A data-driven approach to feature space selection for robust micro-endoscopic image reconstruction

Pascal Bourdon and David Helbert

XLIM Research Institute (UMR CNRS 7252), Université de Poitiers
Bât. SP2MI, Téléport 2, 11 Bd Marie et Pierre Curie, BP 30179. F-86962 Futuroscope Chasseneuil Cedex, France
pascal.bourdon@univ-poitiers.fr

Context

Multi-photon microendoscopy allows real-time imaging of cellular-level morphology with high contrast, high penetration depth, minimal photo-toxicity and limited sensitivity to the diffusion of biological materials [Duc15]. However, the acquisition of large tissue areas requires a process of multi-view reconstruction from a sequence of low-field-of-view images. Noise, illumination changes and geometric distortions induced by hand motion and optics are making the visual matching of said images a challenging task for reconstruction. In this article, we propose a new on-line image feature space selection strategy for displacement field estimation.

Visual matching using convex optimization

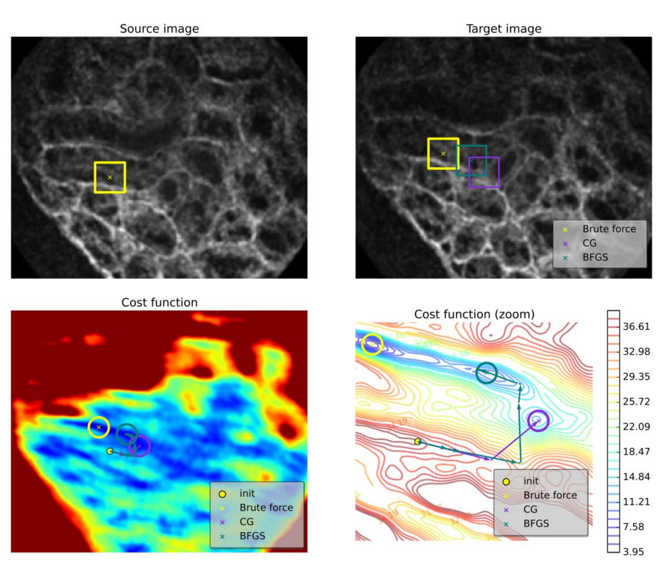
Given a small enough deformation or hand displacement between two consecutive images, we want to identify a geometric transform $\mathbf{W}(\mathbf{x}, \boldsymbol{\theta})$ such that:

$$I(\mathbf{x}, t + \partial t) = I(\mathbf{W}(\mathbf{x}, \boldsymbol{\theta}_\mathbf{t}), t)$$

Image registration then consists in estimating a set of parameters $\boldsymbol{\theta}_\mathbf{t}$ which satisfy this equation *i.e.* :

$$\tilde{\boldsymbol{\theta}}_\mathbf{t} = \arg\min_{\boldsymbol{\theta}_\mathbf{t}}(\xi(\boldsymbol{\theta}_\mathbf{t}))$$
$$\xi(\boldsymbol{\theta}_\mathbf{t}) = \sum_{\mathbf{x} \in \Upsilon} \| I(\mathbf{x}, t) - I(\mathbf{W}(\mathbf{x}, \boldsymbol{\theta}_\mathbf{t}), t + \partial t) \|_2^2.$$

Where objective function $\xi(\boldsymbol{\theta}_\mathbf{t})$ is considered convex and smooth enough to provide a good estimation of $\boldsymbol{\theta}_\mathbf{t}$. However it is not the case with our microendoscopic image data :



Data-driven feature space selection

Suppose there exists a set of image feature spaces $\mathbf{F}(\mathbf{x}, t)$ for which $\xi(\boldsymbol{\theta}_\mathbf{t})$ is actually convex. We could use a linear combination ($\boldsymbol{\beta}$-weighted) of said features to define a new cost function:

$$\hat{\boldsymbol{\theta}}_\mathbf{t}(\boldsymbol{\beta}) = \arg\min_{\boldsymbol{\theta}_\mathbf{t}}\left(\xi_{\boldsymbol{\beta}}(\boldsymbol{\theta}_\mathbf{t})\right) = \hat{\boldsymbol{\theta}}_\mathbf{t}$$
$$\xi_{\boldsymbol{\beta}}(\boldsymbol{\theta}_\mathbf{t}) = \sum_{\mathbf{x} \in \Upsilon} \left\| \boldsymbol{\beta}^T . (\mathbf{F}(\mathbf{x}, t) - \mathbf{F}(\mathbf{W}(\mathbf{x}, \boldsymbol{\theta}_\mathbf{t}), t + \partial t)) \right\|_2^2$$

where $\hat{\boldsymbol{\theta}}_\mathbf{t}$ are ground truth parameters recovered from brute-force search over the original image domain.

