Pseudo-polar based estimation of large translations rotations and scalings in images

Yosi Keller
Department of Mathematics
Yale Univrsity
New Haven, CT, USA
yosi.keller@yale.edu

Amir Averbuch
Department of Computer Science
Tel-Aviv University
Tel-Aviv, Israel

Moshe Israeli
Department of Computer Science
Technion Institute of Technology
Haifa, Israel

Abstract

One of the major challenges related to image registration is the estimation of large motions without prior knowledge. This paper presents a Fourier based approach that estimates large translation, scale and rotation motions. The algorithm uses the pseudo-polar transform to achieve substantial improved approximations of the polar and log-polar Fourier transforms of an image. Thus, rotation and scale changes are reduced to translations which are estimated using phase correlation. By utilizing the pseudo-polar grid we increase the performance (accuracy, speed, robustness) of the registration algorithms. Scales up to 4 and arbitrary rotation angles can be robustly recovered, compared to a maximum scaling of 2 recovered by the current state-of-the-art algorithms. The algorithm utilizes only 1D FFT calculations whose overall complexity is significantly lower than prior works. Experimental results demonstrate the applicability of these algorithms.

1 Introduction

Image registration plays a vital role in many image processing applications such as video compression [1], video enhancement [2] and scene representation [3] to name a few. This problem was analyzed using various computational techniques, such as pixel domain Gradient methods [2], correlation techniques [15] and discrete Fourier (DFT) domain algorithms [6, 11]. Gradient methods based image registration algorithms are considered to be the state-of-the-art. They may fail unless the two images are misaligned by only a moderate motion. Fourier based schemes, which are able to estimate relatively large rotation, scaling and translation, are often used as bootstrap for more accurate gradient methods. The basic notion related to Fourier based schemes is the *shift property* [18] of the Fourier transform which allows robust estimation of translations using the *normalized phase-correlation* algorithm [6, 9, 10]. Hence, in order to account for rotations and scaling, the image is transformed into a polar or log-polar Fourier grid (referred to as the Fourier-Mellin transform). Rotations and scaling are reduced to translations in these representations and can be estimated using phase-correlation.

In this paper we propose to iteratively estimate the polar and log-polar DFT using the pseudo-polar FFT (PPFFT) [19]. The resulting algorithm is able to robustly register images rotated by arbitrary angles and scaled up to a factor of 4. It should be noted that the maximum scale factor recovered in [11] and [16] was 2.0 and 1.8, respectively. In particular, the proposed algorithm does not result to interpolation in either spatial or Fourier domain. Only 1D FFT operations are used, making it much faster and especially suited for real-time applications.

The rest of paper is organized as follows: Prior results related to FFT based image registration are given in Section 2, while the proposed algorithm, is presented in Section 3. Experimental results are discussed in Section 4 and final conclusions are given in Section 5.

2 Previous related work

2.1 Translation estimation

The basis of the Fourier based motion estimation is the *shift property* [18] of the Fourier transform. Denote by

$$\mathfrak{F}\left\{f\left(x,y\right)\right\} \triangleq \widehat{f}\left(\omega_x,\omega_y\right) \tag{1}$$

the Fourier transform of $f\left(x,y\right)$. Then,

$$\mathfrak{F}\left\{f\left(x+\Delta x,y+\Delta y\right)\right\} = \widehat{f}\left(\omega_x,\omega_y\right)e^{j\left(\omega_x\Delta x+\omega_y\Delta y\right)}. \tag{2}$$

Equation 2 can be used for the estimation of image translation [6, 10]. Assume the images $I_1\left(x,y\right)$ and $I_2\left(x,y\right)$ have some overlap that

$$I_1\left(x+\Delta x,y+\Delta y\right) = I_2\left(x,y\right). \tag{3}$$

Proceedings of the IEEE Workshop on Motion and Video Computing (WACV/MOTION'05)
0-7695-2271-8/05 $ 20.00 IEEE

