SUB-PIXEL REGISTRATION OF NOISY IMAGES

Volodya Grancharov, W. Bastiaan Kleijn

KTH (Royal Institute of Technology)
Sound and Image Processing Lab
100 44 Stockholm, Sweden
{Volodya.Grancharov, Bastiaan.Kleijn}@ee.kth.se

Alexander Georgiev

Stockholm University
Department of Biochemistry and Biophysics
106 91 Stockholm, Sweden
Alexander.Georgiev@dbb.su.se

ABSTRACT

The accurate registration of images observed in additive noise is a challenging task. The noise increases the number of misregistered regions, and decreases the accuracy of subpixel registration. To address this problem, we propose an intensity-based algorithm that performs registration based only on regions that are least affected by noise. We select these regions with a signal-to-noise ratio estimate that is obtained from an initial, less-accurate registration. Our simulations demonstrate that the proposed noise-adaptive scheme significantly outperforms the conventional registration approach.

1. INTRODUCTION

The goal of image registration is to map all points in one image plane to positions in a second plane. Image registration has applications in computer vision, such as image matching for stereo vision, pattern recognition, and motion analysis. It is also used in areas such as environmental monitoring, weather forecasting, geographic information systems, super-resolution, and computer tomography [1, 2].

The standard problem addressed in the image registration literature can be formulated as the optimization of the deformation function G that gives the optimal registration of a *target image* $I^t(x, y)$ onto a *reference image* $I^r(x', y')$, with respect to a cost function C:

$$G^{\text{opt}} = \arg\underset{G}{\text{opt}}\ C(I^r, G(I^t)). \tag{1}$$

In practice, images are often observed in noisy conditions that significantly reduce the precision of the alignment. The conventional approach to deal with additive noise is to smooth both the reference and target images. The level of blur is not related to the noise level, but is of a pre-defined constant level, affecting the registration accuracy. The resulting accuracy of registration in noise has been described in [3]. Attempts to eliminate the effect of noise disturbance by modifying the cost function can be found in [4]. To the best of our knowledge, no general study on the influence of noise or on the efficacy of noise attenuation algorithms on registration exists.

Equation (1) is the starting point of most image registration studies. However, the *single image registration* problem stated above does not occur often, in practice. We usually observe a set of noisy images $\{I^0, I^1, \ldots, I^{L-1}, I^L\}$, select a reference image among them, and search for a geometrical transform to register the rest of the images onto the selected one. Without loss of generality we select $I^0(x, y)$ to be the reference image. We assume that both the reference and target images are generated from the same image scene $I(x, y)$. We have for $k = 1 \ldots L$:

$$\begin{aligned} I^0(x, y) &= I(x, y) + N^0(x, y) \\ I^k(x, y) &= G^{-1}_{\mathbf{M}_k}\left(I(x, y) + N^k_1(x, y)\right) + N^k_2(x, y), \end{aligned} \tag{2}$$

where $N^0(x, y)$, $N^k_1(x, y)$, and $N^k_2(x, y)$ are zero-mean random noise disturbances, and search for the set of L parameterized geometrical deformations $G_{\mathbf{M}_k}$ that register the target images onto reference image.

Typically, the space of geometric deformations is constrained to a predefined parametric family of deformations and one searches for a set of optimal parameters. We assume that the image is warped by a global affine transform described by the affine matrix $\mathbf{M}$.

A critical choice in the design of a registration algorithm is the selection of the image cost function and the image representation. Gradient-based shift estimation techniques, e.g., [5] often exhibit degraded performance under noisy conditions due to the fact that the gradient operator amplifies noise [6]. For moment-based registration techniques, e.g., [7], the noise tolerance is also weak since noise can lead to imperfect moment estimates and large errors in parameter determination. These observations motivated us to use intensity-based registration algorithm.

We select as cost function for the optimization procedure the squared sum of intensity differences (SSD):

$$C_{\text{SSD}} = \|\mathbf{I}^0 - G_{\mathbf{M}_k}(\mathbf{I}^k)\|^2. \tag{3}$$

We observed that this criterion is as robust to noise as measures such as normalized cross-correlation and mutual information [8], but at a lower computational cost.

