Hindawi Publishing Corporation
EURASIP Journal on Image and Video Processing
Volume 2008, Article ID 763254, 12 pages
doi:10.1155/2008/763254

Research Article

Robust Color Image Superresolution: An Adaptive M-Estimation Framework

Noha A. El-Yamany and Panos E. Papamichalis

Department of Electrical Engineering, School of Engineering, Southern Methodist University, P.O. Box 750338, Dallas, TX 75275, USA

Correspondence should be addressed to Noha A. El-Yamany, elyamany@ieee.org

Received 2 August 2007; Revised 19 November 2007; Accepted 7 February 2008

Recommended by Shoji Tominaga

This paper introduces a new color image superresolution algorithm in an adaptive, robust M-estimation framework. Using a robust error norm in the objective function, and adapting the estimation process to each of the low-resolution frames, the proposed method effectively suppresses the outliers due to violations of the assumed observation model, and results in color superresolution estimates with crisp details and no color artifacts, without the use of regularization. Experiments on both synthetic and real sequences demonstrate the superior performance over using the L_2 and L_1 error norms in the objective function.

1. INTRODUCTION

Image super-resolution (SR) is a popular research area for producing high-resolution (HR) images with better details. The approach taken is to combine the information in a sequence of low-resolution (LR) images which have subpixel shifts with respect to each other. Most image SR algorithms assume a mathematical model for the imaging process, which could have generated the sequence of LR frames from the *unknown* HR image. However, these models are only approximations to reality, and model violations often occur because of the approximate nature of the model itself, because of inaccuracies in its parameter estimation (such as blur and motion parameters) and because of accidental scene changes. These model violations even small in number can be detrimental to SR estimation.

Robust statistics [1–4] has emerged as a family of theories and techniques for estimation while dealing with deviations from the idealized model assumptions. In particular, robust M-estimation has been found very effective in many computer vision applications such as optical flow estimation [5], robust denoising [6], and robust anisotropic diffusion [7]. The reader is referred to [8] for a review of the applications of robust statistics in computer vision. Robust M-estimation has been explored recently in SR reconstruction. Capel [9] used Huber functions in the prior term in the context of MAP (maximum *a posteriori*) estimation. El-Yamany et al. [10] developed an adaptive M-estimation scheme using the robust Lorentzian error norm in the data fidelity term, without regularization. Patanavijit et al. [11, 12] also demonstrated the use of the Lorentzian error norm in both the data fidelity and regularization terms of the objective function. In this paper, we attempt to address the problem of color image SR in a robust, adaptive M-estimation framework. The proposed approach was first introduced in [13].

2. PROBLEM FORMULATION

A number of observation models have been proposed in the literature for image SR reconstruction [9–28]. In this paper, we employ the following observation model for color image SR

$$\mathbf{Y}_{i,k} = DHF_{i,k}\mathbf{X}_i + \mathbf{Z}_{i,k}, \quad k = 1, 2, \ldots, L, \ i = R, G, B, \qquad (1)$$

where L is the number of LR frames. $\mathbf{X}_i$ and $\mathbf{Y}_{i,k}$ are the ith color component of the *unknown* HR image, and the kth LR frame, respectively. The matrices D and H represent the down-sampling operator and the point spread function (PSF) of the optical system, respectively. In this paper, we assume that we are dealing with one optical system with

space-invariant response, and hence D and H are assumed to be the same for all the LR frames. $F_{i,k}$ is the warping matrix that represents the motion between the LR frames and the *unknown* HR frame. $\mathbf{Z}_{i,k}$ represents the system noise. It is worth mentioning that in (1), the LR and HR images are represented as vectors obtained from the 2D images by lexicographical ordering. Following the observation model in (1) and recasting the SR problem in the generalized M-estimation framework, the color SR output is the solution of the following minimization:

$$\begin{aligned}\mathbf{X}^* &= \arg\min_{\mathbf{X}} \sum_i \sum_{k=1}^{L} \rho(DHF_{i,k}\mathbf{X}_i - \mathbf{Y}_{i,k}) \\ &= \arg\min_{\mathbf{X}} \sum_i \sum_{k=1}^{L} \rho(\mathbf{E}_{i,k}), \quad i = R, G, B,\end{aligned} \tag{2}$$

where $\mathbf{X} = [\mathbf{X}_R\ \mathbf{X}_G\ \mathbf{X}_B]^T$, $\mathbf{E}_{i,k}$ is the vector of the projection errors corresponding to the ith component of the kth LR frame, and ρ is an even-symmetric function which has a unique minimum at zero and satisfies the following condition:

$$\frac{\partial}{\partial \mathbf{X}_i} \sum_{k=1}^{L} \rho(\mathbf{E}_{i,k}) = 0 \Longrightarrow \sum_{k=1}^{L} (DHF_{i,k})^T \psi(\mathbf{E}_{i,k}) = 0, \tag{3}$$

where $\rho(\mathbf{E}) = \sum_j \rho(e_j)$ and $\psi(\mathbf{E}) = [\psi(e_1)\ \cdots\ \psi(e_j)\ \cdots]^T$. $\rho(e_j)$ is a function applied to the element e_j of the vector $\mathbf{E}$. $\psi(e)$ is the first derivative of ρ with respect to e, and is referred to as the *influence function* [1–3].

The robustness of image SR reconstruction has been addressed recently in the literature [9–13, 17, 18, 21–27, 29, 30]. In the context of M-estimation, [17, 18] have addressed the solution of (2) as a least-square (LS) estimation problem using the L_2 error norm, and the SR estimate is the solution of the minimization

$$\begin{aligned}\mathbf{X}^* &= \arg\min_{\mathbf{X}} \sum_i \sum_{k=1}^{L} \rho_2(\mathbf{E}_{i,k}) \\ &= \arg\min_{\mathbf{X}} \sum_i \sum_{k=1}^{L} ||\mathbf{E}_{i,k}||_2^2, \quad i = R, G, B,\end{aligned} \tag{4}$$

where $\rho_2(\cdot)$ refers to the L_2 error norm. The solution of (4) can be found iteratively using the gradient descent algorithm as follows:

$$\begin{aligned}\mathbf{X}_i^{n+1} &= \mathbf{X}_i^n - \beta \sum_{k=1}^{L} (DHF_{i,k})^T \psi_2(\mathbf{E}_{i,k}^n) \\ &= \mathbf{X}_i^n - \beta \sum_{k=1}^{L} (DHF_{i,k})^T \mathbf{E}_{i,k}^n, \quad i = R, G, B,\end{aligned} \tag{5}$$

where $\psi_2(\cdot)$ refers to influence function of the L_2 error norm, and β is a step-size parameter. However, LS estimation exhibits a poor performance in the presence of model violations. The reason for the nonrobustness of the L_2 error norm lies in its influence function (its first derivative ψ_2). As shown in Figure 1 and in (5), ψ_2 is linear, assigning larger weights to larger errors, and hence it is very vulnerable in the face of outliers.

