STEREO RADARGRAMMETRY IN SOUTH-EAST ASIA USING TERRASAR-X STRIPMAP DATA

Xueyan He, Timo Balz*, Lu Zhang, Mingsheng Liao

State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, China
Luoyu Road 129, 430079 Wuhan
balz@lmars.whu.edu.cn

Commission VII

KEY WORDS: Radar, SAR, Stereo, Radargrammetry, TerraSAR-X

ABSTRACT:
Stereo radargrammetry is a mature technique for deriving height information from SAR image pairs. Because stereo radargrammetry is less sensible to temporal decorrelation, it can provide better results than interferometric SAR in certain situations. Using TerraSAR-X stripmap stereo pairs, digital surface models (DSM) with good height accuracy can be generated. We use SRTM data as our initial DEM and a pyramid layer based approach for our radargrammetric processing. Our results are relative noisy with a high error standard deviation. Nevertheless, with a mean overall error of less than 3 m in height the final DSM is quite precise.

1. INTRODUCTION

There are various ways for creating digital elevation models (DEM) from synthetic aperture radar (SAR) data. The most common technique is SAR interferometry (InSAR). InSAR uses the phase difference between the backscattered signals received at slightly different positions. From the phase differences the height of the backscattering object can be determined. InSAR is a very precise technique, but especially spaceborne InSAR suffers from errors caused by the atmosphere. The repeat pass interferometry, which is commonly used with today's sensors, as we are waiting for the TanDEM-X data, is very sensitive to temporal decorrelation. In repeat pass InSAR the time between two data acquisitions is rather large, for example 11 days for TerraSAR-X. In this time the situation on the ground can change, especially over vegetated areas, and the data decorrelates.

This massively influences the usability of InSAR over strongly vegetated areas. Our test area around Kuala Kangsar in Malaysia is a strongly vegetated and mountainous area. The nearby mountains reach up to 1500 m, while Kuala Kangsar resides at around 40 m height. Under these conditions, repeat pass InSAR with TerraSAR-X does not provide good results. Most of the scene is strongly decorrelated.

But InSAR is not the only method of creating DEMs from SAR data. Stereo radargrammetry (StereoSAR) is far less affected by the atmosphere and by temporal decorrelation. StereoSAR is therefore, although in general less precise than InSAR, more suitable for this area.

StereoSAR with TerraSAR-X can deliver quite precise results (Raggam *et al.* 2010). In our approach we use the rational polynomial coefficient (RPC) model (Tao & Hu 2001) for the calculation of the 3D geo-coordinates of detected homologous points. In Section 2, the approach for StereoSAR with high-resolution TerraSAR-X data is explained. In Section 3, we describe the usage of the RPC model for stereo radargrammetry. Our experimental results are shown in Section 4 and finally conclusions are drawn.

2. STEREO RADARGRAMMETRY WITH TERRASAR-X

Comparable to InSAR, StereoSAR needs at least two images. But the images need to have a rather large acquisition angle difference. To use StereoSAR with spaceborne systems, the systems must be able to acquire data under different incidence angles. Before the launch of the new generation of high-resolution SAR satellites, Radarsat data was used for StereoSAR (e.g. Toutin 2000). The new satellite generation is also capable of acquiring data to be used for StereoSAR and in our experiments we used a pair of TerraSAR-X stripmap data. StereoSAR can be acquired in a same-side or opposite-side configuration. In the following we always assume a same-side configuration. In a same-side configuration the images are acquired from parallel tracks under different incident angles.

The geometry of a SAR image is different in range and in azimuth direction. In azimuth direction the location of an object in a SAR image depends on the Doppler Effect and is typically precise if the object is not moving. In range direction the location of an object in a SAR image depends on the distance between the object and the sensor. When an object is seen from different angles, this, so called, relief displacements will be different on each SAR image. The resulting difference between the image coordinates of homologous points in a stereo pair is called parallax (Leberl 1990).

