Satellite Image Registration for Attitude Estimation with a Constrained Polynomial Model

Régis Perrier, Élise Arnaud, Peter Sturm, Mathias Ortner

▶ **To cite this version:**

Régis Perrier, Élise Arnaud, Peter Sturm, Mathias Ortner. Satellite Image Registration for Attitude Estimation with a Constrained Polynomial Model. ICIP 2010 - 17 IEEE International Conference on Image Processing, Sep 2010, Hong Kong, China. pp.925-928, 10.1109/ICIP.2010.5650997 . inria-00514896

HAL Id: inria-00514896

https://inria.hal.science/inria-00514896

Submitted on 3 Sep 2010

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.

SATELLITE IMAGE REGISTRATION FOR ATTITUDE ESTIMATION WITH A CONSTRAINED POLYNOMIAL MODEL

Régis PERRIER[*], *Elise ARNAUD*[†], *Peter STURM*[*], *Mathias ORTNER*[‡]

[*]INRIA Rhône Alpes, [†]Université Joseph Fourier LJK, [‡] EADS Astrium
[*]655 avenue de l'Europe, 38330 Montbonnot, France

ABSTRACT

Satellite image registration has been investigated for several years. Nevertheless, little attention has been paid to the linear geometry of the satellite's imaging sensor, often consisting of several pushbroom cameras. Each pushbroom camera captures 1-D image and uses straight motion of the satellite to build a 2-D image. Yet, attitude variations of the satellite during the aquisition process can lead to significant distortions in the 2-D image. In this paper, we expose the problem and present a constrained image registration method to estimate the satellite's attitude variations, and thus correct the distorted images. We use a Lucas Kanade framework and a piecewise polynomial model under constraints to deduce the registration equation. The performances of our algorithm are shown on two satellite datasets.

Index Terms— satellite, image registration, pushbroom camera, piecewise polynomial model

1. INTRODUCTION

Satellite image processing has been widely investigated over the last decades in several applications: image registration, super-resolution, image fusion and 3-D reconstruction [1]. In most cases, those algorithms do not take care of the geometry of the camera, considering images as being taken by a classical pinhole camera. Yet in remote sensing, pushbroom cameras are preferred over pinhole cameras for their ability to provide high resolution images at a lower cost [2].

The pushbroom camera is a linear sensor; mounted on a platform which moves orthogonally to the camera axis (as can be seen in figure 1(a)), the sensor sweeps out a region of space taking 1-D images at regular time instants. Stiching them gives a complete 2-D image of the observed scene. During the acquisition process of the 1-D images, the stability of the imaging platform is crucial to build a distortion-free 2-D image. Until now, the attitude[1] of the satellite was assumed to be constant in most applications [2, 3]. However, the recent manufacture of smaller satellites with higher sampling resolution has weakened this assumption resulting in possible warps in images. Figure 1(b) shows a synthetic example of deformations which can occur when the pushbroom camera is tilting around its x, y and z-axes. We define those rotations as the roll, the pitch and the yaw respectively.

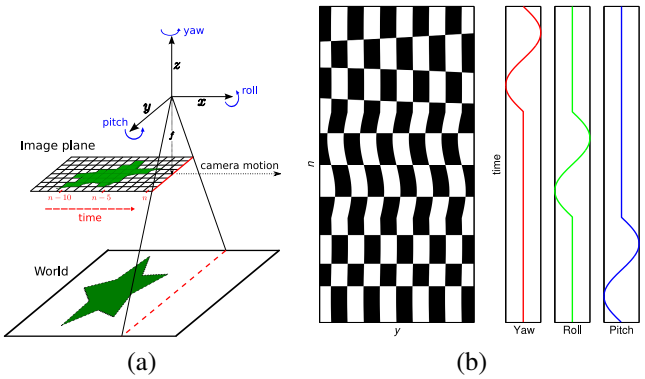

Fig. 1. (a) Pushbroom acquisition principle: the camera is moving straight along the x axis and recording 1-D images over time denoted by n; y is the camera axis and z the orthogonal axis to the image plane. We define the orientation of the camera with the yaw (rotation about z), the roll (rotation about x) and the pitch (rotation about y). (b) Example of warps in a regular checkerboard when the pushbroom camera is tilting around its 3 rotation axes.

