MULTI-MODAL IMAGE REGISTRATION

Amanda Gaudreau-Balderrama

Boston University
Department of Electrical and Computer Engineering
8 Saint Mary's Street
Boston, MA 02215
www.bu.edu/ece

May 25, 2012

ECE-2012-04

Contents

List of Figures

List of Tables

1 Background and Introduction

The mechanisms behind most neurodegenerative diseases are poorly understood. The lack of suitable non-invasive methods for diagnosing and characterizing conditions such as Alzheimer's disease and traumatic brain injury have hindered the ability to effectively treat and monitor these illnesses. Understanding and quantifying the pathological effects of these conditions could potentially facilitate the development of drugs targeted towards specific biological areas and phenomenon. Metals such as Zinc and Lithium have been used in treatment of neurological diseases, but the mechanisms of action of such drugs are still unknown. A technique known as metallomic spectral imaging (MSI) enables the generation of a 2-D metallomic map of a given sample via optical or mass spectrometry. These images provide quantitative data for characterizing the elemental distribution of biological samples with extremely high spatial resolution (on the order of tens of microns).

Ideally, relevant features of the MSIs will be identified and quantitatively analyzed. For biological specimens, relevant features usually correspond to anatomical areas of interest. Due to the complexity of the MSI data acquisition process, the resulting images are characterized by blurred edges, unexpected/unpredictable features, and unique noise patterns. The varying properties of each individual elemental distribution within a sample as well as the other limitations in the quality of the data (i.e. low signal-to-noise ratio, inability to compare MSI to ground truth) make it difficult to perform reliable analysis using the MSIs alone. For this reason, photographic images can be used to guide MSI analysis.

In the proposed project, photos will be registered to their MSI realizations. Spatial properties and features identified in the photos can be used to adjust and analyze properties of the MSIs. This multimodal registration-segmentation technique is vital for correlation and feature analysis across distinct samples and could be used for anatomical labeling and accurate spatial representation of the data. Using staining techniques or "expert" segmentation, brain histology and structures can be identified and used to precisely localize the corresponding regions in the MSIs using the photo images. Further, techniques developed for this multi-modal image registration can be applied between other image types (such as registration of an MSI to a labeled atlas cartoon, for instance). Challenges include differences in capture regions, data dynamic range, relative data magnitudes, and sampling lattices between the MSIs and the photo.

2 Literature Review: Medical Image Registration

Multi-modal image registration is an important aspect of medical image analysis. Different modalities, such as Magnetic Resonance Imaging (MRI), Computed Tomography (CT), and Positron Emission Tomography (PET), show unique tissue features at different spatial resolutions. Whether registering images across modalities for a single patient or registering across patients for a single modality, registration is an effective way to combine information from different images into a normalized frame of reference. Registered datasets can be used to provide information about the structure, function, and pathology of the organ or individual being imaged.

Registration can be thought of as the optimization of a similarity metric over the set of possible transformations. This alludes to the three necessary aspects which must be defined for image registration: a transformation model, a similarity metric, and an optimization method [10]. Depending on the desired robustness and accuracy of the registration outcome, different definitions for these aspects may be used. The transformation model can be either rigid or non-rigid. If the images to be registered are the product of a technique which introduces only rigid distortions, then a combination of translation, rotation and scaling,

which together define affine motion, would be a sufficient transformation model. Non-rigid models, such as perspective and parametric, are used to describe more general object deformations between images.

Similarity metrics which are most widely used in medical image registration include normalized cross correlation [10], correlation ratio [9], and mutual information [3][14]. Normalized cross correlation and correlation ratio assume a functional dependence between the intensity values of the two images, where the former assumes a linear relationship and the latter permits any form of functional relationship [2]. Mutual information, on the other hand, is a metric adapted from information theory and is a statistical measure which indicates the amount of information one random variable (image) gives about another [3]. Improvements to the mutual information metric have been proposed by incorporating spatial information, such as gradient magnitudes and orientations [7][6] and other statistical spatial dependencies [16].

An optimization method must be selected to search the parameter space for a set of transformation parameters which optimally align the two images according to the specified similarity metric. Non-derivative and derivate based optimization methods exist which iteratively select parameters such that the similarity metric is minimized. In the case of normalized cross correlation and mutual information, this will mean minimization of the negative value their functional definitions. Optimization methods typically converge on local extrema since the hypersurface defined by the similarity metric tends to be high-dimensional [10] and is not always a convex function [15]. For this reason, initial estimates or initial approximate alignment are necessary for successful registration when using non-global optimization methods.

Of equal importance is the interpolation method used. The transformation of a moving image to a fixed image (also called test/target, floating/reference) often results in non-integer grid locations specified for the moving image. This requires interpolation to compute the moving image's value at the specified non-integer locations. Netsch *et al.* [5] concluded that quadratic and cubic interpolators produce significantly better registration results over linear interpolators, but that higher order interpolators do not perform significantly better.

3 Problem Statement: MSI and Photo Registration

As discussed in Section 2, a transformation model, similarity metric, and optimization method were selected. Since the registration was performed using the MATLAB function `imregister`[1], the similarity metric and optimization methods were predefined using the `imregconfig` function. Registration of the MSI and photo was done using an affine transformation model with bicubic interpolation. Optimal registration was determined by maximization of mutual information (as defined by Mattes *et al.* [4]) using a (1 + 1) evolution strategy [11]. Because the exact operation of the function is not documented and the m-file could not be accessed[2], the general theory regarding mutual information and referenced optimization strategy will be provided. Details regarding implementation in MATLAB will be given in Section 4.

3.1 Affine Transformation Model

Consider two registered data sets, $\mathcal{F}(\boldsymbol{p})$ and $\mathcal{M}^r(\boldsymbol{p})$ sampled on $\boldsymbol{p} \in \Re^2$. The data sets may represent two imaging modalities of the same underlying object. Given that $\mathcal{F}(\boldsymbol{p})$ and $\mathcal{M}(\boldsymbol{q})$ are observed, we can define

[1] Part of the Image Processing Toolbox

[2] Implementation of optimization algorithm and similarity metric are contained in a file `regmex.mexw64` which cannot be viewed since the source file is unavailable

the relationship between $\mathcal{M}^r(\boldsymbol{p})$ and $\mathcal{M}(\boldsymbol{q})$ as follows:

$$\mathcal{M}^r(\boldsymbol{p}) = \mathcal{M}\left(A(\boldsymbol{q})\right) \qquad \Leftrightarrow \qquad \mathcal{M}(\boldsymbol{q}) = \mathcal{M}^r(A^{-1}(\boldsymbol{p})) \tag{1}$$

where A is an affine transformation. The fixed image, $\mathcal{F}$, is defined on image coordinates $\boldsymbol{p}$ and the moving image, $\mathcal{M}$, is defined on image coordinates $\boldsymbol{q}$. $\mathcal{M}^r$ is the registered moving image. The goal is to find $\hat{A} \approx A$ (or equivalently its inverse) which provides a mapping from $\boldsymbol{q}$ to $\boldsymbol{p}$ (see Eqn (2)) such that it maximizes the mutual information (see Eqn (6)). The affine transformation A is the product of four geometric transformations: translation (in x and y), rotation, scaling (in x and y) and skew. The relationship between the geometric parameters $\boldsymbol{\alpha} = \{t_x, t_y, \theta, s_x, s_y, k\}$ and the transformation parameters $\boldsymbol{a} = [a_1, a_2, a_3, a_4, a_5, a_6]^T$, as well as the individual transform matrices are given in Eqn (3).

