

Implicit similarity: a new approach to multi-sensor image registration

*Y. Keller^{1,2}, A. Averbuch¹

¹School of Computer Science

Tel Aviv University, Tel Aviv 69978, Israel

²Math department

Yale University, Connecticut, USA

Phone: +1-203-432-4345

Email: yosi.keller@yale.edu

Submitted to the IEEE Transactions on Image Processing

April 2004

Abstract

This paper presents a new approach to the registration of significantly dissimilar images, acquired by sensors of different modalities. The algorithm introduces a robust matching criterion by aligning the locations of gradient maxima. The alignment is achieved by iteratively maximizing the magnitudes of the intensity gradients of a set of pixels in one of the images, where the set is initialized by the gradient maxima locations of the second image. No explicit similarity measure which uses the intensities of both images is used. The computation utilizes the full spatial information of the first image and the accuracy and robustness of the registration, depend only on it. By embedding the scheme in a ‘coarse to fine’ formulation, we were able to robustly estimate affine and projective global motions, even when the images were characterized by complex space varying intensity transformations. Compared to prior schemes, when one of the images is of considerably better quality than the other (noise, blur etc.), the algorithm is significantly immuner to noise and has a higher convergence range. We demonstrate these properties via experiments on synthesized and real image sets.

Keywords: Global motion estimation, multimodal images, multi-sensor images, gradient methods, image alignment

EDICS Category={2-ANAL, 2-MOTD}

1 Introduction

The registration of images acquired by sensors of different modalities is of special interest to remote sensing and medical imaging applications, as the information gained from such sets of images is of complementary nature. Proper fusion of the data obtained from separate images

requires accurate spatial alignment. This issue was extensively studied in the context of remote-sensing (SAR, FLIR, IR and optical sensors) [1, 2] and medical image registration (CT, MRI, UltraSound, MRA, DSA, CTA, SPECT, PET, fMRI, EEG, MEG, pMRI, fCT, EIT, MRE) [3, 4, 5]. Due to the different physical characteristics of various sensors, the relationship between the intensities of matching pixels is often complex and unknown a priori. Features present in one image might appear only partially in the other image or do not appear at all. Contrast reversal may occur in some image regions while not in others, multiple intensity values in one image may map to a single intensity value in the other image and vice versa. Furthermore, imaging sensors may produce considerably dissimilar images of the same scene when configured with different imaging parameters.

Geometrical matching techniques [2, 6, 7] align geometrical primitives such as feature points, contours and corners using vectorial representations. These matchings are invariant to intensity changes once the geometrical primitives are detected in both input images. Since the geometrical primitives are matched one by one, these algorithms can not be used within a global optimization scheme, where the registration is derived by optimizing a functional over the entire set of features. This results in poor global alignment, sensitivity to outliers and the need to apply random sampling and cross-validation techniques such as RANSAC [8].

Thus, for the monomodal as well as the multimodal case, the general approach to image registration consists of defining an *explicit similarity measure* [9], a functional $\phi = \phi(I_1, I_2)$ which is applied to the intensities of the input images I_1 and I_2 and measures their similarity. The choice of the similarity measure corresponds to an assumption regarding the relationship between the intensities of the images (e.g., functional, statistical, etc.). The registration is then derived by optimizing ϕ with respect to the motion parameters vector p . For instance, for monomodal images, ϕ is often chosen to be the L_2 norm

$$p^* = \arg \min_p \sum (I_1 - I_2(p))^2 \quad (1.1)$$

where $I_2(p)$ is a replica of I_2 warped according to p and the summation is conducted over the overlapping parts of I_1 and $I_2(p)$. For significantly dissimilar (multi-sensor) images, solving Eq. (1.1) does not result in image registration and other similarity measures such as correlation coefficient [1], correlation ratio [10], mutual information [3, 4, 11] and statistical divergence [12] are used. Further robustness can be achieved, by registering transformed representations of I_1 and I_2 , which are invariant to certain brightness dissimilarities. These representations include feature points [13], edge-maps [6], oriented edge-maps [1] and edge contours [2]. Matching is achieved by applying gradient methods [6], robust gradient methods [1], chain-code correlation [2] or geometrical hashing [7] to the invariant representations. The computation of invariant representations results in important image information being lost and the optimization process might fail to converge.

The Mutual Information (MI) similarity measure [3, 4, 11, 14] originates from information theory and is considered the state-of-the-art registration algorithm for multi modality images. It is a measure of statistical dependency between two data sets. Given two random variables X and Y , their MI is

$$MI(X, Y) = \sum_{i,j} P(X_i, Y_j) \log \left(\frac{P(X_i, Y_j)}{P(X_i) P(Y_j)} \right) \quad (1.2)$$

where $P(X)$ is the probability of X and $P(X, Y)$ is the joint probability of X and Y . $MI(I_1, I_2)$ was shown to be maximized for aligned images [3]. The optimization can be implemented by

either Powell’s method [14] or a faster coarse-to-fine Marquardt–Levenberg scheme [15].

However, MI based matchings can result in misregistration [14], due to several issues. First, the MI measure is highly non-convex with respect to the transformation parameters p . Thus, the optimization procedure can be easily attracted to local maxima. Moreover, the non-convexity of the criterion is enhanced by the interpolation used in sub-pixel accurate motion estimation [16]. Second, when the images are of low resolution, contain little information, or there is only a small overlapping region, they become less statistically dependent and the maximization of the MI measure becomes ill-posed. Third, the assumption of a global statistical relationship between the image intensities may be violated by the presence of various image artifacts. For instance, intensity inhomogeneities are found in several image modalities [5], while in remote-sensing applications, the image structures being matched result from different physical properties.

The work of this paper, preliminarily presented in [17], suggests an implicit similarity measure to address the limitations of the algorithms mentioned above. Thus, no functional $\phi(I_1, I_2)$ which uses the intensities of **both** images is used. Instead, the algorithm aligns the set of pixels having large gradient magnitudes in both images in two steps: First, the pixels are detected in I_1 (using **only** I_1 ’s intensity value). Second, the pixels are aligned to I_2 using iterative optimization, where the optimization is carried out using **only** I_2 ’s intensity values. Hence, the convergence only depend on the properties of I_2 which can be chosen to be the better quality input image. This approach allows feature matching using a global optimization scheme, where in prior works, only one of these attributes was achieved. Furthermore, the second step utilizes the full spatial information of I_2 , improving the convergence compared to schemes based on invariant representations [2, 6, 13].

