Efficient image restoration for active mosaic imaging

Nicolas Lermé, François Malgouyres, Dominique Hamoir, Emmanuelle Thouin

IMPROVED IMAGE RESTORATION FOR MOSAIC ACTIVE IMAGING

Nicolas Lermé[†], François Malgouyres[‡], Dominique Hamoir[⋆], Emmanuelle Thouin[⋆,•]

[†] Institut Mines-Télécom, Télécom ParisTech, CNRS LTCI, Paris, France
[‡] IMT-UMR5219, Université de Toulouse, CNRS, Toulouse, France
[⋆] ONERA, F-31055 Toulouse, France
[•] Université de Toulouse, ISAE, F-31055 Toulouse, France

ABSTRACT

In this paper, we focus on the restoration of images in mosaic active imaging. This emerging imaging technique consists in acquiring a mosaic of images (laser shots) by focusing a laser beam on a small portion of the target object and subsequently moving it to scan the whole field of view. In order to restore the whole image from such a mosaic, a prior work proposed a simplified forward model assuming a prior on the distribution of images and the acquisition parameters as well as an algorithm alternating between the estimation of the restored image and the estimation of these parameters. The novelty of this paper is twofold: (i) we show that the results obtained from this earlier work can be improved when the laser shots are closer to each others and (ii) that faster convergence can be achieved for estimating the acquisition parameters.

***Index Terms*—** active imaging, laser imaging, image reconstruction, image estimation, BFGS, graph-cut.

1. INTRODUCTION

In flash laser imaging, the target object is illuminated with a very short laser flash. A time-gated camera synchronized with the laser is used to detect and select the photons received within a brief time-gate (typically few nano to micro seconds), after a chosen delay (typically 10^{-4} to 10^{-7} seconds) has elapsed. This temporal selection allows to eliminate photons back-scattered by the foreground and the background. Generally, the field of view of the camera is fully illuminated by a Nd:YAG laser and acquired at about 10 Hz. In mosaic active imaging, a 10 kHz fiber laser is used instead. As the repetition-rate is larger by three orders of magnitude, the energy per pulse is lowered by the same ratio. To maintain the same signal-to-noise ratio, only a reduced part of the field of view is illuminated at each laser flash. This results in the successive acquisition of typically 100 to 1000 elementary images [1] (laser shots) on a square tiling, subject to multiple degradations [2, 3] (see Figure 1). The object of interest typically have metric dimensions and lies between 10 m and 20 km from the imaging system. The applications can be either terrestrial or airborne and typically concern surveillance.

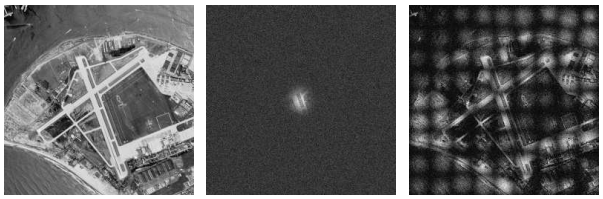

Fig. 1: Acquisition process in mosaic laser imaging. From left to right: the ideal image we want to estimate, a reduced part of the field of view with a single laser flash (laser shot), the image composed of all laser shots where each pixel is assigned its maximum intensity over all of them.

Restore the observed scene from the laser shots is a difficult inverse problem to solve. To our best knowledge, [4] is the first attempt to give a solution to this problem. They model laser shots as isotropic Gaussians and assume a Total Variation (TV) prior on the distribution of images and a Gaussian prior on the acquisition parameters. A two-stages iterative algorithm is proposed, alternating between (i) the estimation of the restored image using graph cuts and (ii) estimation of these parameters using a standard gradient descent.

The novelty of this paper is twofold: (i) we show that the results obtained from [4] can be improved when the laser shots are closer to each others and (ii) that faster convergence can be achieved for estimating the acquisition parameters.

The rest of this paper is organized as follows. We first remind in Section 2, the simplified model of the imaging process and the restoration algorithm of [4]. In Section 3, we evaluate the impact of using a different tiling of the laser shots on the image estimate. Next, we compare the convergence of several differentiable methods for estimating the acquisition parameters. Finally, we compare the results obtained with [4].

