Automatic Registration of Satellite Images

Leila M. G. Fonseca[1]   Max H. M. Costa[2]

[1] National Institute for Space Research - INPE
CP 515, 12227-010 Sao José dos Campos, SP, Brazil
leila@dpi.inpe.br
[2] State University of Campinas - UNICAMP
Campinas, SP, Brazil
max@decom.fee.unicamp.br

Abstract. Image registration is one of the basic image processing operations in remote sensing. With the increase in the number of images collected every day from different sensors, automated registration of multi-sensor/multi-spectral images has become an important issue. A wide range of registration techniques has been developed for many different types of applications and data. Given the diversity of the data, it is unlikely that a single registration scheme will work satisfactorily for all different applications. A possible solution is to integrate multiple registration algorithms into a rule-based artificial intelligence system, so that appropriate methods for any given set of multisensor data can be automatically selected. The objective of this paper is to present an automatic registration algorithm which has been developed at INPE. It uses a multiresolution analysis procedure based upon the wavelet transform. The procedure is completely automatic and relies on the grey level information content of the images and their local wavelet transform modulus maxima. The algorithm was tested on SPOT and TM images from forest, urban and agricultural areas. In all cases we obtained very encouraging results.

Keywords: wavelet transform, image registration, matching, satellite image.

1 Introduction

Image registration is the process of matching two images so that corresponding coordinate points in the two images correspond to the same physical region of the scene being imaged. It is a classical problem in several image processing applications where it is necessary to match two or more images of the same scene [8].The registration process is usually carried out in three steps. The first step consists of selection of features on the images. Next each feature in one image is compared with potential corresponding features in the other image. A pair of features with similar attributes are accepted as matches and are called control points (CPs). Finally the parameters of the best transformation which models the deformation between the images are estimated using the CPs obtained in the previous step.

In remote sensing applications, users generally use manual registration which is not feasible in cases where there is a large amount of data. Thus, there is a need for automated techniques that require little or no operator supervision. Since the performance of a methodology is dependent on application specific requirements, sensor characteristics, and the nature and composition of the imaged area, it is unlikely that a single registration scheme will work satisfactorily for all different applications. Integration of multiple registration algorithms into a rule-based artificial intelligence system which can analyze the image data, and select an appropriate set of techniques for processing, appears to be a feasible alternative. Information such as the data type, present features in the imaged scene, registration accuracy, image variations, and noise characteristics could be provided by the user to assist in this process.

A prototype system which integrates some automatic remotely sensed image registration is currently under development at INPE (National Institute for Space Research). These algorithms will be integrated within a geographic information and image processing system - SPRING [1], which has been developed at INPE. As an initial part of this project an algorithm for automatic satellite image registration based on a multiresolution wavelet transform analysis has been implemented.

A description of the mentioned algorithm and presentation of some satellite image registration results are the objectives of this paper. The technique is similar to that described in [16] but differs in some aspects. In the process of feature selection the algorithm uses feature points which are detected from the local modulus max-

ima of the wavelet transform. The correlation coefficient is used as a similarity measure and only the best pairwise fitting among all pairs of feature points are taken as control points. A consistency checking step is also involved to eliminate mismatches. This way we have a reliable initial guess for the registration transformation which is a crucial phase in the process. A 2-D affine transformation is used to model the deformation between the images and their parameters are estimated in a coarse-to-fine manner.

The registration algorithm is very simple and easy to apply because it needs basically one parameter. Because the matching is carried out only on the selected feature points and in a coarse-to-fine manner, a significant amount of computation is saved in comparison to traditional pixel-by-pixel searching methods. Due to the fact that the registration procedure uses the grey level information content of the images in the matching process it is more adequate to register images of the same sensor or with similar spectral bands. In spite of this, it has demonstrated technical feasibility for many images of forest, urban and agricultural areas from Thematic Mapper (TM) and SPOT sensors taken in different times.

The organization of the paper is as follows: Section 2 gives an introduction to wavelet transform issues; Section 3 discusses the basic steps used in the registration algorithm; Section 4 shows the preliminary experimental results of registering images taken from Landsat and SPOT satellites and Section 5 summarizes the work.