Equation 3 is Fourier transformed, so that

$$\widehat{I}_1(\omega_x, \omega_y)\, e^{j(\omega_x \Delta x + \omega_y \Delta y)} = \widehat{I}_2(\omega_x, \omega_y) \qquad (4)$$

and

$$\frac{\widehat{I}_2(\omega_x, \omega_y)}{\widehat{I}_1(\omega_x, \omega_y)} = e^{j(\omega_x \Delta x + \omega_y \Delta y)}. \qquad (5)$$

Thus, the translation parameters $(\Delta x, \Delta y)$ can be estimated in the spatial domain [6, 7] by taking the inverse FFT of Eq. 5:

$$Corr(x, y) \triangleq \mathfrak{F}\;{}^{1}\left\{e^{j(\omega_x \Delta x + \omega_y \Delta y)}\right\} = \delta(x \quad \Delta x, y \quad \Delta y) \qquad (6)$$

and by finding the position of the maximum value of the correlation function $Corr(x, y)$:

$$(x, y) = \arg\left\{\max_{(x,y)} \{Corr(x, y)\}\right\}. \qquad (7)$$

In order to compensate for possible affine intensity changes Eq. 5 is reformulated as:

$$\begin{aligned} \widetilde{Corr}(\omega_x, \omega_y) &\triangleq \frac{\widehat{f_1}(\omega_x, \omega_y)\, \widehat{f_2}^{*}(\omega_x, \omega_y)}{\left|\widehat{f_1}(\omega_x, \omega_y)\right| \left|\widehat{f_2}(\omega_x, \omega_y)\right|} \qquad (8) \\ &= e^{j(\omega_x \Delta x + \omega_y \Delta y)} \end{aligned}$$

where * denotes the complex conjugate.

2.2 Polar Fourier representations

The polar Fourier representation (Fourier-Mellin transform) is used to register images that are misaligned due to translation, rotation and scale [11, 16]. Let $I_1(x, y)$ be a translated, rotated and scaled replica of $I_2(x, y)$

$$\begin{aligned} I_1(x, y) = I_2(&s \cdot x \cdot cos\theta_0 + s \cdot y \cdot sin\theta_0 + \Delta x, \\ &s \cdot x \cdot sin\theta_0 + s \cdot y \cdot cos\theta_0 + \Delta y) \qquad (9) \end{aligned}$$

where θ_0, s and $(\Delta x, \Delta y)$ are the rotation angle, scale factor and translation parameters, respectively. The DFT of Eq. 9 is

$$\widehat{I}_2(r, \theta) = e_1^{j(\omega_x \Delta x + \omega_y \Delta y)} \frac{1}{s^2} \widehat{I}_1\left(\frac{r}{|s|}, \theta + \theta_0\right). \qquad (10)$$

Hence, M_1 and M_2, the magnitudes of $\widehat{I}_1$ and $\widehat{I}_2$, respectively, are related by rotation and scaling around the DC component

$$M_2(r, \theta) = \frac{1}{s^2} M_1\left(\frac{r}{|s|}, \theta + \theta_0\right). \qquad (11)$$

Therefore, rotation and scale changes can be recovered first, regardless the translation parameters. Using a polar or log-polar DFT, rotation and scaling are reduced to translations, which can be robustly recovered using the phase-correlation procedure. Using Eq. 11 to estimate the rotation angle θ results in an ambiguity by a factor of π [11] in the estimation of the rotation angle. This ambiguity can be resolved by applying both hypotheses θ and $\theta + \pi$ and recover the translational motion $(\Delta x, \Delta y)$ and the correlation peak of each hypothesis. The rotation hypothesis and translation values, which are related to the highest correlation peak, are chosen as the result.

