



HAL
open science

Numerical study of an optimization problem for mosaic active imaging

Nicolas Lermé, François Malgouyres, Emmanuelle Thouin, Dominique Hamoir

► **To cite this version:**

Nicolas Lermé, François Malgouyres, Emmanuelle Thouin, Dominique Hamoir. Numerical study of an optimization problem for mosaic active imaging. 2014. hal-00935725v2

HAL Id: hal-00935725

<https://hal.science/hal-00935725v2>

Preprint submitted on 14 Feb 2014 (v2), last revised 12 Jun 2014 (v3)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

NUMERICAL STUDY OF AN OPTIMIZATION PROBLEM FOR MOSAIC ACTIVE IMAGING

Nicolas Lermé[†], François Malgouyres[‡], Emmanuelle Thouin^{*•}, Dominique Hamoir^{*}

[†] Institut Mines-Télécom, Télécom ParisTech, CNRS LTCI, Paris, France

[‡] IMT CNRS UMR5219, Université de Toulouse, Toulouse, France

^{*} ONERA - The French Aerospace Lab, F-31055 Toulouse, France

[•] Université de Toulouse, ISAE, F-31055 Toulouse, France

ABSTRACT

In this paper, we focus on the restoration of an image 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. To restore the whole image from such a mosaic, a prior work proposed a simplified forward model describing the acquisition process. It also provides a prior on the acquisition parameters. Together with a prior on the distribution of images, this leads to a MAP estimate alternating between the estimation of the restored image and the estimation of these parameters. The novelty of the current paper is twofold: (i) We provide a numerical study and argue that faster convergence can be achieved for estimating the acquisition parameters (ii) we show that the results from this earlier work are improved when the laser shots are acquired according to a more compact pattern.

Index Terms— active imaging, laser imaging, image reconstruction, differentiable optimization, 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 (few nano to micro seconds), after a chosen delay (10^{-4} to 10^{-7} seconds) has elapsed. This temporal selection allows to eliminate photons back-scattered by the foreground (e.g. fog, dust or vegetation) 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 1–10 kHz fiber laser is used instead, expected to offer higher average powers and plug-efficiencies within a few years. 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 at 1–10 kHz (laser shots) [1] subject to multiple degradations [2, 3], that tile as a mosaic to



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 from all laser shots where each pixel is assigned with its maximum intensity over all of them.

build the full-frame image at 10 Hz (see Figure 1). The object of interest typically have metric dimensions (e.g. buildings, vehicles, personnel) and lies between 10 m and 20 km from the imaging system. The applications can be either terrestrial or airborne and concern surveillance/target identification.

Restoring the observed scene from the laser shots is a difficult inverse problem. To our best knowledge, [4] is the first attempt to give a solution to this problem. They model laser shots as isotropic Gaussians, assume a Gaussian prior on the acquisition parameters and a Total Variation (TV) prior on the distribution of images. This choice is motivated by the ability of the TV prior to properly restore images [5] and fast minimization algorithms [6, 7, 8]. In [4], a two-stage iterative algorithm is proposed, alternating between (i) the estimation of the restored image using graph cuts and (ii) the estimation of the acquisition parameters using a standard gradient descent.

The novelty of this paper is twofold: (i) we provide a numerical study for the estimation of the acquisition parameters and argue that faster convergence can be achieved and (ii) show that the results from [4] are improved when the laser shots are acquired according to a more compact pattern.

The rest of this paper is organized as follows. First, we remind in Section 2 the forward model of the imaging process and the restoration algorithm of [4]. In Section 3, we discuss advanced optimization methods for efficiently estimating acquisition parameters. We briefly describe two tilings of the laser shots in Section 4. Finally, we compare the accuracy and the convergence between all these elements in Section 5.