2 Wavelet Transforms

2.1 Definition

Wavelets are functions that satisfy certain mathematical requirements and are used in representing data or other functions.

Let $\psi(x) \in L^2(\mathbb{R})$ be a complex valued function. If the function $\psi(x)$ satisfies:

$$\int_{-\infty}^{\infty} \psi(x)dx = 0 \qquad (1)$$

it is called a basic wavelet or "mother wavelet." The families of functions

$$\psi_{a,b}(x) = \mid a \mid^{-1} \psi(\frac{x-b}{a}) \; with \; a,b \in \mathbb{R}, a \neq 0 \quad (2)$$

generated from the "mother wavelet" under the operations of dilations (or scaling) by a and translations in time by b, are called wavelets.

The continuous wavelet Transform of a function $f(x) \in L^2(\mathbb{R})$ is given by the convolution

$$W_a[f(x)] \quad = \quad \mid a \mid^{-1} \int_{-\infty}^{\infty} f(u)\psi(\frac{x-u}{a})du$$

$$= \quad f * \psi_a(x). \qquad (3)$$

where $\psi_a(x) = (1/a)\psi(x/a)$.

The wavelet can be interpreted as the impulsive response of a band-pass filter and the wavelet transform of a function as a convolution of this function with the dilated filter. Thus the processing can be done at different scales or resolutions. If we look at the data with a large "window" (large a) we notice gross features. Similarly, if we look at the data with a small "window" (small a), we notice small features [9].

In practice the scale a has to be discretized. For a particular class of wavelets, the scale a can be sampled along a dyadic sequence $a = 2^j$ with $j \in Z$, without modifying the overall properties of the transform [15]. The transforms corresponding to dyadic values of a are called discrete wavelet transform (DWT),

$$W_{2^j}[f(x)] = f * \psi_{2^j}(x) . \qquad (4)$$

The wavelet transform of 1-D can be easily extended to 2-D case. Interested readers can find more information on wavelets and the wavelet transform in [4], [7], [9] and [12].

2.2 Wavelet Transform Modulus Maxima

Let us call a smoothing function $\phi(x, y)$, the impulse response of a 2-D low-pass filter. The first order derivative of $\phi(x, y)$ decomposed in two components along the x and y directions, respectively, are

$$\psi^1(x, y) \quad = \quad \frac{\partial \phi(x, y)}{\partial x}$$
$$\psi^2(x, y) \quad = \quad \frac{\partial \phi(x, y)}{\partial y} , \qquad (5)$$

and these functions can be used as wavelets.

For any function f , the wavelet transform at scale $a = 2^j$ defined with respect to these two wavelets has two components [14]:

$$W_{2^j}^1[f(x, y)] \quad = \quad f * \psi_{2^j}^1(x, y)$$
$$= \quad f * (2^j \frac{\partial}{\partial x} \phi_{2^j}(x, y))$$
$$= \quad 2^j \frac{\partial}{\partial x}(f * \phi_{2^j})(x, y)$$
$$W_{2^j}^2[f(x, y)] \quad = \quad f * \psi_{2^j}^2(x, y)$$
$$= \quad f * (2^j \frac{\partial}{\partial y} \phi_{2^j}(x, y))$$
$$= \quad 2^j \frac{\partial}{\partial y}(f * \phi_{2^j})(x, y) . \qquad (6)$$

Therefore, these two components of the wavelet transform are proportional to the coordinates of the gradient

vector of $f(x, y)$ smoothed by $\phi_{2^j}(x, y)$. They characterize the singularities along x and y directions, respectively [14].

We define the function

$$M[f(2^j, x, y)] = \Big(\big| W_{2^j}^1[f(x, y)] \big|^2$$
$$+ \big| W_{2^j}^2[f(x, y)] \big|^2 \Big)^{1/2}, \qquad (7)$$

which is the modulus of the wavelet transform at the scale 2^j. One can prove that for wavelets defined by Eq.(6), $M[f(2^j, x, y)]$ is proportional to the magnitude of the gradient vector.