2.3 The pseudo-polar FFT

The proposed registration algorithm is based on the existence of a fast, accurate and invertible discrete pseudo-polar FFT (PPFFT) [19]. An FFT where the evaluation frequencies lie in an oversampled set of non-angularly equispaced points which we call the pseudo-polar (PP) grid. This grid is shown in Fig. 1. The PPFFT includes fast forward and inverse transforms and a quasi-Parseval relation. A special feature of this approach is that it involves only 1-D equispaced FFT's. In particular, there is no need for re-gridding or interpolation. The grid resembles the polar grid, while having a fast and accurate computation scheme.

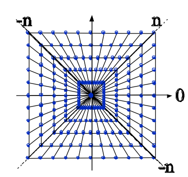


Figure 1: The grid on which the pseudo-polar FFT is evaluated. The grid resembles the polar grid.

3 The proposed registration algorithm

This section presents the proposed image registration algorithms, which use the pseudo-polar FFT (PPFFT) to reduce rotations and scalings to translations in the polar and log-polar domains, respectively.

3.1 Simultaneous estimation of rotation and translation

Let $I_1(i, j)$ and $I_2(i, j)$ be the input images to be registered. $I_1^{(n)}(i, j)$ denotes the fact that I_1 evolves throughout the registration process, while $I_2(i, j)$ remains unchanged. We assume that initially $I_1^{(0)}(i, j) = I_1(i, j)$ $(n = 0)$. The rotation and translation estimation algorithm operates as follows:

1. Let (m_1, l_1) and (m_2, l_2) be the sizes of $I_1(i,j)$ and $I_2(i,j)$, respectively. Then, at iteration $n = 0$, $I_1(i,j)$ and $I_2(i,j)$ are zero padded such that
$$m_1 = l_1 = m_2 = l_2 = 2^k, k \in \mathbb{Z}. \tag{12}$$
2. The magnitudes $M_1^{PP}(\theta_i, r_j)$ and $M_2^{PP}(\theta_i, r_j)$ of the PPFFTs of $I_1^{(n)}(i,j)$ and $I_2(i,j)$ are calculated, respectively.
3. The polar DFTs ,magnitudes $\widehat{M}_1^{Polar}(\theta_i, r_j)$ and $\widehat{M}_2^{Polar}(\theta_i, r_j)$ of $I_1^{(n)}(i,j)$ and $I_2(i,j)$ are substituted by $M_1^{PP}(\theta_i, r_j)$ and $M_2^{PP}(\theta_i, r_j)$ respectively.
4. The translation along the $\overrightarrow{\theta}$ axis of $M_1^{PP}(\theta_i, r_j)$ and $M_2^{PP}(\theta_i, r_j)$ is estimated using phase correlation. The result is denoted by $\Delta\theta_n$.
5. Let θ_n be the accumulated rotation angle estimated at iteration n
$$\theta_n \triangleq \sum_{i=0}^{n} \Delta\theta_i = \theta_{n-1} + \Delta\theta_n.$$
Then, the input image $I_1(i,j)$ is rotated by θ_n (around the center of the image) using the FFT based image rotation algorithm described in [4]. This rotation scheme is accurate and fast since only 1D FFT operations are used
$$I_1^{(n+1)}(\theta, r) = I_1^{(0)}(\theta + \theta_n, r),\ n = 1, \dots$$
The rotation can be conducted around any pixel. We recommend to use the central pixel of $I_1(i,j)$ such that the bounding rectangular of the rotated image will be as small as possible.
6. Steps 2-5 are reiterated until the angular refinement term $\Delta\theta_n$ is smaller than a predefined threshold ε_θ, i.e. $|\Delta\theta_n| < \varepsilon_\theta$, or a predefined number of iterations $n_{\max}$ is reached.
7. The rotated images $I_1^{(n)}(i,j)$ and $I_2(i,j)$ are then registered by phase correlation and the rotation ambiguity is resolved by the procedure presented in section 2.2.

