Image reconstruction of non-planar scenes from compressed multi-view measurements

Gilles Puy
Ecole Polytechnique Fédérale de Lausanne (EPFL)
CH-1015 Lausanne, Switzerland

Pierre Vandergheynst
Ecole Polytechnique Fédérale de Lausanne (EPFL)
CH-1015 Lausanne, Switzerland

***Abstract*—We extend the domain of application of a method developed recently for joint reconstruction of l images, representing the same scene, from few multi-view measurements. While this method was initially designed for planar scenes, we show here that parametric smooth transformations can be used to handle more challenging scene structures. Our algorithm estimates one reference image common to all viewpoints, l complementary images modeling details in the scene that are not always visible, and few transformation parameters modeling the inter-correlation between the observations. The algorithm is an alternating descent method built to minimize a non-convex objective function and which produces a sequence converging to one of the critical points of this function.**

I. Introduction

We have recently described a novel method to reconstruct jointly a set of l images of a scene from few linear measurements obtained at different viewpoints [1]. The correlation between images is modeled using global parametric transformations, such as homographies, and the proposed algorithm accurately estimates both the images and the transformation parameters, while being robust to occlusions.

Unfortunately, the global transformations considered above are not well adapted for non-planar scenes. In this abstract, we exploit the flexibility of the framework [1] and show how smooth parametric transformations allow us to handle more challenging scene structures.

II. Method description

As in [1], we suppose that we have l observations $\boldsymbol{y}_1, \dots, \boldsymbol{y}_l \in \mathbb{R}^m$ of a given scene. We assume that the j^{th} observations is obtained by linear projection of the image $\mathsf{S}(\boldsymbol{\theta}_j)\boldsymbol{x}_0 + \boldsymbol{x}_j$ onto m measurement vectors. The vector $\boldsymbol{x}_0 \in \mathbb{R}^n$ represents a reference image common to all observations, $\boldsymbol{\theta}_j \in \mathbb{R}^q$ some transformation parameters modeling the difference of appearance of $\boldsymbol{x}_0$ at each viewpoint, $\mathsf{S}(\boldsymbol{\theta}_j) \in \mathbb{R}^{n\times n}$ the interpolation matrix that transforms $\boldsymbol{x}_0$, and $\boldsymbol{x}_j \in \mathbb{R}^n$ parts of the scene that are not always visible. Let $\mathsf{A}_j \in \mathbb{R}^{m\times n}$ be the j^{th} measurement matrix, the measurement model satisfies

$$\boldsymbol{y}_j = \left[\begin{array}{cc} \mathsf{A}_j\mathsf{S}(\boldsymbol{\theta}_j) & \mathsf{A}_j \end{array}\right] \left[\begin{array}{c} \boldsymbol{x}_0 \\ \boldsymbol{x}_j \end{array}\right] + \boldsymbol{n}_j, \quad \text{with } j = 1, \dots, l. \tag{1}$$

where $\boldsymbol{n}_j \in \mathbb{R}^m$ represents additive measurement noise.

We consider here smooth geometric transformations modeled by polynomials of the 2D spatial cartesian coordinates, see, e.g., [2]. Let $(\boldsymbol{u}_1, \boldsymbol{u}_2) \in \mathbb{R}^{n\times 2}$ be the coordinates at which $\boldsymbol{x}_0$ is sampled. The transformation of $\boldsymbol{x}_0$ between viewpoints is modeled using functions $\tau_{\boldsymbol{\theta}_j}$ that map the reference coordinates to $\tau_{\boldsymbol{\theta}_j}(\boldsymbol{u}_k) = \boldsymbol{u}_k + \mathsf{P}(\boldsymbol{u}_1, \boldsymbol{u}_2)\,\boldsymbol{\theta}_j^k \in \mathbb{R}^n$, $k = 1, 2$, where $\mathsf{P}(\boldsymbol{u}_1, \boldsymbol{u}_2) \in \mathbb{R}^{n\times q/2}$ contains the first $q/2$ vectors of the canonical polynomial basis, and $\boldsymbol{\theta}_j^{\intercal} = ((\boldsymbol{\theta}_j^1)^{\intercal}, (\boldsymbol{\theta}_j^2)^{\intercal})$.

To estimate the images and the transformation parameters from the observations, we use the algorithm presented in [1], whose design was inspired by [3], which produces a convergent sequence to a critical point of $\|\Psi^{\intercal}\bar{\boldsymbol{x}}_0\|_1 + \sum_{1\leqslant j\leqslant l}\|\Psi^{\intercal}\bar{\boldsymbol{x}}_j\|_1 + \kappa\,\|\mathsf{A}_j\mathsf{S}(\bar{\boldsymbol{\theta}}_j)\bar{\boldsymbol{x}}_0 + \mathsf{A}_j\bar{\boldsymbol{x}}_j - \boldsymbol{y}_j\|_2^2 + i_{\Theta_j}(\bar{\boldsymbol{\theta}}_j)$ ($\bar{\cdot}$ are the optimization variables), where $\Psi \in \mathbb{R}^{n\times n}$ is the Haar wavelet basis, $\kappa > 0$ is a regularizing parameter, $(\Theta_j)_{1\leqslant j\leqslant l}$ are convex subsets of $\mathbb{R}^q$, and i_{Θ_j} is the indicator function of Θ_j.

Fig. 1. Left to right: before registration; after registration with affine transformations; with smooth transformations; SNR vs. m/n for the proposed method with model 1 (dashed red), 2 (continuous black), 3 (continuous red), and the BP method (dashed black). The curves represent the mean SNR over 10 simulations, and the vertical lines the error at 1 standard deviation.

III. Experiments and results

We test the proposed method using 5 images of the same non-planar scene taken from different viewpoints[1]. We generate 5 measurement vectors using the compressed sensing technique [4]. Fig. 1 shows the curves of the reconstruction SNR as a function of m/n obtained with our method and a 1) translation, 2) affine transformation and 3) smooth transformation model with 56 parameters, as well as by solving the Basis Pursuit (BP) problem for each image independently. The use of smooth transformations leads to better reconstruction qualities. We also run the same experiments on 5 other images[2] and show the superposed initial images before and after registration with the affine and smooth transformations estimated from $0.3n$ measurements. One can remark that the smooth transformation model provides the best alignment of the background and parts of the foreground (see the branches at the top of the image).

We have presented a method for image reconstruction of non-planar scenes from few multi-view measurements using smooth parametric transformations to model the correlation between observations. Our algorithm is an alternating descent method that produces a sequence converging to a critical point of a non convex functional. Experiments confirm the benefit of using such transformations.

References

[1] Puy et al., "Robust image reconstruction from multi-view measurements," *SIAM J. Imaging Sci.*, submitted, arXiv:1212.3268, 2012.

[2] Kybic et al., "Fast parametric elastic image registration," *IEEE Trans. Image Process.*, vol. 12(11), pp. 1427–1442, 2003.

[3] Attouch et al., "Convergence of descent methods for semi-algebraic and tame problems: proximal algorithms, forward backward splitting, and regularized gauss-seidel methods," *Math. Program.*, 2011.

[4] Puy et al., "Universal and efficient compressed sensing by spread spectrum and application to realistic fourier imaging techniques," *EURASIP J. Adv. Signal Process.*, vol. 2012(6), 2012.

This work was partly funded by the Hasler Foundation (project number 12080).

[1] Mini Cooper dataset: web.media.mit.edu/~gordonw/SyntheticLightFields/.
[2] Garden dataset: media.xiph.org/video/derf/.