2. MATHEMATICAL MODELING

For an integer $N > 0$, we denote the set of all pixels by $\mathcal{P} = \{1, \dots, N\}^2$ and the number of laser shots by the integer $K > 0$. 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 of a Gaussian. For every $p \in \mathcal{P}$, a laser shot is modeled as an isotropic Gaussian (called illumination dome) defined by

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

where $\|\cdot\|$ 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 sensor 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 Gaussian white noise of standard deviation $\sigma > 0$, and

$$\begin{aligned} M_{(\theta_k)_{1 \leq k \leq K}} : \mathbb{R}^{\mathcal{P}} &\longrightarrow \mathbb{R}^{K\mathcal{P}}, \\ u &\longmapsto \left((G_{\theta_k}(p)u_p)_{p \in \mathcal{P}} \right)_{1 \leq k \leq K}. \end{aligned} \quad (1)$$

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

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

where $\beta, \sigma, \sigma_c, \sigma_w, \bar{w}$ and $(\bar{c}_k)_{1 \leq k \leq K}$ are known parameters and $TV(u)$ denotes the TV of u . Notice that for any $(\theta_k)_{1 \leq k \leq K} \in J^K$, the function $u \mapsto F(u, (\theta_k)_{1 \leq k \leq K})$ is convex. However, when $u \in \mathbb{R}^{\mathcal{P}}$ is fixed, the function $(\theta_k)_{1 \leq k \leq K} \mapsto F(u, (\theta_k)_{1 \leq k \leq K})$ is non-convex. The function F is non-convex and we cannot a priori provide guarantees that we compute a true minimizer of F . We have however not observed any convergence problem in practice. This is likely because the algorithm is well initialized. A two-stage

iterative process is designed in [4]. Its sketch is described in Algorithm 1: it alternates 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 step (i) is solved exactly (modulo a provided quantization step) using graph cuts while the step (ii) is solved approximately by using a standard gradient descent with an Armijo step size rule until some accuracy ε_e is reached. The parameter ε_e decays with the iteration number n in the Algorithm 1. In this way, the estimation of the acquisition parameters is progressively more accurate as n increases. For $v \in \mathbb{R}^{K\mathcal{P}}$, $u \in \mathbb{R}^{\mathcal{P}}$ and $(c_k, w_k)_{1 \leq k \leq K} \in J^K$, we obtain

$$\begin{aligned} \frac{\partial F}{\partial c_{k,i}} &= \frac{c_{k,i} - \bar{c}_{k,i}}{\sigma_c^2} + \frac{1}{\sigma^2 w_k^2} \sum_{p \in \mathcal{P}} (p_i - c_{k,i}) D_p^k, \\ \frac{\partial F}{\partial c_{k,j}} &= \frac{c_{k,j} - \bar{c}_{k,j}}{\sigma_c^2} + \frac{1}{\sigma^2 w_k^2} \sum_{p \in \mathcal{P}} (p_j - c_{k,j}) D_p^k, \\ \frac{\partial F}{\partial w_k} &= \frac{w_k - \bar{w}_k}{\sigma_w^2} + \frac{1}{\sigma^2 w_k^3} \sum_{p \in \mathcal{P}} \|p - c_k\|^2 D_p^k, \end{aligned} \quad (3)$$

with $D_p^k = E_p^k u_p [E_p^k u_p - v_p^k]$ and $E_p^k = e^{-\frac{\|p - c_k\|^2}{2w_k^2}}$. Notice that the above partial derivatives w.r.t. θ_k do not contain variables of $\theta_{k'}$ for $k' \neq k$. This implies that the Hessian matrix of F is block diagonal with blocks of 3×3 .

Algorithm 1 Algorithm for approximating a minimizer of F

INPUTS: $v \in \mathbb{R}^{K\mathcal{P}}, \beta, \sigma, \sigma_c, \sigma_w, \bar{w}$ and $(\bar{c}_k)_{1 \leq k \leq K}$

OUTPUTS: Image estimate u^*

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

3. ADVANCED OPTIMIZATION METHODS FOR ESTIMATING ACQUISITION PARAMETERS

In this section, we discuss possible choices of methods for efficiently estimating the acquisition parameters in the Algorithm 1 (see Section 2). First, we remind that these parameters are estimated using a Standard Gradient Descent (SGD) 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 (see (2) and (3) in Section 2) requires $O(KN^2)$ operations, the worst-case complexity of the SGD per iteration is $O(TKN^2)$ and requires a memory storage of $O(K)$. 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 achieve a

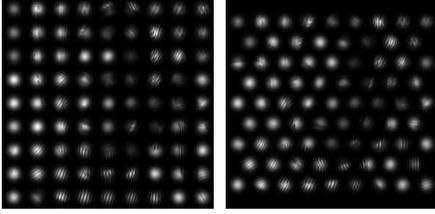


Fig. 2: From left to right: square and hexagonal tilings of laser shots, generated by setting $\bar{w} = 5$ and approximately null values for σ , σ_c and σ_w . Each pixel in these images is assigned with its maximum intensity over all $k \in \{1, \dots, K\}$.

better global convergence with a rate of $O(1/t^2)$, while keeping the same time complexity and memory storage as SGD.