Mallat and Hwang [14] use the same approach as Canny [2] to detect the edge points of $f(x, y)$ at scale a. They detect the points where the magnitude of the gradient is locally maximum in the direction where the gradient vector points to. At each scale 2^j, the modulus maxima of the wavelet transform are defined as points (x, y) where the modulus image $M[f(2^j, x, y)]$ is locally maximum, along the gradient direction.

3 Registration Algorithm

In this section, we discuss the registration algorithm. Let us call the image to be warped the *sensed* image and the image to which the other image will be reduced the *reference* image. As in [6] we consider two cases for the image registration algorithm: 1) the images have the same ground resolution (pixel size); 2) the images are taken from different sensors and have different ground resolutions. In the second case the image with the highest resolution is reduced to the lower resolution. This procedure is very simple since one knows a priori the spatial resolution of satellite images and the process of reducing the resolution can be realized automatically.

Next we describe each step involved in the registration process.

3.1 Feature Point Detection

After reducing the images to the same spatial resolution we compute the discrete multiresolution wavelet transform (L levels) of the two images. For this computation we use the algorithm proposed in [13], whose implementation is computationally efficient.

The wavelet decomposition of an image is similar to a quadrature mirror filter decomposition with the low-pass filter L and its mirror high-pass filter H [7]. We call LL, LH, HL and HH the four images created at each level of decomposition, as in [10], where this decomposition is also used for the purpose of image registration.

The next phase aims to identify features that are present in both images in each level of the decomposi-

tion. Here we use the modulus maxima of the wavelet transform to detect sharp variation points which correspond to edge points in the images. The LH and HL subbands at each level of the wavelet transform are used to estimate the image gradient. For the wavelet decomposition we use the filters given in [3].

A thresholding procedure is applied on the wavelet transform modulus image in order to eliminate non significant feature points. Then, a point (x, y) is recorded only if

$$M[f(2^j, x, y)] > \tau_{2^j} , \qquad (8)$$

where $\tau_{2^j} = \alpha(\sigma_{2^j} + \mu_{2^j})$, α is a constant whose initial value is defined by the user and σ_{2^j} and μ_{2^j} are the standard deviation and mean of the wavelet transform modulus image at level 2^j, respectively. The parameter α controls the number of feature points selected for the matching. Since the number of feature points increases in the finer resolutions the parameter α is also increased in the higher levels in order to select the most significant feature points in the images.

3.2 Initial Point Matching

The actual feature point matching is achieved by maximizing the correlation coefficient over small windows surrounding the points within the LL subbands of the wavelet transform. Let the LL subbands of the sensed and reference images be f_s and f_r, respectively. The correlation coefficient is given by:

$$C_{f_s f_r}(x, y, X, Y) = \frac{1}{\sigma_s \sigma_r w_c{}^2} \sum_{i=-w_c/2}^{w_c/2} \sum_{j=-w_c/2}^{w_c/2}$$
$$\{f_s(x + i, y + i) - \mu_s\}\{f_r(X + i, Y + i) - \mu_r\}, \quad (9)$$

where μ_s, σ_s, μ_r and σ_r are the local means (average intensity values) and standard deviation of the sensed and reference images, respectively, and $w_c{}^2$ is the area of matching window.

The initial matching is performed on the lowest resolution images and is determined by the best pairwise fitting between the feature points in the two images. Let $P_s = \{f_s(x_i, y_j), i = 1, \cdots, N_s\}$ and $P_r = \{f_r(X_j, Y_j), j = 1, \cdots, N_r\}$ be the set of feature points detected in the sensed and reference images, respectively. Let T_c denote the threshold value above which two feature points are considered similar. The point $f_r(X_k, Y_k)$ is the most similar feature point to $f_s(x_l, y_l)$ if

$$C_{f_s f_r}(x_l, y_l, X_k, Y_k) = \max_{1 \le j \le N_r} C_{f_s f_r}(x_l, y_l, X_j, Y_j).$$
$$(10)$$