2. MATHEMATICAL MODELING

In this section, we summarize the imaging process and the restoration algorithm detailed in [4]. For an integer $N > 0$, we denote the set of all pixels by $\mathcal{P} = \{1, \dots, N\}^2$ and the

number of elementary images by $K > 0$. For every index $k \in \{1, \dots, K\}$, we denote by $\theta_k = (c_k, w_k) \in J$ with $J = (\mathbb{R}^2 \times \mathbb{R}_+^*)$ the parameters (size and location) of the light spot (or illumination dome) in each of these images. For every $p \in \mathcal{P}$, a dome G_{θ_k} is modeled as an isotropic Gaussian

$$G_{\theta_k}(p) = \exp\Big(- \frac{\|p - c_k\|^2}{2{w_k}^2} \Big), \quad \text{for } 1 \leq k \leq K, \qquad (1)$$

and where $\|.\|$ denotes the Euclidean norm. Let us denote by $v = (v^k)_{1 \leq k \leq K}$ with $v^k \in \mathbb{R}^{\mathcal{P}}$ the observed data (laser shots) and $u \in \mathbb{R}^{\mathcal{P}}$ the ideal image (i.e. the one that would have been obtained with an ideal captor and illumination). The proposed simplified forward model is

$$v = M_{(\theta_k)_{1 \leq k \leq K}} u + n,$$

where $n = (n^k)_{1 \leq k \leq K}$ with $n^k \in \mathbb{R}^{\mathcal{P}}$ is an additive white Gaussian noise of standard deviation σ, and

$$\begin{array}{rccl} M_{(\theta_k)_{1 \leq k \leq K}} : & \mathbb{R}^{\mathcal{P}} & \longrightarrow & \mathbb{R}^{K\mathcal{P}}, \\ & u & \longmapsto & \big((G_{\theta_k}(p) u_p)_{p \in \mathcal{P}}\big)_{1 \leq k \leq K}. \end{array} \qquad (2)$$

Once the acquisition parameters $(\theta_k)_{1 \leq k \leq K}$ are fixed, this model is linear. Due to some perturbations, these parameters need however to be estimated. We consider that the parameters c_k and w_k are independent random variables and assume that the former follow a Gaussian law of mean $\overline{c}_k$ and standard deviation σ_c while the latter follow a Gaussian law of mean $\bar{w}$ and standard deviation σ_c, $\forall k \in \{1, \dots, K\}$. We also assume a TV prior on u and that it is independent of the parameters $(\theta_k)_{1 \leq k \leq K}$. Based on these assumptions, the Maximum A Posteriori (MAP) of u and $(\theta_k)_{1 \leq k \leq K}$ can be calculated. Given a fixed $v \in \mathbb{R}^{K\mathcal{P}}$ and for any $u \in \mathbb{R}^{\mathcal{P}}$ and $(\theta_k)_{1 \leq k \leq K} \in J$, we denote the minimized function by

$$\begin{aligned} F(u, (\theta_k)_{1 \leq k \leq K}) = & \frac{\|M_{(\theta_k)_{1 \leq k \leq K}} u - v\|^2}{2\sigma^2} + \beta TV(u) \\ & + \sum_{k=1}^{K} \frac{\|c_k - \overline{c}_k\|^2}{2\sigma_c^2} + \sum_{k=1}^{K} \frac{|w_k - \overline{w}|^2}{2\sigma_w^2}, \end{aligned} \qquad (3)$$

where β, σ, σ_c, σ_w, $\overline{w}$ and $\overline{c}_k$ are known parameters and $TV(u)$ denotes the TV of u. Since F is non-convex, only a local minimizer of F can be computed in a reasonable amount of time. A two-stages iterative process is designed in [4] (see Algorithm 1) alternating between (i) the estimation of the restored image (line 3) and (ii) the estimation of the acquisition parameters (line 4), until some accuracy ε_a is reached. The former step is solved using graph cuts while the latter step is solved using a standard gradient descent with Armijo step size rule [1] until some accuracy

$$\varepsilon_e = (\varepsilon_e^{max} - \varepsilon_e^{min}) \exp\Big(- \frac{n}{\sigma_{\varepsilon_e}} \Big) + \varepsilon_e^{min}, \qquad (4)$$

varying between ε_e^{min} and ε_e^{max}, is reached. Thus, the estimation of acquisition parameters is expected to be progressively more accurate as the Algorithm 1 iterates. It is not difficult to see that such an algorithm computes a local minimizer of F. However, when σ_c and σ_w are small enough, we expect the global minimizer to be close to $(\overline{c}_k, \overline{w})_{1 \leq k \leq K}$ and so a good property when using this initialization.