The paper is organized as follows: section 2 presents the proposed algorithm and its convergence is analyzed in section 3. Experimental results are given in section 4.

2 The proposed algorithm

In this section we formulate the proposed registration algorithm illustrated in Fig. 1. Let I_1 and I_2 be the input images with some overlap, denote

$$F(S_2(p)) \triangleq \sum_{v_i \in S_2(p)} |\nabla I_2(v_i)|^2 \quad (2.1)$$

where S_2 is a set of pixels in I_2 , p are the motion’s parameters and $v_i = (x_i, y_i)$. Maximizing $F(S_2(p))$ with respect to p implies that the set S_2 consists of dominant image edges. The maximization is performed iteratively, where S_2^0 , the initial estimate is given by

$$S_2^0 = S_2(p_0) \quad (2.2)$$

and

$$S_2(0) = S_1.$$

S_1 is a set of high gradient pixels detected in I_1 , p_0 is the initial estimate of the motion parameters mapping I_1 to I_2 and $S_2(p_0)$ is the set of pixels in I_2 computed by warping S_1 according to p_0 .

The algorithm matches pixels having high gradient magnitudes in both images and their initial detection in I_1 is straight forward. p^* , the vector of parameters where

$$p^* = \arg \max_p F(S_2(p)) \quad (2.3)$$

maps the set S_1 to $S_2(p^*)$ and since the intensities values of only one of the image are used at each step, without using an explicit similarity measure (similar to $\phi(I_1, I_2)$), p^* is an *implicit solution* of the registration of I_1 and I_2 .

The underlying assumption, that pixels with large gradient magnitudes can be found in both images, stems from the fact that such pixels were shown to correspond to primal structures in natural images [18]. In particular, due to the inherent resolution limitations of imaging sensors, multisensor and multimodality images are often untextured (similar to the ones shown in Fig. 1), thus, the similarity between such images is strictly related to the locations of their edges.

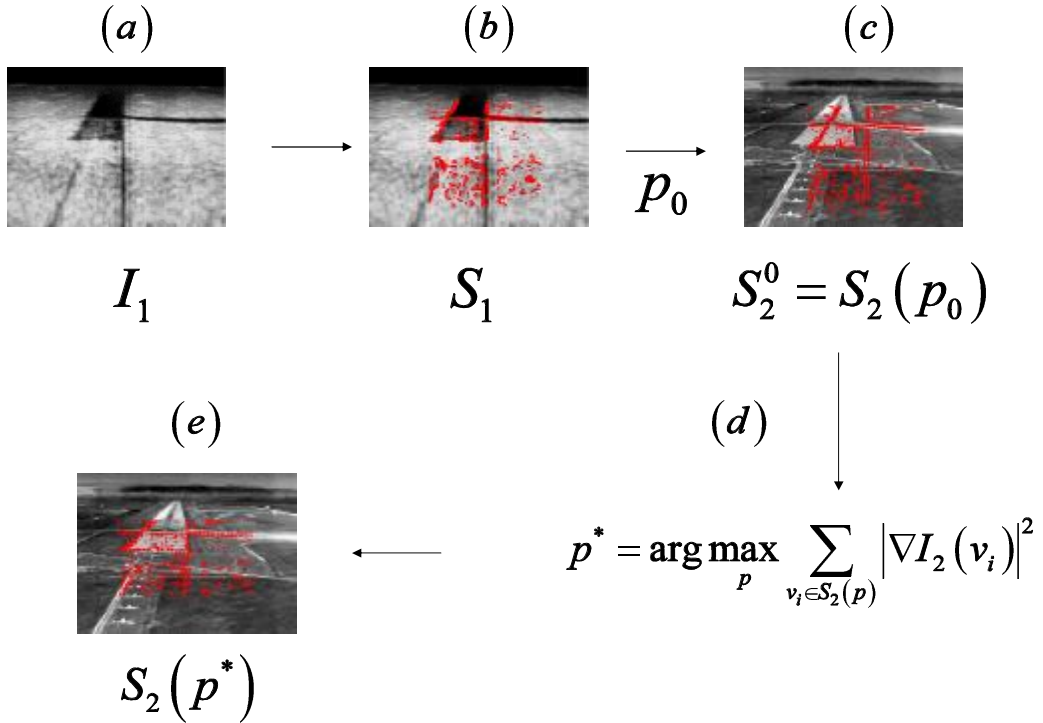


Figure 1: The flow of the multi-sensor registration algorithm. (a) Input image I_1 . (b) S_1 , the set of pixels with the largest gradient magnitude is overlaid in red. (c) S_1 is warped towards I_2 using the initial estimate of the motion p_0 and denoted $S_2(p_0)$. (d) $S_2(p_0)$ is used to initialize an iterative maximization of the gradient magnitude of $S_2(p)$. (e) For the maximal gradient energy of $S_2(p)$, p^* corresponds to the motion between I_1 and I_2 .

2.1 Initial pixel set detection

The initial pixel set S_1 is computed by detecting the set of pixels in I_1 with the highest gradient magnitudes where the gradients are computed using centralized finite differences. Contrary to higher level geometric primitives such as contours and corners, no preset thresholds are needed. For instance, there is no need to determine the parameter k for the Harris corner detector [19] or chain together edge pixels to identify contours. $|S_1|$, the number of pixels used is typically 25% of the total number of pixels in I_1 . An estimation of the optimal set size $|S_1|$ is derived in Section 3.

An accurate global transformation can be found only if the selected points are properly distributed across the image and are not clustered along a few edges. In order to ensure a uniform distribution, I_1 is divided into 10×10 regions, where in each region, the top 25% of the pixels are chosen. Choosing a different number of regions did not change the registration results and 10×10 regions were used in all of the our experiments.