To increase the robustness of SR estimation, Farsiu et al. [20–22] proposed the use of the L_1 error norm as a robust alternative to the L_2 error norm, and the SR estimate is the solution of the following minimization:

$$\begin{aligned}\mathbf{X}^* &= \arg\min_{\mathbf{X}} \sum_i \sum_{k=1}^{L} \rho_1(\mathbf{E}_{i,k}) \\ &= \arg\min_{\mathbf{X}} \sum_i \sum_{k=1}^{L} ||\mathbf{E}_{i,k}||_1, \quad i = R, G, B,\end{aligned} \tag{6}$$

where $\rho_1(\cdot)$ refers to the L_1 error norm. The solution of (6) can also be found iteratively using the gradient descent algorithm as follows:

$$\begin{aligned}\mathbf{X}_i^{n+1} &= \mathbf{X}_i^n - \beta \sum_{k=1}^{L} (DHF_{i,k})^T \psi_1(\mathbf{E}_{i,k}^n) \\ &= \mathbf{X}_i^n - \beta \sum_{k=1}^{L} (DHF_{i,k})^T \operatorname{sign}(\mathbf{E}_{i,k}^n), \quad i = R, G, B,\end{aligned} \tag{7}$$

where $\psi_1(\cdot)$ refers to influence function of the L_1 error norm. As shown in Figure 1 and in (7), the L_1 influence function, ψ_1, is the sign function, and so 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 because of its bounded influence. However, because of its constant-valued influence, which does not distinguish between small and large errors, the resulting SR estimate often suffers from various artifacts.

To improve the robustness of the LS estimation, Ivanovski et al. [27] proposed a pixel-level selection strategy for outlier rejection. In [27], a similarity measure is used to determine the usability of each pixel in the SR estimation process. Then a hard threshold is empirically selected, and a pixel is discarded from the estimation process if the measure evaluated at the pixel location is greater than the selected threshold. However, the performance of the algorithms in [27] depends on the chosen threshold. In addition, because of the pixel-level selection strategy, different pixel locations receive different contributions from neighboring LR frames. These unequal contributions, which depend on the pixels in the LR frames that satisfy the similarity condition, result in newly introduced boundaries between regions of different contributions as shown in [27]. In [26], a frame-level selection strategy was proposed for outlier rejection. In this strategy, the cross correlation between a reference frame and a warped frame is computed, and if the correlation score at a point is below a certain threshold, the corresponding warped pixels are ignored in the SR reconstruction. However as in [27], the performance of the reconstruction algorithm in [26] depends on the selected hard threshold. If a small threshold is chosen, a good frame could be rejected, thus reducing the amount of available information that could have

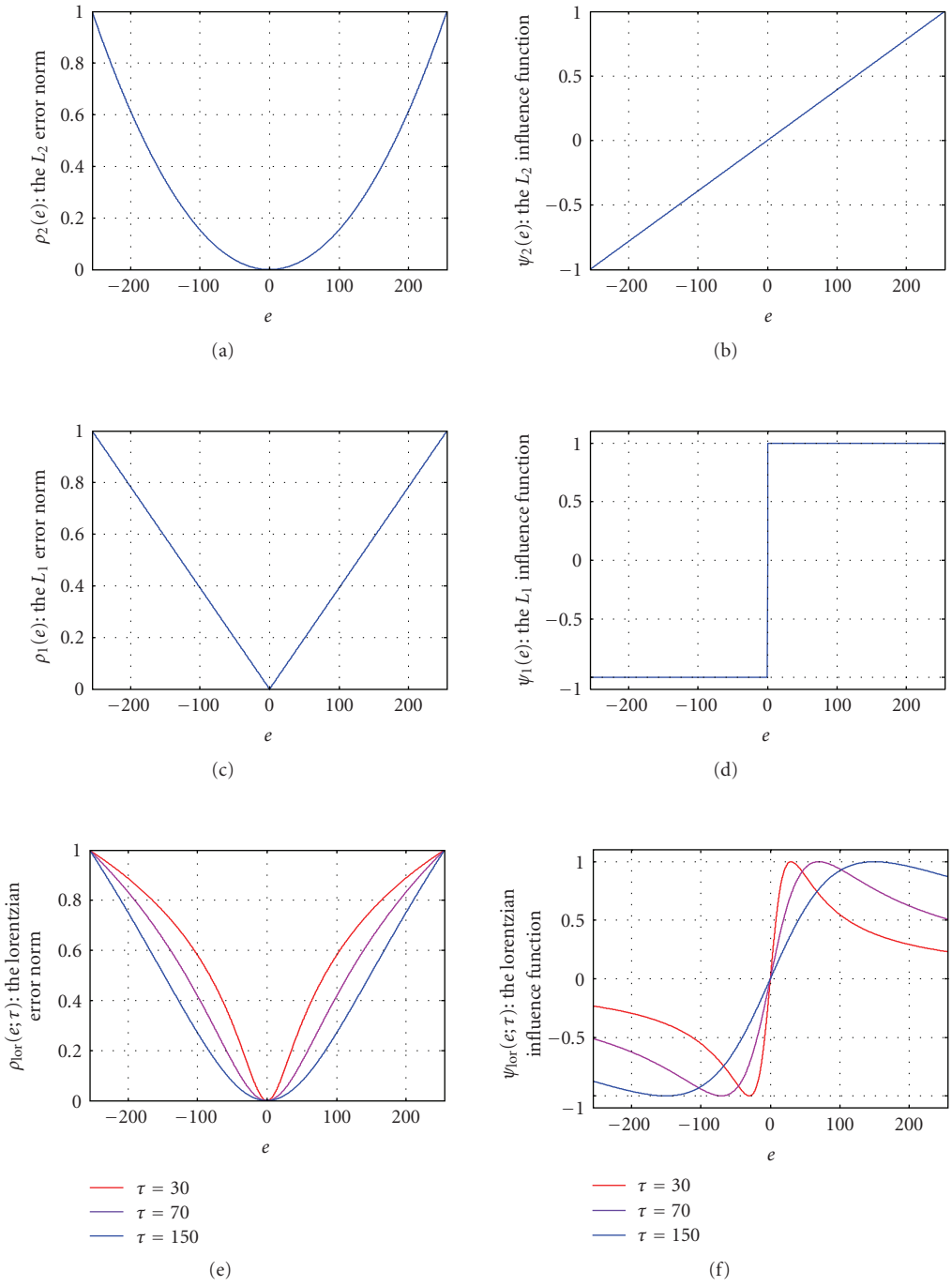

Figure 1: *Plot of the L_1, L_2, and Lorentzian error norms and their corresponding influence functions.* (All functions are normalized to demonstrate the relative weight assigned to different error values.)

been taken advantage of to enhance the resolution of the target frame. On the other hand, if a large threshold is selected, the contributions from incorrectly registered frames can be detrimental to the reconstruction process, leading to poor SR estimates.

Most of the methods that have been developed recently in the literature for color image SR in the context of M-estimation [28–31] use the L_2 or the L_1 error norm in the data fidelity term of the objective function, and incorporate an additional color regularization term to help in dealing with the outliers and ill-posedness of the SR problem. However, the performance of such SR reconstruction schemes is highly affected by the choice of the regularization function and the corresponding regularization parameter as well, not to mention the additional computational complexity associated with the joint optimization of both the data fidelity and regularization terms of the objective function. In this paper, we attempt to address the problem of robust color image SR

in an M-estimation framework, without the use of color regularization in the objective function. The proposed approach was first introduced in [13].