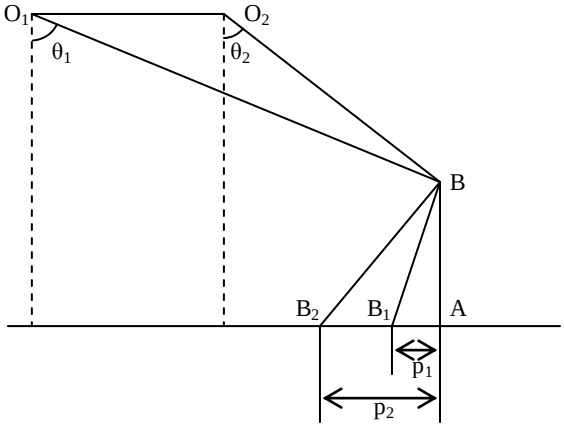


Figure 1. Same-side stereo configuration

* corresponding author

This is schematically shown in Figure 1. Object *B* is displaced differently depending on the incidence angles θ_1 and θ_2. By measuring the displacements p_1 and p_2 the relative height *h* of the object can be calculated with Eq. 1.

$$h = \frac{p_1 - p_2}{\cot\theta_2 \pm \cot\theta_1} \quad (1)$$

The difficult and time consuming part of the StereoSAR method is the search of homologous points in both images. In StereoSAR this is rather difficult because of the speckle noise and the large differences in geometry and radiometry between the two SAR images acquired under different incidence angles.

Our search of homologous points is based on pyramid layers of the images. The search starts at the highest pyramid level, subsequently refining the search using lower pyramid levels. There are many different search criterions suggested for SAR (Tupin and Nicolas 2002), but in our experiments we achieved the best results using the two-dimensional normalized correlation (see also Fayard *et al.* 2007)

$$r = \frac{\sum_i \sum_j (r_{ij} - \bar{r})(m_{ij} - \bar{m})}{(n-1)s_r s_m} \quad (2)$$

where *r* is the reference image and *m* is the match image, $\bar{r}$ and $\bar{m}$ are the mean values of the reference and match image values inside the correlation window, s_r and s_m are the standard deviation of the values inside the correlation window and *n* is the number of pixels in the correlation window. In higher pyramid levels a smaller correlation window is chosen, but in lower pyramid levels the correlation window gets bigger.

One problem of the pyramid level based approach is that small errors on high pyramid levels can propagate and cause large errors and outliers in the final result. In previous experiments we found outliers with height errors of more than 250 m in a rather flat terrain of our test area (Balz *et al.* 2009) caused by this.

The SRTM dataset can be used to improve the accuracy of the process by providing reliable initial values for the iterative search process. Furthermore, the SRTM data can also be used to filter large outliers.

3. USING THE RPC MODEL FOR STEREO RADARGRAMMETRY

A sensor model is established to relate the image coordinates and corresponding object coordinates. Rigorous physical sensor models accurately represent this relationship, but using them is very time-consuming when positioning each pixel by a rigorous sensor model. Moreover, the sensor parameters are needed in rigorous sensor model, which may violate the confidential rules of commercial companies. Although the rigorous sensor models are more accurate, the development of generic sensor models with high fitting accuracy, real-time processing ability and sensor-independent features is very useful.

The rational polynomial coefficients (RPC) model is a typical generic sensor model which describes the relationship between image space and object space by ratios of polynomial functions. It has been successfully applied in geo-coding of high-resolution optical imagery, such as IKONOS, QuickBird, GeoEye-1, etc. (Volpe, 2004; Fraser & Hanley, 2005), and has become a standard component for processing optical data.

Our investigation indicates that the RPC model for TerraSAR-X data has an excellent fitting accuracy (Wei *et al.* 2010). The RPC model uses two main forms for the forward (Eq. 3-4) and two main forms for the inverse transformation (Eq. 5-6).

$$r = \frac{p_1(X,Y,Z)}{p_2(X,Y,Z)} \quad (3)$$

$$c = \frac{p_3(X,Y,Z)}{p_4(X,Y,Z)} \quad (4)$$

$$X = \frac{p1(r,c,Z)}{p2(r,c,Z)} \quad (5)$$

$$Y = \frac{p3(r,c,Z)}{p4(r,c,Z)} \quad (6)$$

r and *c* are the normalized row and column indices in image space, *X*, *Y* and *Z* are the normalized 3D object coordinates. The purpose of normalization is to improve the numerically stability of the equations. The image coordinates and object coordinates are both normalized to values between -1 and 1 using

$$\begin{cases} r = \dfrac{r_{org} - r_{offset}}{r_{scale}} \quad c = \dfrac{c_{org} - c_{offset}}{c_{scale}} \\ X = \dfrac{X_{org} - X_{offset}}{X_{scale}} \quad Y = \dfrac{Y_{org} - Y_{offset}}{Y_{scale}} \quad Z = \dfrac{Z_{org} - Z_{offset}}{Z_{scale}} \end{cases} \quad (7)$$

where the subscript *org* means original coordinates, and *offset* and *scale* are parameters for normalization.

