

Coupled Multi-Frame Super-Resolution Reconstruction and Motion Estimation

Eric Ng and Mehran Ebrahimi

Faculty of Science, University of Ontario Institute of Technology, Oshawa, ON

eric.ng@uoit.net, mehran.ebrahimi@uoit.ca



Introduction

A demand for higher quality and higher resolution images will always exist whether it is medical imaging, surveillance, or simply personal leisure. The process of recovering a high-resolution (HR) image from a set of distorted (i.e. deformed, blurry, noisy, etc.) low-resolution (LR) images is known as *super-resolution*. The *super-resolution* problem will require the reconstruction of the ideal HR image and an estimation of the motion between LR images. The proposed algorithm will attempt to recover the HR image and motion estimation simultaneously using an inverse model followed by a discretize-then-optimize approach. Elastic regularization will be used for non-parametric motion estimation, and total variation regularization for the super-resolved image. Due to the problem being ill-posed (i.e. many number of solutions), parameters will be introduced as constraints to limit the possible solutions.

Mathematical Formulation

Consider a set of m low-resolution images represented by the continuous mappings $y = \{y_1, \dots, y_m\}$, $y_i: \mathbb{R}^2 \rightarrow \mathbb{R}$, the goal is to find an ideal high-resolution image f and a set of transformations $u = \{u_1, \dots, u_m\}$, $u_i: \mathbb{R}^2 \rightarrow \mathbb{R}^2$, such that $\mathcal{H}f[u_i]$ is similar to y_i for all $i = 1, \dots, m$, where \mathcal{H} is a linear degradation operator comprising of a blur operator and downsampling operator. In this study, f is represented by the mapping $f: \Omega \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ where Ω is the area of interest of the image.

Problem Statement. Given a set of m low resolution ($p \times q$) images, $y = \{y_1, \dots, y_m\}$, find a high resolution ($zp \times zq$, $z \in \mathbb{N}$ is a zooming factor) image f and a set of transformations $u = \{u_1, \dots, u_m\}$ that minimizes the objective functional

$$\mathcal{J}[u, f] = \sum_{i=1}^m \mathcal{D}[y_i, \mathcal{H}f[u_i]] + \sum_{i=1}^m \alpha_i \mathcal{S}[u_i - u^{ref}] + \beta \mathcal{Q}[f].$$

Here, $\mathcal{D}[y_i, \mathcal{H}f[u_i]]$ is a distance measure between y_i and $\mathcal{H}f[u_i]$, $\mathcal{S}[u_i - u^{ref}]$ is the elastic regularizer, and $\mathcal{Q}[f]$ is the total variation penalty on the computed image f imposed by Neumann boundary conditions. $\alpha = \{\alpha_1, \dots, \alpha_m\} \in \mathbb{R}^m$ and $\beta \in \mathbb{R}$ are regularization parameters. It is assumed that $u^{ref}(x) = x$ where x is an evenly spaced grid. Sum of squared differences (SSD) will be used to compute distance measure. This can be summarized as follows:

$$\mathcal{D}^{SSD}[y_i, \mathcal{H}f[u_i]] = \frac{1}{2} \int_{\Omega} (y_i - \mathcal{H}f[u_i])^2 dx \quad (1)$$

$$\mathcal{S}(u) = \frac{1}{2} \int_{\Omega} \mu < \nabla u, \nabla u > + (\lambda + \mu) (\nabla \cdot u)^2 dx \quad (2)$$

$$\mathcal{Q}[f] = \int_{\Omega} \sqrt{|\nabla f(x)|^2 + \epsilon} dx \approx \int_{\Omega} |\nabla f(x)| dx \quad (3)$$

Discretization. The discretization of the images will be on the pixel level to prevent loss of information. Let x represent a discretization of Ω , and let D , S and Q represent the discretized counterparts of the distance measure (\mathcal{D}), elastic (\mathcal{S}) and total variation (\mathcal{Q}) regularizers respectively. H is the discretized degradation operator, and corresponding images and grids will now be represented by the discrete vectors y_i , f , and u_i . The discretized objective functional will be

$$\begin{aligned} \mathcal{J}[u, f] &= \sum_{i=1}^m D[y_i, H[Pu_i]] + \sum_{i=1}^m \alpha_i S[u_i - u^{ref}] + \beta Q[f] \\ &= \frac{1}{2} \sum_{i=1}^m \|y_i - H[Pu_i]\|^2 + \sum_{i=1}^m \alpha_i S[u_i - u^{ref}] + \beta Q[f]. \end{aligned}$$

Note that the linear operator P was introduced. This is required as elastic regularization is designed to operate on staggered grids, where coordinates are obtained by the edges of a grid. However, pixel intensity are recorded on a grid's cell-centers. The operator P will convert any

staggered grid used for distance measure to a cell-centered grid in order to preserve consistency.