$$\boldsymbol{p} = A\boldsymbol{q} \tag{2}$$

$$A_{\boldsymbol{\alpha}} = \underbrace{\begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix}}_{translation} \cdot \underbrace{\begin{bmatrix} \theta_c & -\theta_s & 0 \\ \theta_s & \theta_c & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{rotation} \cdot \underbrace{\begin{bmatrix} 1 & k & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{skew} \cdot \underbrace{\begin{bmatrix} s_x & 0 & 0 \\ 0 & s_y & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{scaling} = \begin{bmatrix} s_x\theta_c & s_y(k\theta_c - \theta_s) & t_x \\ s_x\theta_s & s_y(k\theta_s + \theta_c) & t_y \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} a_1 & a_2 & a_3 \\ a_4 & a_5 & a_6 \\ 0 & 0 & 1 \end{bmatrix} \tag{3}$$

where

t_x	=	positive value shifts image to the left
t_y	=	positive value shifts image up
θ	=	rotation angle, measured counterclockwise from the x-axis ($\theta_c = \cos(\theta)$ and $\theta_s = \sin(\theta)$)
k	=	shear factor along the x-axis = tan(skew angle) (the skew angle is measured from the y-axis)
s_x	=	change of scale in x direction
s_y	=	change of scale in y direction

It is important to note that the above expression for the affine transformation matrix A in Eqn (3) corresponds to a transformation with respect to the center pixel, (x_c, y_c). This transformation can easily be written as a transformation applied to the top-left pixel with a different definition for the translation parameters (B). This derivation is shown in Eqn (4). It should also be noted that while transformations were applied using a custom written function, `transform_image`, the built in MATLAB transformation function uses distinct definitions for the semantic meaning of the geometric parameters.

$$\begin{aligned}
\boldsymbol{p} &= A \cdot \boldsymbol{q}_{(x_c,y_c)} = \begin{bmatrix} a_1 & a_2 \\ a_4 & a_5 \end{bmatrix} \cdot \begin{bmatrix} q_x - x_c \\ q_y - y_c \end{bmatrix} + \begin{bmatrix} a_3 + x_c \\ a_6 + y_c \end{bmatrix} = \begin{bmatrix} a_1q_x + a_2q_y + a_3 - a_1x_c - a_2y_c + x_c \\ a_4q_x + a_5q_y + a_6 - a_4x_c - a_5y_c + y_c \end{bmatrix} \\
\boldsymbol{p} &= \begin{bmatrix} b_1 & b_4 & 0 \\ b_2 & b_5 & 0 \\ b_3 & b_6 & 1 \end{bmatrix}^T \boldsymbol{q}_{(1,1)} = \underbrace{\begin{bmatrix} a_1 & a_2 & a_3 - a_1x_c - a_2y_c + x_c \\ a_4 & a_5 & a_6 - a_4x_c - a_5y_c + y_c \\ 0 & 0 & 1 \end{bmatrix}}_{=B^T} \cdot \begin{bmatrix} q_x \\ q_y \\ 1 \end{bmatrix} \\
B &= \begin{bmatrix} a_1 & a_4 & 0 \\ a_2 & a_5 & 0 \\ a_3 - a_1x_c - a_2y_c + x_c, & a_6 - a_4x_c - a_5y_c + y_c, & 1 \end{bmatrix} = \begin{bmatrix} b_1 & b_4 & 0 \\ b_2 & b_5 & 0 \\ b_3 & b_6 & 1 \end{bmatrix} \\
B_{\boldsymbol{\alpha}} &= \begin{bmatrix} s_x\theta_c & s_x\theta_s & 0 \\ s_y(k\theta_c - \theta_s) & s_y(\theta_c + k\theta_s) & 0 \\ t_x + x_c + s_yy_c(\theta_s - k\theta_c) - s_x\theta_cx_c, & t_y + y_c - s_yy_c(\theta_c + k\theta_s) - s_x\theta_sx_c, & 1 \end{bmatrix}
\end{aligned} \tag{4}$$

3.2 Mutual Information

Mutual information (MI) is an entropy-based metric adapted from information theory which represents the amount of shared information two random variables have about one another in terms of bits. Typically, the entropy definition used is the Shannon entropy given by $H = -\sum_i p_i \log_2 p_i$, but other entropy definitions can be used (such as Jumarie or Rényi). Several interpretations of MI provide different perspectives to the same underlying statistical measure [8]. Mutual information, J, between the overlapping region of two random variables (images) C and D can be interpreted as follows:

1. $J(C, D) = H(D) - H(D|C)$: indicates "the amount by which the uncertainty about D decreases when C is given: the amount of information C contains about D" [8]
2. $J(C, D) = H(C) + H(D) - H(C, D)$: theoretically $H(C)$ and $H(D)$ should remain constant despite registration, however, the joint entropy will change. Through this interpretation we see that minimization of joint entropy, another similarity metric used in medical image registration, promotes maximization of J. The independent entropies prevent dominance of trivial solutions (such as only aligning background).
3. $J(C, D) = \sum_{c,d} p(c, d) \log_2 \left(\frac{p(c,d)}{p(c)p(d)}\right)$: measures the dependence between random variables.

Registration of $\mathcal{F}$ and $\mathcal{M}$ maximizes MI so that when they align, the amount of information they contain about each other is maximal. One challenge associated with using MI in derivative based optimization methods is that the metric must be posed as a continuous function in order to enable differentiation of the function. This requires both the images and the estimated probability distributions to be continuous, which can be achieved by: using a B-spline basis to represent $\mathcal{M}$, using a Parzen window to estimate the joint probability distribution $p_{\mathcal{F},\mathcal{M}}$, and using cubic B-splines for implementing deformations [4][13][12]. Derivation of the differentiation of the MI is given in Appendix A.

The definition used for MI in `imregister` is given by [4]:

$$J\left(\mathcal{F}(\boldsymbol{p}), \mathcal{M}(A \cdot \boldsymbol{q})\right) = \sum_{\ell \in L_{\mathcal{M}}} \sum_{\kappa \in L_{\mathcal{F}}} p_{\mathcal{F},\mathcal{M}}(\ell, \kappa; \boldsymbol{a}) \log_2 \left(\frac{p_{\mathcal{F},\mathcal{M}}(\ell, \kappa; \boldsymbol{a})}{p_{\mathcal{F}}(\kappa) p_{\mathcal{M}}(\ell; \boldsymbol{a})} \right) \tag{5}$$

where $L_{\mathcal{F}}$ and $L_{\mathcal{M}}$ are the discrete sets of intensities associated with the fixed and moving images, respectively, and $p_{\mathcal{F},\mathcal{M}}$, $p_{\mathcal{F}}$, and $p_{\mathcal{M}}$ are the joint, fixed marginal, and moving marginal probability distributions, respectively.