Algorithm 1 Algorithm for approximating a minimizer of F.

1. Initialize $(\theta_k^0)_{1 \leq k \leq K} = (\overline{c}_k, \overline{w})_{1 \leq k \leq K}$
2. **while** $\|u^n - u^{n-1}\| \leq \varepsilon_a$ **do**
3. $\quad u^n \in \operatorname{argmin}_{u \in \mathbb{R}^{\mathcal{P}}} F(u, (\theta_k^n)_{1 \leq k \leq K})$.
4. $\quad (\theta_k^{n+1})_{1 \leq k \leq K} \in \operatorname{argmin}_{(\theta_k)_{1 \leq k \leq K} \in J} F(u^n, (\theta_k)_{1 \leq k \leq K})$.
5. **endwhile**

3. EXPERIMENTAL RESULTS

3.1. Applicative framework and implementation details

The experiments of the subsequent sections are conducted on simulated data with images of size 256×256 (i.e.$N = 256$). Realistic values are however used (expressed in pixels) for the parameters σ_c, σ_w, $\overline{w}$ and $\overline{c}_k$, $\forall k \in \{1, \dots, K\}$. In this setting, we set $\sigma_c = 0.81$, $\sigma_w = 0.07$ and $\overline{w} = 16.2$. The way of how the parameters $\overline{c}_k$ are fixed is discussed in Section 3.2.

For estimating the restored image, we use the max-flow implementation of [5] and the dyadic parametric scheme of [6]. For estimating the acquisition parameters, the accuracy ε_e (see (4) in Section 2) is set using $\varepsilon_e^{min} = 5 \times 10^{-3}$, $\varepsilon_e^{max} = 0.5$ and $\sigma_{\varepsilon_e} = 2.0$. For the accuracy of the Algorithm 1, we found that setting $\varepsilon_a = 1.0$ is a good tradeoff between convergence and accuracy [2]. Whenever it is possible, the regularizer β (see (3) in Section (3)) is set by minimizing the Mean Square Error (MSE) [3] between the image estimate and the ideal image using Golden Section Search [7]. Also, we do not provide detailed computing times since we believe they are not representative of an optimized version (the restoration of an image typically requires between 1 and 6 minutes on a dodeca core Intel Xeon 3.47 GHz). The estimation of the restored image typically represents 10% of the overall computation time. Notice that the pixel intensity in the ideal observed image ranges in $[0, 1]$ and is coded on 8 bits, meaning that the noise levels are scaled accordingly.

In what follows, we provide experiments for a single noise level $\sigma = 0.1$ using the above values of parameters, except if different ones are mentioned. In practice, the noise level σ is generally unknown. In such a situation, a reliable estimate of this parameter can be for instance obtained using [8].

[1] For lack of space, the partial derivatives of F are only available in [4].

[2] In particular, that corresponds to an error of one grayscale level for all pixels between two successive image estimates.

[3] A description of the MSE and PSNR measures is available at `http://megawave.cmla.ens-cachan.fr/stuff/guid3/node256.html#fmse`.

3.2. Impact of tilings on the image estimate

In this section, we evaluate how two different tilings of the illumination domes affect the image estimate computed by the Algorithm 1 (see Section 2). To be as fair as possible, we impose that the size and location of these domes essentially do not very by setting $\sigma_c = 10^{-5}$ and $\sigma_w = 10^{-5}$. Due to the particular values of these parameters and the level of accuracy ε_a, we have $w_k \simeq \bar{w}$ and $c_k \simeq \bar{c}_k$, $\forall k \in \{1, \ldots, K\}$. The penalty parameter and the noise level are resp. set with $\beta = 3 \times 10^4$ and $\sigma = 10^{-5}$. To accentuate the difference between the tilings, we also set $\overline{w} = 5$. Let us now describe these two tilings. For the first one, the domes form a square tiling of size $K = K' \times K'$ (see [4]). For any $i, j \in \{0, \ldots, K'-1\}^2$, the expected center $\bar{c}_{i \times K' + j}$ is defined as

$$\left(\frac{N}{K'} \left(\frac{1}{2} + i \right), \frac{N}{K'} \left(\frac{1}{2} + j \right) \right). \tag{5}$$