Optimization. Limited-memory Broyden-Fletcher-Goldfarb-Shanno algorithm (ℓ -BFGS) will be the choice of optimization scheme as it is known to be well balanced between rate of convergence and computational requirements. ℓ -BFGS requires the Jacobian of the objective functional to be explicitly defined, hence the partial derivatives are computed:

$$\begin{aligned} \frac{\partial J}{\partial u_i} &= \frac{\partial D}{\partial r_i} \frac{\partial f}{\partial (Pu_i)} \frac{\partial (Pu_i)}{\partial u_i} + \alpha_i \frac{\partial S}{\partial u_i} \\ &= (r_i^T)(-H) \left(\frac{\partial f}{\partial (Pu_i)} \right) (P) + \alpha_i \frac{\partial S}{\partial u_i} \end{aligned}$$

for all $i = 1, \dots, m$, and

$$\begin{aligned} \frac{\partial J}{\partial f} &= \sum_{i=1}^m \frac{\partial D}{\partial f} \frac{\partial r_i}{\partial f} + \beta \frac{\partial Q}{\partial f} \\ &= \sum_{i=1}^m (r_i^T)(-H) + \beta dQ[f] \end{aligned}$$

where

$$r_i = y_i - Hf[Pu_i]$$

Experimental Results

The first experiment involves two 32×32 images with a black circle in a white background of radii 15 and 13. This experiment is to observe the behaviour when the algorithm is applied to a simple object/shape. The experiment was done using the parameters $\alpha = 500$, $\beta = 200$, $\lambda = 0$, $\mu = 1$, $\epsilon = 0.1$ with 100 iterations. The test results are shown in Figure 1 below.