3.2 Simultaneous estimation of rotation, translation and scaling

Let $I_1(i,j)$ and $I_2(i,j)$ be the input images to be registered. $I_1^{(n)}(i,j)$ denotes the fact that I_1 evolves throughout the registration process, while $I_2(i,j)$ remains unchanged. We assume that initially $I_1^{(0)}(i,j) = I_1(i,j)$ $(n = 0)$. The scaling, rotation and translation estimation algorithm operates as follows:

1. Let (m_1, l_1) and (m_2, l_2) be the sizes of $I_1^{(n)}(i,j)$ and $I_2(i,j)$, respectively. Then, at iteration $n = 0$, $I_1^{(0)}(i,j)$ and $I_2(i,j)$ are zero padded such that
$$m_1 = l_1 = m_2 = l_2 = 2^k, k \in \mathbb{Z}. \tag{13}$$
The padding is needed to abide by the limitations of the PPFFT described in [19].
2. The magnitudes $M_1^{PP}(\theta, r)$ and $M_2^{PP}(\theta, r)$ of the PPFFTs of $I_1^{(n)}(i,j)$ and $I_2(i,j)$ are calculated, respectively.
3. Let
$$base^{length\left(M_2^{PP}(\theta_i, r_j)\right)} = length\left(M_2^{PP}(\theta_i, r_j)\right) \tag{14}$$
be the logarithmic base used to define the log domain in the $\overrightarrow{r}$ axis. The log-polar DFTs of $I_1^{(n)}(i,j)$ and $I_2(i,j)$ are approximated by $M_1^{PP}(\theta, r)$ and $M_2^{PP}(\theta, r)$, respectively, using the following procedure:
$$\begin{aligned} \widehat{M}_1^{Log\ \ Polar}(i,j) &= M_1^{PP}\left(i, \lfloor base^j \rfloor\right) \\ \widehat{M}_2^{Log\ \ Polar}(i,j) &= M_2^{PP}\left(i, \lfloor base^j \rfloor\right) \end{aligned} \tag{15}$$
where $\lfloor x \rfloor$ denotes the integral part of x $(x > 0)$. The polar axis is approximated using the same procedure as in section 3.1, while the radial axis is approximated using nearest-neighbor interpolation.
4. The relative translation between $\widehat{M}_1^{Log\ \ Polar}(i,j)$ and $\widehat{M}_2^{Log\ \ Polar}(i,j)$ is recovered by a 2D phase correlation on the $\overrightarrow{\theta}$ and $\overrightarrow{r}$ axes.
5. Let $\Delta\theta_n$ and Δr_n be the rotation angle and the scaling value estimated at iteration n, respectively. Then, the input image $I_1(x,y)$ is rotated (around the center of the image) [4] and then scaled using DFT domain zero padding
$$I_1^{(n+1)}(\theta, r) = I_1^{(0)}(\theta + \theta_n, r \cdot r_n) \tag{16}$$
where
$$\begin{aligned} \theta_n &= \sum_{i=0}^{n} \Delta\theta_i = \theta_{n-1} + \Delta\theta_n \\ r_n &= \prod_{i=0}^{n} \Delta r_i = r_{n-1} \cdot \Delta r_n. \end{aligned} \tag{17}$$

6. Steps 1-3 are reiterated until either the updated estimates are smaller than predefined thresholds $|\Delta\theta_n| < \varepsilon_\theta$ and $\Delta r_n < \varepsilon_r$ or a predefined number of iterations $n_{\max}$ is reached.

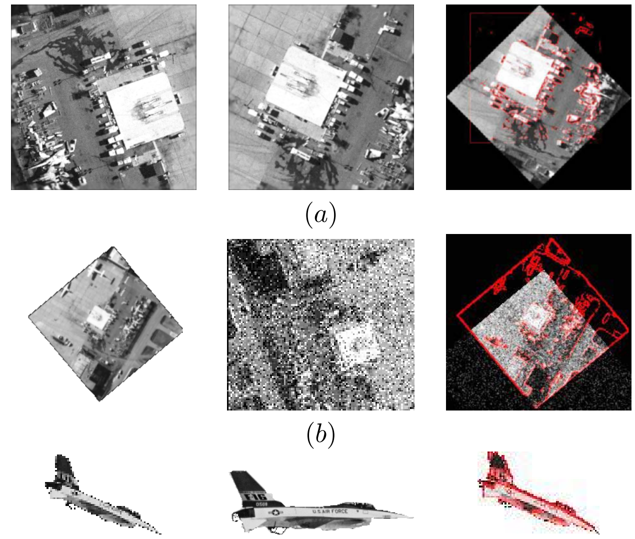
7. The rotation and scale parameters θ_n and r_n, respectively, are used as inputs to the spatial domain registration algorithm presented in section 3.1 step 7.