2.2 Iterative optimization

In this section we present a numerical scheme for the solution of Eq. (2.3). $F(p)$ is non-linear with respect to p and the motion parameters are computed using Newton's iterative optimization method [20]. The iterative refinement is given by

$$p_{n+1} = p_n + (H^n)^{-1} \nabla_p F_n, \quad (2.4)$$

where $\nabla_p F_n$ and H^n are the gradient and Hessian of F , respectively, with respect to p at iteration n . $\nabla_p F_n(v_i)$ and $H^n(v_i)$, the gradient and Hessian evaluated at a pixel v_i are computed by the derivative chain rule. Thus, for the affine and projective motion models we have [21]

$$\nabla_p F_n = \sum_{v_i \in S_1(p)} X_i^T \nabla_v F_n(v_i) \quad (2.5)$$

and

$$H_p = \sum_{v_i \in S_1(p)} X_i^T H_v^n(v_i) X_i \quad (2.6)$$

where $\nabla_v F_n(v_i)$ and $H_v^n(v_i)$ are the gradient and Hessian of F_n , respectively, with respect to (x, y) . For the affine and projective motion models, X is given by

$$X_{Affine} \triangleq \begin{bmatrix} x_i & y_i & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & x_i & y_i & 1 \end{bmatrix} \quad (2.7)$$

and

$$X_{Projective} \triangleq \begin{bmatrix} x_1 & y & 1 & 0 & 0 & 0 & x_i^2 & x_i y_i \\ 0 & 0 & 0 & x_i & y_i & 1 & x_i y_1 & y_1^2 \end{bmatrix}. \quad (2.8)$$

Eq. (2.4) is iterated until either of the following convergence criteria is met:

1. A maximal predefined number of iteration is reached.
2. The update of the parameters in Eq. (2.4) becomes smaller than a predefined threshold ε

$$\|p_{n+1} - p_n\| < \varepsilon.$$

2.3 Multiscale extension

In order to improve the convergence properties, a resolution pyramid is constructed. The alignment algorithm starts at the coarsest resolution scale of the pyramid, then follows the subsequent levels in a coarse-to-fine approach. The computation given in Section 2.2 is repeated at each resolution scale where the result of each scale serves as an initial estimate of the solution of the finer scale. Finally, when the procedure stops at the finest resolution scale, the final motion parameters are obtained.

The set S_1 is initially detected at the finest resolution scale, then, given the set at a certain resolution scale, the coarser level set \tilde{S}_1 is computed using the following steps:

1. The coordinates of each pixel in S_1 are scaled down by the scale factor S and the new set is denoted \hat{S}_1 .
2. A local search in an area of $S \times S$ around each pixel in \hat{S}_1 is conducted to find the pixel with the largest gradient magnitude. That pixel is inserted into \tilde{S}_1 .
3. \tilde{S}_1 is scanned to eliminate duplicate pixels.

The coarse-to-fine refinement process allows the alignment process to lock on a single motion denoted *dominant motion* [21], even when multiple motions are present. This property is essential when outlier feature points are present.

3 Convergence properties

The convergence properties of the proposed algorithm are similar to those of gradient methods [21]. Hence, a sufficient convergence condition can be derived by considering the refinement term in Eq. (2.4)

$$\delta p_n \triangleq (H^n)^{-1} \nabla_p F_n. \quad (3.1)$$

The asymptotic result ($n \rightarrow \infty$) of the optimization does not depend on H^n , which can be any positive-definite matrix and whose choice, only affects the convergence rate. Hence, the The asymptotic result depends on $\nabla_p F_n$. Define the following pixel sets, depicted in Fig. 2.

- S_2 All the pixels in I_2 .
- S_2^n A subset of S_2 - the set of pixels used at iteration n in I_2 to evaluate Eq. (2.1).
- S_G^n A subset of S_2^n - the set of pixels in S_2^n having large gradient values corresponding to true alignment.
- $\overline{S_G^n}$ A subset of S_2^n - the set of pixels in S_2^n having small gradient values corresponding to false alignment.

$E \{ \nabla_p F_n \}$, the expectancy of $\nabla_p F_n$ was shown in Appendix A to be

$$E \{ \nabla_p F_n \} \sim (E \{ S_G^n \} - E \{ \overline{S_G^n} \}) P(v_2 \in S_G^n) (1 - P(v_1 \in S_G^n)) \quad (3.2)$$

and in order to achieve robustness to outliers and image noise we require $|E \{ \nabla_p F_n \}| \gg 0$.

For given input images, the optimal value of $P(v_2 \in S_G^n)$ can be numerically estimated by evaluating Eq. (3.2). Figure 3 shows $|E \{ \nabla_p F_n \}|$ as a function of $P(v_2 \in S_G^n)$ for several images

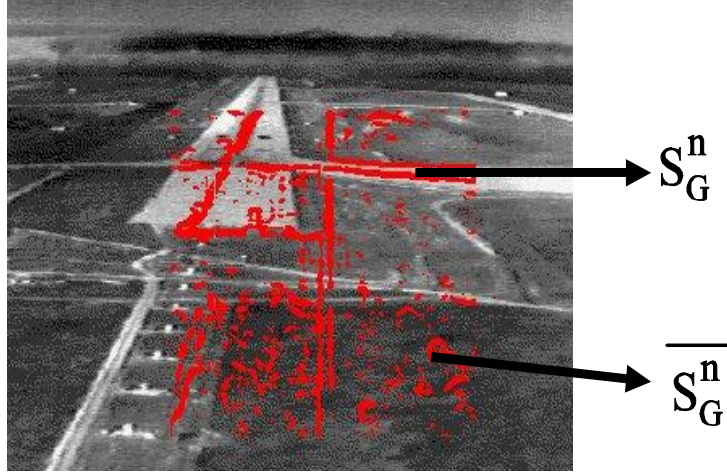


Figure 2: Definition of the pixel sets used in the convergence analysis. S_2 is the set of all the pixels in the image. S_2^n is the set of pixels marked in red. These pixels are used by the registration algorithm. S_G^n are the pixels within S_2^n which have high gradient magnitudes. The rest of the pixels in S_2^n are referred to as $\overline{S_G^n}$.

and setting $\frac{|S_G^n|}{|S_2|} = 25\%$ turns out to be a reasonable choice. While $P(v_2 \in S_G^n)$ can not be set directly, we note that

$$P(v_2 \in S_G^n) = \frac{|S_G^n|}{|S_2|} \leq \frac{|S_2^n|}{|S_2|} \quad (3.3)$$

and an upper limit to S_G^n can be placed by choosing the size of S_2^n . Furthermore, it follows that for a given value of $|E\{\nabla_p F_n\}|$ the higher the contrast in I_2 between $E\{S_G^n\}$ and $E\{\overline{S_G^n}\}$, the better the convergence properties.