3. THE PROPOSED ALGORITHM

To improve the robustness of SR reconstruction, we propose the use of robust error norms in the data fidelity term of the objective function, in particular, the robust error norms which correspond to the class of M-estimators known as *redescending M-estimators* [1–4]. For these estimators, the influence function ψ increases up to a point, which is referred to as the outlier threshold, after which it starts to decrease (redescend) as the error grows. Because of this behavior, large errors falling beyond the outlier threshold are assigned weights that decrease as the errors increase, thus providing a soft outlier suppression or rejection if the weights are too small. Of all redescending M-estimators, we are particularly interested in estimators whose influence functions are differentiable and have only one parameter, which will be determined from the available observations as shown later. Examples of these estimators are the Lorentzian, Geman & McClure, and Tukey's biweight [1–8]. In this paper, only the use of the Lorentzian estimator is demonstrated. The Lorentzian error norm is defined as

$$\rho(e;\tau) = \log\left[\frac{e^2+\tau^2}{\tau^2}\right], \tag{8}$$

where e and τ are the error and the outlier threshold parameter, respectively. The Lorentzian influence function (which is proportional to the first derivative of (8)) is given by

$$\psi(e;\tau) = \frac{2\tau e}{e^2+\tau^2}. \tag{9}$$

Note that the influence function in (9) is scaled, by multiplying by τ, to have a maximum weight of unity independent of the outlier threshold value. This normalization is particularly important in the proposed adaptive formulation to ensure that all influence functions have the same weight at their respective outlier thresholds. Figure 1 depicts plots of the Lorentzian error norm and its influence function for three different values of τ. These plots show how the influence function decreases faster for smaller τ, assigning lower weights to the errors falling beyond the outlier threshold.

3.1. The objective function

Recasting the problem of color image SR in a robust M-estimation framework using redescending M-estimators, the SR estimate is the solution of the following minimization:

$$\begin{aligned}\mathbf{X}^* &= \arg\min_{\mathbf{X}} \sum_i \sum_{k=1}^{L} \rho\left(DHF_{i,k}\mathbf{X}_i - \mathbf{Y}_{i,k};\tau\right) \\ &= \arg\min_{\mathbf{X}} \sum_i \sum_{k=1}^{L} \rho\left(\mathbf{E}_{i,k};\tau\right), \quad i = R,G,B,\end{aligned} \tag{10}$$

where ρ is the Lorentzian (robust) error norm whose influence function is redescending [1–4], and τ is the corresponding outlier threshold.

Violations to the assumed mathematical model for the imaging process in (1) result in large projection errors ($\mathbf{E}_{i,k}$), which are referred to as *outliers*, and can be detrimental to the reconstruction process if the estimation procedure does not suppress or eliminate their contributions. The choice of the outlier threshold τ for a redescending M-estimator plays a vital role in dealing with these outliers. As shown in Figure 1, the errors falling beyond τ are assigned smaller weight as the error grows, thus providing soft outlier suppression, or rejection if the weight is too small. Also for smaller τ, the influence function decays faster, assigning smaller weights to the errors greater than the outlier threshold. Because of this behavior of redescending influence functions, choosing a fixed outlier threshold for all the LR frames (and all color components) would not be appropriate. If a small τ is selected, the contributions from all frames (inliers and outliers) will be suppressed or rejected, leading to a poor SR estimate because of the insufficient information to increase the spatial resolution of the target LR frame. On the other hand, if a large τ is selected, the outliers will contribute to the estimation procedure, resulting in a poor SR estimate suffering from various artifacts.

To appreciate this fact, Figures 2(f) and 2(g) display the SR estimates for $\tau = 10$ and $\tau = 70$, respectively, for the SMU Helmet sequence experiment detailed in Section 4.1. From these results, it is shown that for a small threshold ($\tau = 10$), the outliers are successfully eliminated, but the resulting SR estimate is blurry and of poor quality because there are no sufficient contributions from all the frames. On the other hand, for a relatively large threshold ($\tau = 70$), the resulting SR estimate suffers from noticeable artifacts due to the outliers.

From these results, it is concluded that an adaptive procedure is necessary to deal effectively with the outliers, yet consider the contributions from all the *good* frames to enhance the resolution of the target frame. In this adaptive procedure, different outlier thresholds are assigned to different LR frames depending on their similarity to the target frame, as discussed later in Section 3.2. Recasting the problem of color image SR in an adaptive, robust M-estimation framework using redescending M-estimators, the SR estimate is the solution of the following minimization:

$$\mathbf{X}^* = \arg\min_{\mathbf{X}} \sum_i \sum_{k=1}^{L} \rho\left(\mathbf{E}_{i,k};\tau_{i,k}\right), \quad i = R,G,B, \tag{11}$$

where $\rho(\mathbf{E}_{i,k};\tau_{i,k})$ is the Lorentzian error norm associated with the ith color component of the kth LR frame, and $\tau_{i,k}$ is the corresponding outlier threshold. It is worth noting that in the adaptive formulation in (11), different sets of outlier thresholds are calculated for each color component in each of the LR frames. This strategy helps in dealing effectively with the outliers that might appear in one color channel/frame and not in the other channels/frames (as shown later in Section 4.2).

The calculation of the outlier threshold for redescending M-estimators is usually done using statistical methods [1–4], or based on problem-dependent choices [5–8, 10–12]. In the proposed SR framework, we have developed an algorithm to

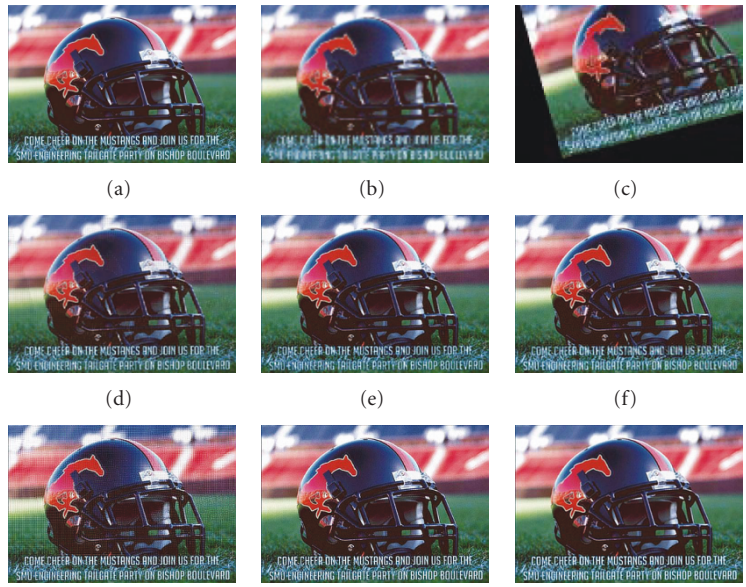

FIGURE 2: 4× *superresolution reconstruction results for the* 16*-frames SMU Helmet sequence*: (a) original HR image (ground truth), (b) LR frame #1, (c) LR frame #4, (d) SR estimate using L_2 error norm + Tikhonov regularization, (e) SR estimate using L_1 error norm + Tikhonov regularization, (f) SR estimate using the Lorentzian error norm with a fixed outlier threshold $\tau = 10$ for all the LR frames. (g) SR estimate using the Lorentzian error norm with a fixed outlier threshold $\tau = 70$ for all the LR frames, (h) SR estimate using the proposed adaptive scheme without regularization, and (i) with Tikhonov regularization.

estimate $\tau_{i,k}$ from the LR frames, which is described in detail in the following subsection.

3.2. Calculation of the outlier thresholds

3.2.1. Calculation of $\tau_{i,1}$

In image SR reconstruction, we are interested in increasing the resolution of a LR reference frame ($\mathbf{Y}_1$) using the information in that frame and other available LR frames ($\mathbf{Y}_k$). In this sense, the projection errors corresponding to the reference frame should all be considered in the estimation process, that is, they are all *inliers*. For 8-bit data, the maximum absolute value for the projection errors is 255 and hence a reasonable choice for the outlier threshold for the reference frame is 255. Therefore, we have set the outlier threshold for the three color components of the reference frame ($\tau_{i,1}$) to 255.