Therefore, the matching process is achieved in the following way. For each point $f_s(x_l, y_l) \in P_s$ all points $f_r(X_j, Y_j) \in P_r$ are examined and its most similar point $f_r(X_k, Y_k)$ is chosen. Next we test whether the achieved correlation is reasonably high. If $C_{f_s f_r}(x_l, y_l, X_k, Y_k) > T_c$, then $f_r(X_k, Y_k)$ is called "the best match" of $f_s(x_l, y_l)$. To verify that the match is consistent in the reverse direction, we test whether the best match of $f_r(X_k, Y_k)$ exists and is $f_s(x_l, y_l)$. If that is the case, both points are matched.

This reversed verification reduces the number of mismatched pairs in the matching process and allows the use of smaller window sizes. Nevertheless, some false matches will inevitably occur. Therefore, a consistency-checking procedure similar to the one used in [11] is performed in order to eliminate incorrect matches and improve registration precision. The procedure is performed recursively in such a way that the most likely incorrect match is deleted first, followed by the next most likely incorrect match, and so on. This first part of the matching process is the crucial phase of the registration process. If the initial registration parameters are invalid the search for a registration transformation goes in a wrong direction, and the correct trend may not be recovered in later steps.

3.3 Image Warping

The above procedure provides a set of reliable matches which are used to determine a warping function that gives the best registration of the LL subbands to the precision available in level L of the wavelet transform.

To model the deformation between the images a 2-D affine transform with the parameters $(s, \theta, \Delta x, \Delta y)$ is used [16]:

$$
\begin{aligned}
X = T_1(x, y) &= s[x\cos(\theta) + y\sin(\theta)] + \Delta x \\
Y = T_2(x, y) &= s[-x\sin(\theta) + y\cos(\theta)] + \Delta y,
\end{aligned}
\tag{11}
$$

where (x,y) and (X,Y) are corresponding points in the sensed and reference image, respectively. This model is commonly used in remote sensing applications and is a good approximation for images taken under similar imaging directions [5], and which have been geometrically corrected (e.g., for Earth curvature and rotation). In the most of remote sensing applications the images have a certain level of geometrical correction which enables the use of this kind of transformation.

3.4 Refinement

The point matching and image warping steps can be performed at progressively higher resolutions in a similar fashion to that described above. At each level $l < L$, the image f_s is transformed using the parameters estimated from lower resolution level $(l + 1)$. In other words, the LL, LH, HL, and HH subbands at level $l + 1$ are used to reconstruct the LL subband at level l, which is then warped by the transformation specified at the previous point matching operation.

Let f_s^t denote the warped (sensed) image. The refinement matching is achieved using the warped image and the set of feature points detected in the reference image. Each feature point $f_r(X_k, Y_k)$ detected in the image f_r at level l is matched to $f_s^t(x_l, y_l)$ if

$$
\begin{aligned}
C_{f_s^t f_r}(x_l, y_l, X_k, Y_k) = &\max_{-w_r/2 \leq m, n \leq w_r/2} \\
&C_{f_s^t f_r}(x_l + m, y_l + n, X_k, Y_k),
\end{aligned}
\tag{12}
$$

where w_r is the size of the refinement matching window.

The traditional measure of registration accuracy is the root mean square error (RMSE) between the matched points after the transformation, defined as:

$$
\begin{aligned}
RMSE = \Bigg(\sum_{i=1}^N &\Big[\big(T_1(x_i, y_i) - X_i\big)^2 \\
&+ \big(T_2(x_i, y_i) - Y_i\big)^2 \Big]/N \Bigg)^{1/2},
\end{aligned}
\tag{13}
$$

where N is the number of matched points. This measure is used as a criterion to eliminate earlier matches which are considered imprecise. Poor matches are sequentially eliminated in a iterative fashion similar to that in [11] until the RMSE value is lower than 0.5 pixel.

At each level the warping parameters are updated considering the refined list of matched points. After processing all levels the final parameters are determined and used to warp the original sensed image.