The estimation of feature weighting vector $\boldsymbol{\beta}$ is a key element of our method. We are looking for a sparse solution to reduce computation times and provide safety against redundant feature descriptions. Hence the addition of a L1-norm constraint:

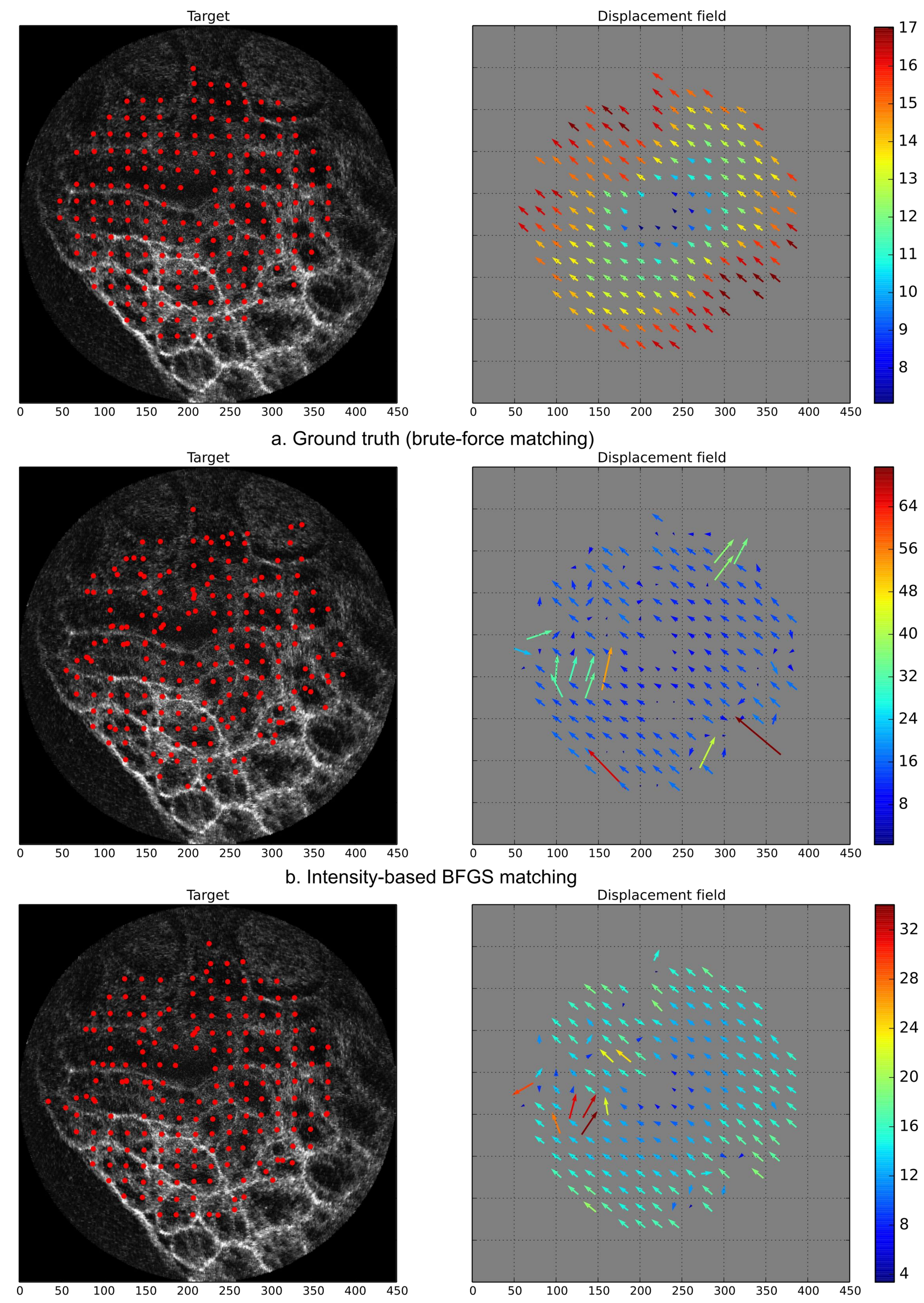
$$\hat{\boldsymbol{\beta}} = \arg\min_{\boldsymbol{\beta}}\left(\alpha \|\boldsymbol{\beta}\|_1 + \sum_{t \in [0, t_0]} \left\| \hat{\boldsymbol{\theta}}_\mathbf{t} - \tilde{\boldsymbol{\theta}}_\mathbf{t}(\boldsymbol{\beta}) \right\|_2^2 \right)$$

In practice the endoscope operator is required to capture a few images for ground truth generation, with brute-force matchings providing the algorithm with $\hat{\boldsymbol{\theta}}_\mathbf{t}$ samples.

[Duc15] Ducourthial et al (2015). Development of a real-time flexible multiphoton microendoscope for label-free imaging in a live animal. Scientific reports, Nature Publishing Group.

Experimental results

Data: 450x450 low-field images taken with a 250µm fiber endoscope. Image features : scale-variant polynomials and Gabor wavelets. Local optimization: BFGS algorithm.



a. Ground truth (brute-force matching)
b. Intensity-based BFGS matching
c. BFGS matching in a learned feature space

Feature space	MSE	SD
Raw pixel intensity	12.67	± 19.40
Multi-scale polynomial, $\alpha = 0.1$	11.05	±16.70
Multi-scale polynomial, $\alpha = 1.0$	10.37	± 14.57
Gabor wavelet basis, $\alpha = 0.1$	11.21	±16.66
Gabor wavelet basis, $\alpha = 1.0$	11.05	±16.70

Conclusion

- A new feature space selection method for displacement field estimation
- On-line learning of ground truth data from non-optimal global methods