and speed of convergence.

***Index Terms*—** Robust M-estimators, super-resolution, information-theoretic divergence

1. INTRODUCTION

Image super-resolution (SR) has attracted a lot of attention recently as a way of producing high-resolution images with better details, by combining the information in a sequence of sub-pixel shifted low-resolution (LR) images. It can be considered as a cheap alternative to costly high-precision optics, which is, however, limited by sensor noise. Super-resolution can have a wide range of applications ranging from consumer electronics, to surveillance and military applications. Most SR algorithms assume a mathematical model for the imaging system, which could have generated the sequence of LR frames from the unknown high-resolution image. However, violations to the assumed model occur because of the approximate nature of the model, inaccuracies in the model parameters (such as blur and/or motion parameters), and/or accidental scene changes. For instance, the model could assume translational motion while, in reality, a more complex motion (e.g. affine) may have taken place. These violations even small in number can be detrimental to SR estimation and result in estimation artifacts.

Robust statistics [1-2] have emerged as a family of theories and techniques for estimation while dealing with deviations from idealized model assumptions. In particular, robust *M*-estimators have been found very effective in many image processing applications such as optical flow estimation [8], robust denoising [9] and robust anisotropic diffusion [10], just to name a few. However, to the best of our knowledge, they have not been applied to the problem of image super-resolution. Motivated by the robustness of redescending *M*-estimators [1-2], we attempt to address the problem of image SR in a robust estimation framework.

2. PROBLEM FORMULATION

In this paper, we use the following observation model considered by many super-resolution algorithms [such as 3-6]:

$$\mathbf{Y}_k = \mathbf{DHF}_k\,\mathbf{X} + \mathbf{Z} \quad (1)$$

where **D**, **H,** and $\mathbf{F}_k$ represent the downsampling, burring and warping operations; respectively. Downsampling is assumed to be by a constant factor in both the x and y directions. This factor will be referred to as *r*, the resolution enhancement factor. **X** is the unknown SR output, $\mathbf{Y}_k$ is the *k*th LR observation (frame), and **Z** is a noise term. Recasting the problem in the generalized *M*-estimation framework, the SR output is the solution of the following minimization problem:

$$\mathbf{X}^* = \underset{\mathbf{X}}{\mathbf{arg\,min}} \sum_{k=1}^{N} \rho\,(\mathbf{DHF}_k\mathbf{X} - \mathbf{Y}_k) = \underset{\mathbf{X}}{\mathbf{arg\,min}} \sum_{k=1}^{N} \rho\,(\mathbf{E}_k) \quad (2)$$

where *N* is the number of LR frames, $\mathbf{E}_k$ is the projection error related to the *k*th LR frame, and ρ is an even-symmetric function, which has a unique minimum at zero and satisfies the condition:

$$\sum_{k=1}^{N} (\mathbf{DHF}_k)^T \psi\,(\mathbf{E}_k) = \mathbf{0} \quad (3)$$

This last condition is the result of minimizing ρ with respect to **X** by setting its partial derivative equal to zero. ψ is the first derivative of ρ with respect to **E**, and is referred to as the ***influence function*** [1].

Many references, such as [3], have addressed the solution of (2) as a least squares (LS) estimation problem using the L_2 error norm $\Rightarrow \mathbf{X}^* = \underset{\mathbf{X}}{\mathbf{arg\,min}} \sum_{k=1}^{N} \rho(\mathbf{E}_k) = \underset{\mathbf{X}}{\mathbf{arg\,min}} \sum_{k=1}^{N} \|\mathbf{E}_k\|_2^2$. The solution can be found iteratively using the steepest descent algorithm as follows:

$$\mathbf{X}^{n+1} = \mathbf{X}^n - 2\beta \sum_{k=1}^{N} (\mathbf{DHF}_k)^T \psi(\mathbf{E}_k^n) = \mathbf{X}^n - 2\beta \sum_{k=1}^{N} (\mathbf{DHF}_k)^T\,\mathbf{E}_k^n \quad (4)$$