3.2.2. Calculation of $\tau_{i,k}$

To compute the outlier thresholds for the rest of the LR frames, a metric that measures the similarity between the reference frame and the kth motion-compensated LR frame ($\widetilde{\mathbf{Y}}_k$) is first computed. This metric is denoted by $d_k = d(\mathbf{Y}_1, \widetilde{\mathbf{Y}}_k)$, where $k = 1, 2, 3, \ldots, L$, and L is the number of LR frames. The outlier threshold for a given LR frame is then calculated as a function of this metric such that if $d_k \to 0$, $\tau_k \to \tau_1$, if $d_k \to \infty$, theoretically, $\tau_k \to 0$ and if $d_k \to d_{\max}$ (upper bound on d), $\tau_k \to \tau_{\min}$ (lower bound on τ). Under these constraints, we consider the following exponential function as a reasonable choice to calculate τ_k from d_k:

$$\tau_k = \tau_1 e^{-\alpha d_k} = 255 e^{-\alpha d_k}. \tag{12}$$

The parameter α in (12) controls the decay of the exponential function, and, given the two constraints above, it is calculated as

$$\alpha = \frac{1}{d_{\max}} \log\left(\frac{255}{\tau_{\min}}\right). \tag{13}$$

The lower bound on the outlier threshold $\tau_{\min}$ is chosen to be an arbitrarily very small number. In the experiments presented in this paper, $\tau_{\min}$ was set to 10^{-8}. For the similarity metric, we have used the normalized average sum of absolute differences (SAD) between $\mathbf{Y}_1$ and $\widetilde{\mathbf{Y}}_k$, which is defined by

$$d_k = \frac{1}{255 \times MN} \sum_{x=1}^{M} \sum_{y=1}^{N} \left| y_1(x, y) - \tilde{y}_k(x, y) \right|. \tag{14}$$

The normalized average SAD has an upper bound of unity, that is, $d_{\max} = 1$. Therefore α is computed as

$$\alpha = \log\left(255 \times 10^8\right) \approx 24. \tag{15}$$

The outlier thresholds for the three color components of the LR frames are thus computed as follows:

$$\tau_{i,k} = 255 e^{-24 d_{i,k}}, \quad i = R, G, B, \ k = 2, 3, \ldots, L, \tag{16}$$

where $d_{i,k}$ is the normalized average SAD between the ith color component of the reference LR frame ($\mathbf{Y}_1$) and the kth

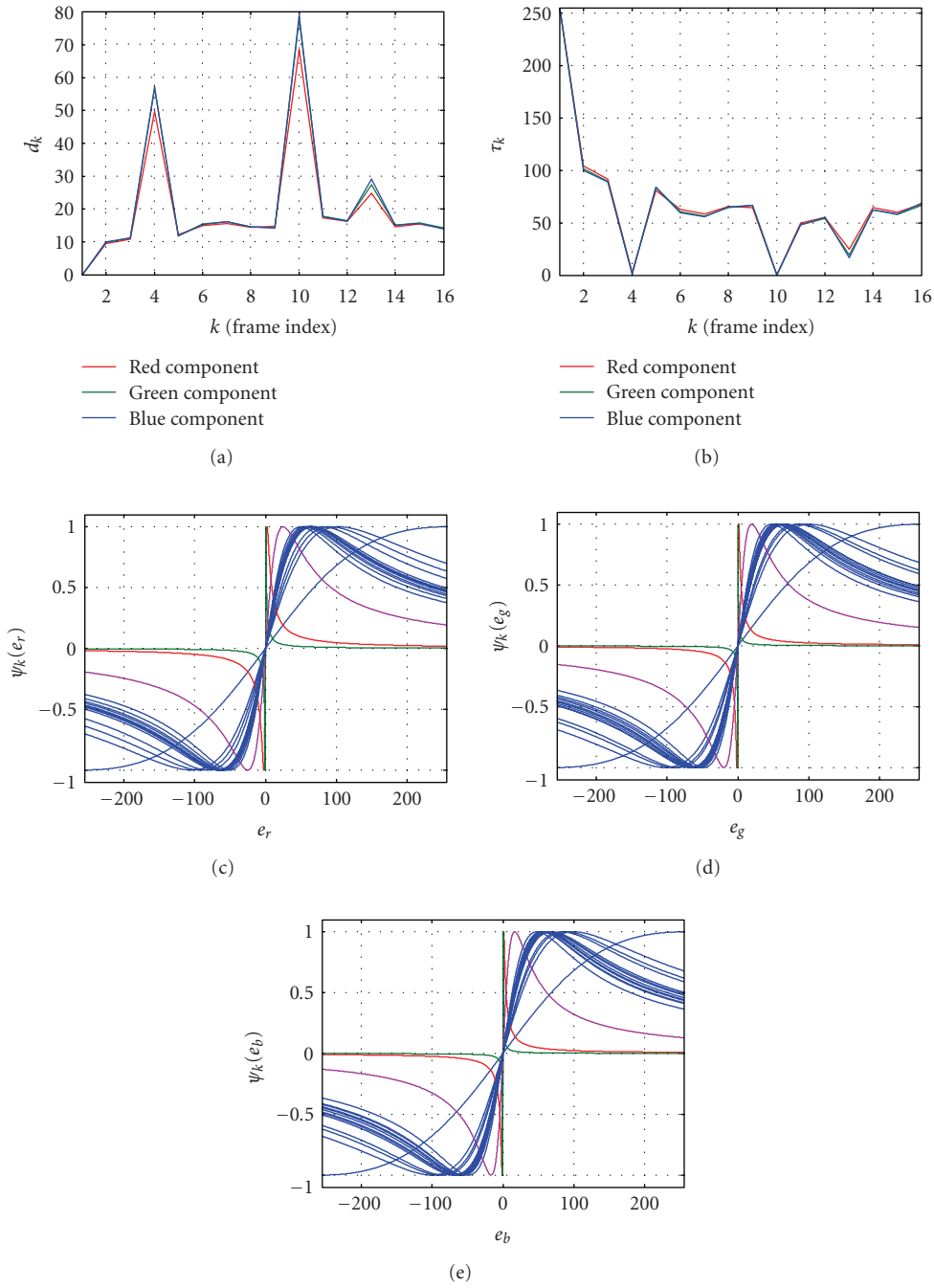

FIGURE 3: *Superresolution reconstruction results of the* 16*-frames SMU Helmet sequence using the proposed approach*: (a) plot of average SAD values (d_k's) for the three color components, (b) plot of the outlier thresholds (τ_k's) for the three color components, and plot of the Lorentzian influence functions (ψ_k's) for (c) the red component, (d) green component, and (e) blue component. From (c) to (e), the red, green, and magenta curves correspond to LR frame #4, LR frame #10, and LR frame #13, respectively.

motion-compensated LR frame ($\widetilde{\mathbf{Y}}_k$). Figure 3 depicts plots of the SAD measure values (d_k's) and the outlier thresholds (τ_k's) for the three color components (R, G, and B) of the SMU Helmet sequence experiment explained in Section 4.1. It is worth mentioning that the average SAD is one possible measure to assess the similarity between the reference frame and each of the motion-compensated LR frames. We have chosen to use this measure because of its low computational complexity, and because it captures the mismatch between the two LR frames.

It is worth mentioning that the proposed algorithm for computing the outlier thresholds is different from those in [26, 27] in terms of the following. First, the outlier thresholds in the proposed scheme are not *hard* thresholds, that is, errors falling beyond a given outlier threshold are not rejected. They are, however, assigned smaller weights as their values

increase, providing gradual outlier suppression, or rejection if the weights are too small (the errors are too large with respect to the threshold). Second, the outlier thresholds are calculated automatically from the available LR frames. Using a distance (similarity) metric between each of the LR frames and the target frame, the outlier thresholds are computed such that the higher the dissimilarity between a given LR frame and the reference frame, the smaller the corresponding outlier threshold.