4 Experimental results

The proposed implicit similarity (IS) algorithm was experimentally evaluated by registering image pairs acquired by medical and remote sensing applications. For the sake of brevity, representative results are reported for each class of images. The optimal size of S_1 was computed according to Eq. (3.2) and a multi-resolution pyramid (Section 2.3) with three resolution scales was used. 100 iterations were used at each resolution scale and the initial pixel set was computed using 10×10 subgrids. The initial estimate of the motion was given by choosing a single pair of matching pixels in both images. The registration accuracy was verified in two ways. First, by overlaying the initial and final pixel sets on both images. Second, at each iteration, I_2 was warped towards I_1 using the accumulative estimated motion and the Mutual Information (MI) between the images was computed. Thus, the dynamics and convergence properties of the IS can be analyzed using the MI as a similarity measure. The images were also registered with a MI based algorithm,

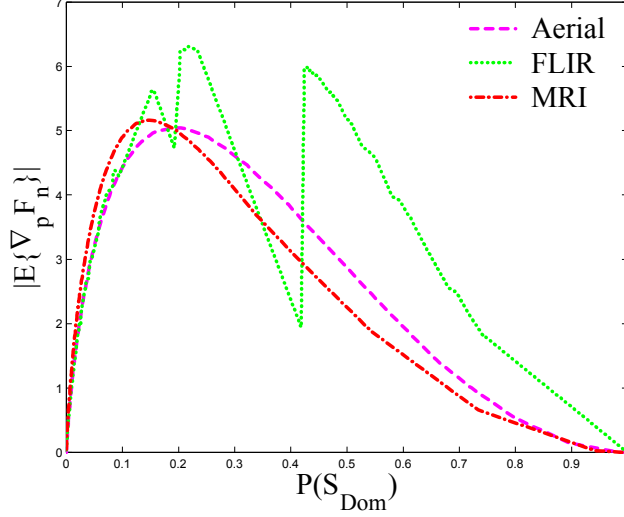


Figure 3: $|E\{\nabla_p F_n\}|$ as a function of the size of the initial pixel set. Choosing the pixel set size to be 25% of the images size, seems to be a reasonable choice for all the images. The images used for the computation of $|E\{\nabla_p F_n\}|$ are shown in Section 4.

where the probability functions were computed by histograms and Powell’s method was used for optimization. Next, we present the registration of some typical multi sensor and multi modality images. Figure 4 shows the accurate registration of MRI and CT images using an affine motion model

$$\begin{aligned} x_1 &= a \cdot x_2 + b \cdot y_2 + c \\ y_1 &= d \cdot x_2 + e \cdot y_2 + f \end{aligned} \quad (4.1)$$

This is a typical example of images usually registered by MI and Fig. 4e shows that the IS maximizes the MI measure as the alignment is computed. By examining Fig. 4e, it follows that most of the registration is achieved at the coarser scales ($S = 1$ and $S = 2$). The MI algorithm had a wider basin of convergence, which is attributed to the fact that the MI is well adapted to such images and its maximization corresponds to image alignment.

A counterexample is given in Figs. 5 and 7 which show the registration of remote sensing imagery. These images are characterized by a space-varying intensity mapping, hence the MI is less characteristic of the image alignment. Figure 5e shows that the relative increase of the MI during the registration was considerably smaller than in the registration of the medical images in Fig. 4e. Hence, the MI’s basin of convergence was significantly smaller compared to that of the SI. Figure 6 shows the registration of the images by the IS using a projective motion model.

The accuracy of the IS was measured by registering the images in Fig. 8. First, the true motion was estimated using a gradient methods based algorithm [21]. Second, the images were registered by the IS algorithm. The results depicted in Figs. 8b and 8c show the high accuracy of the IS when I_2 (the image used in the optimization step) is of high quality, these were the most accurate results over the entire test set. In particular, we note that for the IS algorithm, there is no difference between the registration of Fig. 8c to either Figs. 8a or 8d. The quality of I_1 is less significant, since it is only used to detect the initial locations of the gradient maxima. On the contrary, the MI diverged while registering Figs. 8c and 8d. Figure 8e shows the inconsistency of the MI measure in this case, the nonzero pixels in 8d are edge indicators and their intensity values are statistically