Since F is twice continuously differentiable and has a block diagonal Hessian of reasonable size, second-order methods are also computationally accessible. Under particular assumptions, the Newton's method converges quadratically when close enough to the local minimizer. However, this method is known to be (i) computationally demanding and (ii) can converge to a local maximum or a saddle point when the functional is not convex (and this is our case). Quasi-Newton methods like Broyden-Fletcher-Goldfarb-Shanno (BFGS) overcomes the above issues of the Newton's method by (i) computing an approximation of the Hessian matrix (or its inverse) and (ii) ensuring that the functional decreases during the iterates when using Wolfe's step size conditions. Moreover, BFGS converges superlinearly when close enough to the local minimizer. However, it would not be numerically acceptable to enforce Wolfe's condition in our case since it involves computations of ∇F . Although we do not have theoretical convergence guarantee, we have therefore to use an Armijo condition. In this case, if $K \ll N^2$, the time complexity of BFGS remains the same as SGD while having a memory storage of $O(K^2)$.

Finally, we have empirically observed that the Hessian is dominated by its diagonal. In such a particular situation, it is reasonable to apply a modified Newton method in which we approximate the Hessian by canceling its diagonal elements. This strategy is named "Diagonal-Scaling" in [10] and we will denote it as Diagonally-Scaled Newton (DSN) in the rest of this paper. Moreover, to ensure that F decreases during the iterates, we replace in the diagonalized Hessian, the negative values by a small constant. Using an Armijo step size rule, the time complexity and the memory storage are the same as SGD. Although we do not have convergence guarantees with DSN, we expect it to be faster than SGD and AGD.

4. TILINGS OF LASER SHOTS

In this section, we briefly present two possible choices for the location of the laser shots: the square tiling and the hexagonal tiling. In the first one (used in [4]), $K = K' \times K'$ laser

shots are tiled on a uniform square grid. In the second one, $K = K'^2 - \lfloor \frac{K'}{2} \rfloor$ laser shots are tiled on a uniform hexagonal grid. For any $K' > 1$, notice that the number of laser shots of a hexagonal tiling is smaller than for a square tiling. The difference between these tilings is illustrated in Figure 2. Notice that we expect less accurate estimates of the borders of the restored image, when acquired with a hexagonal tiling.

5. EXPERIMENTAL RESULTS

5.1. Applicative framework and implementation details

The experiments detailed in the next section 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 , σ_c , σ_w , \bar{w} , K' and \bar{c}_k , $\forall k \in \{1, \dots, K\}$. In this setting, we set $\sigma = 0.1$, $\sigma_c = 0.81$, $\sigma_w = 0.07$, $\bar{w} = 16.2$ and $K' = 9$. Possible choices for the parameters \bar{c}_k are discussed in Section 4. Notice that the pixel intensity in the ideal image ranges in $[0, 1]$ and is coded on 8 bits.

For estimating the restored image, we use the max-flow implementation of [11] and the dyadic parametric scheme of [12]. For estimating the acquisition parameters, the accuracy ε_e decays between 0.1 and 0.01. For the accuracy of the Algorithm 1, we found that setting $\varepsilon_a = 9.61 \times 10^{-7}$ is a good tradeoff between convergence and accuracy. Its exact value corresponds to a per-pixel error of 9.8×10^{-4} between two successive image estimates. The penalty parameter β (see (2) in Section (2)) is set to minimize the Mean Square Error (MSE)¹ between the image estimate and the ideal one using Golden Section Search [13]. Also, we do not provide detailed computing times since we believe they are not representative of an optimized version of the Algorithm 1. A simple improvement with this regards would consist in extracting from each image v^k a small window containing the laser shots. With the current C++ implementation, the restoration of an image typically requires between 1 and 6 minutes on a dodeca core Intel Xeon 3.47 GHz. Also, the estimation of the restored image typically represents 10% of the overall computation time.