where β is a step size parameter. However, LS estimation exhibits a poor performance in the presence of model violations, as shown in [5]. In fact, in [4] Zomet *el al* adopted a post-processing step to remove estimation artifacts after the LS solution is obtained. The reason behind the non-robustness of the L_2-estimator lies in its influence function (i.e., its first derivative ψ). As shown in the equation above, ψ is linear and increases without bound assigning large weights to large errors. As a first step toward robustifying SR estimation, Farsiu *et al* [5-6] proposed the use of the L_1-estimator as an alternative to LS estimation $\Rightarrow \mathbf{X}^* = \underset{\mathbf{X}}{\mathbf{arg\,min}} \sum_{k=1}^{N} \rho(\mathbf{E}_k) = \underset{\mathbf{X}}{\mathbf{arg\,min}} \sum_{k=1}^{N} \|\mathbf{E}_k\|_1$. The solution can also be found using the steepest descent algorithm [5] as follows:

$$\mathbf{X}^{n+1} = \mathbf{X}^n - \beta \sum_{k=1}^{N} (\mathbf{DHF}_k)^T \psi(\mathbf{E}_k^n) = \mathbf{X}^n - \beta \sum_{k=1}^{N} (\mathbf{DHF}_k)^T\, \mathbf{sign}(\mathbf{E}_k^n) \quad (5)$$

As shown, the L_1 influence function ψ is the Signum function. Therefore, all errors (small or large) are assigned the same weights either 1 or -1, depending only on their sign. The L_1 error norm is definitely more robust than the L_2 in the presence of outliers. However, because of its constant-valued influence function, the resulting SR solution suffers from various artifacts. This is especially the case when the problem is underdetermined (i.e. there are fewer LR frames than the required minimum number of frames to fill-in the missing pixels in the HR image), and a regularization term is required to stabilize the solution.

3. THE PROPOSED SR ESTIMATION ALGORITHM

To improve the robustness of SR estimation, we propose the use of an objective function based on a specific class of robust *M*-estimators, which have redescending influence functions (redescending *M*-estimators [1]). For these estimators, the influence curve ψ increases up to a given point referred to as the outlier threshold (τ) after which it starts to decrease (redescends) as the error grows. Because of this behavior, large errors falling beyond the outlier threshold are assigned weights that decrease as the error increases, thus providing a soft outlier rejection rule. Of all the redescending *M*-estimators, we are particularly interested in estimators whose influence functions have only one parameter (τ), which will be determined from the available observations as shown later. Examples of these estimators are the Lorentzian, Geman and McClure and Tukey's biweight functions [1]. In this paper, only the Lorentzian estimator is demonstrated due to space limitations. Fig.1. (a) and (b) depict an example of a redescending estimator ρ (the Lorentzian estimator) and the corresponding influence function ψ for $\tau = 50$.

3.1. The Robust Objective Function

Recasting the problem of super-resolution using redescending *M*-estimators, the SR output is given by

$$\mathbf{X}^* = \underset{\mathbf{X}}{\mathbf{arg\,min}} \sum_{k=1}^{N} \rho_k\ (\mathbf{E}_k;\tau_k) \quad (6)$$

where ρ_k is the objective function term corresponding to the *k*th LR frame and τ_k is the corresponding outlier threshold. The Lorentzian estimator [1] has an error norm given by $\rho(e;\tau) = \log(\tau^2 + e^2/\tau^2)$. For redescending *M*-estimators a one-step (one iteration) estimation using Newton's algorithm is possible, provided that both the initial guess and the outlier threshold are robustly selected [1-2]. Unfortunately, these two conditions cannot be met in general and an iterative scheme to solve (6) is inevitable. Two computational issues related to redescending *M*-estimators [1], are addressed however; 1) the non-uniqueness of the solution (because of the redescending ψ , which goes to zero for very large error values as shown in Fig.1.b) and 2) the non-convexity of the cost function in (6). To overcome the first issue, a proper initial guess of **X** is estimated such that the resulting projection errors are not large enough to drive ψ to zero. In our experiments, we have estimated the initial guess through bilinear interpolation of the first LR frame. And to avoid getting trapped in local minima, we chose to apply the graduated nonconvexity (GNC) algorithm [11] in conjunction with successive over-relaxation (SOR) to minimize (6). To find the solution of (6) iteratively, the proposed update equation is:

$$\mathbf{X}^{n+1} = \mathbf{X}^n - \sum_{k=1}^{N} (\mathbf{DHF}_k)^T \frac{\omega}{T_k}\ e^{-\alpha\, L_k}\ \boldsymbol{\psi}(\mathbf{E}_k^n;\tau_k) \quad (7)$$

where ω is the SOR parameter ($1 < \omega < 2$) and T_k is the maximum value of the second derivative of ρ_k , which is $2/\tau_k^2$ for the Lorentzian estimator. The term $(\mathbf{DHF}_k)^T$ is the backprojection operator from the low-resolution grid to the high-resolution one. The term $\boldsymbol{\psi}(\mathbf{E}_k;\tau_k)$ is referred to as the local influence vector because for the *i*th location on the low-resolution grid, the Lorentzian influence function is evaluated for $e_{k,i}$ (*i*th value in $\mathbf{E}_k$). The factor $e^{-\alpha\, L_k}$ is explained below.

3.2. Estimation of the Outlier Thresholds τ_k

After the initial SR estimate $\mathbf{X}^0$ is obtained (through bilinear interpolation of the first (reference) LR frame), the projection errors $\mathbf{E}_k^0 = \mathbf{DHF}_k\, \mathbf{X}^0 - \mathbf{Y}_k$ are estimated. For the reference frame, the outlier threshold τ_1 is set to the maximum absolute error value in $\mathbf{E}_1^0$. The reason it is selected this way is that the first LR frame is considered as a frame without any outliers and its corresponding projection errors are all accepted toward the estimation of the SR output. This is a reasonable assumption since in image SR the goal is usually to enhance the resolution of the first frame. To estimate the outlier threshold for each of the other LR frames, we propose an information-theoretic divergence based strategy, which is described in what follows. First, the distribution (normalized histogram) of the projection error related to each of the LR frames $h_k = h\left(\mathbf{E}_k^0\right)$ is computed. Then, a divergence (i.e., dissimilarity) measure between the error distribution of each of the LR frames and the reference frame error distribution is calculated. We have chosen the *L*-divergence proposed in [12] because it is symmetric, bounded and it does not require absolute continuity and common support of the distributions [12]. A soft decision rule for calculating the outlier threshold for each frame, based on the divergence measure is then adopted. The outlier threshold for the *k*th frame is calculated as $\tau_k = \tau_1 e^{-\alpha L_k}$, $k = 2,3,....,N$, where L_k is the divergence (dissimilarity) between the two distributions h_k and h_1, and is calculated as in [12]. L_k has a minimum value of zero [12]. Therefore, if the two distributions are identical, τ_k will be equal to τ_1. The constant α controls the decay of the weighting exponential and is calculated from

$$\tau_{\mathbf{min}} = \tau_1 e^{-\alpha L_{\mathbf{max}}} \Rightarrow \alpha = \mathbf{log}(\tau_1 / \tau_{\mathbf{min}}) / L_{\mathbf{max}}$$

where $\tau_{\mathbf{min}}$ is the minimum outlier threshold value assigned at the maximum divergence $L_{\mathbf{max}}$. The *L*-divergence has an upper bound of 2 as shown in [12]. However, this maximum is never attained unless the two distributions are completely dissimilar. Through extensive experiments, we found that reasonable values for $L_{\mathbf{max}}$ and $\tau_{\mathbf{min}}$ are 1 and 0.01, respectively. The factor $e^{-\alpha L_k}$ in (7) is included to guarantee that all ψ_k have the same maximum value at their corresponding τ_k (as shown in Fig.1.e). The strategy outlined above to estimate the outlier thresholds, embeds a soft outlier rejection rule in the robust formulation in (7). It is worth noting that the proposed strategy only assumes stationary scenes (no moving objects). This assumption can, however, be relaxed by pursuing a region-based version.