Though this problem was already known in airborne imagery [4], it has been noticed only recently on satellite images [5]. Indeed, the spacecraft is continuously exposed to space turbulences and vibrations of its engines, making unknown satellite motions more probable. The use of inertial sensors measurements has been suggested to estimate attitude variations and get a distortion-free image [4]. While this can be applied in airborne imagery, space constraints restrict the use of such sensors for cost reasons. Moreover, the sampling rate of inertial sensors (4 to 16Hz) is usually much lower than the sampling rate of each pushbroom camera (around 2500Hz), leading to an information loss on the satellite's attitude.

Observation satellites have several pushbroom cameras of different radiometric modalities; figure 2 presents a typical satellite focal plane with 3 pushbroom cameras. All images are deformed due to the unknown motions of the focal plane. There is no way that one can retrieve the vibrations using only

This work was funded by EADS Astrium (European Aeronautic Defence and Space Company).

[1]usual name for the orientation of the air and space vehicle in flight dynamics science defined by the yaw, the roll, and the pitch

one 2-D image by correlating each line with its neighbors. Considering all 2-D images, one can try to register them pairwise to estimate the deformations. This has been suggested in [5], but their approach uses local correlation on features points which keeps from considering the problem globally. Thus they cannot consider a multi-frame image alignment as described in [6], yielding a sub-optimal estimate.

In this paper, we present an original image registration method to recover the attitude of the satellite using only image. We use a direct method, also known as Lucas Kanade based registration algorithm [7], to globally register images in a same coordinate system. We model the attitude variations with a piecewise polynomial function under constraints. This yields a fast and simple algorithm with no regularizing parameter to tune.

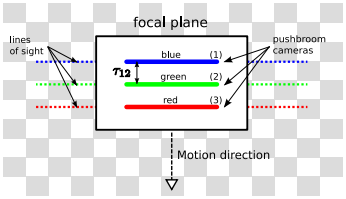

Fig. 2. Focal plane geometry of an observation satellite with 3 pushbroom cameras: blue, green and red (respectively 1,2 and 3). What is seen by the camera 2 at time n will be seen by the camera 1 at time $n + \tau_{12}$ in the absence of attitude variations.

2. ATTITUDE ESTIMATION

In this paper, we denote the set of 3 images (red, blue, green) as $\mathbf{I} = \{I_1, I_2, I_3\}$; radiometric values of images belong to a same set $\mathcal{V}$. Each image is defined by its pixel coordinates $[n, y]$ in the pixel set $\mathcal{S}$, with $n \in [1, N]$ being the discretised time and N the total number of time samples (which is also equivalent to the number of lines in the acquired images). Following figure 2, we call τ_{ij} the time gap which spaces the observation of a same scene part between pushbroom sensors i and j. In the perfect case when the imaging platform is not vibrating, we expect to have the following relation:

$$I_i(n, y) - f_{ij}(I_j(n + \tau_{ij}, y)) \sim \mathcal{N}(0, \sigma^2) \quad (1)$$

where $f_{ij} : \mathcal{V} \to \mathcal{V}$ is a linear function which corrects the radiometric differencies between images i and j, and σ^2 the variance of a zero mean i.i.d. Gaussian noise over all pixels of images. To achieve a more compact formulation, let $I_j^{\tau_{ij}}$ being the time shifted image I_j such that $I_j^{\tau_{ij}}(n, y) = I_j(n + \tau_{ij}, y)$ and $\mathbf{y} = [n, y]^T$ the vector of pixel coordinates, thus:

$$I_i(\mathbf{y}) - f_{ij}(I_j^{\tau_{ij}}(\mathbf{y})) \sim \mathcal{N}(0, \sigma^2). \quad (2)$$

Due to the motion of the focal plane around its 3 rotations axes, all images are warped depending on $\theta(n) \in \Theta$, the unknown attitude of the satellite at time n; it is a (3×1) vector whose components are respectively the yaw $\theta_{ya}(n)$, the roll $\theta_r(n)$ and the pitch $\theta_p(n)$. To avoid redundancy in equations, we will refer to $\theta_\alpha(n)$ in section 2.2 where α denotes separately either the yaw, the roll and the pitch. We call $\boldsymbol{\theta}$ the $(3N \times 1)$ vector that gathers all attitudes for all time instants:

$$\boldsymbol{\theta} = [\theta_{ya}(1) \dots \theta_{ya}(N), \theta_r(1) \dots \theta_r(N), \theta_p(1) \dots \theta_p(N)]^T. \quad (3)$$