For the second one, the domes form an hexagonal tiling of size $K = K'^2 - \lfloor \frac{K'}{2} \rfloor$, vertically centered in the image. For any $i \in \{0, \ldots, K'-1\}$ and any $j \in \{0, \ldots, K' - 1 - mod(i,2)\}$, the expected center $\bar{c}_{i \times K' - \lfloor \frac{i}{2} \rfloor + j}$ is defined as

$$\left(\frac{1}{2} \left(N - (K'-1)d \right) + i \times d, \frac{N}{K'} \left(\frac{1}{2 - mod(i,2)} + j \right) \right), \tag{6}$$

where $d = \frac{N\sqrt{3}}{2K'}$ is the distance separating two successive rows of domes and $mod(x, y)$ is the remainder of x/y. Notice that for any fixed $K' > 1$, the number of domes of an hexagonal tiling is always smaller than for a square tiling.

The impact of the above tilings on the image estimate is shown in Figure 2. Each pixel in the left and middle columns is resp. assigned with its maximum intensity over all $k \in \{1, \ldots, K\}$. The image estimate is depicted in the right column. Compared to a square tiling, an hexagonal tiling allows to recover more details between domes while using a smaller number of domes. This is due to the fact that the domes are closer from each other in the hexagonal tiling. In the latter, less details are however generally recovered near borders since the number of domes can be smaller than in the square tiling for some rows. By discarding these borders (i.e. the pixels for which their Tchebychev distance to the borders is greater or equal than $\frac{N}{2K'}$), the benefit of using an hexagonal tiling against a square tiling is confirmed by a significantly larger Peak Signal-to-Ratio Noise (PSNR) and smaller MSE when comparing the image estimate to the ideal one.

3.3. Performance for estimating acquisition parameters

The goal of this section is to compare the performance of several differentiable methods for estimating the acquisition parameters using an hexagonal tiling of illumination domes (see (6) in Section 3.2). First, we remind that these parameters are estimated using a Standard Gradient Descent (SGD)

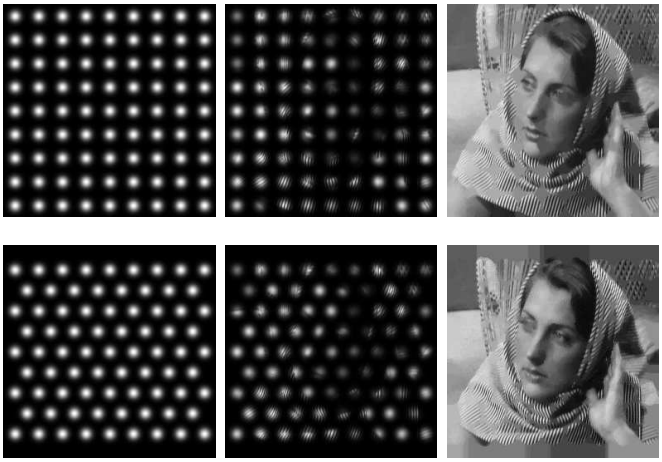

Fig. 2: Impact of a square tiling (top row, see (5)) against an hexagonal tiling (bottom row, see (6)) on the image estimate. Each pixel in the left and middle columns resp. is assigned with its maximum intensity over all $k \in \{1, \ldots, K\}$ in illumination domes and laser shots. The right column represents the image estimate. In this experiment, we set $\sigma_c = 10^{-5}$, $\sigma_w = 10^{-5}$, $\sigma = 10^{-5}$, $\overline{w} = 5$ and $\beta = 3 \times 10^4$.

with an Armijo step size rule in [4]. Let us denote by T the desired number of iterations in this rule. Since the computation of F and its gradient requires $O(KN^2)$ operations, the worst-case complexity of the SGD per iteration is $O(TKN^2)$. Under particular assumptions, it also has a global convergence rate of $O(1/t)$ (t is the number of iterations) and converges linearly when close enough to the local minimizer. Among first-order methods, the Nesterov's Accelerated Gradient Descent (AGD) algorithm [9] can however achieves a better global convergence with a rate of $O(1/t^2)$, while keeping the same complexity as SGD.