Using equations (2) and (3), the problem of multi-channel image registration becomes

$$\mathbf{M}^{\text{opt}}_k = \arg\min_{\mathbf{M}_k}\ C_{\text{SSD}}(I^0, G_{\mathbf{M}_k}(I^k)),\ k = 1 \ldots L. \tag{4}$$

We describe a new, noise-robust two-step registration algorithm as the solution to the problem defined in equation (4). The purpose of the first step is a coarse registration to provide an accurate local estimate of the signal-to-noise ratio (SNR). The second high-resolution registration step is based only on the regions least affected by noise, as indicated by the SNR results of the first step. As shown by the results, our algorithm leads to highly accurate image registration.

2. NOISE-ROBUST REGISTRATION ALGORITHM

In this section, we first describe an initial registration that is insensitive to noise. In subsection 2.2 we describe how to use this initial registration to estimate the local SNR and then use the SNR to make a more accurate second registration.

2.1. Initial Registration Algorithm

The basic registration algorithm has as first step the selection of control points. The control points are matched between the images and the corresponding parameters of the mapping functions are determined. This operation is based on the known coordinates of control points. Next, the target image is geometrically transformed by means of a mapping function.

The *Transform Model* used in this presentation is a general affine transform. It can address scaling, transformation, rotation, skew, and aspect ratio. The affine mapping from old (x, y) to new pixel coordinates (x', y') is defined as:

$$G^{\text{Affine}}: \left| \begin{array}{lcl} x' & = & m_{11}x + m_{12}y + m_{13} \\ y' & = & m_{21}x + m_{22}y + m_{23} \end{array} \right.$$

For the sake of presentation simplicity we define the affine transformation as $\mathbf{z}' = \mathbf{M}\mathbf{z}$, where $\mathbf{z} = [x\ y\ 1]^T$, $\mathbf{z}' = [x'\ y'\ 1]^T$, and the affine transform matrix is of the form:

$$\mathbf{M} = \begin{pmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ 0 & 0 & 1 \end{pmatrix}.$$

The first step in the registration algorithm is the *control points* extraction. The idea to register only regions of the image with rich texture is of even greater importance in noisy image registration, since the selected high energy regions are less affected by the noise. In the proposed algorithm we use the Harris corner detector [9] for extraction of control points from the reference image. The set of control points is a subset of all points in the image, and they are denoted as $\{\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_N\}$ in the target image and $\{\mathbf{z}'_1, \mathbf{z}'_2, \ldots, \mathbf{z}'_N\}$ in the reference image.

It is not appropriate to track a single pixel, but a small window around each control point. An appropriate window size should be determined so that it is large enough to be statistically significant and stable, but sufficiently small to minimize the local geometric variation within the image window (cf. Fig. 1). The only method yielding the globally optimal solution is an exhaustive search over the entire image that is computationally not feasible. We use a multi-resolution hill-climbing optimization [10]. We generate Gaussian pyramids [11], on both reference and target image, see Fig. 2. The image registration starts with a coarse resolution, and when the optimal parameters are found, they are used as an initial guess for the next level. The estimates of the parameters gradually improve with increasing resolution. One advantage of pyramidal approach is that computational cost is greatly reduced. In addition, the resolution pyramid regularizes the optimization problem by causing the error surface to be smoother at a coarser resolution.

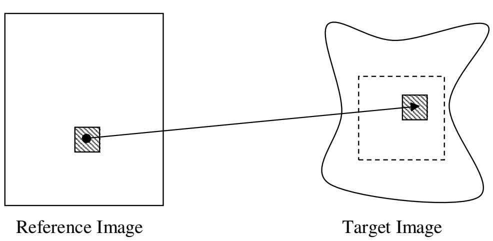

Fig. 1. Sliding window around a control point (small filled rectangular) and its exploration area (large dashed contour).

Once the minimum is located at pixel level, we perform *error interpolation for subpixel accuracy*. The subpixel shift is the distance between the minimum of the interpolation function and the closest integer pixel. Bicubic interpolation is used on a 5 x 5 grid around the estimated minimum.