3.3. The update equation

To find the solution of (11), one might use Newton's algorithm. However, the influence function of redescending M-estimators is bounded (as shown in Figure 1), and its derivative is not always positive and goes to zero at infinity, which makes using Newton's algorithm unreliable and convergence is not guaranteed [3]. Therefore, we have chosen to use the gradient descent method as discussed in what follows.

Using the iterative gradient descent algorithm, the update equation minimizing (11) for the ith color component is

$$\mathbf{X}_i^{n+1} = \mathbf{X}_i^n - \eta_i \sum_{k=1}^{L} \nabla\rho\left(\mathbf{E}_{i,k}^n; \tau_{i,k}\right), \quad i = R, G, B, \tag{17}$$

where η_i is the step-size parameter corresponding to the ith color component. Following the derivation in the appendix, (17) can be written as

$$\mathbf{X}_i^{n+1} = \mathbf{X}_i^n - \eta_i \sum_{k=1}^{L} \left(DHF_{i,k}\right)^T \boldsymbol{\psi}_{i,k}^n, \quad i = R, G, B, \tag{18}$$

where $\boldsymbol{\psi}_{i,k}^n$, is a vector whose jth element is $\psi(e_{j,i,k}^n; \tau_{i,k})$, the *Lorentzian* influence function evaluated at $e_{j,i,k}^n$ (the jth element in $\mathbf{E}_{i,k}^n$).

The choice of the step-size parameter η_i plays an important role in the convergence behavior of the gradient descent method. If the step size is too large, divergence will occur, while if the step size is too small, the rate of convergence may be very slow [32]. Choosing a constant step size is the simplest approach. However, constant step-size selection is only useful in cases where an appropriate step-size value is known or can be determined fairly easily [32]. For twice differentiable robust error norms, such as the *Lorentzian*, a proper constant step-size selection can be obtained using the method of simultaneous over-relaxation (SOR) [5, 6, 33]. The SOR step size is defined as $\eta = \omega/T$, where $0 < \omega < 2$ and T is an upper bound on the second partial derivative of the $\rho(e; \tau)$ with respect to e. The exact choice of ω only affects the rate of convergence. In the proposed approach, ω is set to 1 and the step size is approximated by $\eta = 1/T \approx \tau/2$, for the *Lorentzian* error norm. To achieve fast convergence, and motivated by the SOR algorithm [5, 6, 33], we used an adaptive step size for each of the L terms in (18), that is,

$$\mathbf{X}_i^{n+1} = \mathbf{X}_i^n - \sum_{k=1}^{L} \eta_{i,k} \left(DHF_{i,k}\right)^T \boldsymbol{\psi}_{i,k}^n, \quad i = R, G, B, \tag{19}$$

where the step-size parameter $\eta_{i,k}$ is computed as

$$\eta_{i,k} = \frac{\tau_{i,k}}{2} \tag{20}$$

for the *Lorentzian* error norm. Having set the adaptive step size as in (20), the convergence is also checked after each iteration, and if the cost function is not improving, then $\eta_{i,k}$ can be reduced by a small amount (e.g., 5%), otherwise it is kept at its current value. In all the experiments we have conducted, however, setting the step size as in (20) has shown fast convergence (typically from 7 to 12 iterations), and so there was no need to reduce the step size calculated in (20). In all the experiments presented in this paper, the initial SR estimate ($\mathbf{X}^0$) was found through bilinear interpolation of the first (reference) LR frame.

4. EXPERIMENTAL RESULTS

In this section, the performance of the proposed SR reconstruction scheme is evaluated and compared to methods using the L_2 and L_1 error norms in the data fidelity term and Tikhonov regularization in the smoothness prior term in the objective function. In all the experimental results presented here for L_2 and L_1 SR estimation, the gradient descent method is used and is applied to each of the three color components (R, G and B) separately to obtain the final SR estimate, as shown below:

$$\begin{aligned} &L_2: \\ &\mathbf{X}_i^{n+1} = \mathbf{X}_i^n - \beta\left\{\sum_{k=1}^{L} \left(DHF_{i,k}\right)^T \mathbf{E}_{i,k}^n + \lambda\Gamma^T\left(\Gamma\mathbf{X}_i^n\right)\right\}, \\ &L_1: \\ &\mathbf{X}_i^{n+1} = \mathbf{X}_i^n - \beta\left\{\sum_{k=1}^{L} \left(DHF_{i,k}\right)^T \operatorname{sign}\left(\mathbf{E}_{i,k}^n\right) + \lambda\Gamma^T\left(\Gamma\mathbf{X}_i^n\right)\right\}, \end{aligned} \tag{21}$$

where Γ is a high-pass operator [20]. In this paper, the Laplacian operator is used, which is approximated by the following symmetric convolution kernel:

$$\gamma = \frac{1}{8}\begin{bmatrix} 1 & 1 & 1 \\ 1 & -8 & 1 \\ 1 & 1 & 1 \end{bmatrix}. \tag{22}$$

Experiments with different values of the step-size parameter (β) and Tikhonov regularization parameter (λ) were conducted, and the one that gave the best visual quality is only presented in the paper (same β and λ were used for all the three color components). For the SR experiments using the proposed scheme without regularization, the algorithm described in Section 3 is followed. For the results using the proposed method with the inclusion of Tikhonov regularization, the following update equation is used:

$$\mathbf{X}_i^{n+1} = \mathbf{X}_i^n - \left\{\sum_{k=1}^{L} \eta_{i,k} \left(DHF_{i,k}\right)^T \psi_{i,k}^n + \lambda\Gamma^T\left(\Gamma\mathbf{X}_i^n\right)\right\}. \tag{23}$$

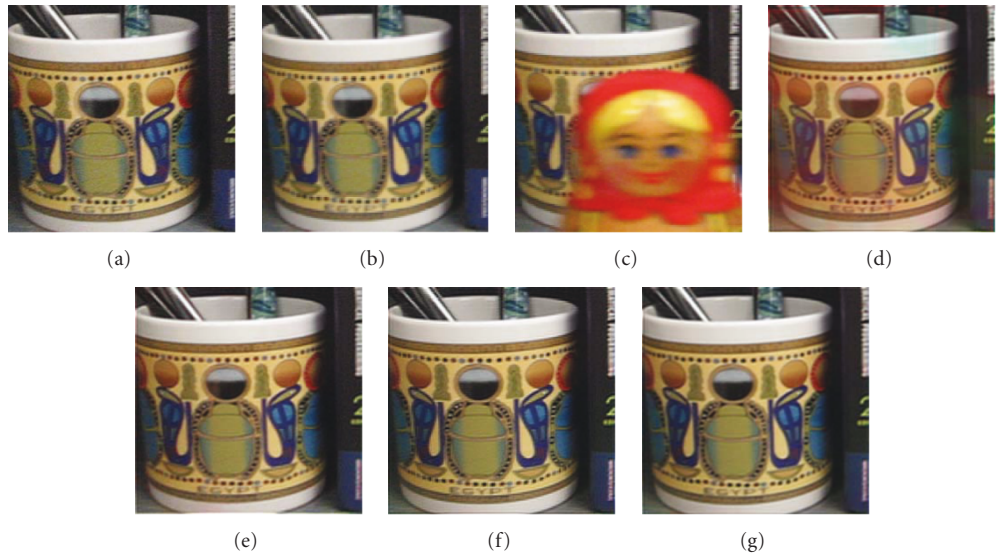

Figure 4: 2× *superresolution reconstruction results for the* 12*-frames Egyptian Mug sequence:* (a) original HR frame #1 (ground truth), (b) LR frame #1, (c) LR frame #5, (d) SR estimate using L_2 error norm + Tikhonov regularization. (e) SR estimate using L_1 error norm + Tikhonov regularization, (f) SR estimate using the proposed algorithm without regularization, and (g) with Tikhonov regularization.