4. EXPERIMENTAL RESULTS

Both synthetic and real experiments are considered in this section. In the synthetic experiment as in [5], we simulated a set of LR frames from the SMU Helmet image through shifting it by all the 16 possible integer shifts in a 4×4 square on the high-resolution grid, blurring by a Gaussian Kernel of size 5×5 with a standard deviation of 1, and then downsampling by a factor of 4 in both directions. Then, frame # 4 is rotated by 20^0 CCW and frame # 10 is zoomed in by a factor of 1.2 (in addition to the shifts). Gaussian noise was also added so that the SNR is 30 dB. During estimation pure translation is assumed and the blurring kernel is assumed to be Gaussian of size 3×3 of standard deviation of 1, to simulate both motion and blur estimation errors. In this experiment, the resolution enhancement factor is 4 (i.e., r = 4). For the real experiment, the Books sequence was captured by a Canon *P*ower*S*hot A400 and 10 frames were used in the estimation. Translational motion is assumed (although the sequence undergoes an affine motion) and the algorithm in [7] is used to estimate the motion vector between each of the LR frames and the first (reference frame). The blurring kernel is assumed to be Gaussian of size 3×3 and a unity standard deviation. In this experiment, r = 2.

Fig.1 and Fig.2 depict the 4× and 2× SR estimation results for the synthetic and real data, respectively. The SOR parameter in (7) is set to 1.5. Steepest descent is pursued to minimize the L_1 and L_2 cost functions (equations (4) and (5)), and the step size β is set to 1. Form these results, it is shown that the L_1 estimator is more robust than the L_2 estimator is, but it results in blurry SR solutions with estimation artifacts. And, a regularization term must be incorporated in the cost function to stabilize the solution. On the other hand, the proposed algorithm results in crisp, artifact-free high-resolution images, without the use of regularization. It is worth noting the proposed scheme converges to the solution only within a few iterations (typically from 6 to 12). For the L_1 and L_2 estimators, the convergence depends on the step size parameter β. For the selected step sizes, the L_1 and L_2 estimators converged in 30 and 40 iterations respectively, in the first experiment and in 25 and 50 iterations respectively, in the second experiment.

5. CONCLUSIONS

In this paper, we introduced a new SR algorithm in a robust estimation framework. Two fundamental features of the algorithm are 1) robustness to model violations (outliers) and 2) fast convergence. The proposed framework outperforms L_1 and L_2 estimation, in the absence of a regularization term in the objective function. We are currently investigating the incorporation of a robust regularization term based on redescending *M*-estimators to address the ill posedness of SR estimation when there is insufficient number of LR frames, which is typical in real video sequences.

6. REFERENCES

[1] F. R. Hampel, E. M. Ronchetti, P. J. Rousseeuw, and W. A. Stahel, *Robust Statistics: The Approach Based on Influence Functions.* Wiley, New York, 1986.

[2] P. J. Huber, *Robust Statistics.* Wiley, New York, 1981.

[3] A. Zomet and S. Peleg, "Efficient Super-Resolution and Applications to Mosaics*", Proceedings of the ICPR'00*, Barcelona, vol. I, pp. 3-7, 2000.

[4] A. Rav-Acha, A. Zomet, S. Peleg, "Robust Super Resolution", *Proceedings of the CVPR'01*, Hawaii, vol. I, pp. 645-650, 2001.

[5] S. Farsiu, D. Robinson, M. Elad, and P. Milanfar, "Fast and Robust Multi-frame Super-resolution", *IEEE Trans. on Image Processing*, vol. 13, no. 10, pp. 1327-1344, 2004.

[6] S. Farsiu, M. Elad, and P. Milanfar, "Multi-Frame Demosaicing and Super-Resolution of Color Images", *IEEE Trans. on Image Processing*, vol. 15, no. 1, pp. 141-159, 2006.