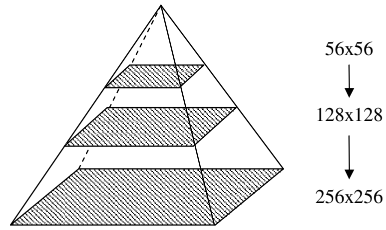

Fig. 2. Example of a three-level pyramid scheme. Control points locations are propagated from the top to the pixel level. The subpixel shift and affine parameters are calculated only on the pixel level.

The affine model parameters are derived from the matched control points. Given a number of corresponding control points in two images, we estimate the parameters for the mapping function. A least-square approach is used to determine the affine matrix, minimizing the criterion:

$$J_{MSE} = \frac{1}{N} \sum_{n=1}^{N} \|\mathbf{z}'_n - \hat{\mathbf{M}}\mathbf{z}_n\|. \tag{5}$$

The estimate of affine matrix that minimizes the above criterion can be calculated as:

$$\hat{\mathbf{M}} = (\mathbf{D}_\mathbf{z}^T \mathbf{D}_\mathbf{z})^{-1} \mathbf{D}_\mathbf{z}^T \mathbf{D}_{\mathbf{z}'}, \tag{6}$$

where the control points are stored in the data matrices $\mathbf{D}_\mathbf{z} = [\mathbf{z}_1 \mathbf{z}_2 \ldots \mathbf{z}_N]$ and $\mathbf{D}_{\mathbf{z}'} = [\mathbf{z}'_1 \mathbf{z}'_2 \ldots \mathbf{z}'_N]$.

Some of the control points may not be properly aligned and have to be rejected as outliers. We use a simple threshold criterion to select only a subset of control points, by rejecting the control points that do not fulfill the criterion:

$$\|\mathbf{z}'_n - \hat{\mathbf{M}}\mathbf{z}_n\| < \Theta, \tag{7}$$

where the pre-defined threshold Θ is set to 0.8. With the outliers removed, we recalculate the affine model (6). Once we have the optimal affine parameters $\mathbf{M}^{\text{opt}}$ at the control points, the target image is warped with Catmull-Rom splines [12]. Catmull-Rom splines resemble a sinc function, but do not introduce excessive ringing, which makes them attractive for image registration.

The proposed registration algorithm is iterative, and it is outlined in Table 1. With each iteration, the sliding window centered around each control point is less in violation of the assumed affine transform.

2.2. Refined Registration; SNR-Based Control-Point Selection

The refined registration step is a repetition of the basic registration, but is based only on control points with a high level of noise immunity. The proper control points selection requires an accurate estimate of the local SNR that is described in the current section.

The output of the initial coarse registration step is a set of geometrically transformed noisy images

$$I^k_* = G_{\mathbf{M}_k}(I^k),\ k = 1 \ldots L, \tag{8}$$

which, together with the reference I^0, we use to obtain an initial estimate of the source scene

$$\hat{I} = \text{median}\{I^0, I^1_*, I^2_*, \ldots, I^L_*\}. \tag{9}$$

Table 1. Summary of the registration algorithm used in both steps. Iterations are terminated when the improvement is below the predefined threshold Υ set to 10^{-6}.

do:

1. Select control points
2. Match control points
   (a) optimize for C_{SSD} measure
   (b) hill-climbing on a multiresolution pyramid as a search strategy
   (c) bicubic interpolate the error surface to obtain the subpixel shift
3. Calculate affine parameters from matched control points
   (a) optimize for J_{MSE}
   (b) reject outliers, before computing the affine parameters
4. Affine warp the target image $I_{i+1}^k = G_{\mathbf{M}_k}(I_i^k)$, with Catmull-Rom splines

while: $|C_{SSD}(I^0, I_i^k) - C_{SSD}(I^0, I_{i+1}^k)| \geq \Upsilon$

The median operator's behavior is used since outliers resulting from misregistration do not affect the registration quality. The denoising effect of the median filtering as a function of the number of noisy observations is shown in Fig. 3. We conclude that with only ten observations we obtain a SNR gain of about 10 dB.