It is worth mentioning that although a number of superresolution reconstruction algorithms that incorporate color regularization tying the color bands have been introduced in the literature [29–31, 34], in this paper, however, we have not included any such regularization terms in the objective function because this will obscure the impact of the influence function of a given M-estimator. Also, the goal is to emphasize the effect of the influence function on the reconstruction process.

In this section, experiments with both synthetic and real sequences are demonstrated, and the notation $r\times$ superresolution is used to denote increasing the resolution of the first (reference) frame in a given sequence by a factor of r in each of the x and y directions.

4.1. 4× SR experiment: the SMU Helmet sequence

In this experiment, a sequence of 16 LR frames is synthetically generated from an HR image (the SMU Helmet image shown in Figure 2(a)) as follows. A set of 16 integer shift pairs (in both the x and y directions) is generated and the HR image is shifted by these shift values. Frames #4 and #10 are rotated and zoomed in, respectively, to create general affine motion. The resulting 16 warped frames are then convolved with a normalized 5×5 Gaussian kernel of zero mean and variance of 0.5, and downsampled by a factor of 4 in both the x and y directions. The MATLAB function *fspecial* is used to generate the Gaussian kernel. A zero-mean Gaussian noise is then added to the resulting LR frames such that each has a signal-to-noise ratio of 25 dB. The MATLAB function *imnoise* is used to add the Gaussian noise. It is worth mentioning that warping, blurring, downsampling, and noise addition are performed on each of the color components (R, G, and B) of the HR image separately. To simulate motion estimation errors, a translational motion model is assumed. In addition, a motion bias of 4 pixels (in the HR grid) is added to the motion vector of frame #13 (to each of its three color components). To simulate blur estimation errors, the PSF is assumed to be a normalized 5×5 Gaussian kernel of zero mean and unity variance. Figures 2(b) and 2(c) show LR frames #1 and #4, respectively.

The SR estimate using the L_2 error norm and Tikhonov regularization is shown in Figure 2(d). β and λ were both set to 0.1. From this result, it is shown that the L_2 estimate suffers from noticeable artifacts due to the outliers (the shadows corresponding to the rotated and zoomed frames appear clearly in the background). This result is not surprising, since the L_2 error norm is vulnerable to the outliers because of its linear influence function that assigns larger weights to larger errors, and hence amplifies their influence in the estimation. The SR estimate using the L_1 error norm and Tikhonov regularization is shown in Figure 2(e). β and λ were set to 2 and 0.02, respectively. From this result, it is shown how using the L_1 error norm has suppressed the outliers compared to using the L_2 error norm. However, as discussed in Section 2, because of its constant valued influence function (± 1), it results in a blurry SR estimate of a relatively poor quality.

The SR reconstruction result using the proposed scheme without regularization is shown in Figure 2(h). From this result, it is observed how the proposed approach has successfully suppressed the effect of the outliers, resulting in artifacts-free SR estimate of crisp details. Figure 2(i) depicts the SR estimate using the proposed scheme with Tikhonov regularization ($\lambda = 0.2$). The use of regularization slightly improved the visual appearance of the SR estimate because of the imposed smoothness constraint.

Figure 3 depicts plots of the average SAD (d_k's), the outlier thresholds (τ_k's) and the Lorentzian influence functions

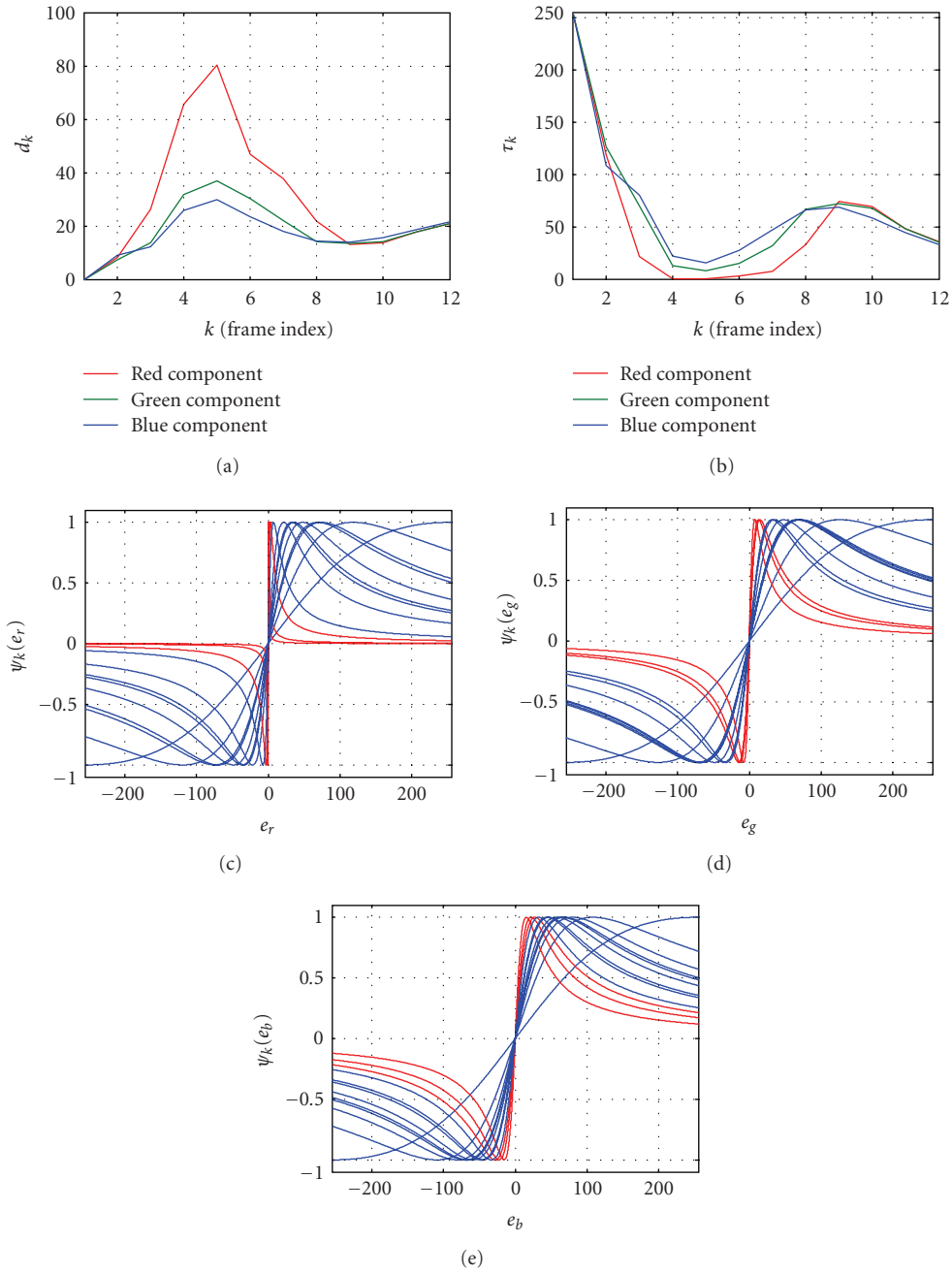

Figure 5: *Superresolution reconstruction results of the* 12*-frames Egyptian Mug sequence using the proposed approach*: (a) plot of average SAD values (d_k's) for the three color components, (b) plot of the outlier thresholds (τ_k's) for the three color components, and plot of the Lorentzian influence functions (ψ_k's) for the (c) red component, (d) green component, and (e) blue component. From (c) to (e), the red curves correspond to LR frames #4, #5, and #6, where the matryoshka doll occludes big parts of the scene.

