AN IMAGE FUSION APPROACH FOR DENOISING SIGNAL-DEPENDENT NOISE

Mrityunjay Kumar, Rodney L. Miller

Research Laboratories
Eastman Kodak Company
Rochester, NY 14650-2118
Email: {mrityunjay.kumar, rodney.miller}@kodak.com

ABSTRACT
In this paper an image fusion approach is proposed for denoising digital images corrupted with signal-independent and signal-dependent noise. In the proposed approach, multiple captures of the same scene of interest are acquired and fused to estimate the original, noise-free image. This approach is motivated by the fact that noise is random in nature; hence, its interaction with the pixels will change with each capture, which in turn can be exploited for denoising purposes. In order to fuse multiple captures, a local affine model is developed to relate these captures and the corresponding original image. Furthermore, total variation (TV) regularization, which preserves discontinuity and is robust to noise, is used to solve the local affine fusion model iteratively to estimate the original image. While the proposed approach requires multiple captures, it is still computationally very fast and the quality of the denoised images clearly indicates the feasibility of the proposed approach.

Index Terms— Signal-dependent denoising, image fusion, total variation regularization, local affine model, multiple captures.

1. INTRODUCTION
The images captured by digital cameras are corrupted by device-specific noise and denoising algorithms are typically used to denoise and improve the signal-to-noise ratio of the captured images. In [1, 2], it has been shown that device-specific noise for CCD/CMOS-based digital cameras contain signal-independent as well signal-dependent components and can be modeled as
\[ I(x, y) = I_o(x, y) + (K_0 + K_1 I_o(x, y)) \eta(x, y) \]  
(1)
where \( I(x, y) \) and \( I_o(x, y) \) are observed noisy and original images, respectively. Furthermore, \( K_0 \) and \( K_1 \) are constants and \( \eta \sim N(0, 1) \). Traditional denoising algorithms assume only signal-independent noise (\( K_1 = 0 \)) [3] and are inadequate to deal with signal-dependent noise of the form in (1). In [2], Hirakawa et al. proposed a total least-square (TLS) based approach for estimating \( I_o(x, y) \) from (1). This approach seems promising but it is computationally expensive and requires precise knowledge of \( K_0 \) and \( K_1 \).

In this paper, multiple captures of the same scene of interest are acquired and fused to estimate the original image. In order to fuse multiple captures, a local affine model is developed to relate these captures and the corresponding original image. Furthermore, total variation (TV) regularization, which preserves discontinuity and is robust to noise [4], is used to solve the local affine fusion model iteratively to estimate the original image. The proposed approach is computationally fast and does not require knowledge of \( K_0 \) and \( K_1 \).

This paper is organized as follows. In Section 2, the local affine fusion model for denoising is described. Section 3 discusses the proposed denoising approach. Section 4 presents the results and is followed by concluding remarks.

2. LOCAL AFFINE FUSION MODEL
Let \( I_1(x, y), I_2(x, y), \ldots, I_n(x, y) \) be the \( n \) captures of the same scene and \( I_o(x, y) \) is the corresponding original, noise-free image for \( 1 \leq x \leq P, 1 \leq y \leq Q \). Furthermore, \( I_k(x, y) \) (\( 1 \leq k \leq n \)) are assumed to be co-registered; refer to [5] for a survey of image registration techniques. From (1),
\[ I_k = I_o + (K_0 + K_1 I_o) \eta_k \]  
(2)
where \( \eta_k \sim N(0, 1) \) is the random noise of the \( k^{th} \) capture and reference to the pixel location \((x, y)\) has been dropped for notational simplicity.

From (2),
\[ I_k = (1 + K_1 \eta_k) I_o + K_0 \eta_k \]  
⇒ \[ I_k = \beta_k I_o + K_0 \eta_k; \quad \beta_k = (1 + K_1 \eta_k) \]  
(3)
\[ \begin{pmatrix} I_1 \\ \vdots \\ I_n \end{pmatrix} = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} I_o + \begin{pmatrix} K_0 \eta_1 \\ \vdots \\ K_0 \eta_n \end{pmatrix} \]  
⇒ \[ I = \beta I_o + \eta \]  
(4)
where \( I = [I_1, \ldots, I_n]^t \), \( \beta = [\beta_1, \ldots, \beta_n]^t \), and \( \eta = [K_0 \eta_1, \ldots, K_0 \eta_n]^t \) and “\( t \)” represents transpose operation.
3. PROPOSED APPROACH

The goal of denoising is to estimate \( I_o \forall (x, y) \) from (5). However, \( \beta \) is also unknown and must be estimated.

\[
\beta_k = (1 + K_1 \eta_k) \text{ and } \eta_k \sim N(0, 1)
\]

\[
E[\beta_k] = 1
\]

where \( E[\bullet] \) is the expectation operator.