Since the functional F (see (3) in Section 2) is twice continuously differentiable and the number of variables is typically of only several hundred, second-order methods are computationally accessible. Under particular assumptions, the Newton's method can for instance locally converges much faster than SGD. However, it is (i) computationally demanding and (ii) can converge towards to a local maximum or a saddle point when the functional is not convex (and this is our case). Broyden-Fletcher-Goldfarb-Shanno (BFGS) overcomes these difficulties by computing an approximation of the inverse of the Hessian matrix and (ii) ensuring that this direction is a descent during the iterates. This method converges superlinearly when close enough to the local minimizer. Using an Armijo step size rule and if $K \ll N^2$, the complexity of BFGS remains the same as the SGD and the AGD.

Based on these observations, from a global point of view, AGD is probably a better candidate in our situation compared to SGD. From a local point of view, BFGS however appears to be a better choice compared to SGD. Without additional

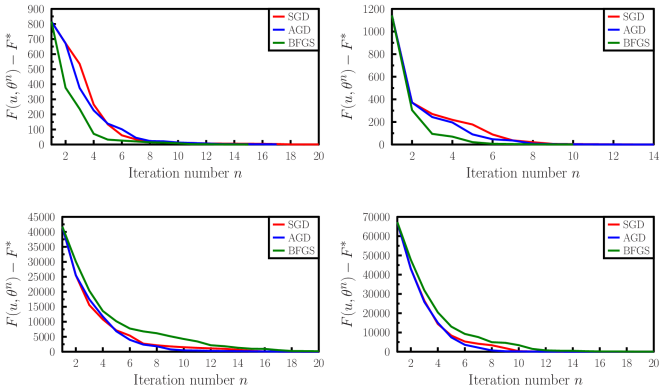

Fig. 3: Performance of the SGD, AGD and BFGS methods for estimating the acquisition parameters with an initialization not so far (top row, $\sigma_c = 0.81$, $\sigma_w = 0.81$) and not so close (bottom row, $\sigma_c = 4$ and $\sigma_w = 2$) from $(\overline{c}_k, \overline{w})_{1 \leq k \leq K}$ for any $k \in \{1, \ldots, K\}$ and for the images "baboon" (left column) and "factory" (right column).

information on F, it is difficult to guess if SGD, AGD or BFGS is the most efficient in our situation. We therefore propose to empirically compare them with two different initializations: for the first one, the parameters σ_c and σ_w are set as described in Section 3.1. For the second initialization, we set $\sigma_c = 4$ and $\sigma = 2$. The remaining parameters are set as described in Section 3.1. For each initialization, we then proceed as follows: first, we run these methods for a large number of iterations by setting $\varepsilon_e = 10^{-4}$. Next, we compute a reference energy F^* taken as the minimum among the energies reached by each of these methods. Finally, we run again these methods using the same parameters, except by setting $\varepsilon_e = 0.05$. Also, the image estimate u is taken as the ideal image. The Figure 3 shows the results of this comparison between these methods. When the true acquisition parameters are expected to be not so close from $(\overline{c}_k, \overline{w})_{1 \leq k \leq K}$ for any $k \in \{1, \ldots, K\}$, BFGS generally converges faster than AGD while the latter only converges slightly faster than SGD. As opposite, when the true acquisition parameters are expected to be not so close from $(\overline{c}_k, \overline{w})_{1 \leq k \leq K}$ for any $k \in \{1, \ldots, K\}$, all methods tend to behave equivalently.