(ψ_k's) for the 16 LR frames of the SMU Helmet sequence. In Figures 3(c)–3(e), the red, green, and magenta curves correspond to the LR frames #4, #10, and #13, respectively. From these plots, it is shown how the average SAD measure captures the mismatch between the three outliers' frames and the reference (first) LR frame. It is also noted that the outlier thresholds corresponding to frame #4 and #10 are considerably small because of their extreme violation of the assumed translational motion model. Whereas the outlier threshold of frame #13 is relatively big because of the relatively small mismatch due to the motion bias. It is also observed that the influence functions corresponding to frames #4 and #10 decay very rapidly assigning almost negligible weight to the projection errors of these two frames, and hence effectively suppressing their effect in the estimation process. The influence function of LR frame #13, however, decays slightly faster than those of the inliers' frames, assigning lower weight to the projection errors of frame #13.

4.2. 2× SR experiment: the Egyptian Mug sequence

In this experiment, a sequence of 12 compressed frames (MJPEG) is captured by a handheld camera, Canon *Power* Shot A400. This sequence follows approximately the global translation motion. Occlusion is introduced by moving a matryoshka doll into the scene in 4 frames (4–7). The LR sequence of frames is obtained from the original HR sequence captured by the camera via downsampling by a factor of 2 in both the x and y directions. Figures 4(a)–4(c) show the original HR frame #1 (ground truth), LR frame #1, and LR frame #5 in which the matryoshka doll occludes the Egyptian-Mug, respectively. It is worth mentioning that occlusion was introduced intentionally in this sequence to simulate accidental scene changes that typically occur in real video sequences. A translational motion model is assumed, and the algorithm in [35] is used to estimate the motion vectors for each of R, G, and B components, separately. The *unknown* camera PSF is assumed to be a normalized 5×5 Gaussian kernel of zero mean and unity variance and the MATLAB function *fspecial* is used to generate this kernel.

The SR estimate using the L_2 error norm and Tikhonov regularization is shown in Figure 4(d). β and λ were set to 0.5, and 0.1, respectively. From this result, it is shown that the L_2 estimate suffers from excessive false color artifacts. These artifacts result from the matryoshka doll whose color is mostly yellow and red. The SR estimate using the L_1 error norm and Tikhonov regularization is shown in Figure 4(e). β and λ were set to 2 and 0.025, respectively. From this result, it is shown how using the L_1 error norm results in a better estimate than that using the L_2 error norm. However, the SR estimate also suffers from noticeable false (reddish) coloring.

The SR reconstruction results using the proposed scheme without and with Tikhonov regularization ($\lambda = 0.15$) are shown in Figures 4(f) and 4(g), respectively. From these results, it is observed how the proposed approach results in a SR estimate of crisp details and no color artifacts, even without the use of color regularization in the objective function.

Figure 5 depicts plots of the average SAD (d_k's), the outlier thresholds (τ_k's) and the Lorentzian influence functions (ψ_k's) for the 12 LR frames of the Egyptian-Mug sequence. In Figures 5(c)–5(e), the red curves correspond to the LR frames #4–#6, in which the matryoshka doll occludes big potions of the scene. From these plots, it is shown how the average SAD measure captures the mismatch between the outliers' frames, in which the matryoshka doll occludes the Mug, and the reference (first) LR frame. It is also noted that the outlier thresholds corresponding to outliers' frames for the red component are considerably smaller than those for the green and blue components, and those for the green component are smaller than those for the blue component. This is because the colors of the matryoshka doll are mostly red and yellow. From these results, it is shown that computing different outlier thresholds for each of the three color components is very effective in dealing with outliers that might appear in one (or more) color component and not in the rest of the color components.

5. SUMMARY

In this paper, a new adaptive M-estimation framework has been presented for robust color image super-resolution. Using a robust error norm in the data fidelity term of the objective function, and adapting the estimation process to each of the low-resolution frames and each of the color components, the proposed method effectively suppresses the outliers due to model violations, and results in color SR images of crisp details and no artifacts, without the use of regularization. Experimental results on both synthetic and real sequences have demonstrated the superior performance of the proposed algorithm over using the L_2 and the L_1 error norms in the objective function. We are currently investigating the extension of the proposed solution to video sequences in which dealing with local outliers will be addressed.

APPENDIX

DERIVATION OF (18)

Let $\mathbf{E} = DHF\mathbf{X} - \mathbf{Y} = A\mathbf{X} - \mathbf{Y}$, where A is $N \times M$. Then $\mathbf{E}$ can be written as

$$\mathbf{E} = \begin{bmatrix} a_{11} & a_{12} & \cdot & a_{1M} \\ a_{21} & a_{22} & \cdot & a_{2M} \\ \cdot & \cdot & \cdot & \cdot \\ a_{N1} & a_{N2} & \cdot & a_{NM} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ x_M \end{bmatrix} - \begin{bmatrix} y_1 \\ y_2 \\ \cdot \\ y_N \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{M} a_{1i}x_i - y_1 \\ \cdot \\ \cdot \\ \sum_{i=1}^{M} a_{Ni}x_i - y_N \end{bmatrix} = \begin{bmatrix} e_1 \\ e_2 \\ \cdot \\ e_N \end{bmatrix}. \tag{A.1}$$

Since $\rho(\mathbf{E})$ is an error norm, it follows that $\rho(\mathbf{E}) = \sum_j \rho(e_j)$ and the derivative of $\rho(\mathbf{E})$ with respect to $\mathbf{X}$ is

$$\nabla\rho(\mathbf{E}) = \begin{bmatrix} \dfrac{\partial\rho(\mathbf{E})}{\partial x_1} & \dfrac{\partial\rho(\mathbf{E})}{\partial x_2} & \cdots & \dfrac{\partial\rho(\mathbf{E})}{\partial x_M} \end{bmatrix}^T. \tag{A.2}$$

Let $\psi(e) = \partial\rho(e)/\partial e$, applying chain rule and using (A.1), (A.2) can be written as

$$\begin{aligned} \nabla\rho(\mathbf{E}) &= \begin{bmatrix} \sum_{j=1}^{N} a_{j1}\psi(e_j) & \cdots & \sum_{j=1}^{N} a_{jM}\psi(e_j) \end{bmatrix}^T \\ &= \mathbf{A}^T \begin{bmatrix} \psi(e_1) & \cdots & \psi(e_N) \end{bmatrix}^T \\ &\Longrightarrow \nabla\rho(\mathbf{E}) = \mathbf{A}^T\psi(\mathbf{E}) = (DHF)^T\psi. \end{aligned} \tag{A.3}$$

Substituting by (A.3) in (17), we get

$$\begin{aligned} \mathbf{X}_i^{n+1} &= \mathbf{X}_i^n - \eta \sum_{k=1}^{L} \nabla\rho(\mathbf{E}_{i,k}^n; \tau_{i,k}) \\ &= \mathbf{X}_i^n - \eta \sum_{k=1}^{L} (DHF_{i,k})^T \boldsymbol{\psi}_{i,k}^n, \quad n = 0, 1, 2, \ldots, \ i = R, G, B. \end{aligned} \tag{A.4}$$

ACKNOWLEDGMENT

This research was supported in part by the US ARL Grant no. W911NF-06-2-0035.