In RPC-based geo-coding, the RPC parameters need to be solved in advance. A virtual regular object grid covering the full extent of the image with several elevation layers is established. Each grid point's corresponding image row and column indices can be calculated using the rigorous physical sensor model and the satellite ephemeris data. After the image coordinates and object coordinates are obtained, the RPC parameters can be estimated using a least-squares solution.

The accuracy of RPC-based geo-location is affected by the computation accuracy of RPC parameters directly. This is the difficult point, because the equations are usually very ill-conditioned. There are two main categories for solving ill-posed equations: biased methods, like ridge trace, GCV, or L-curve, and unbiased methods, like the iteration method by correcting characteristic value (IMCCV). The ridge trace method is quite time-consuming with a low accuracy. The GCV method is sometimes not converging, which is a major drawback. The L-curve method is fast with a high accuracy. But the method is a little less accurate than the IMCCV, while the IMCCV takes a lot of time for the iterative improvement and relies heavily on the initial data. Therefore, we use the L-curve results as the initial data for the IMCCV. This approach can get very accurate results fast.

Given a over-determined linear system

$$AX = L \quad (8)$$

and using the least square solution, the results of *X* will be

$$X = (A^T PA)^{-1} A^T PL \quad (9).$$

Unfortunately, in this way we usually cannot obtain stable results. To deal with this problem, the ridge estimate is used

$$X(k) = (A^T PA + kE)^{-1} A^T PL \quad (10),$$

where P is the weight matrix and k is the regularization parameter. In order to determine for which k the X-value can get the best result, the L-curve method selects different k-values and calculates the corresponding X. So a group of points are obtained:

$$(\eta(k), \xi(k)) = (\lg \|AX - L\|_k, \lg \|X\|_k) \quad (11).$$

This curve normally is shaped like the letter 'L'. The optimal value of the regularization parameter k is considered to be the one corresponding to the corner of the 'L', i.e. the point with maximum curvature.

The principle of IMCCV method is different from that of the L-curve method. It replaces the ordinary least-square solution with

$$(A^T PA + E)\hat{X} = A^T PL + E \quad (12).$$

Since both sides of the equation have unknowns, we can only solve by using iterative methods.

$$\hat{X}^{(k)} = (B^T PB + E)^{-1}(B^T PL + \hat{X}^{(k-1)}) \quad (13)$$

In our test we use the result of L-curve as the initial value of IMCCV method. The result will be calculated by solving Eq. 13 iteratively until the termination condition is satisfied.

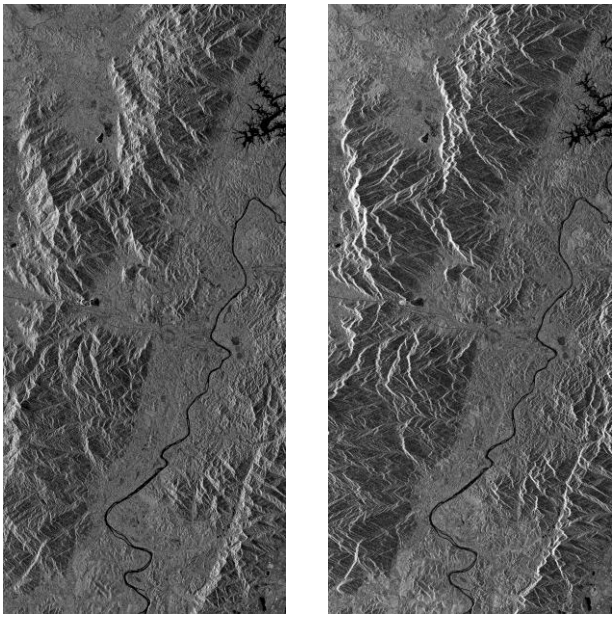
Once the RPC parameters are obtained, we can use them to geocode our data. The forward form of the RPC model is adopted. For each object points with known latitude and longitude and height fetched from our DEM, the corresponding row and column indices can be calculated. After resampling, the height of each point can be acquired.

4. EXPERIMENTAL RESULTS

Our test area is around Kuala Kangsar, Malaysia (N4°29′8′′-N5°04′38′′, E100°39′33′′-E101°3′25′′). The area is rather flat along the river, but is flanked by high mountains on both sides. The area is strongly vegetated and the elevation ranges from 40 m up to 1500 m. A TerraSAR-X stripmap image pair acquired on September 13 and September 18, 2008, with incidence angles of 21.4° and 42.7° at the scene centers, is used in our experiments. The amplitude images of the scenes are depicted in Figure 2.