Let $W : \mathcal{S} \times \Theta \to \mathcal{S}$ be a warp function that maps pixel coordinates to a new position depending on the attitude of the satellite. From equation (2), what we are truly observing is:

$$I_i\big(W(\mathbf{y}; \theta(n))\big) - f_{ij}\big(I_j^{\tau_{ij}}\big(W(\mathbf{y}; \theta(n+\tau_{ij}))\big)\big) \sim \mathcal{N}(0, \sigma^2), \quad (4)$$

where it is important to notice that both images are warped by θ, but at different time instants.

2.1. Image registration formulation

A popular way to solve equation (4) is to use direct methods which make optimal use of all pixels in the images [7]. This is well suited in our context as warps are of few pixels and registration needs a sub-pixel accuracy. We define $T_{\tau_{ij}}$ as an operator which shifts time samples of $\boldsymbol{\theta}$ by a factor of τ_{ij}. This operator is a $(3N \times 3N)$ sparse matrix which equals for row $n \in [1, 3N]$ and column $m \in [1, 3N]$ to:

$$T_{\tau_{ij}}(n, m) = \begin{cases} 1 & \text{for } n = m - \tau_{ij}, \text{and } n \notin]N - \tau_{ij}, N] \\ & \quad \cup]2N - \tau_{ij}, 2N] \cup]3N - \tau_{ij}, 3N] \\ 0 & \text{elsewhere} \end{cases}$$

The global registration equation to minimize is:

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}}{\operatorname{argmin}} \sum_{i,j;i \neq j} \sum_{\mathbf{y} \in \mathcal{S}} \Big(I_i\big(W(\mathbf{y}; \boldsymbol{\theta})\big) - f_{ij}\big(I_j^{\tau_{ij}}\big(W(\mathbf{y}; T_{\tau_{ij}}\boldsymbol{\theta})\big)\big)\Big)^2. \quad (5)$$

As opposed to the classical Lucas Kanade formulation, there is no reference frame here as both images are warped. Considering a reference frame would lead us to the estimation of a **relative** estimate of the warps that would imply a convolution process to recover the **absolute** attitude. This is a major drawback of [5]. In our case, we define a method to directly estimate the **absolute** attitude. Let's also note that we use a **multi-frame** registration formulation to estimate $\boldsymbol{\theta}$, this has been seen to be more efficient for motion estimation [6].

Direct estimation of $\boldsymbol{\theta}$ in equation (5) is highly ill-posed if no constraints are defined. One could add a prior term to this equation, but the choice of the regularising function is tricky. Moreover, the computation of the regularising parameter which tunes the belief between the image term and the prior term is expensive and has no trivial solution. Going further into the origins of the problem, we know that the attitude variations mainly originate in the engines of the satellite. Such a vibratory process usually has a stationary form in time, so that smoothness is a trustful property of the signal.

2.2. Constrained polynomials

To model the vibrations, we use a piecewise polynomial function where polynomials are linked together by constraints.

Polynomials are denoted by $\mathbf{\Phi_{k,\alpha}}$, $\boldsymbol{k} \in [\mathbf{1}, \boldsymbol{M}]$. $\boldsymbol{M}$ is the number of polynomials needed to fit the attitude variations, and $\boldsymbol{n_k}$ for $\boldsymbol{k} \in [\mathbf{1}, \boldsymbol{M-1}]$ the time instants where constraints are defined between two successive polynomials. This can be related to splines [8], except that our knots are spaced by several time samples as this can be seen on figure 3.

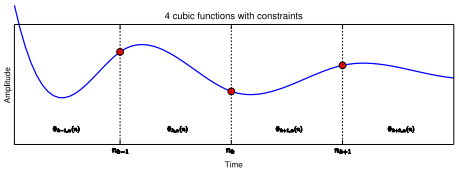

Fig. 3. Example of a piecewise polynomial function consisting of 4 polynomials ($\mathbf{\Phi_{k-1,\alpha}}, \mathbf{\Phi_{k,\alpha}}, \mathbf{\Phi_{k+1,\alpha}}, \mathbf{\Phi_{k+2,\alpha}}$) with constraints defined on $\boldsymbol{n_{k-1}}$, $\boldsymbol{n_k}$ and $\boldsymbol{n_{k+1}}$. The red dots show the location of constraints on the global function whereas vertical dashed lines delimit each local polynomial. In this example, we choose polynomials of order 3 (cubic) with constraints to the second derivative.