4 Experimental Results

In order to test the algorithm and demonstrate its feasibility for different type of images preliminary results are presented in this section. We have used images from SPOT and Landsat-TM satellites in the experiments. All images tested are 512x512 pixels and were extracted from larger images to reduce processing time and disk usage.

The principal parameters to be specified by the users are the number of levels of wavelet decomposition (L) and the value of α in Eq.(8). The wavelet decomposition is carried out up to the third level. For all test images the process was realized with the parameters $\alpha = 2$, $w_c = 7$, $w_r = 3$, and $T_c = 0.7$. The sensed images are warped using bilinear interpolation.

Experimental results are depicted in Figures 1 through 4. Each figure displays a column of images with the sensed image on top (a), the reference image in the middle (b), and the registration result in the bottom (c). The images shown in the figures are enhanced for purpose of display.

Figure 1 shows the registration of two images taken from TM sensor, band 5 on different dates, 07/18/94 (TM945A) and 09/09/90 (TM905A). They correspond to an agricultural region near Itapeva, Sao Paulo. The image TM945A was taken as the reference one and TM905A as the sensed one. The images in Figure 2 are as in Figure 1 but the reference image in Figure 2(b) (TM945RA) is georeferenced.

Figure 3 shows the registration of two images from the urban area of Sao Paulo. A SPOT image, band 3, dated of 08/08/95 (SP953U), was reduced to a 30 m pixel size and taken as the sensed image. A Landsat-TM image, band 4, dated of 06/07/94 (TM944U), was taken as the reference image.

Finally, Figure 4 shows the registration of Amazon region images taken from TM sensor, band 5, in different dates, 06/07/92 (TM925F) and 07/15/94 (TM945F). The image TM945F was taken as the reference one and TM925F as the sensed one.

The parameters of the transformation and the number of control points for the test images shown in Figures 1-4 are listed in Table I, where the unit for Δx and Δy is in pixels and for θ is in degrees. The registration error (RMSE) is less than one pixel for the test images in Figures 1, 3 and 4. For the images in Figure 2 the registration error is about one pixel.

The processing time depends on the image. Broadly speaking the process described in this paper takes around 1 minute on a SUN SPARC 20 workstation.

5 Conclusion

We have described an automatic algorithm for the registration of satellite images. If the images have different spatial resolution the highest resolution image is reduced to the lowest resolution one and the processing is accomplished as for images of the same resolution. The method is simple and does not need to set up a large number of parameters. The algorithm is performed at progressively higher resolution, which allows for faster implementation and higher registering precision. It is more adequate to register images taken from the same sensor or with similar spectral bands. Nevertheless, it worked well for images taken at different times and from urban, forest and agricultural areas which are typical to remote sensing applications.

TABLE I

Images	θ	s	Δx	Δy	Number of CPs
TM905A TM945A	0.02	0.99	87.6	-77.7	150
TM905A TM945RA	10.35	1.02	9.3	-83.1	75
SP953U TM944U	7.48	0.99	-70.9	-56.2	308
TM925F TM945F	0.08	0.99	36.5	-182.6	150

References

[1] G. Câmara, R. C. M. Souza, U. M. Freitas, and J. Garrido. Spring: Integrating remote sensing and gis by object oriented data modeling. *Computers & Graphics*, 20(3):395–403, May-June 1996.

[2] J. F. Canny. A computational approach to edge detection. *IEEE Trans. on Pattern Anal. and Machine Intell.*, 8(6):679–698, Nov. 1986.

[3] C. Cheong, K. Aizawa, T. Saito, and M. Hatori. Subband image coding with biorthonormal wavelets. *EICE Trans. Fundamentals of Eletronics, Communications and Computer Sciences*, E75-A:871–881, July 1992.

[4] C. Chui. *An Introduction to wavelets*. C.Chui (editor), Academic Press, New York, 1992.

[5] K. Dana and P. Anandan. Registering of visible and infrared images. In *Proceedings of the SPIE Conference on Architecture, Hardware and Forward-Looking Infrared Issues in Automatic Target Recognition*, volume 1957, Orlando, FL, USA, 12-13, April 1993.