3.4. Accuracy

In this section, we study the quality of the image estimate provided by the Algorithm 1 (see Section 2) for an hexagonal tiling (see Section 3.2) and using the BFGS algorithm (see Section 3.3). The remaining parameters are set as described in Section 3.1. Let us describe now the experimental setting. For each image, we independently generate 10 sequences of laser shots and illumination domes using the above parameters. Next, we apply the Algorithm 1 on each sequence, measure the PSNR between the image estimate and the ideal image, and compute statistics from all the measures. Also, we

Image	PSNR (dB)	Δ PSNR (dB)
baboon	23.89 ± 0	$+0.116$
barbara	25.21 ± 0	$+0.131$
peppers	28.16 ± 0	$+0.081$
cameraman	28.32 ± 0	$+0.441$
lena	27.95 ± 0	$+0.278$
man	26.07 ± 0	$+0.283$
boat	26.83 ± 0	$+0.289$
factory	26.99 ± 0	$+0.153$

Table 1: Accuracy of the restoration algorithm for a noise level $\sigma = 0.1$. The remaining parameters are set as follows: $\sigma_c = 0.81$, $\sigma_w = 0.07$ and $\overline{w} = 16.2$. The PSNR measures are calculated over 10 runs and rounded to the nearest value.

Fig. 4: Reconstruction of the images “barbara” (top row) and “factory” (bottom row) with a noise level $\sigma = 0.1$. The remaining parameters are set as $\overline{w} = 16.2$, $\sigma_c = 0.81$ and $\sigma_w = 0.07$. The left column correspond to the available data where each pixel is assigned its maximum intensity over all laser shots. The middle and right columns correspond resp. to the image estimate and the ideal image.

compare the obtained results with those reported in [4] by providing the difference of mean PSNR (denoted as Δ PSNR). If Δ PSNR is positive, the results obtained in this setting are considered to be better. The results of all these experiments are summarized in Table 1 and illustrated in Figure 4. In Figure 4, the selected images correspond to the sequence for which the MSE between the image estimate and the ideal image is minimum. Also, we provide for each image the ideal one, the image estimate as well as the partial laser shots. Let us now briefly comment these results. Despite a substantial noise level, we observe in Figure 4 that the Algorithm 1 behaves globally well. Large flat areas are well denoised while thin structures and details are well preserved, even between illumination domes where the knowledge about data is more incertain. Although the amplitude of the Δ PSNR measure is quite low, it is positive for all the presented images.

4. REFERENCES

[1] D. Hamoir, "Procédé et système d'imagerie active à champ large : Method and system for active imaging with a large field," Patent WO 2010119225, October 2010.

[2] L. Hespel, M.-T. Velluet, A. Bonnefois, N. Rivière, M. Fracès, D. Hamoir, B. Tanguy, B. Duchenne, and J. Isbert, "Comparison of a physics-based BIL simulator with experiments," in *Society of Photo-Optical Instrumentation Engineers, International Symposium on Photoelectronic Detection and Imaging*, 2009, pp. 73822T–73822T–10.

[3] N. Rivière, L. Hespel, M.-T. Velluet, Y.-M. Frédéric, P. Barillot, and F. Hélias, "odeling of an active burst illumination imaging system: comparison between experimental and modelled 3d scene," in *Society of Photo-Optical Instrumentation Engineers, International Symposium on Photoelectronic Detection and Imaging*, 2010, vol. 7382, pp. 783509–783509–11.

[4] N. Lermé, F. Malgouyres, D. Hamoir, and E. Thouin, "Bayesian image restoration for active mosaic imaging," *Inverse Problems and Imaging Journal*, November 2012.

[5] Y. Boykov and V. Kolmogorov, "An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision," *Pattern Analysis And Machine Intelligence*, vol. 26, no. 9, pp. 1124–1137, 2004.

[6] A. Chambolle and J. Darbon, "On total variation minimization and surface evolution using parametric maximum flows," *International Journal of Computer Vision*, vol. 84, no. 3, pp. 288–307, 2009.

[7] J. Kiefer, "Sequential minimax search for a maximum," *Proceedings of the American Mathematical Society*, vol. 4, no. 3, pp. 502–506, 1953.

[8] D. Donoho and J. Johnstone, "Ideal spatial adaption via wavelet shrinkage," *Biometrika*, vol. 81, no. 3, pp. 425–455, 1994.

[9] Y. Nesterov, "A method for solving a convex programming problem with convergence rate $o(1/k^2)$," *Soviet Mathematics Doklady*, vol. 27, pp. 372–376, 1983.