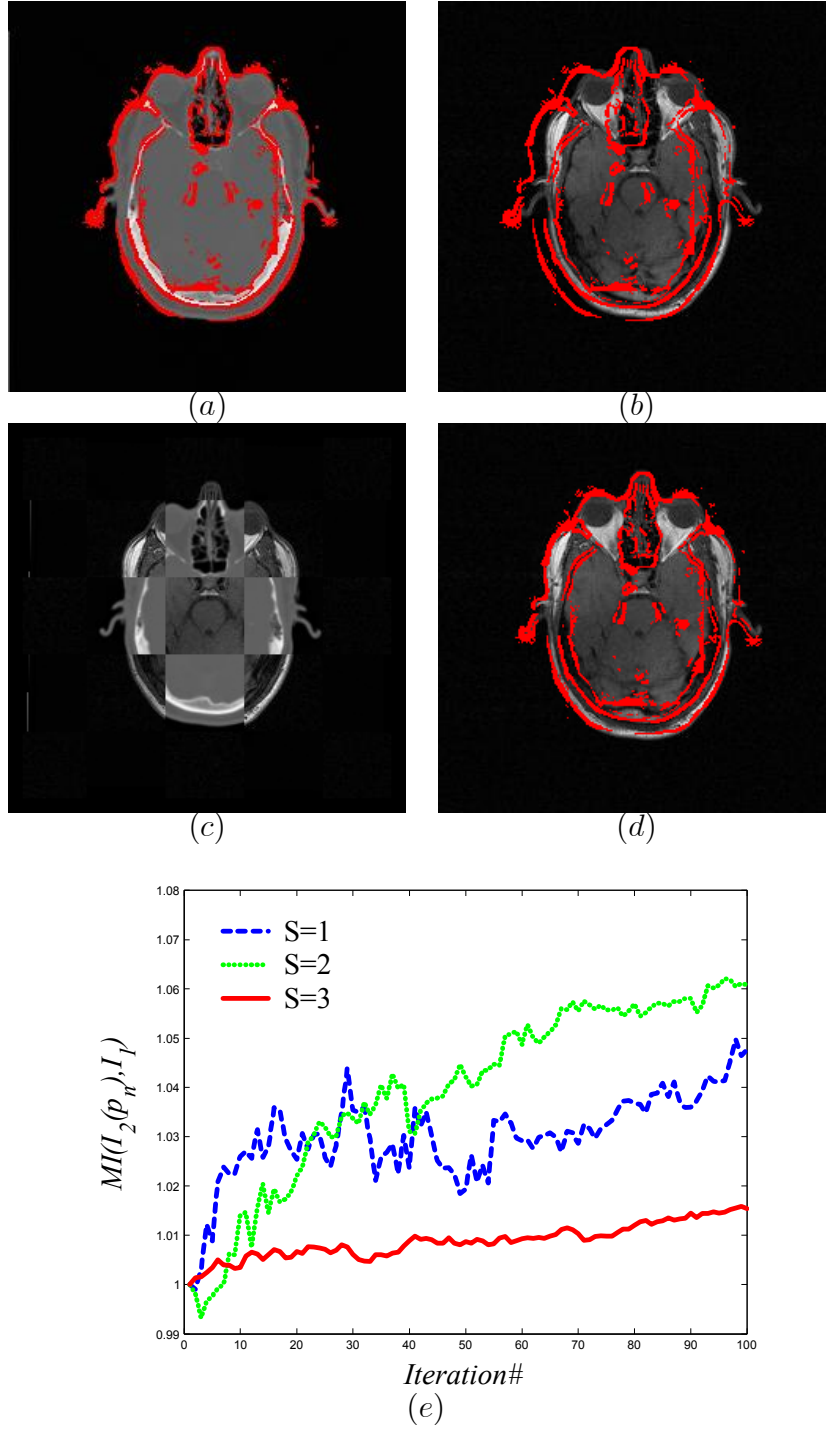


Figure 4: Registration results of CT and MRI multi-modality images using the affine motion model: (a) original CT image overlaid with the set of high-gradient pixels. (b) Initial estimate of the overlaid set in the MRI image. (c) Composite image. (d) Final alignment of the pixel set after registration. (e) The relative change in the MI measure for the different resolution scales. $S = 3$ is the finest resolution scale.

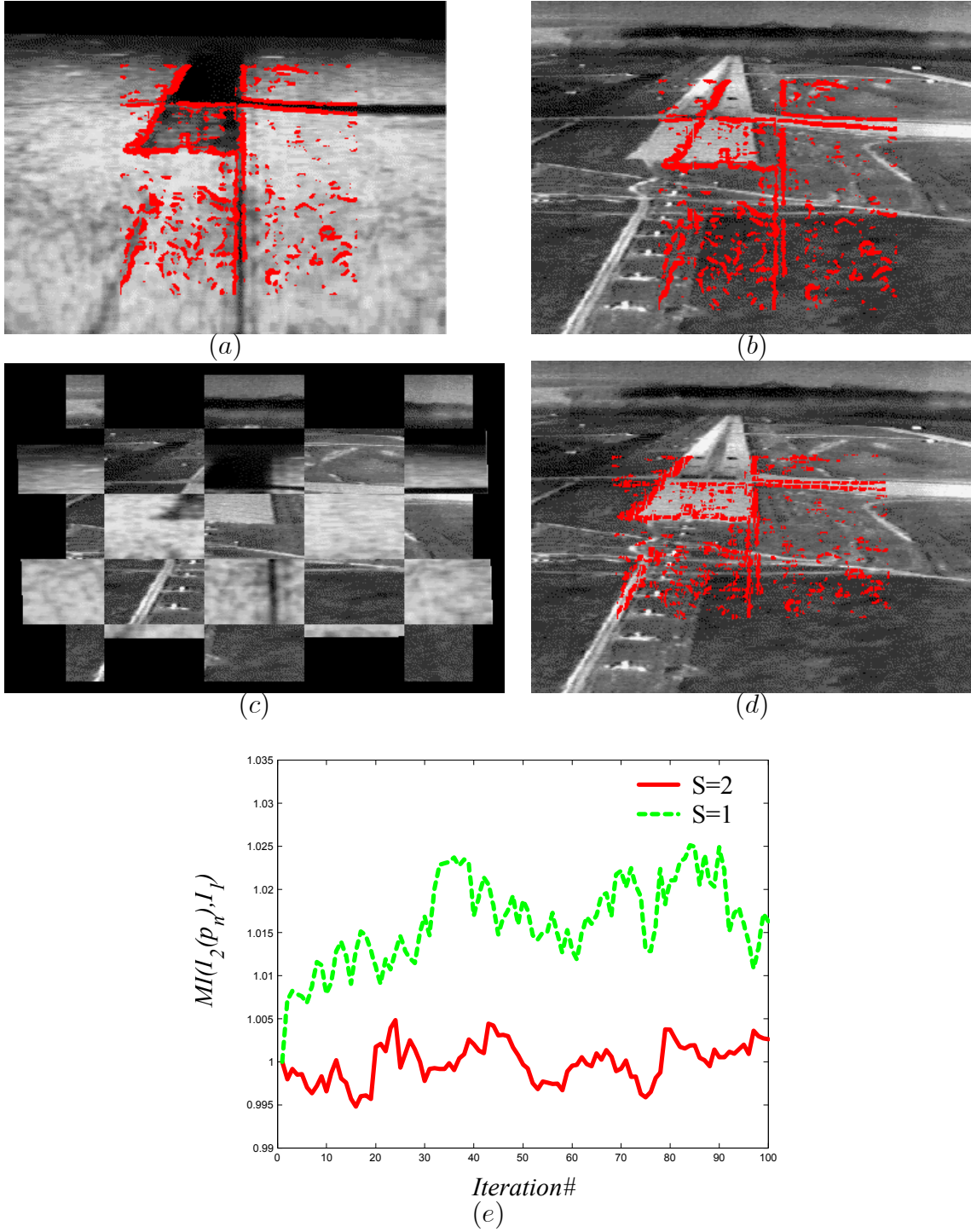


Figure 5: Registration results of SAR and FLIR multi-sensor images using the affine motion model: (a) original SAR image overlayed with the set of high-gradient pixels . (b) Initial estimate of the overlayed set in the FLIR image. (c) Composite image. (d) Final alignment of the registered pixel set. (e) The relative change in the MI measure for the different resolution scales. $S = 2$ is the finest resolution scale.