[6] J. P. Djamdji, A. Bijaoui, and R. Maniere. Geometrical registration of images: the multiresolution approach. *Photogrammetric Engineering and Remote Sensing*, 59(5):645–653, May 1993.

[7] N. J. Fliege. *Multirate Digital Signal Processing:multirate systems, filter banks, wavelets.* John Wiley & Sons, 1994.

[8] L. M. G. Fonseca and B. S. Manjunath. Registration techniques for multisensor remotely sensed imagery. *Photogrammetric Engineering and Remote Sensing*, 62(9):1049–1056, Sept. 1996.

[9] A. Graps. An introduction to wavelets. *IEEE Computational Science and Engineering*, 2(2):50–61, Summer 1995.

[10] J. Le Moigne. Parallel registration of multi-sensor remotely sensed imagery using wavelet coefficients. In *Proceedings of the SPIE - The International Society for Optical Engineering*, volume 2242, pages 432–43, Orlando, FL, USA, 5-8, April 1994.

[11] H. Li, B. S. Manjunath, and S. K. Mitra. A contour-based approach to multisensor image registration. *IEEE Transactions on Image Processing*, 4(3):320–334, March 1995.

[12] S. G. Mallat. Multifrequency channel decompositions of images and wavelet models. *IEEE Transactions on Acoustics, Speech, and Signal processing*, 37(12):2091–2110, Dec. 1989.

[13] S. G. Mallat. A theory for multi-resolution signal decomposition: the wavelet representation. *IEEE Trans. on Pattern Anal. and Machine Intell.*, 11(7):674–693, July 1989.

[14] S. G. Mallat and W. H. Hwang. Singularity detection and processing with wavelets. *IEEE Transactions on Information Theory*, 38(2):617–643, March 1992.

[15] S. G. Mallat and S. Zhong. Characterization of signals from multiscale edges. *IEEE Transactions on Pattern Anal. Mach. Intell.*, 14(7):710–732, July 1992.

[16] Q. Zheng and R. A. Chellappa. computational vision approach to image registration. *IEEE Transactions on Image Processing*, 2(3):311–326, July 1993.

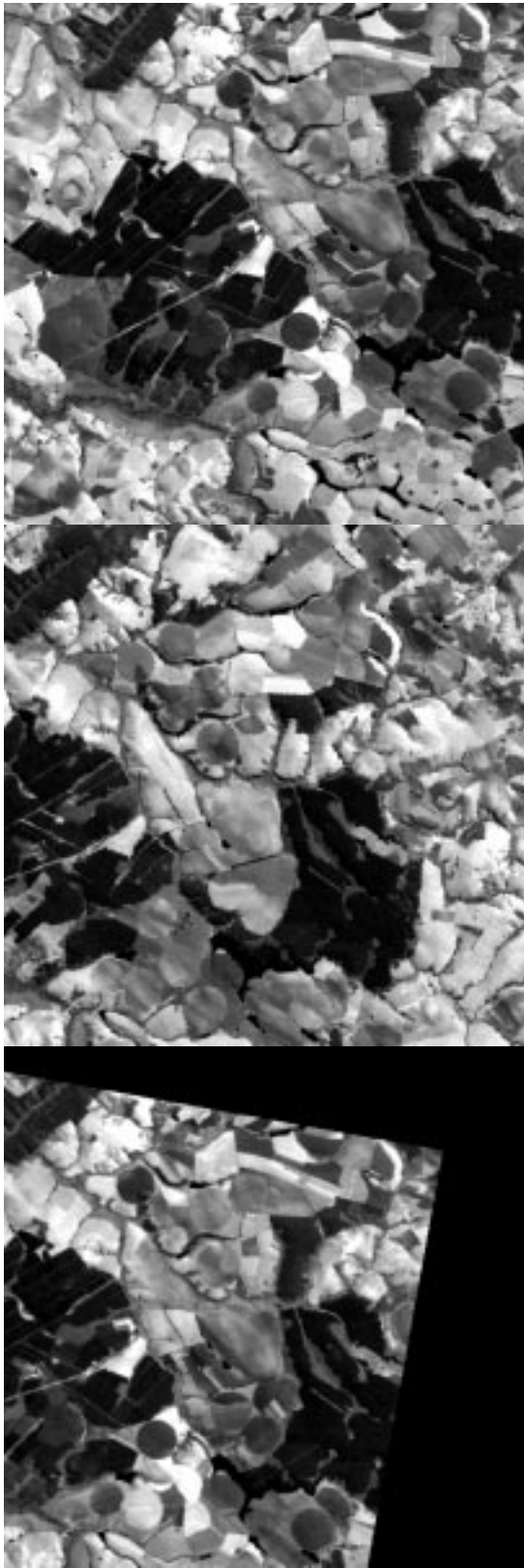Figure 1: Images from an agricultural area: (a)TM905A; (b)TM945A; (c)registering (a) with (b).