The vibrations can be described with $\boldsymbol{M}$ polynomials:

$$\boldsymbol{\theta_\alpha(n)} = \sum_{\boldsymbol{k=1}}^{\boldsymbol{M}} \boldsymbol{w_k(n)\Phi_{k,\alpha}(n)}, \tag{6}$$

where the logical weights $\boldsymbol{w_k(n)}$ are defined such that:

$$\boldsymbol{w_k(n)} = \begin{cases} \mathbf{1} & \text{for } \boldsymbol{n} \in [\boldsymbol{n_{k-1}}, \boldsymbol{n_k}[\\ \mathbf{0} & \text{elsewhere.} \end{cases} \tag{7}$$

The polynomials of order $\boldsymbol{P}$ are defined as:

$$\boldsymbol{\Phi_{k,\alpha}(n)} = \sum_{\boldsymbol{p=0}}^{\boldsymbol{P}} \boldsymbol{a_{p,k,\alpha} n^p}. \tag{8}$$

$\boldsymbol{a_{p,k,\alpha}}$ are the polynomial coefficients which fully characterise the attitude variations. We call $\mathbf{a}$ the $\boldsymbol{(3MP \times 1)}$ vector which stacks all the coefficients for the attitude estimate. We define the constraints between two polynomials in the same spirit as splines for all $\boldsymbol{n_k}$ with $\boldsymbol{k} \in [\mathbf{1}, \boldsymbol{M-1}]$:

$$\boldsymbol{\Phi^{(p)}_{k,\alpha}(n_k)} = \boldsymbol{\Phi^{(p)}_{k+1,\alpha}(n_k)} \text{ for } \boldsymbol{p} \in [\mathbf{0}, \boldsymbol{P-1}]. \tag{9}$$

This way, continuity is imposed for the derivatives of the polynomial up to order $\boldsymbol{P-1}$. As for splines, there is only one degree of freedom per time segment $[\boldsymbol{n_k}, \boldsymbol{n_{k+1}}[$ [8]. Notice that equation (9) defines linear equality constraints with $\mathbf{a}$. We call $\boldsymbol{C}$ the $\boldsymbol{(3(M-1)P \times 3MP)}$ sparse matrix such that equation (9) can be written in the following matrix form:

$$\boldsymbol{C\mathbf{a} = \mathbf{0}}, \tag{10}$$

with $\mathbf{0}$ the null vector of size $\boldsymbol{(3(M-1)P \times 1)}$. Also, concatenating linear equations (6), (7) and (8) yields:

$$\boldsymbol{\theta = H\mathbf{a}}, \tag{11}$$

where $\boldsymbol{H}$ is a sparse matrix of size $\boldsymbol{(3N \times 3MP)}$ which combines $\boldsymbol{w_k(n)}$ of equation (6) and $\boldsymbol{n^p}$ of equation (8). We can reformulate the image registration problem of equation (5) as a non linear problem with linear equality constraints:

$$\hat{\mathbf{a}} = \underset{\mathbf{a}}{\mathbf{argmin}} \sum_{i,j;i\neq j} \sum_{\mathbf{y}\in\mathcal{S}} \Big(\boldsymbol{I_i}(\boldsymbol{W}(\mathbf{y};\boldsymbol{H}\mathbf{a})) - \boldsymbol{f_{ij}}\left(\boldsymbol{I_j^{\tau_{ij}}}\left(\boldsymbol{W}(\mathbf{y};\boldsymbol{T_{\tau_{ij}}H}\mathbf{a})\right)\right)\Big)^{\mathbf{2}}, \text{ such that } \boldsymbol{C\mathbf{a}} = \mathbf{0}. \tag{12}$$

This global non linear equation is minimised iteratively. The equality constraints are used to reduce the problem to a lower dimensionality with a QR factorisation, then the remaining variables are estimated with a least square procedure.