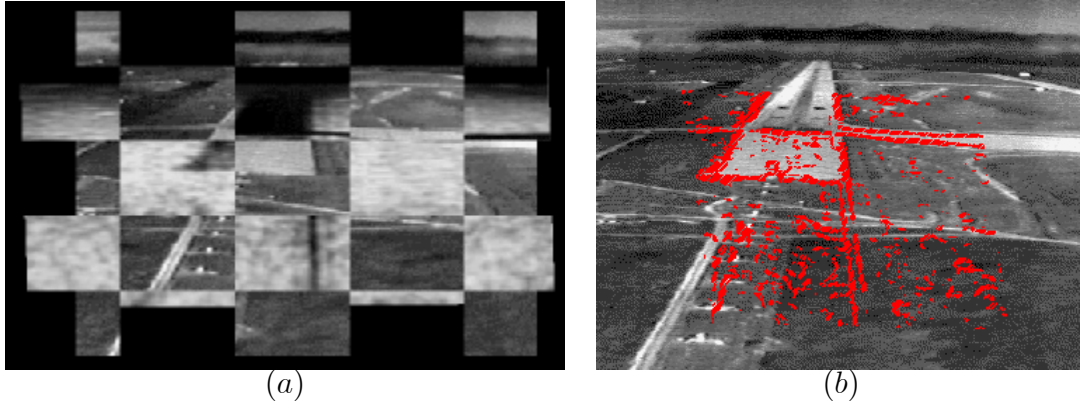


Figure 6: Registration results of SAR and FLIR multi-sensor images using the projective motion model: (a) Composite image. (b) Final alignment of the pixel set after registration. The initial pixel set and its initial projection are the same as in Figure 5.

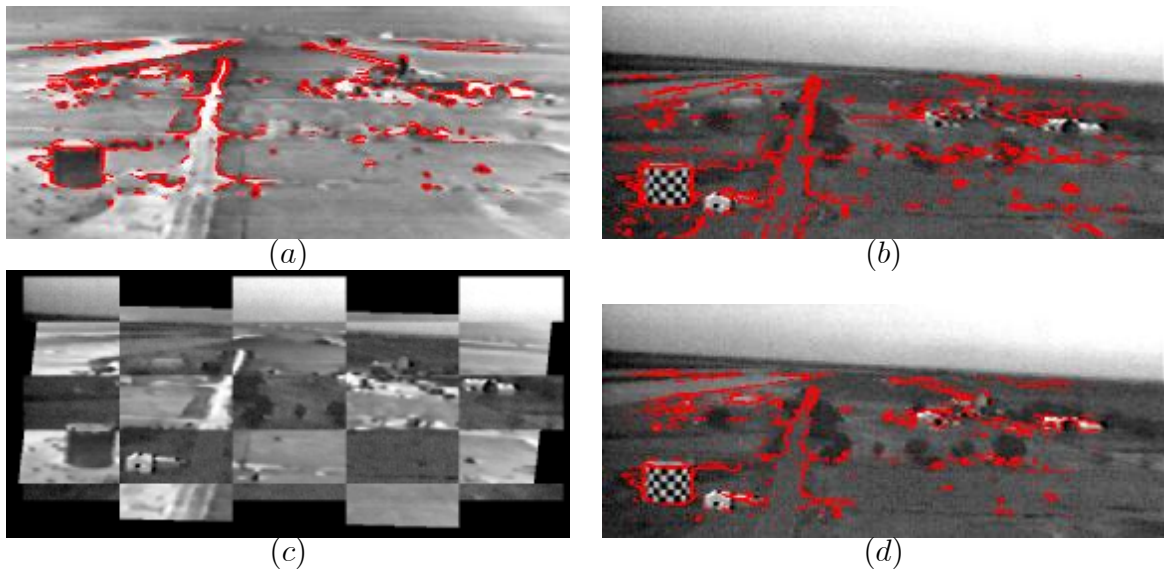


Figure 7: Registration results of EO and IR multi-modality images presented in [1] using the affine motion model: (a) original IR image overlaid with the set of high-gradient pixels. (b) Initial estimate of the overlaid set in the EO image. (c) Composite image. (d) Final alignment of the pixel set after registration.

uncorrelated to the values of the corresponding pixels in 8c. To conclude, the IS algorithm was able to register a wide class of multi modality and multi sensor imagery. For high quality images the MI was shown to have better convergence properties. The IS outperforms the MI in the registration of poor quality images, where the MI similarity measure might prove to be unrelated to the actual similarity between the images. The IS is especially suitable for the alignment of image pairs where one of the images is of considerably better quality than the other.

5 Concluding Remarks

In this paper we presented a novel approach to multi-modality image registration. The algorithm uses an implicit similarity measure which is invariant to intensity dissimilarities. The measure is optimized by a fast multiscale gradient scheme. The convergence properties of the IS were analyzed and an approximation of the optimal size of the initial pixel set size was derived. The experimental results demonstrate the applicability of the IS, where it was shown to be especially suited for the registration of image sets, where one of the images is of considerably better quality (noise, blur etc.) than the other. In such scenarios, it significantly outperforms prior algorithms such as MI. Future work will involve the extension of the IS to non-rigid registration by maximizing the gradient energy with respect to an elastic, locally varying motion model similar to the work in [22].