The test data was processed using our own radargrammetric processor. The results are compared to a DSM created by Infoterra's Pixel Factory™ (Infoterra 2009), to a publically available DEM generated by Infoterra (2010), to GPS ground control points provided by Infoterra, as well as to the SRTM DEM.

Table 1 shows the parameters we used for our experiments. The search for homologous points starts at the pyramid level 5 with 40 meters pixel size, subsequently refining the search using lower pyramid levels until pyramid level 2 with 5 meters pixel size. In higher pyramid levels a smaller correlation window is chosen, but in lower pyramid levels the correlation window gets bigger.



Reference image
Sep. 18, 2009
© DLR/Infoterra

Match image
Sep. 13, 2009
© DLR/Infoterra

Figure 2. TerraSAR-X stereo pair of Kuala Kangsar, Malaysia

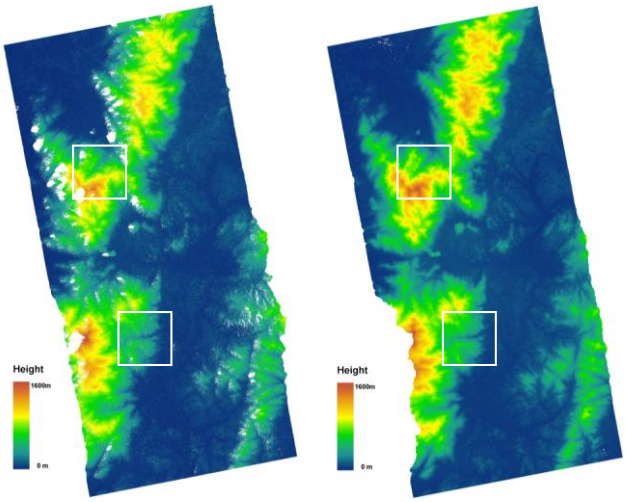
The search size in Y direction is set to be 3 because we found that the azimuth parallaxes are very small. As we can see in Table 1, the mean correlation value is smaller in lower pyramid levels. This is due to the increasing level of noise in the lower pyramid level images. With a mean correlation of only 0.3 in the 2nd pyramid level, we can assume the results to be noisy and less reliable.

Table 1. Correlation Parameters

Level	Pixel size	Corr. Window Size	Search Size in Range	Search Size in Azimuth	Mean Correlation
5	40m	5x5	11	3	0.68
4	20m	7x7	11	3	0.54
3	10m	9x9	17	3	0.41
2	5m	11x11	17	3	0.30

Figure 3 shows the DSM calculated by Infoterra's Pixel Factory™ and by our radargrammetric processor respectively. As we can see, some pixels with no height information are masked out, appearing white in the Infoterra's DSM. In our DSM, all pixels are reconstructed except a part in the lower left corner, where no height information can be obtained.

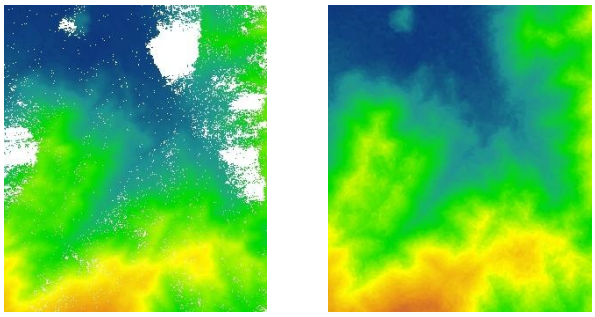
The DSM created using Infoterra's Pixel Factory™ has a mean absolute height error of about -1.7 m with a standard deviation of 8.4 compared to 26 GPS ground control points. Our DSM was generated fully automatically without using any ground control points. In this way we got a mean error of -44.2 m and standard deviation of 22.88 m. After using one control point, located in the middle of the reference image (100.947 E / 4.79557 N), to correct our results in elevation direction, we got a mean error of 2.79 m. Comparing our DSM with Infoterra's DSM, the mean error for the full image is 7.07 m and standard deviation is 19.97 m. Compared with the newest DEM available from Infoterra (2010) we have a mean error of -12.64 with a standard deviation of 18.84. We also compared our DSM with the SRTM DSM and got a mean error of 2.8 m with a standard deviation 13.6 m.