In practice, the $\boldsymbol{f_{ij}}$ function introduced in section 2 to correct radiometric differences is linear such that:

$$\boldsymbol{f_{ij}}(\boldsymbol{I_j}(\mathbf{y})) = \boldsymbol{r_{ij,1}} + \boldsymbol{r_{ij,2}}\boldsymbol{I_j}(\mathbf{y}) \tag{13}$$

We estimate radiometric coefficients $[\boldsymbol{r_{ij,1}}, \boldsymbol{r_{ij,2}}]$ for each pair of images, and once for all at the beginning of the algorithm with a least square method. The warp function $\boldsymbol{W}$ has a tractable analytical formulation for the three rotation angles if the observed scene is planar. This assumption is approximately true in some cases and the analytical warp shows nice results. In other cases, we use numerical derivatives correlated with a Digital Elevation Model (DEM) data to estimate the attitude variations. We choose the second option in the following results.

3. EXPERIMENTS AND DISCUSSION

We present experimental results on two satellite datasets which have been simulated by ASTRIUM. The ground truth on attitude variations is available, but the simulation process is totally unknown to us. It creates real-life condition data acquisition and can be considered as difficult as could be real data. We use a Matlab implementation on a Core2 duo at 3GHz with 3.8GiB to process the data. In both cases, cameras 1-2, and 2-3 are spaced respectively by 40 and 20 time samples in accordance to the focal plane in figure 2 ($\boldsymbol{\tau_{12}}$=40 and $\boldsymbol{\tau_{23}}$=20). All images are of size (2564×900) pixels.

The first dataset is composed of multispectral data with blue, green and red images. We set the polynomials order $\boldsymbol{P}$ to 3 and their size to 15 time samples. The algorithm took 230 seconds to converge in 10 iterations; the standard deviation of the error compared to ground truth is below $\frac{\mathbf{15}}{\mathbf{100}}$ either for the roll and the pitch (see fig. 4(a)). The second dataset is composed of monomodal images. This case is considered to be a challenging one as the vibrations contain very low frequency components. We set the polynomial order $\boldsymbol{P}$ to $\mathbf{3}$ and their size to 20 time samples. The algorithm took 410 seconds to converge in 20 iterations; the standard deviation of the error compared to ground truth is below $\frac{\mathbf{20}}{\mathbf{100}}$ either for the roll and the pitch (see fig. 4(b)). Let us remark that no estimation of the yaw is given in both experiments. The yaw causes deformations lower than $\frac{\mathbf{2}}{\mathbf{100}}$ pixel on each vertical side of the images. In fact, the deformations are dominated by geometrical distortions induced by the roll and pitch.