5.2. Accuracy and convergence

We evaluate both the benefit of the advanced methods for estimating the acquisition parameters (see Section 3) as well as the use of a hexagonal tiling against a square tiling (see Section 4). We remind that in preprint [4], these parameters are estimated with SGD and that a square tiling is used. The evaluation is conducted on the same images as [4] and using the parameters provided in Section 5.1. Let us now describe the experimental setup for the square (resp. hexagonal) tiling. For each image, we generate 3 sequences of $K = 81$ (resp.

¹MSE and PSNR measures are described at <http://megawave.cmla.ens-cachan.fr/stuff/guid3/node256.html#fmse>.

Image	PSNR_H	PSNR_S
baboon	$23.92 \pm 4 \times 10^{-2}$	$23.77 \pm 2.5 \times 10^{-2}$
barbara	$25.19 \pm 3.1 \times 10^{-2}$	$25.01 \pm 5.3 \times 10^{-2}$
peppers	$28.22 \pm 7 \times 10^{-2}$	$28.02 \pm 6.1 \times 10^{-2}$
cameraman	$28.25 \pm 3.6 \times 10^{-2}$	$27.98 \pm 6.2 \times 10^{-2}$
lena	$27.92 \pm 4 \times 10^{-2}$	$27.65 \pm 3.9 \times 10^{-2}$
man	$26.05 \pm 3.6 \times 10^{-2}$	$25.89 \pm 3.3 \times 10^{-2}$
boat	$26.77 \pm 2.5 \times 10^{-2}$	$26.51 \pm 2.1 \times 10^{-2}$
factory	$27.03 \pm 2.1 \times 10^{-2}$	$26.82 \pm 2.3 \times 10^{-2}$

Table 1: Accuracy of the Algorithm 1 when using a hexagonal tiling (PSNR_H) against a square tiling (PSNR_S) with SGD for estimating the acquisition parameters. The measures are expressed in dB and rounded to the nearest value.

$K = 77$) laser shots and illumination domes. The sequences are then restored by applying Algorithm 1 and using SGD, AGD, BFGS and DSN for estimating the acquisition parameters. For each sequence, we measure the Peak Signal-to-Noise Ratio (PSNR) between the image estimate and the ideal image, denoted as PSNR_S (resp. PSNR_H). To be as fair as possible between tilings, these measures are restricted to the center of the image (i.e. the pixels for which their Tchebychev distance to the borders is greater or equal than $\frac{N}{2K}$). For each restored sequence, we also compute the total number of iterations used to estimate the acquisition parameters by summing the number of iterations required by all the estimation required by Algorithm 1. The results of these experiments are summarized in Table 1 and Table 2 and illustrated in Figure 3.

In Table 1, we provide statistics PSNR_S and PSNR_H for each image. For each tiling, since the measures are nearly the same for all optimization methods, we only provide those obtained using SGD. For all images, PSNR_H is slightly larger than PSNR_S . In words, the hexagonal tiling leads to better image estimates while using a smaller number of laser shots. Results for a subset of images of Table 1 are shown in Figure 3. The selected images correspond to the sequence for which the MSE between the image estimate and the ideal image is the smallest. We provide for each image the ideal one, the image estimate as well as the image where each pixel is assigned with its maximum intensity over all laser shots. Despite a substantial noise level, we observe that the Algorithm 1 behaves globally well. Large flat areas are well denoised while thin structures are well preserved, even between domes where the knowledge about data is less accurate.

In Table 2, we provide the mean of the total number of iterations required for estimating the acquisition parameters. Since these measures are nearly the same for both tilings, we only provide those obtained with the hexagonal tiling. We also provide the average condition number of the Hessian matrix at the minimizer obtained with SGD. For all images, second-order methods exhibit faster convergence than first-order ones, even for moderately ill-conditioned problems.