3.3 Optimization

The optimal parameters of the transformation, $\hat{\boldsymbol{a}}$, are determined according to the maximization of mutual information as follows:

$$\hat{\boldsymbol{a}} = \arg\max_{\boldsymbol{a}} \; J\left(\mathcal{F}(\boldsymbol{p}), \mathcal{M}(A \cdot \boldsymbol{q})\right) = \arg\max_{\boldsymbol{a}} \sum_{\ell \in L_{\mathcal{M}}} \sum_{\kappa \in L_{\mathcal{F}}} p_{\mathcal{F},\mathcal{M}}(\ell, \kappa; \boldsymbol{a}) \log_2 \left(\frac{p_{\mathcal{F},\mathcal{M}}(\ell, \kappa; \boldsymbol{a})}{p_{\mathcal{F}}(\kappa) p_{\mathcal{M}}(\ell; \boldsymbol{a})} \right) \tag{6}$$

where $J\left(\mathcal{F}(\boldsymbol{p}), \mathcal{M}(A \cdot \boldsymbol{q})\right)$ is defined in Eqn (5). This problem is solved using the (1+1) evolutionary strategy (ES) proposed by Styner *et al.* [11]. In summary, each optimization step begins with a set of parameter vectors $\{\hat{\boldsymbol{a}}_1, \hat{\boldsymbol{a}}_2, \ldots, \hat{\boldsymbol{a}}_i\}$, called a population, where each parameter vector $\hat{\boldsymbol{a}}_i$ represents an individual. At the beginning of each optimization step, the population will have i individuals. Each

individual is randomly mutated by a vector of multidimensional normal distributions, with mean $\mu = \hat{\boldsymbol{a}}_{parent}$ and covariance matrix Σ^2, to produce a new individual (child). A new population of size $2i$ is formed by combining the children and the parents. The fitness of each individual in the population is then assessed according to $J(\mathcal{F}(\boldsymbol{p}), \mathcal{M}(A_{\hat{\boldsymbol{a}}_i}\boldsymbol{q}))$ such that individuals with the highest fitness (largest J) are retained for the following optimization step. The covariance matrix Σ^2 is adapted at each optimization step according to whether the fitness of the individuals in the population increased or decreased. For a detailed account of the (1+1) ES, the reader is referred to [11, Appendix B, pp.163-4]. This method boasts an ability to step out of non-optimal minima due to the random nature of the parameter variation. The number of individuals used MATLAB's implementation of this optimizer is not given.

Oftentimes optimization is enhanced by the use of image pyramids for hierarchical image registration. While the documentation for `imregister` states that the function does perform multiresolution optimization, the function's documentation does not detail the method used. Presumably, it implements a scheme similar to that used in the `impyramid` function, which computes a Gaussian pyramid reduction or expansion of the input image. The flow of multiresolution optimization using a 3-level pyramid is shown in Figure 1 [1, Figure 1]. Multiresolution schemes increase the chance that the initial parameter estimate will be within a convergence basin of the global minimum of the similarity metric, J [1].

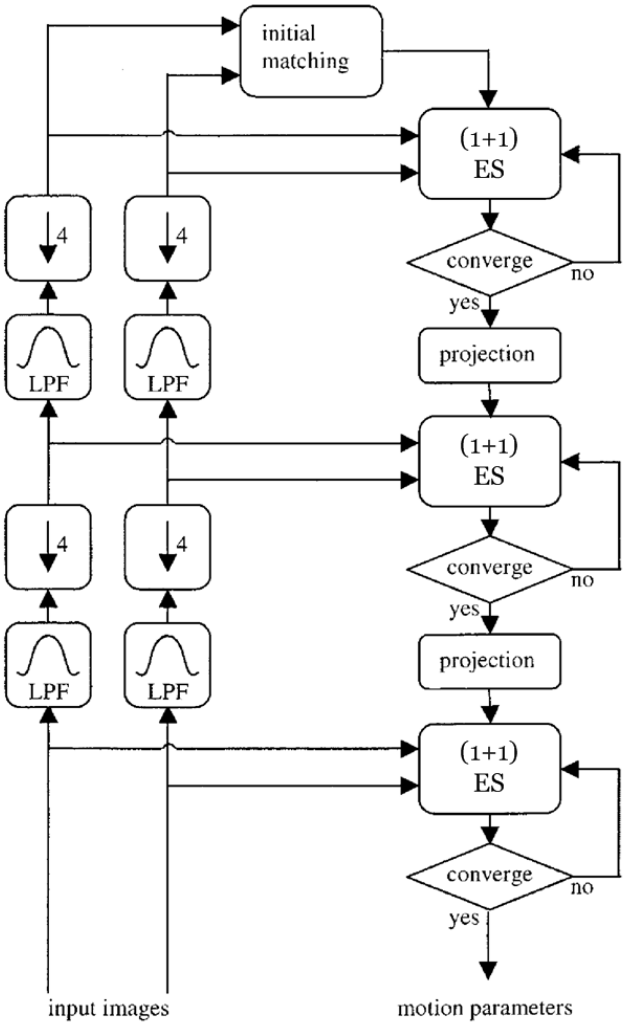

Figure 1: Multiresolution optimization using a 3 level pyramid (copied from [1], Figure 1)

4 Implementation: MSI and Photo Registration in MATLAB

As mentioned in Section 3, the MATLAB function `imregister` was used with the optimizer and metric specified according to the command `[optimizer, metric] = imregconfig('multimodal')`. The resulting variable definitions and field values are given on Table 1. Generally, a registration function would take as inputs the moving and fixed images, and would give as an output the optimal parameters which align the moving image to the fixed image. While attempting to utilize the `imregister` function, it was found that the output transformation matrix, T, was structured differently than the expected transformation matrix A . The discrepancy between expected and actual outputs is depicted in Figure 2. Details regarding characterization of the `imregister` function will be given in Section 4.1. Section 4.2 addresses the operating assumptions for the image registration process and Section 4.3 provides the algorithm. The MATLAB function order and hierarchy is described in Section 4.4.

Var. Name	Data Type	Fields	Value
`optimizer`	`OnePlusOneEvolutionary`	`GrowthFactor`	1.05
		`Epsilon`	$1.5e-6$
		`InitialRadius`	0.00625
		`MaximumIterations`	1000
`metric`	`MattesMutualInformation`	`NumberOfSpatialSamples`	500
		`NumberOfHistogramBins`	50
		`UseAllPixels`	1

Table 1: Definitions of output structures from `imregconfig('multimodal')`

4.1 Characterizing `imregister` Output

When developing the registration function flow, two photos of copper grids were input into `imregister`. One of the images had undergone a known transformation dictated by geometric parameters $\boldsymbol{\alpha}$. Initially, transformation was restricted to translational and rotational motion. For these geometric transformations, the output of `imregister`, T, correctly gave the inverse transformation $T = A^{-1}$ such that $\mathcal{F}(\boldsymbol{p}) \approx \mathcal{M}^r(\boldsymbol{p}) = \mathcal{M}(T \cdot \boldsymbol{q})$. Once scale and skew were varied, Figure 2 clearly shows that $T \neq A^{-1}$. By individually varying each parameter in $\boldsymbol{\alpha}$, the relationship between T and $\boldsymbol{\alpha}$ was gradually exposed.