4 Experimental results

The registration algorithm was tested on the 128×128 *Lena* and *Airfield* images, which were rotated, scaled and translated to create the test pairs shown in Figs. 2 and 3 similar to [11, 15]. The images were registered using the rotation and rotation/scale algorithms presented in sections 3.1 and 3.2, respectively. The results are graphically presented by overlaying the edges of one of the images on the other using the estimated motion parameters, while the corresponding numerical results are presented in tables 1 and 2. The translation was estimated using the phase-correlation algorithm discussed in section 2.1 whose accuracy is limited to integral translation values. We note that the choice of integral scaling factors for the experiments is of no special significance for the performance of the proposed algorithm, which was able to register any scaling factor within its dynamic range that is discussed later.

Table 1 and Fig. 2 present the registration results of the *Airfield* image which contains man made objects characterized by sharp edges surrounded by smooth regions. The algorithm was able to robustly register the images in Fig. 2a having a scaling factor of 2 and a large rotation of 100°. This test was repeated for various rotation angles resulting in a similar accurate registrations. Figures 2b and 2c present a situation of partial alignment under large rotation. The registration algorithm was able to register the images since the DFT's magnitude is the Periodogram [18] and thus it is an approximation of the image's power spectrum [18]. These statistical properties are invariant to partial alignment when the non-corresponding image parts have similar statistical properties. The robustness property was examined using the images in Fig. 2c, where significant noise was added to one of the images and registration was conducted using rotation and rotation/scale registration modes. In both cases the translation is evaluated afterward. When the rotation mode is used, we were able to register images in both noisy and noise-free scenarios, while the algorithm diverged when registering the images using the rotation/scale mode.

The numerical results in table 1 show that when the algorithm converges it achieves high accuracy. No consistent misalignments ("local minima") were observed. The scale/rotation mode was found to be less stable resulting in the algorithm's failure to estimate the motion in Fig. 2b using this mode. When the algorithm converged, convergence was achieved within 3-4 iterations for all image sets while consuming 5-7 seconds of computational time on a 1.5MHz P4 Win2000 computer, where the algorithm was implemented using non-optimized C++.



(a) (b) (c)

Figure 2: Registration results of *Airfield* and *F16* images. In each row, the left and center images were registered. The edges of the images in the center column are overlayed on the left images to illustrate the registration accuracy. The results are shown in the right column. (a) Scale = 1, Angle = 100°. (b) Robustness was tested by adding noise $(\mu = 0, \sigma = 40)$ to the images related by Scale=1, Angle = 130°. (c). Registration of a segmented F16 image. Scale = 3, Angle = 23°.

The images in Fig. 3 were used to test the algorithm for registration of images with large scaling factors. Prior algorithms [11] reported to recover scalings up to a factor of 2. The proposed algorithm was successfully tested for scalings factors up to 4. The graphical registration results shown in Fig. 3 are accurate, where the numerical results are presented in table 2. Figures 3a, 3c and 3d demonstrate the algorithm's ability to recover large image scalings. We were not able to consistently recover scalings larger than 4. Computation times were similar to the results in Table 1 and Fig. 2. We conclude that the proposed algorithm is able to register images with scalings up to 4 with good accuracy. Within this range any rotation angle can be estimated while translations are estimated.