	SGD	AGD	BFGS	DSN	H^*
baboon	27	22.7	16.7	20.3	24.8
barbara	30.3	28.7	18.3	20.7	40.7
peppers	29.7	32.7	19.7	21	56.1
cameraman	31.3	30	19.7	18	197.1
lena	27	25.7	17	17.7	39.5
man	27.7	28	17.3	15.3	77.3
boat	23	23	16.7	20	56.5
factory	25.7	28.7	18.3	16.7	67.4

Table 2: Convergence of Algorithm 1 using advanced methods for estimating acquisition parameters and for a hexagonal tiling. For each image and method, we provide the mean of the total number of iterations required for estimating these parameters. For each image, we provide the mean condition number \bar{H}^* of Hessian at the minimizer obtained with SGD.

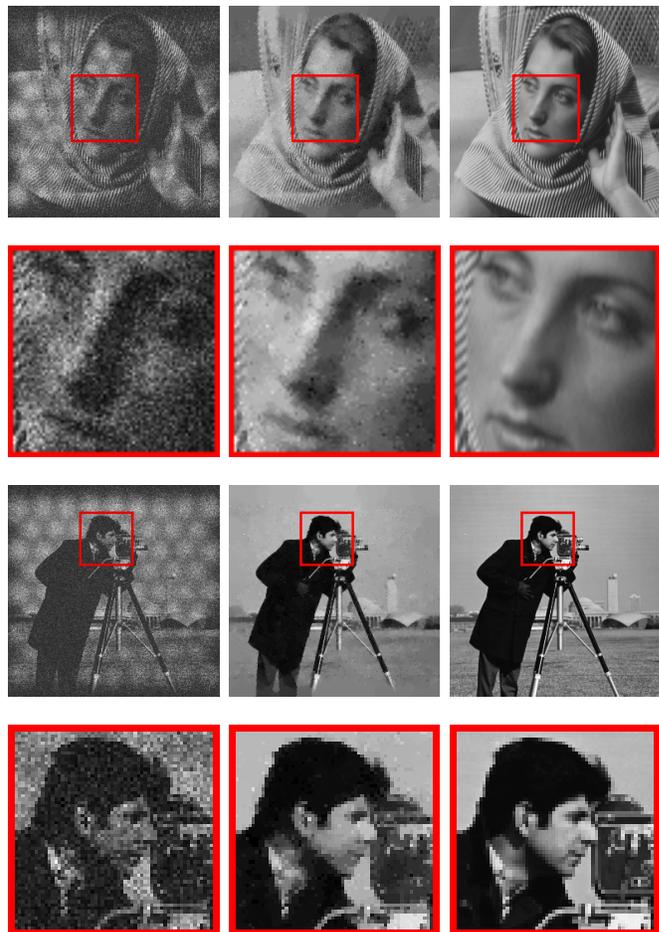


Fig. 3: Reconstruction of the images “barbara” (upper-half) and “cameraman” (lower-half) with Algorithm 1, using a hexagonal tiling. In the left column, each pixel of the image is assigned with its maximum intensity over all laser shots. The middle and right columns are resp. the image estimate and the ideal one. Detailed views of all these images are also provided in the second and fourth rows.

6. 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, “Modeling 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,” *Accepted in Inverse Problems and Imaging Journal*, November 2012.
- [5] L.I. Rudin, S. Osher, and E. Fatemi, “Nonlinear total variation based noise removal algorithms,” *Physica D*, vol. 60, pp. 259–268, 1992.
- [6] J. Darbon and M. Sigelle, “Image restoration with discrete constrained total variation part I: Fast and exact optimization,” *Journal of Mathematical Imaging and Vision*, vol. 26, no. 3, pp. 261–276, 2006.
- [7] A. Chambolle, “An algorithm for total variation minimization and applications,” *Journal of Mathematical Imaging and Vision*, vol. 20, no. 1-2, pp. 89–97, January–March 2004.
- [8] J. Bect, L. Blanc-Féraud, G. Aubert, and A. Chambolle, “A L1-unified variational framework for image restoration,” *Lecture Notes in Computer Science*, vol. 3024, pp. 1–13, 2004, Proceedings of European Conference on Computer Vision.
- [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.
- [10] D.P. Bertsekas, *Nonlinear Programming*, Athena Scientific, second edition, 2003.
- [11] 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.
- [12] 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.
- [13] J. Kiefer, “Sequential minimax search for a maximum,” *Proceedings of the American Mathematical Society*, vol. 4, no. 3, pp. 502–506, 1953.