Figure 2: Images from an agricultural area: (a)TM905A; (b)TM945RA; (c)registering (a) with (b).

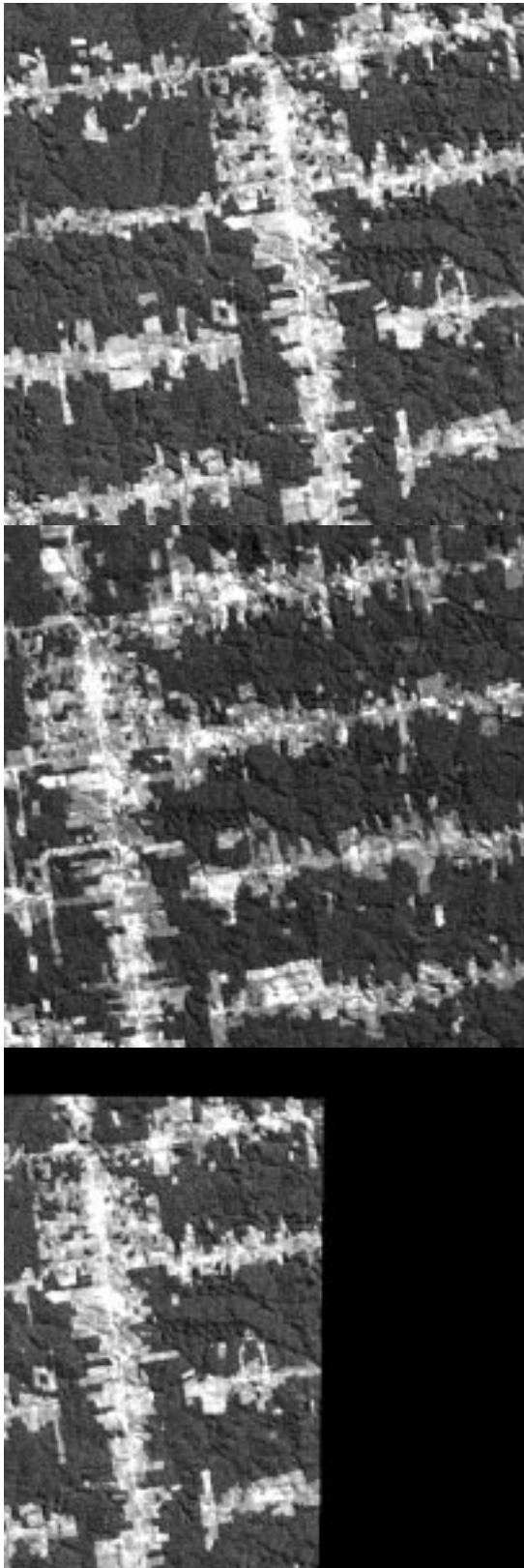Figure 3: Images from an urban area:(a)SP953U; (b)TM944U; (c)registering (a) with (b).

Figure 4: Amazon region images: (a)TM925F; (b)TM945F; (c)registering (a) with (b).