The transformation matrix as a function of $\boldsymbol{\alpha}$ (which is defined with respect to the center pixel), $T_{\boldsymbol{\alpha}}$, was found to be the product of the individual transformation matrices as given in Eqn (7). Since `imregister` internally uses `imtransform`, the output T specifies a transformation with respect to the top-left pixel (Eqn (8)).

$$T_{\boldsymbol{\alpha}}^{x_c,y_c} = \underbrace{\begin{bmatrix} \theta_c & \theta_s & 0 \\ -\theta_s & \theta_c & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{rotation} \cdot \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ k & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{skew} \cdot \underbrace{\begin{bmatrix} s_y & 0 & 0 \\ 0 & s_x & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{scaling} \cdot \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ t_x & t_y & 1 \end{bmatrix}}_{translation} = \begin{bmatrix} s_y(\theta_c + k\theta_s), & s_x\theta_s & 0 \\ s_y(k\theta_c - \theta_s), & s_x\theta_c & 0 \\ t_x & t_y & 1 \end{bmatrix} \tag{7}$$

$$T_{\boldsymbol{\alpha}}^{1,1} = \begin{bmatrix} s_y(\theta_c + k\theta_s), & s_x\theta_s & 0 \\ s_y(k\theta_c - \theta_s), & s_x\theta_c & 0 \\ t_x + x_c - s_y x_c(\theta_c + k\theta_s) + s_y y_c(\theta_s - k\theta_c), & t_y + y_c - s_x\theta_s x_c - s_x\theta_c y_c, & 1 \end{bmatrix} = \begin{bmatrix} t_1 & t_4 & 0 \\ t_2 & t_5 & 0 \\ t_3 & t_6 & 1 \end{bmatrix} \tag{8}$$

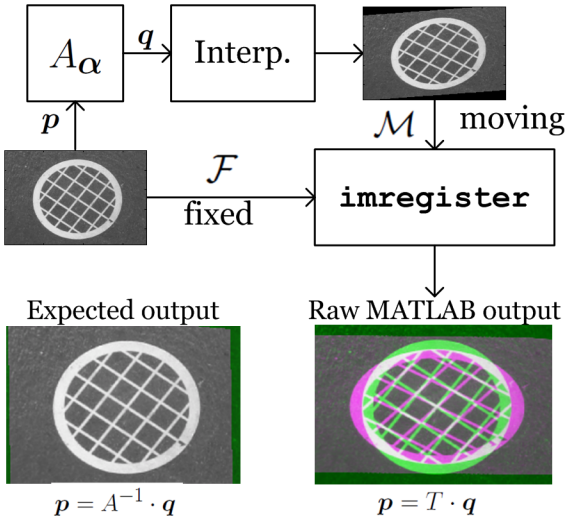

Figure 2: Example showing discrepancy between expected and actual output from `imregister`

From Eqn (8), we have 6 equations (t_i = function of $\boldsymbol{\alpha}$) and 6 "unknowns" ($\boldsymbol{\alpha}$), so the value of each geometric parameter can be solved for explicitly as follows:

$$\begin{aligned} t_4 &= s_x \sin(\theta) \\ t_5 &= s_x \cos(\theta) \end{aligned} \quad \Rightarrow \quad \boxed{s_x = \frac{t_5}{\cos(\theta)}} \quad \Rightarrow \quad t_4 = t_5 \frac{\sin(\theta)}{\cos(\theta)} = t_5 \tan(\theta) \quad \Rightarrow \quad \boxed{\theta = \tan^{-1}\left(\frac{t_4}{t_5}\right)}$$

$$\begin{aligned} t_1 &= s_y(\theta_c + k\theta_s) \\ t_2 &= -s_y(\theta_s - k\theta_c) \end{aligned} \quad \Rightarrow \quad \boxed{s_y = \frac{t_1}{\theta_c + k\theta_s}} \quad \Rightarrow \quad t_2 = \frac{-t_1(\theta_s - k\theta_c)}{\theta_c + k\theta_s} \quad \Rightarrow \quad \boxed{k = \frac{t_2\theta_c + t_1\theta_s}{t_1\theta_c - t_2\theta_s}}$$

$$\boxed{t_x = t_3 + \frac{N}{2}\left(s_y(\theta_c + k\theta_s) - 1\right) - \frac{M}{2} s_y(\theta_s - k\theta_c)} \qquad \boxed{t_y = t_6 + \frac{M}{2}\left(s_x\theta_c - 1\right) + \frac{N}{2} s_x\theta_s} \tag{9}$$

where $[M, N] = \texttt{size}(I)$. As shown in Figure 3, by incorporating a block in the system which takes as an input the solution transformation matrix T from `imregister` and outputs the inverse of the original transformation, the registered moving image, $\mathcal{M}^r(\hat{\boldsymbol{p}})$, closely matches the fixed image, thereby empirically verifying the interpretation of T.

4.2 Assumptions

Two primary assumptions were made in order to successfully register photos and MSI. Due to the fact that the optimizer does not perform global optimization, one of the key assumptions dictating the success of the final registration is that the MSI and the photo have similar orientations and capture regions. An example of the registration output when this assumption is violated (Experiment 5f) shows that the solution converges on a sub-optimal minima, resulting in a clearly incorrect mapping from $\boldsymbol{q}$ to $\boldsymbol{p}$ (Figure 9). Another key assumption is that the moving image yields itself well to interpolation. As suggested in the algorithm below, the MSI always serves as the fixed image for two reasons. For one, the spatial intensity of the MSI tend to be much less uniform than that of the photos. This is because MSI are more susceptible to noise which is inherent in the acquisition process. Additionally, maximal MSI intensity values can be significantly greater in magnitude (up to 10^4 times) than those of photos. In line with this assumption, it was also empirically found that maintaining a "reasonable" aspect ratio (ar) was important. For instance,

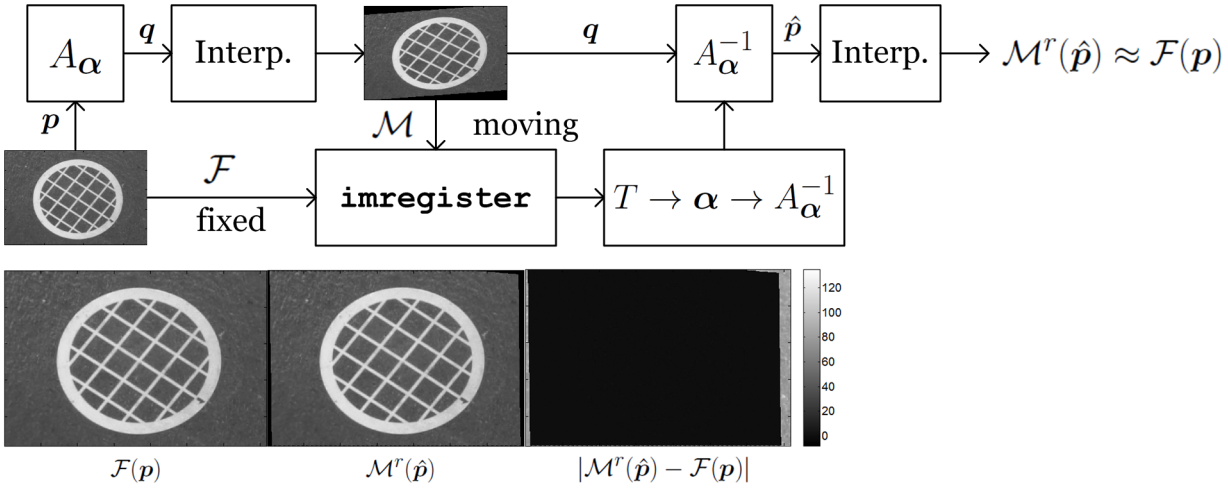

Figure 3: Functional registration block diagram

when a photo with an original *ar* of 2:3 was rescaled to an *ar* of 1:12 (Experiment 3f, Figure 6), the registration could not be performed. However, when the *ar* was maintained (Experiment 3, Figure 7), registration was successful.