[7] J.R. Bergen, P. Anandan, K.J. Hanna, and R. Hingorani. "Hierarchical model-based motion estimation", *In Proc. 2nd European Conference on Computer Vision*, pp. 237-252, 1992.

[8] M. J. Black and P. Anandan, "The robust estimation of multiple motions: Parametric and piecewise-smooth flow fields", *Computer Vision and Image Understanding*, CVIU, 63(1), pp. 75-104, 1996.

[9] Tamer Rabie, "Robust Estimation Approach for Blind Denoising", *IEEE Trans. on Image Processing*, vol. 14, no. 11, pp. 1755-1765, 2005.

[10] M. J. Black, G. Sapiro, D H. Marimont and D. Heeger, "Robust anisotropic diffusion", *IEEE Trans. on Image Processing,* vol.7, no.3, pp. 421-432, 1998.

[11] A. Blake and A. Zisserman, Eds., *Visual Reconstruction*. Cambridge, MA: MIT Press, 1987.

[12] J. Lin, "Divergence Measures Based on the Shannon Entropy", *IEEE Trans. on Information Theory*, vol.37, no.1, pp.145-151, 1991.

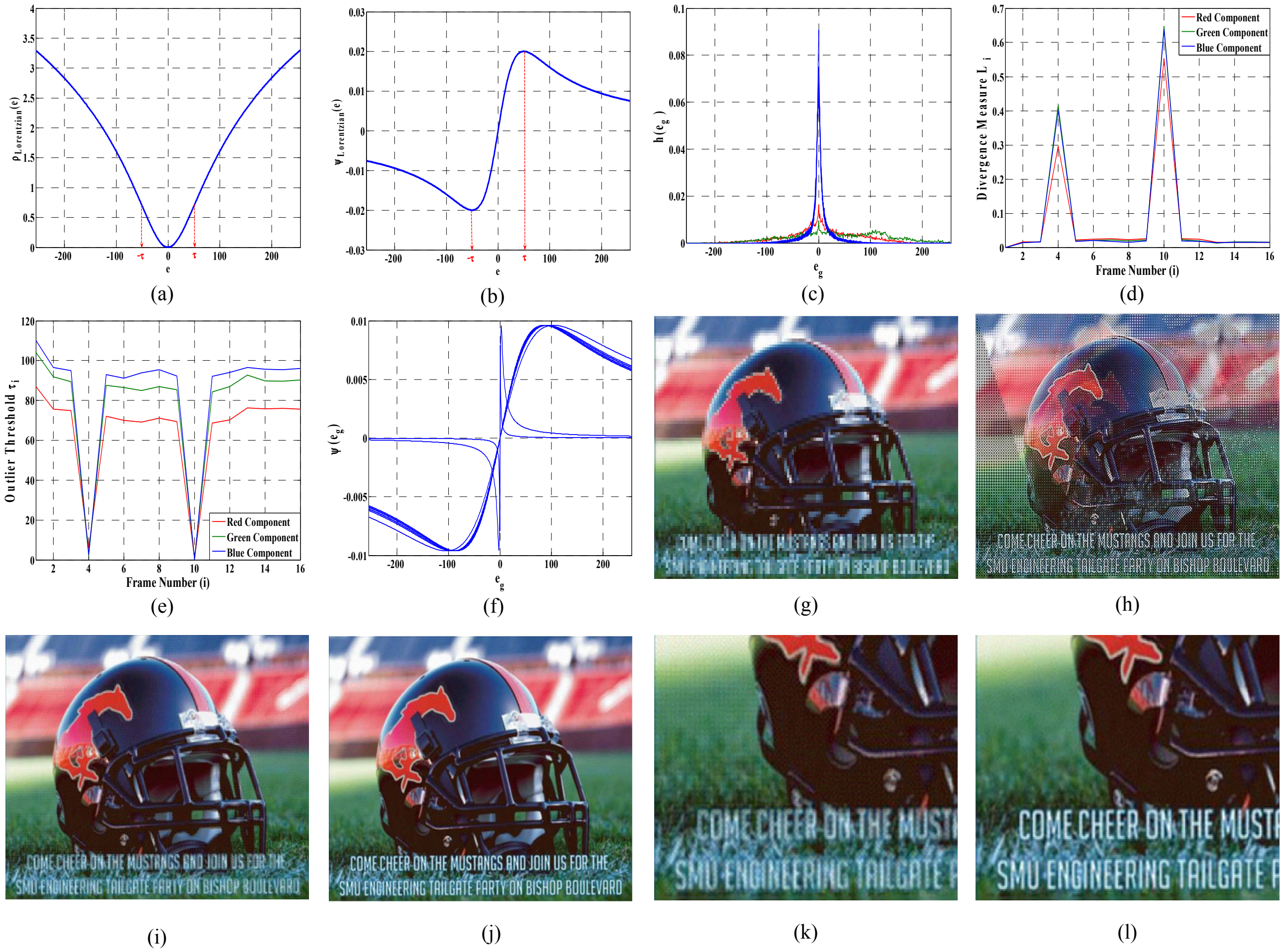

Fig.1 4× Super-resolution estimation results for SMU Helmet image

(a) Lorentzian estimator for $\tau = 50$, (b) Lorentzian influence function for $\tau = 50$, (c) Error distributions of the green component, (d) L-divergence values for the three color components, (e) Outlier thresholds for the three color components, (f) Influence curves for the green component, (g) First LR frame, (h) L_2 estimate, (i) L_1 estimate, (j) Lorentzian estimate, (k) Bottom left corner of (i), (l) Bottom left corner of (j)

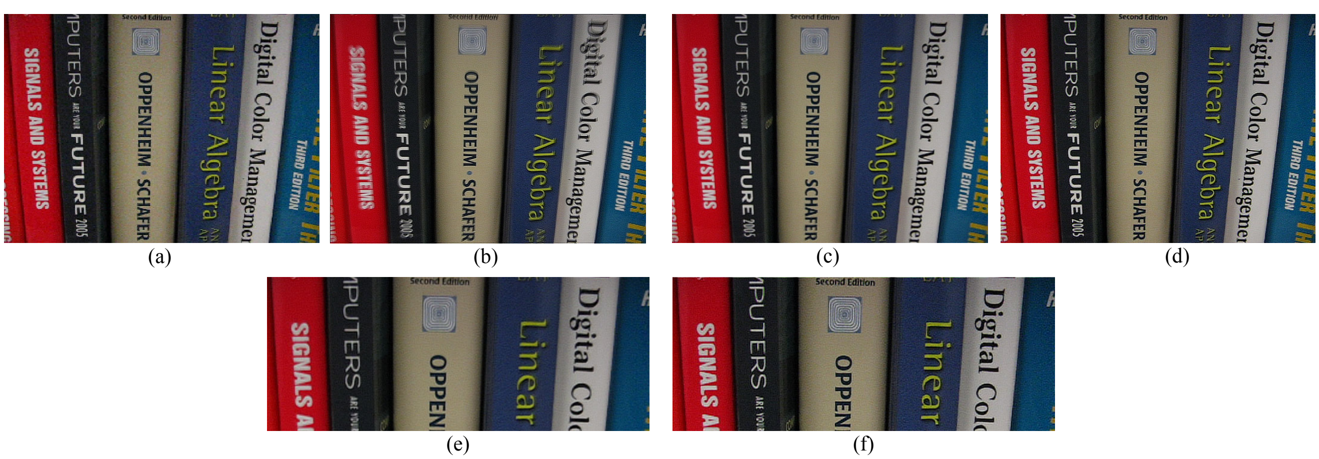

Fig.2 2×Super-resolution estimation results for the Books sequence

(a) First LR frame, (b) L_2 estimate, (c) L_1 estimate, (d) Lorentzian estimate, (e) Top part of (c), (f) Top part of (d)