Images	Correct $(Scale, \theta)$	Registration mode	Computed $(Scale, \theta, \Delta x, \Delta y)$
a	$(1, 100°)$	scale,rotation,translation	$(1, 100.65°)$
b	$(1, 130°)$	rotation,translation	$(1, 129.72°)$
c	$(3, 23°)$	rotation,translation	$(3, 23.22°)$

Table 1: Registration results of the image sets presented in Figure 2 using the pseudo-polar based algorithm. The images are rotated and scaled by small scaling factors. The results demonstrate the algorithm's robustness even when noise is present and there is a partial alignment.

Image sequence	Registration mode	Correct $(Scale, \theta)$	Computed $(Scale, \theta)$
a	scale,rotation,translation	$(4.0, 105°)$	$(4.0, 103.78°)$
b	rotation,translation	$(1, 150°)$	$(1, 150.11°)$
c	scale,rotation,translation	$(1, 135°)$	$(1.03, 135.2°)$
d	scale,rotation,translation	$(3, 33°)$	$(3.01, 32.9°)$

Table 2: Registration results of the *Lena* images presented in Figure 3 using the pseudo-polar based algorithm.

5 Conclusions

In this paper we proposed a FFT based image registration algorithm, which was shown to be able to recover large rotations and scaling factors. This algorithm which uses the pseudo-polar FFT [19] enhance the current state-of-the-art image registration algorithms. The overall complexity is dominated by FFT operations and is $O\left(n^2 \log n\right)$. The algorithm can achieve close to real-time performance by using machine specific optimized FFT implementations which are widely available.

References

[1] A. Tekalp, "Digital Video Processing", Prentice Hall, 1995.

[2] M. Irani and S. Peleg, "Motion Analysis for Image Enhancement: Resolution, Occlusion and Transparency". Journal of Visual Communication and Image Representation, Vol. 4, No. 4, pp.324-335, December 1993.

[3] M. Irani, P. Anandan, and S. Hsu. "Mosaic based representations of video sequences and their applications", International Conference on Computer Vision, pages 605-611, 1995.

[4] M. Unser, P. Thvenaz, and L. Yaroslavsky, "Convolution-based interpolation for fast, high-quality rotation of images" IEEE Trans. Image Processing, vol. 4, pp. 1371–1381, Oct. 1995.

[5] S. Kruger and A. Calway, "Image Registration Using Multiresolution Frequency Domain Correlation". In British Machine Vision Conference, pages 316–325. British Machine Vision Association, September 1998

[6] C. D. Kuglin and D. C. Hines. "The phase correlation image alignment method", IEEE Conference on Cybernetics and Society, pp. 163–165, September 1975.

[7] H. Foroosh; J.B. Zerubia, M. Berthod, "Extension of phase correlation to subpixel registration", IEEE Trans. on Image Processing vol. 11 no. 3 , March 2002, pp. 188 -200, March 2002.

[8] M. Peyman, "Projection-based, Frequency-Domain Estimation of Superimposed Translational Motions", Journal of the Optical Society of America: A, Optics and Image Science, vol. 13, no. 11, pp. 2151-2162, Nov. 1996.

[9] Y. Chou and H. Hang, "A New Motion Estimation Method using Frequency Components", Journal of Visual Communication and Image Representation, Vol. 8, No. 1, pp. 83-96, 1997.

[10] H. Stone, M. Orchard and E. Chang, "Subpixel Registration of Images", 33rd Asilomar Conference on Signals, Systems, and Computers, October 1999

[11] S. Reddy and B. N. Chatterji. "An FFT-based technique for translation, rotation, and scale-invariant image registration". IEEE Trans. on Image Processing, 3(8):1266–1270, August 1996.

[12] A. Tabernero, J. Portilla,R. Navarro, “Duality of Log-Polar Image Representations in the Space and Spatial-frequency Domains”, IEEE Trans. on Signal Processing, vol. 47, No. 9, pp. 2469-2479, September 1999.