4.3 Algorithm

The following algorithm was used for MSI-photo registration:

1. Resize images so they have matching dimensions (if *ar* of the MSI is less than 1:5, then resize to photo dimensions)
2. Define `imregister` inputs
   (a) `fixed = MSI; moving = photo`
   (b) Motion model: `'affine'`
   (c) `[optimizer, metric] = imregconfig('multimodal')`
3. Run `[moving_reg,T] = imregister(fixed,moving,'affine',optimizer,metric);`
   - Note that `moving_reg` is a meaningless output that is later redefined (see Figure 2, Raw MATLAB Output)
4. Use `T` to estimate the geometric parameters, $\boldsymbol{\alpha}_{est}$
5. Use $\boldsymbol{\alpha}_{est}$ to generate the transformation matrix B
6. Apply B to photo (`moving`) to give output of registered moving image, $\mathcal{M}^r$

4.4 MATLAB Function Hierarchy

Experiments were run using a central function called `EC720reg`. Variable inputs are allowed into this function to allow for versatility, but the function contains the necessary information such that by running

the command below, the outputs shown in Figures 4-12 will be output.

```
[reginfo, fixed_reg ,moving_reg] = EC720reg([],MSInum,'fnum',photonum,'exp',expnum,);
```

where `MSInum` is the MSI number ($\mathcal{F}_{\#}$ on Table 2), `photonum` is the photo number ($\mathcal{M}_{\#}$ on Table 3), and `expnum` corresponds to the experiment number on Table 4. From this function, the functional flow is as follows:

- `[alpha_est, fixed_reg, moving_reg] = register_image(fixed,moving);`
  - `[optimizer, metric] = imregconfig('multimodal');`
  - `[moving_reg,T] = imregister(fixed,moving,'affine',optimizer,metric);`
  - `[Aest,Mest,iMest] = alpha2tmat(alpha_est,ysize,xsize);`
  - `moving_reg = transform_image(moving,Mest,'transtype','trans');`
- `imoverlay1x4(fnum,reginfo.I1.original, moving_reg, datasetnum,fixed)`

5 Experimental Results

Registration between MSI and photo was successfully applied to 7 data sets[3], where successful registration was determined qualitatively and marked by an increase in mutual information. In all experiments, $\mathcal{F}$ = `fixed` = MSI and $\mathcal{M}$ = `moving` = photo. Relevant MSI and photo information can be found on Tables 2 and 3 respectively. Experiments are numbered on Table 4 and experimental outputs are depicted in Figures 4-12. The moving image used in experiment 5, $\mathcal{M}_2^*$, corresponds to a manual rotation of the original image $\mathcal{M}_2$ by $-60°$ with respect to the center ($\mathcal{M}_2^* = \mathcal{M}_2(A_{-60°} \cdot \boldsymbol{q})$).

5.1 Reference Tables

$\mathcal{F}_{\#}$	Filename	Dim.	Element	t_a (s)	v_{scan} (μ/s)	d_{spot}
$\mathcal{F}_1$	20120322_OESCugrid1ta	170×184	Cu3273	1	20	20
$\mathcal{F}_2$	20120320_OESCugrid40umps	170×510	Cu3273	0.2	40	20
$\mathcal{F}_3$	20120321_OESCugrid0,1ta	70×862	Cu3273	0.1	50	50
$\mathcal{F}_4$	20120322_OESCugrid0,75ta	70×86	Cu3273	0.75	75	50
$\mathcal{F}_5$	20120425_OESheye	260×275	K_7664	1	50	50

Table 2: Metallomic spectral image information

[3]Modifications to step 1 of the algorithm given in Section 4.3 were sometimes required. Additionally, when the assumption that MSI and photo have similar orientation and capture region was violated (Exp 6f), manual transformation of photo was required.

$\mathcal{M}_{\#}$	Filename	Dim.
$\mathcal{M}_1$	PI_CuG1_orig	480×718
$\mathcal{M}_2$	PI_CuG2_orig	480×718
$\mathcal{M}_3$	heye1	972×1164
$\mathcal{M}_4$	eye_atlas	311×400

Table 3: Photo information

Exp #	$\mathcal{F}_{\#}$	$\mathcal{M}_{\#}$	Resized Dim.	$J(\mathcal{F},\mathcal{M})$	$J(\mathcal{F},\mathcal{M}^r)$	Reg. time (s)	$\boldsymbol{\alpha}_{est} = [t_x, t_y, \theta, s_x, s_y, k]$
1	$\mathcal{F}_1$	$\mathcal{M}_1$	$\mathcal{F}_1$	0.388	1.097	6.36	[1.46, 0.88, -1.34, 0.77,0.95,-0.01]
2	$\mathcal{F}_2$	$\mathcal{M}_1$	$\mathcal{F}_2$	0.652	0.777	17.72	[-4.65, 1.35, -0.22, 0.85,0.95,-0.01]
3f	$\mathcal{F}_3$	$\mathcal{M}_1$	$\mathcal{F}_3$	0.756	0.756	1.11	[0.00, 0.00, 0.00, 1.00,1.00,0.00]
3	$\mathcal{F}_3$	$\mathcal{M}_1$	$\mathcal{M}_1$	0.441	0.683	73.26	[-8.40, 6.61, -1.18, 0.90,0.98,0.00]
4	$\mathcal{F}_4$	$\mathcal{M}_1$	$\mathcal{F}_4$	2.219	2.310	2.41	[-2.53, 1.52, -4.30, 1.00,0.99,-0.02]
5f	$\mathcal{F}_1$	$\mathcal{M}_2$	$\mathcal{F}_1$	0.343	0.189	2.86	[38.05, -32.94, 12.58, 0.26,0.25,0.38]
5	$\mathcal{F}_1$	$\mathcal{M}_2^*$	$\mathcal{F}_1$	2.385	2.946	6.55	[-5.15, -0.14, -0.98, 0.71,1.04,0.06]
6	$\mathcal{F}_5$	$\mathcal{M}_3$	$\mathcal{F}_5$	1.085	1.446	13.42	[27.12, -10.68, -6.32, 0.76,0.82,0.01]
7	$\mathcal{F}_5$	$\mathcal{M}_4$	$\mathcal{F}_5$	0.940	0.965	12.57	[12.65, -2.93, -4.61, 0.70,0.91,-0.05]