REFERENCES

[1] P. J. Huber, *Robust Statistics*, Wiley Series in Probability and Statistics, John Wiley & Sons, New York, NY, USA, 2003.

[2] F. R. Hampel, E. M. Ronchetti, P. J. Rousseeuw, and W. A. Stahel, *Robust Statistics: The Approach Based on Influence Functions*, Wiley Series in Probability and Statistics, John Wiley & Sons, New York, NY, USA, 2005.

[3] R. A. Maronna, D. R. Martin, and V. J. Yohai, *Robust Statistics: Theory and Methods*, Wiley Series in Probability and Statistics, John Wiley & Sons, New York, NY, USA, 2006.

[4] N. Sebe and M. S. Lew, *Robust Computer Vision: Theory and Applications*, Springer, Berlin, Germany, 2003.

[5] M. J. Black and P. Anandan, "The robust estimation of multiple motions: parametric and piecewise-smooth flow fields," *Computer Vision and Image Understanding*, vol. 63, no. 1, pp. 75–104, 1996.

[6] T. Rabie, "Robust estimation approach for blind denoising," *IEEE Transactions on Image Processing*, vol. 14, no. 11, pp. 1755–1765, 2005.

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

[8] P. Meer, D. Mintz, A. Rosenfeld, and D. Y. Kim, "Robust regression methods for computer vision: a review," *International Journal of Computer Vision*, vol. 6, no. 1, pp. 59–70, 1991.

[9] D. Capel, *Image Mosaicing and Superresolution*, Springer, Berlin, Germany, 2004.

[10] N. A. El-Yamany, P. E. Papamichalis, and W. R. Schucany, "A robust image superresolution scheme based on redescending M-estimators and information-theoretic divergence," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '07)*, vol. 1, pp. 741–744, Honolulu, Hawaii, USA, April 2007.

[11] V. Patanavijit and S. Jitapunkul, "A Lorentzian stochastic estimation for a robust iterative multiframe superresolution reconstruction with Lorentzian-Tikhonov regularization," *EURASIP Journal on Advances in Signal Processing*, vol. 2007, Article ID 34821, 21 pages, 2007.

[12] 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," in *Proceedings of IEEE International Conference on Image Processing (ICIP '07)*, vol. 5, pp. 393–396, San Antonio, Tex, USA, September-October 2007.

[13] 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 of *Proceedings of SPIE*, pp. 1–12, San Jose, Calif, USA, January 2008.

[14] S. C. Park, M. K. Park, and M. G. Kang, "Superresolution image reconstruction: a technical overview," *IEEE Signal Processing Magazine*, vol. 20, no. 3, pp. 21–36, 2003.

[15] M. Elad and A. Feuer, "Restoration of a single superresolution image from several blurred, noisy, and undersampled measured images," *IEEE Transactions on Image Processing*, vol. 6, no. 12, pp. 1646–1658, 1997.

[16] M. Elad and Y. Hel-Or, "A fast superresolution reconstruction algorithm for pure translational motion and common space-invariant blur," *IEEE Transactions on Image Processing*, vol. 10, no. 8, pp. 1187–1193, 2001.

[17] A. Zomet and S. Peleg, "Efficient superresolution and applications to mosaics," in *Proceedings of the 15th International Conference on Pattern Recognition (ICPR '00)*, vol. 1, pp. 579–583, Barcelona, Spain, September 2000.

[18] A. Zomet, A. Rav-Acha, and S. Peleg, "Robust superresolution," in *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '01)*, vol. 1, pp. 645–650, Kauai, Hawaii, USA, December 2001.

[19] N. Nguyen, P. Milanfar, and G. Golub, "A computationally efficient superresolution image reconstruction algorithm," *IEEE Transactions on Image Processing*, vol. 10, no. 4, pp. 573–583, 2001.

[20] S. Farsiu, D. Robinson, M. Elad, and P. Milanfar, "Robust shift and add approach to superresolution," in *Applications of Digital Image Processing XXVI*, vol. 5203 of *Proceedings of SPIE*, pp. 121–130, San Diego, Calif, USA, August 2003.

[21] S. Farsiu, M. D. Robinson, M. Elad, and P. Milanfar, "Fast and robust multiframe superresolution," *IEEE Transactions on Image Processing*, vol. 13, no. 10, pp. 1327–1344, 2004.

[22] S. Farsiu, D. Robinson, M. Elad, and P. Milanfar, "Advances and challenges in superresolution," *International Journal of Imaging Systems and Technology*, vol. 14, no. 2, pp. 47–57, 2004.

[23] E. S. Lee and M. G. Kang, "Regularized adaptive high-resolution image reconstruction considering inaccurate subpixel registration," *IEEE Transactions on Image Processing*, vol. 12, no. 7, pp. 826–837, 2003.

[24] M. C. W. Zibetti and J. Mayer, "Outlier robust and edge-preserving simultaneous superresolution," in *Proceedings of IEEE International Conference on Image Processing*, pp. 1741–1744, Atlanta, Ga, USA, October 2006.

[25] M. Trimeche, R. C. Bilcu, and J. Yrjänäinen, "Adaptive outlier rejection in image superresolution," *EURASIP Journal on Applied Signal Processing*, vol. 2006, Article ID 38052, 12 pages, 2006.

[26] W.-Y. Zhao and H. S. Sawhney, "Is superresolution with optical flow feasible?" in *Proceedings of the 7th European Conference on Computer Vision-Part I (ECCV '02)*, vol. 2350 of *Lecture Notes in Computer Science*, pp. 599–613, Copenhagen, Denmark, May 2002.

[27] Z. A. Ivanovski, L. Panovski, and L. J. Karam, "Robust superresolution based on pixel-level selectivity," in *Visual Communications and Image Processing*, vol. 6077 of *Proceedings of SPIE*, pp. 1–8, San Jose, Calif, USA, January 2006.

[28] C. A. Segall, A. K. Katsaggelos, R. Molina, and J. Mateos, "Superresolution from compressed video," in *Superresolution Imaging*, S. Chaudhuri, Ed., pp. 211–242, chapter 9, Kluwer Academic Publishers, Dordrecht, The Netherlands, 2001.

[29] S. Farsiu, M. Elad, and P. Milanfar, "Multiframe demosaicing and superresolution of color images," *IEEE Transactions on Image Processing*, vol. 15, no. 1, pp. 141–159, 2006.

[30] S. Farsiu, M. Elad, and P. Milanfar, "Video-to-video dynamic superresolution for grayscale and color sequences," *EURASIP Journal on Applied Signal Processing*, vol. 2006, Article ID 61859, 15 pages, 2006.

[31] B. C. Tom and A. K. Katsaggelos, "Resolution enhancement of monochrome and color video using motion compensation," *IEEE Transactions on Image Processing*, vol. 10, no. 2, pp. 278–287, 2001.

[32] D. P. Bertsekas, *Nonlinear Programming*, Athena Scientific, Belmont, Mass, USA, 1999.

[33] A. Blake and A. Zisserman, *Visual Reconstruction*, MIT Press, Cambridge, Mass, USA, 1987.

[34] N. R. Shah and A. Zakhor, "Resolution enhancement of color video sequences," *IEEE Transactions on Image Processing*, vol. 8, no. 6, pp. 879–885, 1999.

[35] J. R. Bergen, P. Anandan, K. J. Hanna, and R. Hingorani, "Hierarchical model-based motion estimation," in *Proceedings of the 2nd European Conference on Computer Vision (ECCV '92)*, vol. 588 of *Lecture Notes in Computer Science*, pp. 237–252, Santa Margherita Ligure, Italy, May 1992.