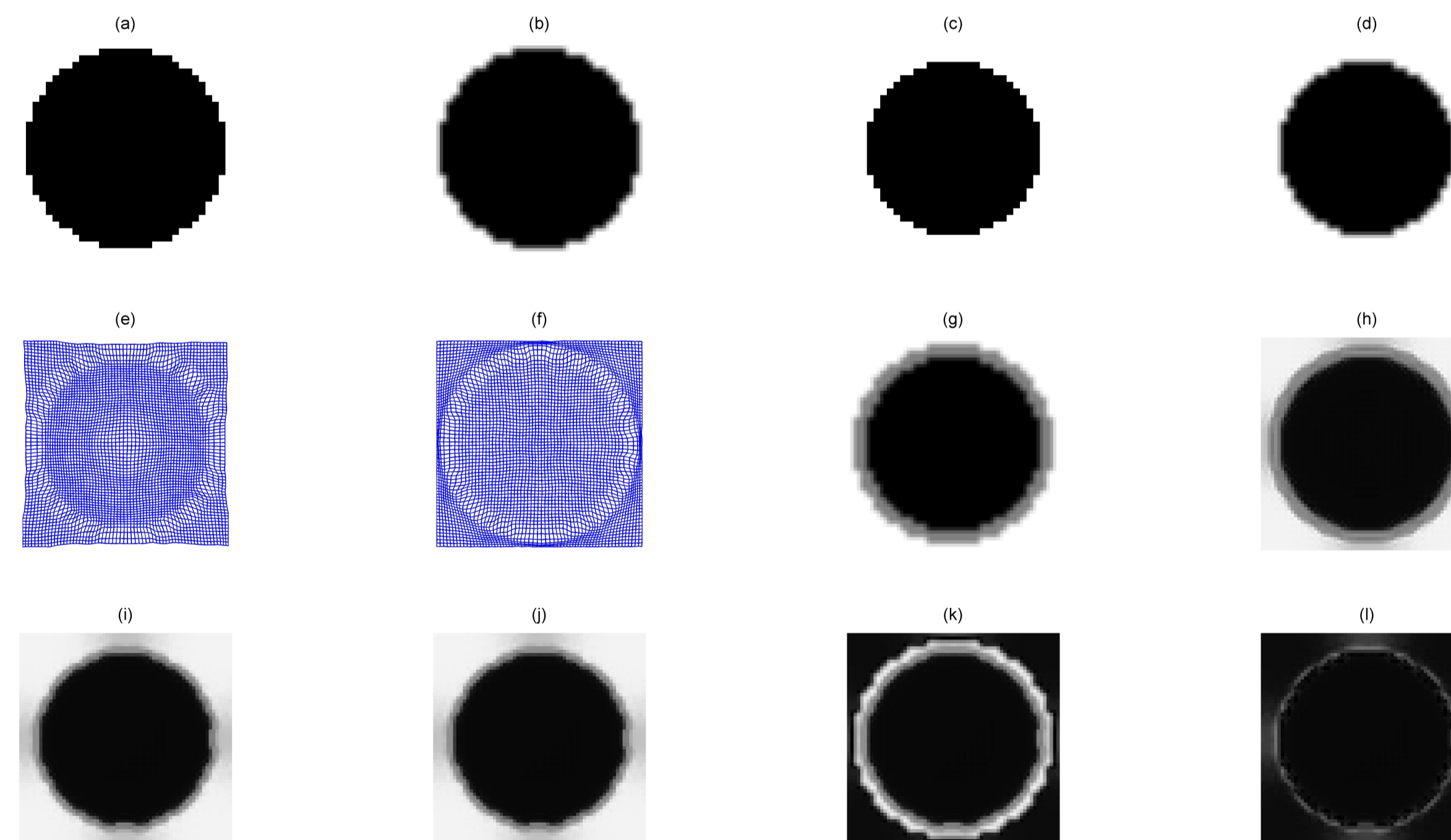


Figure 1: Experimental results using circles of radii 15 and 13 with a zooming factor of 2: (a) y_1 (Source frame #1) (b) y_1 with doubled resolution through bilinear interpolation (c) y_2 (Source frame #2) (d) y_2 with doubled resolution through bilinear interpolation. (e) u_1 (f) u_2 (g) Average of y_1 and y_2 (h) Computed image f (i) $f[Pu_1]$ (j) $f[Pu_2]$ (k) $|y_1 - Hf[Pu_1]|$ (l) $|y_2 - Hf[Pu_2]|$

The second experiment uses a 60-frame dataset obtained from a surveillance camera. Figure 2 displays a comparison between noiseless and noisy frames using the parameters $\alpha = 300$, $\beta = 75$, $\lambda = 0$, $\mu = 1$, $\epsilon = 0.1$ with 100 iterations. The bottom row are results obtained from the same image sequence with Gaussian noise of $\sigma = 1$ applied. Frame 26 of 60 is used as a reference frame for visual comparison.