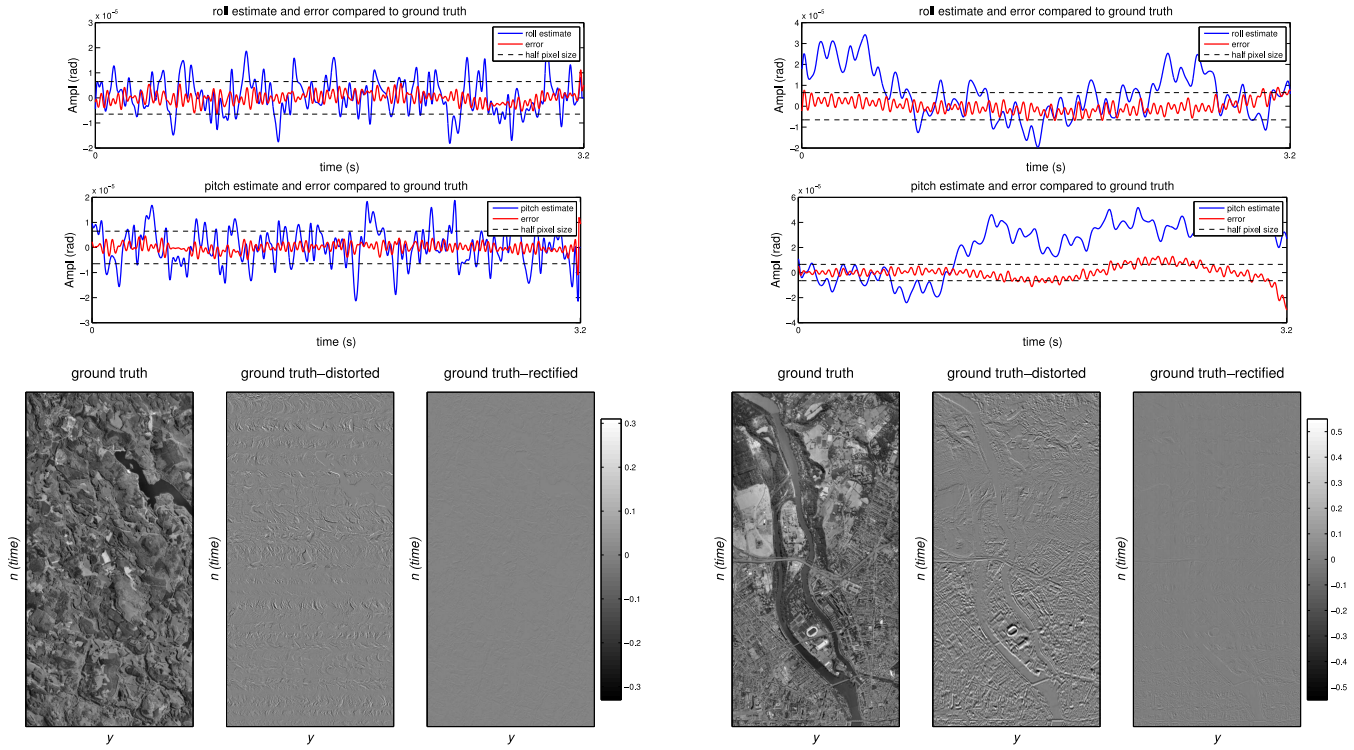

Fig. 4. (a) Results of the first dataset (b) Results of the second dataset: the two top figures show the final attitude estimate and its error compared to the ground truth for the roll and the pitch. Below, the left image is an original image patch of the third camera of size $(\mathbf{1000} \times \mathbf{500})$, following figures show error images before and after image registration compared to ground truth.

The results on figures 4(a) and 4(b) show performances and some limits of our algorithm. Most of the frequencies are recovered, and the registration process is able to rectify images with an accuracy much below than the pixel unit. Constrained polynomials are suited to the estimation of such vibration processes mainly because of their stationary properties. However in each case, the residual error has a high frequency component. This is mainly linked to the way the polynomial can fit the vibrations in a given time window. For larger time windows, the polynomial will recover low frequency components. For shorter time windows, it has the ability to fit high frequency vibrations but eq. (12) has more degrees of freedom and can become ill-posed. To a lesser extent, the polynomial degree also influences the attitude estimate. Ways to choose the size of the time window, as well as the polynomial degree, are currently investigated.

The originality of our work is to present a global image registration technique for images acquired by pushbroom cameras on board of a satellite. This can solve the problem of attitude variation of the satellite during the process of acquisition. Although in general, the registration problem is highly ill-posed in this case, the algorithm does not need any regularisation due to the built-in realistic assumption of stationarity.

4. REFERENCES

[1] Z. Liu R. S. Blum, *Multi-sensor image fusion and its applications*, Taylor and Francis, 2005.

[2] G. Petrie, "Airborne pushbroom line scanners: An alternative to digital frame scanners," *Geoinformatics*, vol. 8, no. 1, pp. 50–57, 2005.

[3] R. Gupta and R. I. Hartley, "Linear pushbroom cameras," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 9, pp. 963–975, 1997.

[4] D. Poli, "General model for airborne and spaceborne linear array sensors," in *International Archives of Photogrammetry and Remote Sensing*, 2002, vol. 34.

[5] F. de Lussy, D. Greslou, and L. Gross Colzy, "Process line for geometrical image correction of disruptive microvibrations," in *International Society for Photogrammetry and Remote Sensing*, 2008, pp. 27–35.

[6] S. Farsiu, M. Elad, and P. Milanfar, "Constrained, globally optimal, multi-frame motion estimation," in *IEEE Workshop on Stat. Signal Proc.*, 2005, pp. 1396 – 1401.

[7] R. Szeliski, "Image alignment and stitching: a tutorial," *Found. Trends. Comput. Graph. Vis.*, vol. 2, no. 1, pp. 1–104, 2006.

[8] M. Unser, "Splines: A perfect fit for signal and image processing," *IEEE Signal Processing Magazine*, vol. 16, no. 6, pp. 22–38, November 1999.