The proposed mathematical formulation to estimate \( I_o \) and \( \beta \) from (5) is stated as follows:

\[
\min_{I_o, \beta} g(I_o, \beta) = \min_{I_o, \beta} \left[ \frac{\lambda}{2} ||\beta I_o - I||_2^2 + \alpha \int_\Omega |\nabla I_o| dxdy + \frac{\gamma}{2} ||\beta||_2^2 \right]
\]

and satisfying constraint \( E[\beta_k] = 1 \). \( ||\beta I_o - I||_2 \) is the \( L_2 \)-norm of the residual term \( (\beta I_o - I) \) and represents the accuracy of the estimation of \( I_o \). The expression \( \int_\Omega |\nabla I_o| dxdy \) is the total variation regularization term, where \( |\nabla I_o| = \sqrt{T_{x o}^2 + T_{y o}^2} \). \( I_{x o} \) and \( I_{y o} \) are the first derivatives of \( I_o \) in the \( x \) and \( y \) directions, respectively. Similarly, \( ||\beta||_2 \) is \( L_2 \)-norm and regularizes the estimation of \( \beta \). The parameters \( \lambda > 0 \), \( \alpha > 0 \), and \( \gamma > 0 \) control the trade-off between the residual term and the regularization terms for the denoised image \( I_o \) and \( \beta \). A numerical scheme to solve (6) is explained next.

3.1. Numerical Scheme

\( g(I_o, \beta) \) is not a jointly convex function in general, however, for a given \( I_o \), \( g(I_o, \cdot) \) is convex with respect to \( \beta \). Similarly, for given \( \beta \), \( g(\cdot, \beta) \) is convex with respect to \( I_o \). Therefore, an alternating minimization (AM) approach is used to estimate \( I_o \) and \( \beta \) from (6). The first-order optimality conditions required to design an AM algorithm in this case are given as follows:

\[
\frac{\partial g(I_o, \beta)}{\partial I_o} = \lambda \beta^t (\beta I_o - I) - \alpha \nabla \bullet \left( \frac{\nabla I_o}{|\nabla I_o|} \right) = 0
\]

and satisfy the boundary condition \( \frac{\partial I_o}{\partial \xi} \bigg|_{\partial \Omega} = 0 \), where “\( \bullet \)”

represents the dot product, \( \partial \Omega \) represents the boundary of \( \Omega \), and \( \xi \) is the outward normal along \( \partial \Omega \).

Similarly,

\[
\frac{\partial g(I_o, \beta)}{\partial \beta} = \lambda I_o (\beta I_o - I) + \gamma \beta = 0
\]

\[
\Rightarrow \beta \bigg|_{\partial \Omega} = \frac{I_o I}{I_o^2 + \frac{\gamma}{\lambda}}.
\]  

The proposed AM algorithm can be stated as follows:

- Start with initial guess \( I_o^0 \)
- for iteration \( i = 0 : N \)
  - Step I: Estimate \( \beta^i \forall (x, y) \) using (8)
    \[
    \beta^i = \frac{I_o^i I}{(I_o^i)^2 + \frac{\gamma}{\lambda}}
    \]
    Impose boundary condition: \( \frac{\partial \beta^i}{\partial \xi} \bigg|_{\partial \Omega} = 0 \)
  - Step II: Solve \( I_o^{i+1} \forall (x, y) \) by applying gradient descent method to (7)
    \[
    I_o^{i+1} = I_o^i + \tau \left[ \lambda (\beta^i)^t (I - \beta^i I_o^i) + \alpha \nabla \bullet \left( \frac{\nabla I_o^i}{|\nabla I_o^i|} \right) \right]
    \]

4. RESULTS

The numerical results to illustrate the feasibility of the proposed approach is presented in this section. The first set of results for \((K_o, K_1) = (25, 0.1)\) is shown in Fig. 1. To simulate the noisy captures, four captures were generated by applying (2) to the original image four times. Due to the page constraints, only one such capture has been shown here in Fig. 1(b). Fig. 1(c)-1(f) show the results of the proposed approach applied to one, two, three, and all four noisy captures, respectively. The average of the noisy captures included in each experiment was used as the initial guess \( I_o^0 \). While images shown in Fig. 1(e) and 1(f) look almost similar, they are more visually pleasing than those shown in Fig. 1(c) and 1(d). Clearly, this indicates the advantage of using multiple captures for denoising. The average of all four captures is shown in Fig. 1(g). Averaging does reduce the noise, however, the proposed approach provides better results. In fact, the proposed approach with two captures seems to outperform averaging. The values of the regularization parameters \((\lambda, \alpha, \tau)\) and the number of iterations \((N)\) for Lena and Barbara images have been summarized in Table 1. Ratio \( \frac{\tau}{\lambda} \) was always set to 0.1. Note that these parameters did not change with the number of captures used in the proposed approach.

A comparison of the proposed approach with the TLS denoising algorithm published in [2] is shown in Fig. 2. The proposed approach seems to perform better than TLS denoising as the latter introduced visually unpleasing distortion.
unlike the proposed approach, TLS denoising requires exact knowledge of $K_0$ and $K_1$. Although TLS denoising requires only a single capture of the noisy scene, it is much more computationally expensive than the proposed approach. On a 2.66 GHz Intel Machine, with 3.0 GB RAM, for a $512 \times 512$ size Lena image, the proposed approach with three captures took 6.7 sec whereas TLS denoising took 38 min for denoising.