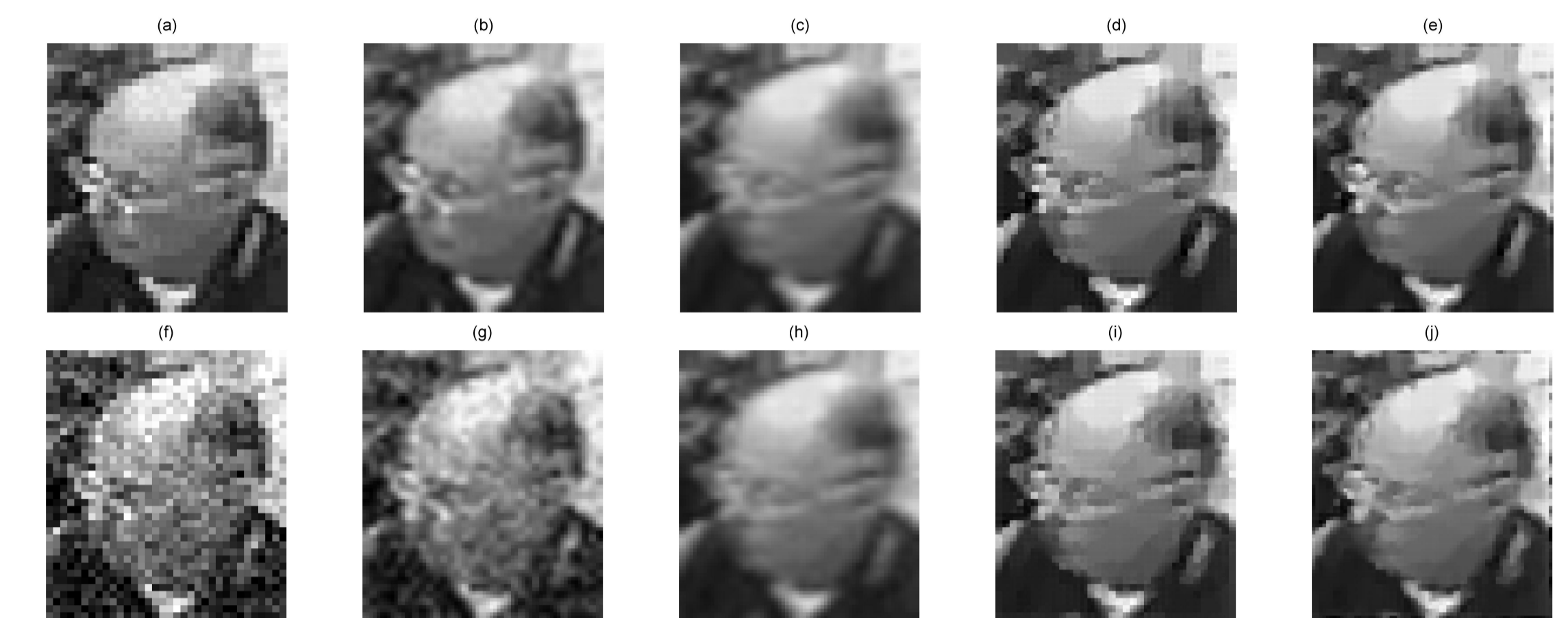


Figure 2: Experimental results using 60 frames of noiseless (top row) and noisy (bottom row) images with a zooming factor of 2: (a, f) y_{26} (Source frame #26). (b, g) y_{26} with doubled resolution through bilinear interpolation. (c, h) Average of the 60 input frames. (d, i) Computed image f . (e, j) $f[Pu_{26}]$

Conclusion and Future Work

In Figure 1, it can be observed that the algorithm was able to smooth out any rough edges that arise from direct interpolation of LR images as show in image (h). In addition, grids (e) and (f) show that the algorithm was able to estimate the motion between individual frames. Ideally, the dissimilarities between the initial images and the ideal image under their corresponding transformations should equal zero (i.e. a pitch black image). Images (k) and (l) show that some small differences still exists between the images and it is believed that this can be improved by further adjusted the parameters used for the tests.

A real life example was used for the second experiment as shown in the series of images in Figure 2. As shown in images (d) and (i), the algorithm was able to produce results, using noisy inputs, that are highly similar to the results produced by its noiseless counterpart. However, the transformed grids for each frame appeared more chaotic for noisy frames compared to noiseless frames. This was expected as it is visually, and computationally more difficult to detect true motion when there are noise in the source images.

Future works will include searching for more suitable values of α and β , as α being too large restricts motion while β being too large blurs the computed image. There is also suspicion that the solution found were not the global minimum, but a local minimum of the objective functional. Techniques such as multi-level image registration may provide better initial conditions to prevent the solution from being trapped in a local minimum. Different optimization schemes such as Steepest Descent (SD) can also be explored to circumvent this issue. Super-resolution with images of different modalities can be explored by using different distance measure methods such as normalized gradient field (NGF), normalized cross-correlation (NCC), and mutual information (MI). The proposed algorithm provides a flexible framework that enables simple implementation of these methods, and can be further expanded to include reconstruction of super-resolved video sequences as well.

References

- [1] M. Ebrahimi and A. Lausch and A.L. Martel. A Gauss-Newton Approach to Joint Image Registration and Intensity Correction. *Computer Methods and Programs in Biomedicine*, Vol. 112(3): pp 398-406, 2013.
- [2] M. Ebrahimi and E. Vrscay and A.L. Martel. Coupled Multi-Frame Super-Resolution with Diffusive Model and Total Variation Regularization. *International Workshop on Local and Non-Local Approximation in Image Processing, LNLA*, pp 62-69, 2009.
- [3] J. Modersitzki. *(FAIR) Flexible Algorithms for Image Registration*. SIAM, 2009.
- [4] J. Nocedal and S. J. Wright. *Numerical Optimization*. Springer, 2nd edition, 2006.
- [5] C. R. Vogel. *Computational Methods for Inverse Problems*. SIAM, 2002.

Acknowledgements

This research was supported in part by Natural Sciences and Engineering Research Council of Canada (NSERC) through an NSERC Undergraduate Student Research Award (USRA) for Eric Ng and an NSERC Discovery Grant for Mehran Ebrahimi.