Table 4: Experimental cross-reference and MI results

5.2 Experimental Figures

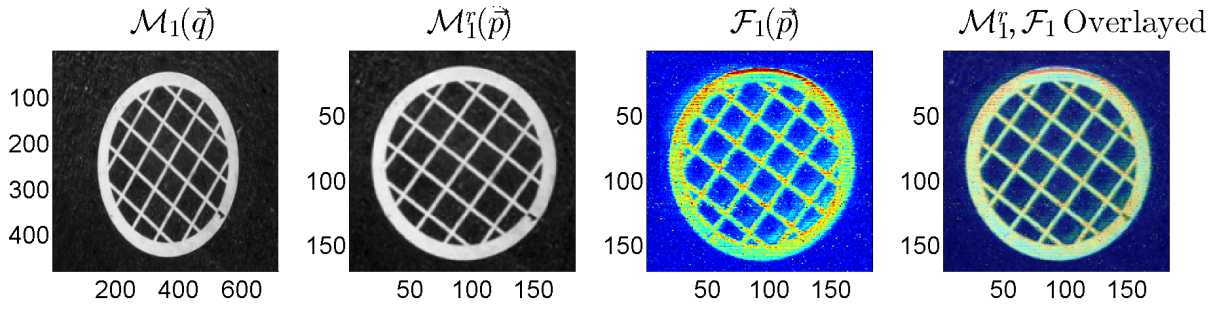

Figure 4: Experiment 1 Result

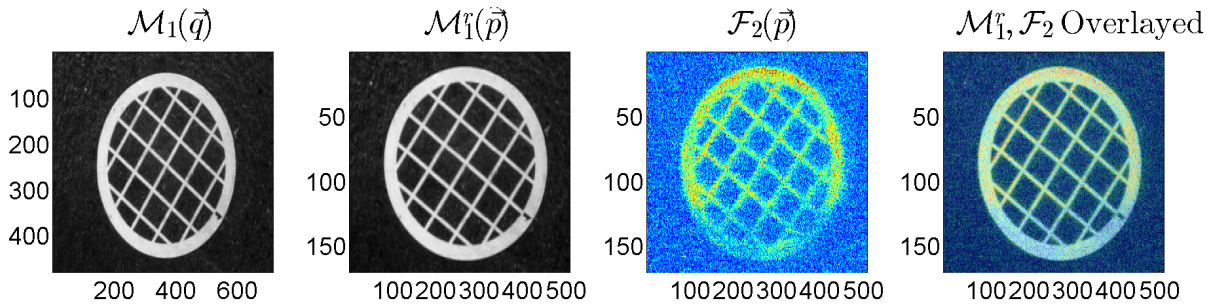

Figure 5: Experiment 2 Result

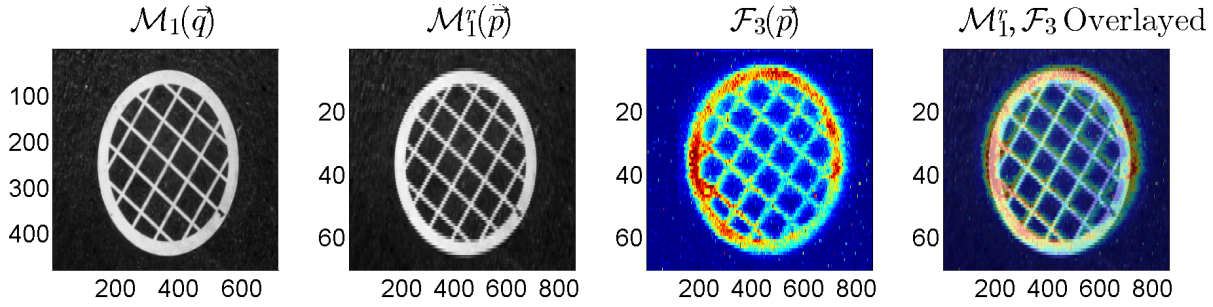

Figure 6: Experiment 3f – Failed Result

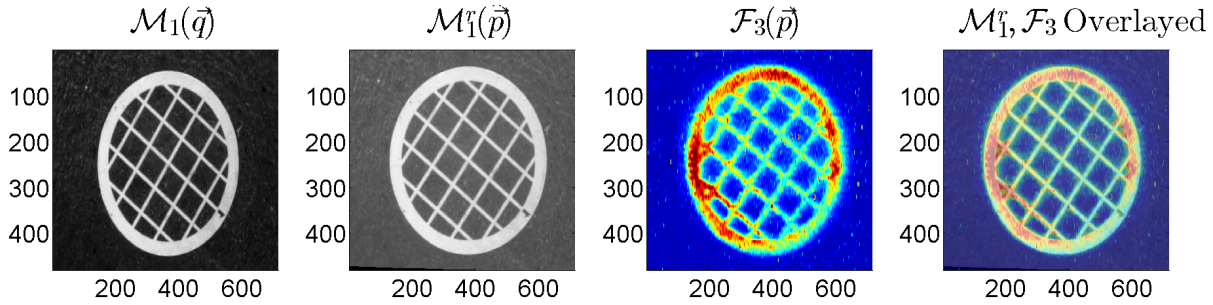

Figure 7: Experiment 3 Result

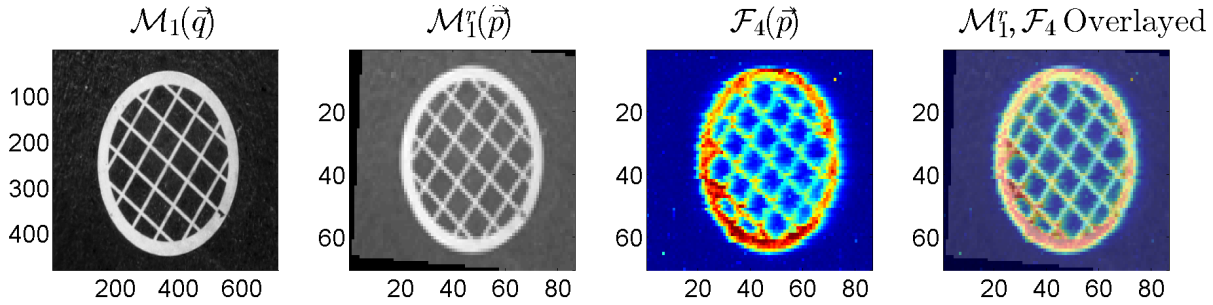

Figure 8: Experiment 4 Result

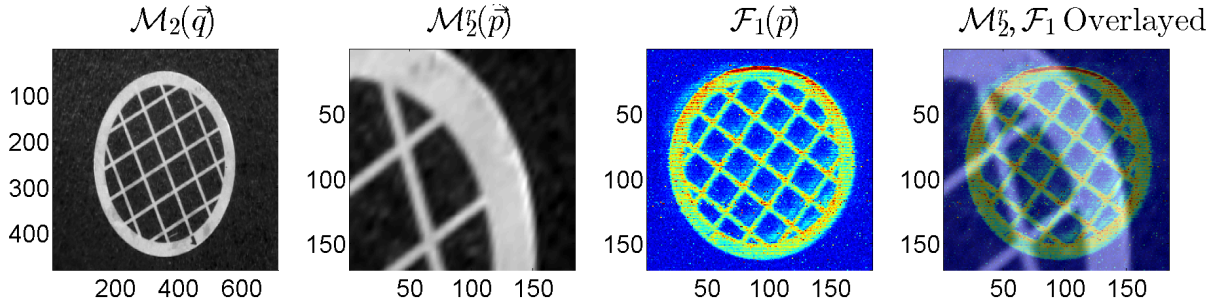

Figure 9: Experiment 5f – Failed Result

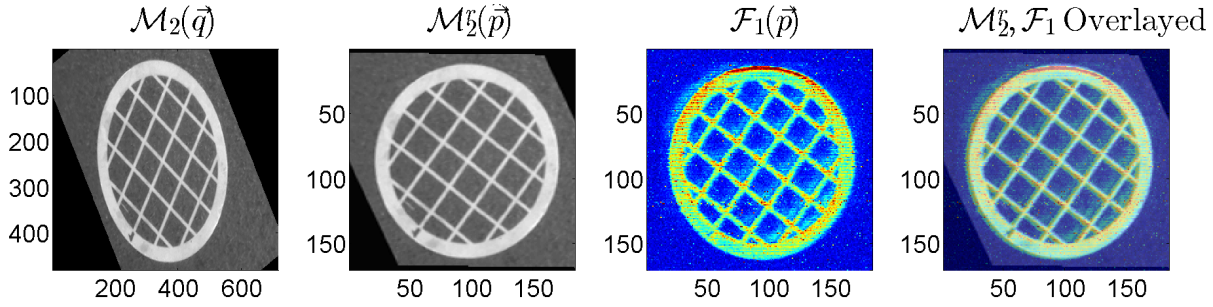

Figure 10: Experiment 5 Result

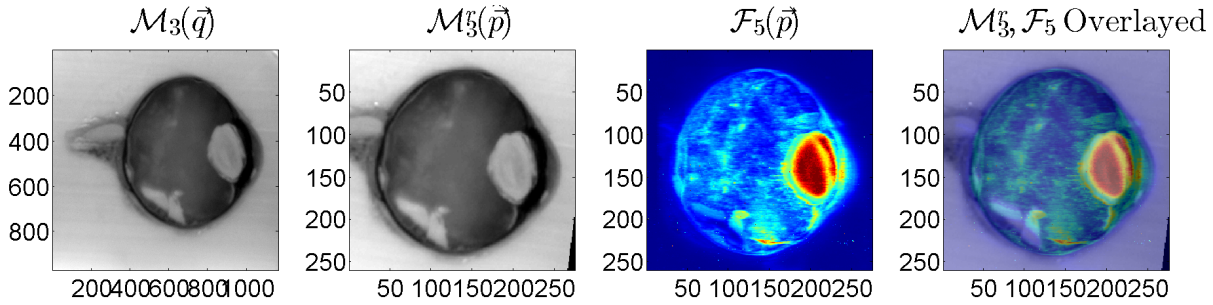

Figure 11: Experiment 6 Result

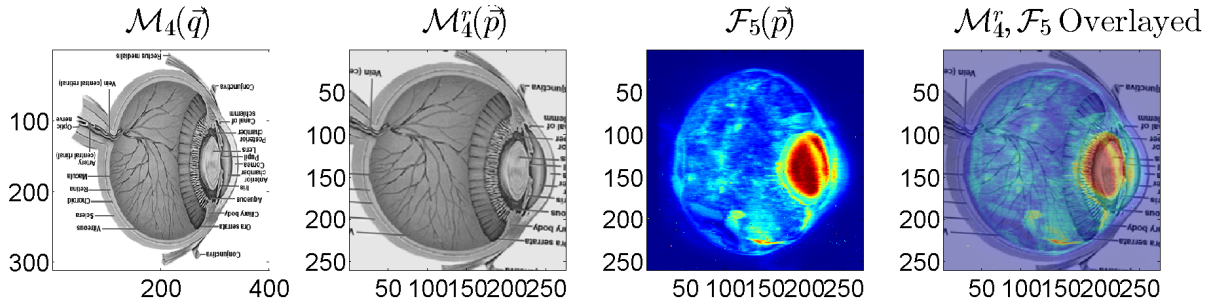

Figure 12: Experiment 7 Result

6 Conclusions

Multimodal registration between MSI and photos of an object was accomplished in this work for four copper grid MSI data sets and one biological data set (human eye). In completing this task, various custom MATLAB functions and a variety of debugging experiments were constructed to leverage the functionality of the existing MATLAB function `imregister`. Through the course of this project I have gained an in-depth knowledge of the common practices in and the aspects required for multimodal medical image registration. Many of the aforementioned challenges associated with MSI-photo registration, such as difference in dynamic range and relative intensity magnitudes, were alleviated by the use of the mutual information similarity metric.

The seven experiments performed exposed a few of the challenges associated with the registration method developed in this work. However, a number of improvements would enhance the practical utility

of the proposed method. First, a more thorough evaluation of the system's failure modes would help make registration between MSI and photo more robust. This effort would be two fold: (1) characterization of the registration result for a variety of MSI-photo pairs ranging in quality, size, and joint intensity relationship; and (2) performing range finding experiments to determine the point at which the initial transformation between two images (photo and transformed photo, for instance) is large enough to induce registration failure. The idea is that by exposing the causes for failed registration between two images, data can be pre-conditioned so as to avoid these failures, or at least to understand why they happen.

Another addition which would greatly enhance the practical utility of the proposed registration method is the incorporation of an automated pre-conditioning function for `fixed` and `moving` images (i.e. reorienting, denoising, selecting rescaling dimensions, ect.). A scheme for performing a coarse registration step before inputting the images into the registration function has been envisioned, but upon subsequent consideration and testing of the coarse registration step, it was found that the conceived method would require further development.