To estimate the quality of the denoised image quantitatively, the structural similarity index (SSIM) proposed in [6] was used. This index measures the similarity between two images and performs better than MSE or PSNR. The range of this index is $[-1, 1]$ and index value 1 indicates that two images are exactly the same. A summary of SSIM for $(K_0, K_1) = (25, 0.01), (25, 0.1), \text{and} (25, 0.2)$, respectively, is presented in Table 2. The original image shown in Fig. 1(a) was used as the reference image to compute SSIM. Note that the SSIM for all four noisy captures was the same. The index value for TLS denoising is slightly higher as compared to the proposed approach, however, the proposed approach performs better visually as supported by Fig. 2.

5. CONCLUSIONS AND FUTURE WORK

An image fusion framework for signal-dependent denoising is presented in this paper. The results on several images indicate the feasibility of the proposed approach. The estimation of step size ($\tau$) and regularization parameters ($\lambda, \alpha, \gamma$) adaptively will be investigated as the part of the future work. In addition, more robust algorithms to estimate the quality of the denoised image will also be investigated.

6. REFERENCES


Fig. 2. Comparison, $(K_0, K_1) = (25, 0.1)$ (a) Proposed approach (3 captures) (b) TLS approach.

Table 1. Summary of regularization parameters and step size.

<table>
<thead>
<tr>
<th>Images</th>
<th>$(K_0, K_1)$</th>
<th>No. Captures</th>
<th>$(\lambda, \alpha, \tau, N)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lena</td>
<td>(25, 0.01)</td>
<td>1/2/3/4</td>
<td>(0.4, 2.0, 0.5, 20)</td>
</tr>
<tr>
<td></td>
<td>(25, 0.1)</td>
<td>1/2/3/4</td>
<td>(0.2, 2.5, 0.5, 25)</td>
</tr>
<tr>
<td></td>
<td>(25, 0.2)</td>
<td>1/2/3/4</td>
<td>(0.1, 2.8, 0.5, 30)</td>
</tr>
<tr>
<td>Barbara</td>
<td>(25, 0.01)</td>
<td>1/2/3/4</td>
<td>(0.4, 1.6, 0.5, 20)</td>
</tr>
<tr>
<td></td>
<td>(25, 0.1)</td>
<td>1/2/3/4</td>
<td>(0.15, 2.2, 0.5, 25)</td>
</tr>
<tr>
<td></td>
<td>(25, 0.2)</td>
<td>1/2/3/4</td>
<td>(0.1, 3.0, 0.5, 30)</td>
</tr>
</tbody>
</table>
Table 2. Performance summary of SSIM.

<table>
<thead>
<tr>
<th>Images</th>
<th>((K_0, K_1))</th>
<th>Noisy</th>
<th>Proposed (1 capture)</th>
<th>Proposed (2 captures)</th>
<th>Proposed (3 captures)</th>
<th>Proposed (4 captures)</th>
<th>Average (4 captures)</th>
<th>TLS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lena</td>
<td>(25, 0.01)</td>
<td>0.2597</td>
<td>0.5062</td>
<td>0.7052</td>
<td>0.7774</td>
<td>0.8046</td>
<td>0.5042</td>
<td>0.8507</td>
</tr>
<tr>
<td></td>
<td>(25, 0.1)</td>
<td>0.1786</td>
<td>0.4138</td>
<td>0.6166</td>
<td>0.7074</td>
<td>0.7454</td>
<td>0.3745</td>
<td>0.8235</td>
</tr>
<tr>
<td></td>
<td>(25, 0.2)</td>
<td>0.1305</td>
<td>0.3503</td>
<td>0.5324</td>
<td>0.6365</td>
<td>0.6875</td>
<td>0.2888</td>
<td>0.7974</td>
</tr>
<tr>
<td>Barbara</td>
<td>(25, 0.01)</td>
<td>0.3892</td>
<td>0.5292</td>
<td>0.6728</td>
<td>0.7434</td>
<td>0.7759</td>
<td>0.6259</td>
<td>0.8576</td>
</tr>
<tr>
<td></td>
<td>(25, 0.1)</td>
<td>0.2874</td>
<td>0.4597</td>
<td>0.5959</td>
<td>0.6557</td>
<td>0.6812</td>
<td>0.5130</td>
<td>0.8087</td>
</tr>
<tr>
<td></td>
<td>(25, 0.2)</td>
<td>0.2168</td>
<td>0.4234</td>
<td>0.5325</td>
<td>0.5822</td>
<td>0.6027</td>
<td>0.4237</td>
<td>0.7610</td>
</tr>
</tbody>
</table>

Fig. 1. Denoising results \( (K_0, K_1) = (25, 0.1) \) (a) Original Image (b) Noisy capture (c) proposed approach (1 capture) (d) proposed approach (2 captures) (e) proposed approach (3 captures) (f) proposed approach (4 captures) (g) average (all four captures).