6 Appendix A

This section analyzes $E \{ \nabla_p F_n \}$ with regards to the probabilities of the pixel sets $\overline{S_G^n}$ and S_G^n , defined in Section 3. Following Eq. (2.5), $E \{ \nabla_p F_n \}$, the expectancy of $\nabla_p F_n$ with respect to the motion parameters is given by

$$E \{ \nabla_p F_n \} = E \{ X_i \nabla_v F_n (v_i) \}.$$

In the optimization procedure, $\nabla_p F_n$ is evaluated at pixels which are projections of the initial pixel set S_1 . Since S_1 is chosen to be spatially uniformly distributed (see Section 2.1), its projection is also uniformly distributed in I_2 . Thus, $\nabla_v F_n (v_i)$ is statistically independent of $E \{ X_i \}$ and

$$E \{ \nabla_p F_n \} = E \{ X_i \} E \{ \nabla_v F_n (v_i) \}.$$

$E \{ \nabla_v F_n (v_i) \}$ is computed by analyzing $E \{ \frac{\partial F_n}{\partial x} \}$ and the results apply to $E \{ \frac{\partial F_n}{\partial y} \}$ due to symmetry. Using a centralized approximation of $\frac{\partial F_n}{\partial x}$ we get

$$2E \left\{ \frac{\partial F_n}{\partial x} \right\} \simeq E \{ F(x+1, y) \} - E \{ F(x-1, y) \}$$

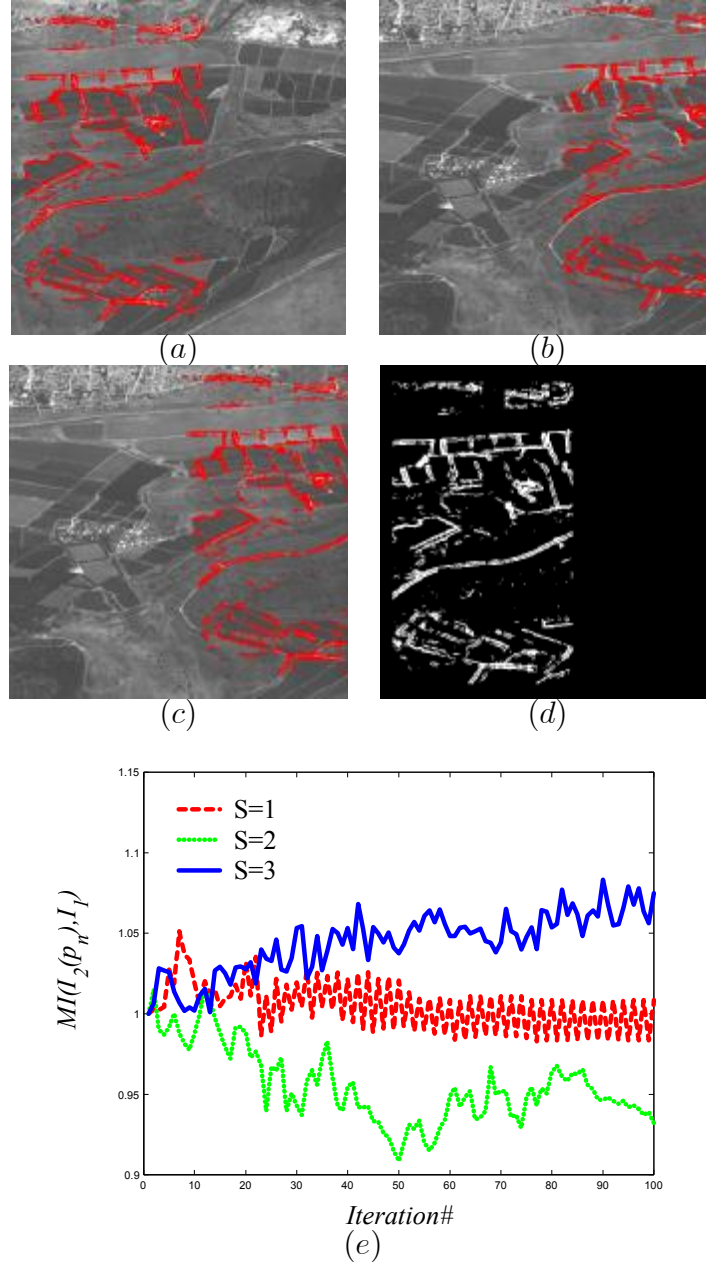


Figure 8: Registration of real aerial images. (a) The pixel set with high gradient magnitude initially detected (this image acts as I_1). (b) The pixel set projected onto the second image. Note the misalignments in the lower and upper parts of the image. (c) The pixel set after alignment. (d) A binary edge map of (a) registered to image (b). For the proposed algorithm there is no difference for the registration of (d) to either (a) or (d). (e) The MI measure for the registration of (d), the binary edge image, with image (b) using the proposed algorithm. Note that the MI is not significantly increased by the registration and that there is a decrease in it at the coarser resolution scales.

and denote $v_1 = (x + 1, y)$ and $v_2 = (x - 1, y)$

$$\begin{aligned}
2E \left\{ \frac{\partial F_n}{\partial x} \right\} &= E \left\{ \frac{\partial F_n}{\partial x} \mid v_1, v_2 \in S_G^n \right\} P(v_1, v_2 \in S_G^n) \\
&\quad + E \left\{ \frac{\partial F_n}{\partial x} \mid v_1, v_2 \in \overline{S_G^n} \right\} P(v_1, v_2 \in \overline{S_G^n}) \\
&\quad + E \left\{ \frac{\partial F_n}{\partial x} \mid v_1 \in S_G^n, v_2 \in \overline{S_G^n} \right\} P(v_1 \in S_G^n, v_2 \in \overline{S_G^n}) \\
&\quad + E \left\{ \frac{\partial F_n}{\partial x} \mid v_1 \in \overline{S_G^n}, v_2 \in S_G^n \right\} P(v_1 \in \overline{S_G^n}, v_2 \in S_G^n).
\end{aligned} \tag{6.2}$$

For v_1 and v_2 of the same class (either S_G^n or $\overline{S_G^n}$)

$$E \left\{ \frac{\partial F_n}{\partial x} \mid v_1, v_2 \in S_G^n \right\} = E \{S_G^n\} - E \{S_G^n\} = 0$$

and

$$E \left\{ \frac{\partial F_n}{\partial x} \mid X_1, X_2 \in \overline{S_G^n} \right\} = E \{\overline{S_G^n}\} - E \{\overline{S_G^n}\} = 0.$$