An additional challenge lies in the inability to verify the registration accuracy, which stems from the fact that no ground truth exists for the data sets used. Because MSI is a novel imaging method, validated techniques for improving image quality or labeling data do no exist. Additionally, labeled or segmented data sets for ground truth verification are not available. This issue could perhaps be surmounted by generating a synthetic data set using a photo to emulate the data characteristics of MSI. Since the transformation between the original photo and the MSI-modeled photo could be tracked, the registration accuracy could presumably be assessed in this way.

The ability to register photos and MSI will be a tremendously useful utility in furthering metallomic spectral imaging research and analysis. Since controlling the quality and acquisition parameters of a photo is simple compared to MSI, segmentation of the object of interest in a photo tends to be more accurate and reliable. If the segmentation curve can be represented as a continuous function (either via parameterization of the curve or by defining the curve as a zero-level set), the transform which registers the photo to the MSI can also be applied to the segmentation curve, thereby passing the segmentation to the MSI. In the coming weeks, the analysis of background and foreground signal statistics, edge blurring, and background masking in MSI will be explored using a joint photo-segmentation/photo-MSI registration technique.

References

[1] F. Dufaux and J. Konrad, "Efficient, robust, and fast global motion estimation for video coding," Image Processing, IEEE Transactions on, vol. 9, pp. 497–501, Mar 2000.

[2] G. Hermosillo, C. Chefd'hotel, K.-H. Herrmann, G. Bousquet, L. Bogoni, K. Chaudhuri, D. R. Fischer, C. Geppert, R. Janka, A. Krishnan, B. Kiefer, I. Krumbein, W. Kaiser, M. Middleton, W. Ou, J. R. Reichenbach, M. Salganicoff, M. Schmitt, E. Wenkel, S. Wurdinger, and L. Zhang, "Image registration in medical imaging: Applications, methods, and clinical evaluation," in Multi Modality State-of-the-Art Medical Image Segmentation and Registration Methodologies (A. S. El-Baz, R. Acharya U, A. F. Laine, and J. S. Suri, eds.), pp. 263–313, Springer New York, 2011.

[3] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, "Multimodality image registration by maximization of mutual information," Medical Imaging, IEEE Transactions on, vol. 16, pp. 187–198, April 1997.