We use the knowledge of the source scene to estimate locally the noise variance $\hat{\sigma}_{N^k}^2$ for the reference and each of the target images. Note that the noise in the reference image is $N^0(x, y)$, and that the effective noise for the warped target images $I_*^k(x, y)$ can be expressed as

$$N^k(x, y) \approx N_1^k(x, y) + G_{\hat{\mathbf{M}}_k}\left(N_2^k(x, y)\right). \qquad (10)$$

Let $\sigma_{I^k}^2$ be the sample variance of the noisy image, estimated from a local neighborhood of each pixel. Then the noise variance is estimated as

$$\hat{\sigma}_{N^k}^2 = \sigma_{I_*^k}^2 - \sigma_{\hat{I}}^2. \qquad (11)$$

Next we define the SNR at each control-point position (as found by the Harris corner detector on the noisy reference image) as

$$\text{SNR}(\mathbf{z}'_n) = 10 \log_{10}\left(\frac{\sigma_{I_*^k}^2(\mathbf{z}'_n)}{\hat{\sigma}_{N^k}^2(\mathbf{z}'_n)}\right), \qquad (12)$$

and assign a SNR label $\{\text{SNR}(\mathbf{z}'_1), \text{SNR}(\mathbf{z}'_2), \ldots, \text{SNR}(\mathbf{z}'_N)\}$ to each control point $\{\mathbf{z}'_1, \mathbf{z}'_2, \ldots, \mathbf{z}'_N\}$. We have found experimentally that it is sufficient to reject control points with a SNR below 5 dB.

Thus, the proposed algorithm consists of two registration steps. The first step performs the basic registration based on the control points $\{\mathbf{z}'_1, \mathbf{z}'_2, \ldots, \mathbf{z}'_N\}$ that is described in section 2.1. The second step is a refined registration that is identical to the first step except that it is based on a new set of control points $\{\mathbf{z}'_1, \mathbf{z}'_2, \ldots, \mathbf{z}'_S\}$ that were selected for their high local SNR. The second registration step is not performed if the mean SNR at the control points, after the initial registration, is above 25 dB.

3. EVALUATION

For the assessment of the proposed algorithm, we used artificially deformed images, thus facilitating the evaluation of the performance

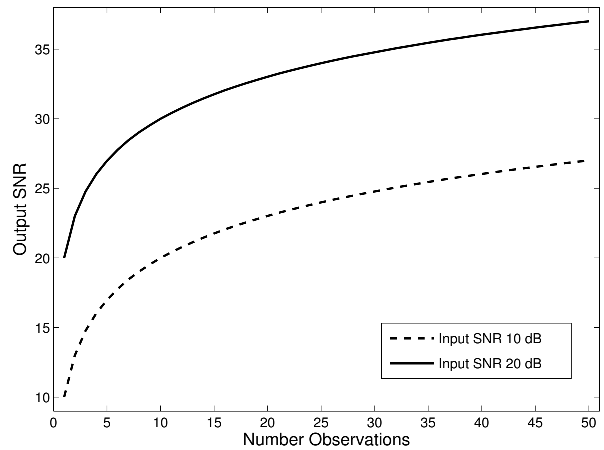

Fig. 3. The denoising effect of noisy images by median filtering, for two input SNR levels. The output SNR is measured between the true $I(x, y)$ and estimated $\hat{I}(x, y)$.

of the algorithm. The similarity measures used for evaluation are the Hilbert-Schmidt norm on the difference between true and estimated affine matrices $C_M = \|\mathbf{M} - \hat{\mathbf{M}}\|$, and the mean squared error at control points $C_{CP} = \frac{1}{N}\sum_{n=1}^{N} \|\mathbf{z}'_n - \mathbf{z}_n\|$. For the simulations we used a popular set of eight images: *Lena, Barbara, Mandrill, Goldhill, House, Peppers, Boats,* and *Cameramen*. All images are of size 256 x 256 pixels.

3.1. Combination of Gaussian and Impulsive Noise

We warp the original image with a deformation belonging to the warp space (general affine transform):

$$G^{\text{Affine}}: \left| \begin{array}{lcl} x' & = & 0.94x - 0.03y + 15.2 \\ y' & = & 0.20x + 0.98y - 11.0, \end{array} \right.$$

and add white Gaussian noise to the target images at 5 dB SNR, combined with impulsive noise with probability 0.02. An example of images used in the simulations is shown in Fig. 4. The results of simulations, presented in Table 2, clearly indicate the advantage of the refined registration step.