Due to symmetry, it suffices to consider one of the last two terms in Eq. (6.2)

$$E \left\{ \frac{\partial F_n}{\partial x} \mid v_1 \in S_G^n, v_2 \in \overline{S_G^n} \right\} = E \{S_G^n\} - E \{\overline{S_G^n}\} > 0. \tag{6.3}$$

Hence,

$$\begin{aligned}
E \left\{ \frac{\partial F_n}{\partial x} \right\} &\approx (E \{S_G^n\} - E \{\overline{S_G^n}\}) P(v_1 \in \overline{S_G^n}, v_2 \in S_G^n) v \\
&= (E \{S_G^n\} - E \{\overline{S_G^n}\}) P(v_2 \in S_G^n) P(v_1 \in \overline{S_G^n}) \\
&= (E \{S_G^n\} - E \{\overline{S_G^n}\}) P(v_2 \in S_G^n) (1 - P(v_1 \in S_G^n))
\end{aligned}$$

and

$$E \{\nabla_p F_n\} \sim (E \{S_G^n\} - E \{\overline{S_G^n}\}) P(v_2 \in S_G^n) (1 - P(v_1 \in S_G^n)). \tag{6.4}$$

References

- [1] M. Irani and P. Anandan, “Robust multi-sensor image alignment,” in *Sixth International Conference on Computer Vision (ICCV’98)*, January 1998, pp. 959–966.
- [2] H. Li, B. S. Manjunath, and S. K. Mitra, “A contour-based approach to multisensor image registration,” *IEEE Trans. on Image Processing*, vol. 4, no. 3, pp. 320–334, March 1995.
- [3] H. A. W. M. Wells III, P. Viola and S. Nakajima, “Multi-modal volume registration by maximization of mutual information,” *Medical Image Analysis*, vol. 1, no. 1, pp. 35–51, 1996.

- [4] J. Pluim, J. Maintz, and M. Viergever, "Image Registration by Maximization of Combined Mutual Information and Gradient Information," *IEEE Transactions on Medical Imaging*, vol. 19, no. 8, pp. 809–814, August 2000.
- [5] J. B. A. Maintz and M. A. Viergever, "A survey of medical image registration," *Medical Image Analysis*, vol. 2, no. 1, pp. 1–36, 1998.
- [6] R. Sharma and M. Pavel, "Registration of video sequences from multiple sensors," in *Proceedings of the Image Registration Workshop*. NASA GSFC, 1997, pp. 361–366.
- [7] A. Gueziec, X. Pennec, and N. Ayache, "Medical image registration using geometric hashing," *IEEE Computational Science & Engineering, special issue on Geometric Hashing*, vol. 4, no. 4, pp. 29–41, October-December 1997.
- [8] M. Fischler and R. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Communication ACM*, vol. 24, no. 6, pp. 381–395, 1981.
- [9] A. Roche, G. Malandain, N. Ayache, and S. Prima, "Towards a Better Comprehension of Similarity Measures used in Medical Image Registration," in *Proc. of 2nd Int. Conf. on Medical Image Computing and Computer-Assisted Intervention (MICCAI'99)*, ser. LNCS, C. Taylor and A. Colchester, Eds., no. 1679. Cambridge, UK: Springer Verlag, Sept. 1999, pp. 555–566.
- [10] A. Roche, G. Malandain, X. Pennec, and N. Ayache, "The correlation ratio as a new similarity measure for multimodal image registration," in *Proc. 1st MICCAI*, ser. Lecture Notes in Computer Science, vol. 1496. Cambridge, MA: Springer Verlag, October 1998, pp. 1115–1124.
- [11] A. Collignon, F. Maes, D. Delaere, D. Vandermeulen, P. Suetens, and G. Marchal, *Information Processing in Medical Imaging*. Dordrecht: Kluwer Academic, 1995, ch. Automated Multi-modality Image Registration Based On Information Theory, pp. 263–274.
- [12] H. K. Yun He, A. Ben Hamza, "A generalized divergence measure for robust image registration," *IEEE Transactions on Signal Processing*, vol. 51, no. 5, pp. 1211–1220, May 2003.
- [13] L. G. Brown, "A survey of image registration techniques," *ACM Computing Surveys*, vol. 24, no. 4, pp. 325–376, December 1992.
- [14] J. M. J.P.W Pluim and M. Viergever, "Mutual-information-based registration of medical images: a survey," *IEEE Transactions on Medical Imaging*, vol. 22, no. 8, pp. 986 – 1004, August 2003.
- [15] P. Thevenaz and M. Unser, "Spline pyramids for inter-modal image registration using mutual information," in *Proceedings of the SPIE Conference on Mathematical Imaging: Wavelet Applications in Signal and Image Processing V*, vol. 3169, San Diego CA, USA, July 30-August 1, 1997, pp. 236–247.

- [16] J. Pluim, J. Maintz, and M. Viergever, “Interpolation artifacts in mutual information-based image registration,” *Computer Vision and Image Understanding*, vol. 77, no. 2, pp. 211–232, February 2000.
- [17] Y. Keller and A. Averbuch, “Implicit similarity: a new approach to multi-sensor image registration,” in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, vol. 2, June 2003, pp. 543–548.
- [18] E. P. Simoncelli and E. H. Adelson, “Noise removal via Bayesian wavelet coring,” in *Third Int’l Conf on Image Proc*, vol. I. Lausanne: IEEE Sig Proc Society, September 1996, pp. 379–382.
- [19] C. G. Harris and M. Stephens, “A combined corner and edge detector,” in *Proc. 4th Alvey Vision Conference, Manchester*, 1988, pp. 147–151.
- [20] P. Gill, *Practical Optimization*. Academic Press, 1982.
- [21] M. Irani and P. Anandan, *Vision Algorithms: Theory and Practice: International Workshop on Vision Algorithms, Corfu, Greece, September 1999. Proceedings*. Springer-Verlag, 1999, ch. All about direct methods, pp. 267–278.
- [22] J. Kybic and M. Unser, “Fast parametric elastic image registration,” *IEEE Transactions on Image Processing*, vol. 12, no. 11, pp. 1427–1442, November 2003.