[13] L. Lucchese, “A Frequency Domain Technique Based on Energy Radial Projections for Robust Estimation of Global 2D Affine Transformations”, Computer Vision and Image Understanding Vol. 82, No. 1, pp. 72–116, Apr 2001.

[14] S. Derrode, F. Ghorbel , “Robust and Efficient Fourier-Mellin Transform Approximations for Gray-Level Image Reconstruction and Complete Invariant Description”, Computer Vision and Image Understanding, Vol. 83, No. 1, pp. 57-78, July 2001.

[15] G. Wolberg and S. Zokai, “Robust Image Registration Using Log-Polar Transform”, Proc. IEEE Intl. Conference on Image Processing, Vancouver, Canada, September 2000.

[16] Q. Chen, M. Defrise, and F. Deconinck. “Symmetric phase-only matched filtering of Fourier-Mellin transforms for image registration and recognition”. IEEE Transactions on Pattern Analysis and Machine Intelligence, 16(12):1156–1168, December 1994.

[17] S.A.Kruger and A.D.Calway, “Image registration using multiresolution frequency domain correlation”, Proc. British Machine Vision Conf, 1998.

[18] B. Porat, “A Course in Digital Signal Processing”, John Wiley Pub.,1997.

[19] A. Averbuch, D.L. Donoho, R.R Coifman, M. Israeli, and Y. Shkolnisky, “Fast slant stack: A notion of Radon transform for data in cartesian grid which is rapidly computable, algebrically exact, geometrically faithful and invertible”, to appear in SIAM Scientific Computing.

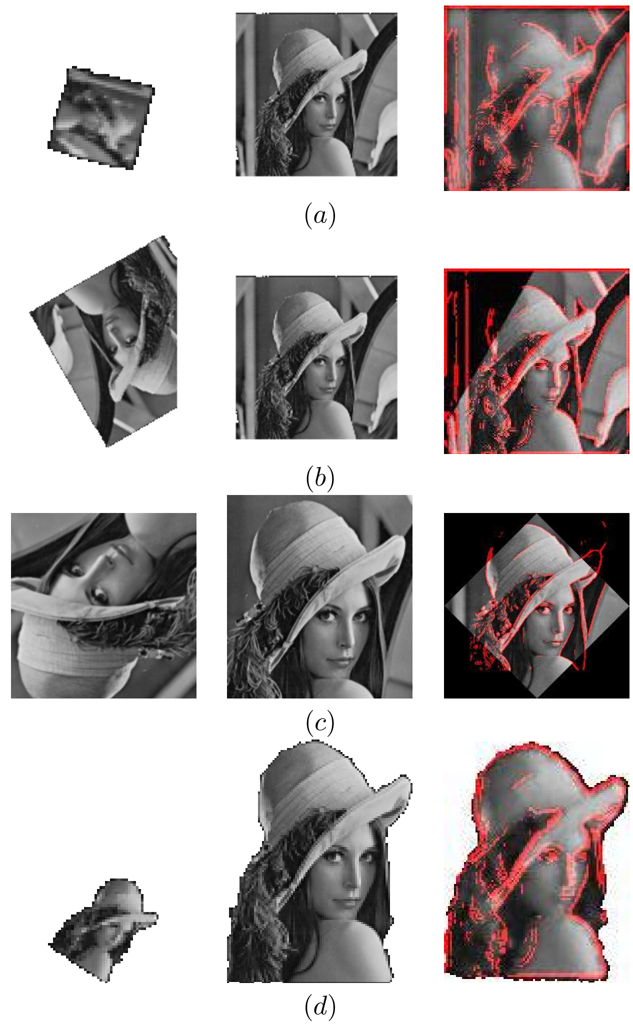


Figure 3: Registration results of the *Lena* image when large scalings factors are present. In each row, the left and center images were registered. The edges of the center image are overlayed on the left image to illustrate the registration accuracy. The output is given in the right column. (a) Scale = 4, Angle = 105°. (b) Images with partial non overlapping areas: Scale=1, Angle = 150°. (c) Scale = 1, Angle = 135°.