[4] D. Mattes, D. R. Haynor, H. Vesselle, T. K. Lewellyn, and W. Eubank, "Nonrigid multimodality image registration," vol. 4322, pp. 1609–1620, SPIE, 2001.

[5] T. Netsch, P. Roesch, J. Weese, A. van Muiswinkel, and P. Desmedt, "Grey-value-based 3D registration of functional MRI time-series: comparison of interpolation order and similarity measure," vol. 3979, pp. 1148–1159, SPIE, 2000.

[6] J. Pluim, Mutual information based registration of medical images. PhD thesis, Utrecht University, The Netherlands, June 2001.

[7] J. Pluim, J. Maintz, and M. Viergever, "Image registration by maximization of combined mutual information and gradient information," Medical Imaging, IEEE Transactions on, vol. 19, pp. 809–814, Aug. 2000.

[8] J. Pluim, J. Maintz, and M. Viergever, "Mutual-information-based registration of medical images: a survey," Medical Imaging, IEEE Transactions on, vol. 22, pp. 986–1004, Aug. 2003.

[9] A. Roche, G. Malandain, X. Pennec, and N. Ayache, "The correlation ratio as a new similarity measure for multimodal image registration," in Medical Image Computing and Computer-Assisted Interventation – MICCAI'98 (W. Wells, A. Colchester, and S. Delp, eds.), vol. 1496 of Lecture Notes in Computer Science, pp. 1115–1124, Springer Berlin / Heidelberg, 1998.

[10] D. Rueckert and P. Aljabar, "Nonrigid registration of medical images: Theory, methods, and applications [applications corner]," Signal Processing Magazine, IEEE, vol. 27, pp. 113–119, July 2010.

[11] M. Styner, C. Brechbuhler, G. Szckely, and G. Gerig, "Parametric estimate of intensity inhomogeneities applied to MRI," Medical Imaging, IEEE Transactions on, vol. 19, pp. 153–165, March 2000.

[12] P. Thevenaz and M. Unser, "Optimization of mutual information for multiresolution image registration," Image Processing, IEEE Transactions on, vol. 9, pp. 2083–2099, Dec 2000.

[13] P. Thevenaz and M. A. Unser, "Spline pyramids for intermodal image registration using mutual information," vol. 3169, pp. 236–247, SPIE, 1997.

[14] P. Viola and W. M. Wells III, "Alignment by maximization of mutual information," International Journal of Computer Vision, vol. 24, pp. 137–154, 1997.

[15] L. Zöllei, J. Fisher, and W. Wells, "An introduction to statistical methods of medical image registration," in Handbook of Mathematical Models in Computer Vision (N. Paragios, Y. Chen, and O. Faugeras, eds.), pp. 531–542, Springer US, 2006.

[16] L. Zöllei, J. Fisher, and W. Wells, "A unified statistical and information theoretic framework for multimodal image registration," in Information Processing in Medical Imaging (C. Taylor and J. Noble, eds.), vol. 2732 of Lecture Notes in Computer Science, pp. 366–377, Springer Berlin / Heidelberg, 2003.

A Differentiation of Mutual Information [4]

Continuous estimates of the distributions are approximated via Parzen windowing of the marginal and joint histograms of the moving and fixed images. With a cubic spline Parzen window ($\beta^{(3)}$) and a zero-order spline Parzen window ($\beta^{(0)}$), joint and marginal discrete probability densities are given by:

$$p(\ell, \kappa; \boldsymbol{a}) = \gamma \sum_{\boldsymbol{p} \in V} \beta^{(0)} \left(\kappa - \frac{f_{\mathcal{F}}(\boldsymbol{p}) - f'_{\mathcal{F}}}{\Delta b_{\mathcal{F}}} \right) \cdot \beta^{(3)} \left(\ell - \frac{f_{\mathcal{M}}(A \cdot \boldsymbol{q}) - f'_{\mathcal{M}}}{\Delta b_{\mathcal{M}}} \right)$$

$$p_{\mathcal{M}}(\ell; \boldsymbol{a}) = \sum_{\kappa \in L_{\mathcal{F}}} p(\ell, \kappa; \boldsymbol{a}) \qquad p_{\mathcal{F}}(\kappa) = \gamma \sum_{\boldsymbol{p} \in V} \beta^{(0)} \left(\kappa - \frac{f_{\mathcal{F}}(\boldsymbol{p}) - f'_{\mathcal{F}}}{\Delta b_{\mathcal{F}}} \right) \tag{10}$$

where

γ	=	normalization factor that ensures $\sum p(\ell,\kappa) = 1, \forall \ell \in L_{\mathcal{M}}$ and $\forall \kappa \in L_{\mathcal{F}}$
$f_{\mathcal{F}}(\boldsymbol{p})$ and $f_{\mathcal{M}}(A \cdot \boldsymbol{q})$	=	samples of fixed and moving images
$f'_{\mathcal{F}}$ and $f'_{\mathcal{M}}$	=	minimum intensity values of fixed and moving images
$\Delta b_{\mathcal{F}}$ and $\Delta b_{\mathcal{M}}$	=	intensity range of each bin
V	=	set of pixel pairs that contribute to the distribution

Using these definitions for the marginal and joint probability distributions, the gradient of MI with respect to the transformation parameters $\boldsymbol{a}$ can be defined as follows:

$$\nabla J = \left[\frac{\partial J}{\partial a_1}, \frac{\partial J}{\partial a_2}, \frac{\partial J}{\partial a_3}, \frac{\partial J}{\partial a_4}, \frac{\partial J}{\partial a_5}, \frac{\partial J}{\partial a_6}\right]^T$$

$$\frac{\partial J}{\partial a_i} = \sum_{\ell \in L_{\mathcal{M}}} \sum_{\kappa \in L_{\mathcal{F}}} \frac{\partial p_{\mathcal{F},\mathcal{M}}(\ell,\kappa;\boldsymbol{a})}{\partial a_i} \log_2 \left(\frac{p_{\mathcal{F},\mathcal{M}}(\ell,\kappa;\boldsymbol{a})}{p_{\mathcal{F}}(\kappa) p_{\mathcal{M}}(\ell;\boldsymbol{a})}\right) \tag{11}$$

$$\frac{\partial p(\ell,\kappa)}{\partial a_i} = \frac{1}{\Delta b_{\mathcal{M}}(\#V)} \sum \beta^{(0)} \left(\kappa - \frac{f_{\mathcal{F}}(\boldsymbol{p}) - f'_{\mathcal{F}}}{\Delta b_{\mathcal{F}}}\right) \cdot \frac{\partial \beta^{(3)}(\xi)}{\partial \xi}\bigg|_{\xi = \ell - \frac{f_{\mathcal{M}}(A\cdot\boldsymbol{q}) - f'_{\mathcal{M}}}{\Delta b_{\mathcal{M}}}} \left(-\frac{d f_{\mathcal{M}}(\boldsymbol{t})}{d\boldsymbol{t}}\bigg|_{\boldsymbol{t} = A\cdot\boldsymbol{q}}\right)^T \cdot \frac{\partial (A \cdot \boldsymbol{q})}{\partial a_i} \tag{12}$$

where $\#V$ is the number of pixels used in the summation. Thevenaz *et al.* [12] gives a much more thorough and explicit derivation of the gradient and the Hessian of MI, however since MATLAB implements MI according to Mattes et al., this derivation has not been provided.