Table 2. Evaluation of squared error at control points, and model parameter accuracy. Results are averaged over eight images.

	Initial Registration	Refined Registration
C_{CP}	0.310	0.204
C_M	0.121	0.099

3.2. Locally Varying Noise

It is common that registration is performed for images with spatially varying distortion. Due to its ability to select only meaningful data regions, our algorithm is particularly robust to this type of distortion.

To evaluate the performance in spatially varying noise, we added white Gaussian noise in a single 64 x 64 pixels rectangle, arbitrarily located within each of the images. The target images were warped

Fig. 4. The images at the Row 1 are noisy reference $I^0(x, y)$ and one of the target images $I^k(x, y)$. At the Row 2 we see the ideal scene $I(x, y)$ and the median of the registered images $\hat{I}^k(x, y)$.

Fig. 5. Row 1: control points localization for "ideal" and global-noise case. Row 2: points localization for two configurations locally varying noise.

with the affine model described in the previous section. The results shown in Table 3 indicate that the refined registration step gives a stronger improvement for this case than the spatially uniform noise case. To explain the simulation results, we present the localization of the control points on the reference image for clean, global and local noise cases in Fig. 5. The localized distortion region captures control points and these are difficult to match, resulting in a large model error. However, since our algorithm marks the noisy regions as low SNR regions, it prevents the selection of control points in these regions and the registration accuracy is similar to that for noise-free images.

Table 3. Evaluation of squared error at control points, and model parameter accuracy. Results are averaged over eight images.

	Initial Registration	Refined Registration
C_{CP}	0.641	0.105
C_M	0.206	0.032

4. DISCUSSION

We presented a robust registration algorithm that can register noisy images with sub-pixel accuracy. Our main contribution is that image regions that are least affected by noise should be selected and tracked in the registration procedure. We showed that this improves the registration of noisy images significantly. If the noise level varies spatially, as is common, the performance improvement is particularly dramatic. While our implementation is for the global affine geometrical transformation, we expect the method to work well for other transformations as well. If more complex models are needed, they can be built on the affine transform [5]. It is expected that the noise sensitivity will be higher for more complex geometric transforms, increasing the benefit of the proposed scheme of SNR adaptive region selection.

5. REFERENCES

[1] L. Brown, "A survey of image registration techniques," *ACM Comput. Surv.*, vol. 24, pp. 326–376, 1992.

[2] B. Zitova and J. Flusser, "Image registration methods: a survey," *Image Vision Comput.*, vol. 21, pp. 977–1000, 2003.

[3] J. Kybic and M. Unser, "Fast parametric elastic image registration," *IEEE Trans. Image Process*, vol. 12, pp. 1427–1442, 2003.

[4] S. Cain, M. Hayat, and E. Armstrong, "Projection-based image registration in the presence of fixed-pattern noise," *IEEE Trans. Image Process*, vol. 10, pp. 1860–1872, 2001.

[5] M. Irani, B. Rousso, and S. Peleg, "Computing occluding and transparent motions," *Int. J. Computer Vision*, vol. 12, pp. 5–16, 1994.

[6] D. Robinson and P. Milanfar, "Bias minimizing filter design for gradient-based image registration," *Signal Processing: Image Communication*, vol. 20, pp. 554–568, 2005.

[7] X. Dai and S. Khorram, "A feature-based image registration algorithm using improved chain-code representation combined with invariant moments," *IEEE Trans. Geosci. Remote Sensing*, vol. 37, pp. 2351–2362, 1999.

[8] P. Viola and W. Wells III, "Alignment by maximization of mutual information," in *Proc. 5th Int. Conf. Computer Vision*, 1995, pp. 16–23.

[9] C. Harris and M. Stephens, "A combined corner and edge detector," in *Proc. 4th Alvey Vision Conf.*, 1988, pp. 147–151.

[10] E. Rich and K. Knight, *Artificial Intelligence*. McGraw Hill, 1991.

[11] A. Rosenfeld, Ed., *Multiresolution Image Processing and Analysis*. Springer Verlag, 1984.

[12] E. Catmull and J. Rom, *A class of local interpolating splines*. in Computer Aided Geometric Design, R. Barnhill and R. Riesenfeld, Eds., Academic Press Orlando, 1974.