DSM calculated by Infoterra's Pixel Factory™ | DSM calculated using the procedure explained above

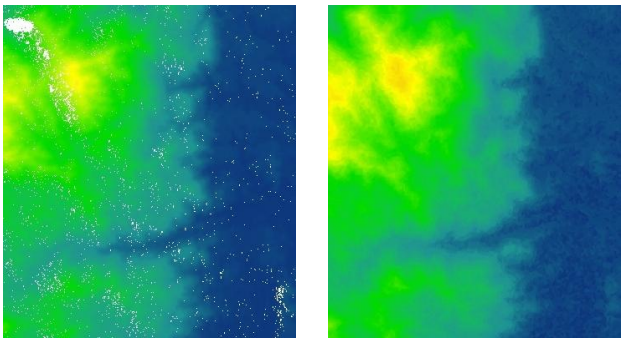
Figure 3. Comparison of the DSMs

The absolute mean error is not very high, but our DSM is quite noisy. The standard deviation of the error is quite high. The GPS control point with the largest absolute error of -68 m is located near the northwestern mountain peak. In Figure 4 and Figure 5 the details of the DSMs marked by white boxes in Figure 3 can be seen.



Subset of the Pixel Factory™ DSM on the mountain ridge | Subset of our DSM on the mountain ridge

Figure 4. Detailed DSM comparison



Subset of the Pixel Factory™ DSM on the mountain ridge | Subset of our DSM on the mountain ridge

Figure 5. Detailed DSM comparison

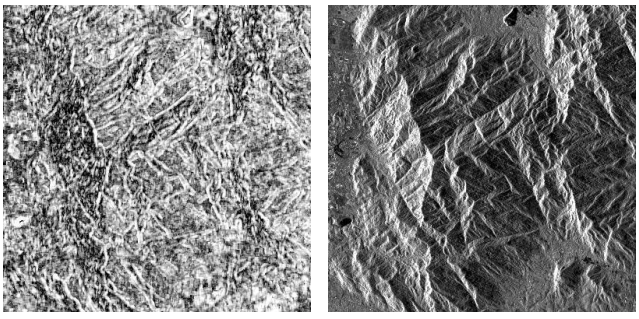
The DSM created by Infoterra's Pixel Factory™ includes large areas with no value in Figure 4, whereas our DSM interpolates all areas. Pixel Factory™ is marking the areas where no or no reliable information is available. This is an important feature of the Pixel Factory™ for mapping applications.

Our DSM is overestimating the heights at the mountain ridges compared to the results from Pixel Factory™. The resulting DSM is noisier, which can be seen very clearly in Figure 5. The mountains are quite comparable to each other, but in the rather flat area on the eastern side of the subset shown in Figure 5 our radargrammetric processor produces very noisy results. This is still the main problem of our approach.

Figure 6 shows the correlation value of the whole scene. The river is very striking with a high correlation value. The lower and flat area near the river also has a high overall correlation, but the results in this area are still noisy. Along the hillside of the mountainous areas the overall correlation is much lower.

Figure 6. Correlation image

In Figure 7(*a*) a subset of the correlation image from Figure 6 can be seen. Figure 7(*b*) shows the amplitude image of the same area. The correlation is high at the mountain ridges and the valleys, but lower at the mountain sides, especially in the layover area. The differences in these areas between the two stereo images are quite high. Interestingly we also find very low coherence in the rather flat areas, where we would assume the image pair to be higher correlated, because these areas are less affected by the strong differences in geometry and backscattering strength caused by the differences in local incidence angles.



(*a*) (*b*)

Figure 7. Correlation image (*a*) and amplitude image (*b*) subset

But in fact, these flat areas are less correlated, because they are covered by uniform vegetation and not many structural elements can be found.
Figure 8 shows the overlay of the correlation image and the amplitude image at the north-western mountain subset shown in Figure 7. The higher correlation values are painted in darker color, which is opposite to the color code used in Figure 6 and Figure 7, but this allows for a more meaningful visualization. The lakes at the left side and at the top are obviously highly correlated and are shown in a very dark green. The picture also indicates that the mountain sides facing the sensor (towards the left side of the image) have low correlation values due to the layover effect. The mountain sides facing in far-range direction (towards the right side of the image) have higher correlation values. We can see this also when analyzing the no-data-areas of the Pixel Factory™ DSMs in Figure 3 and Figure 4. The large no-data-areas are located in near-range of a mountain.

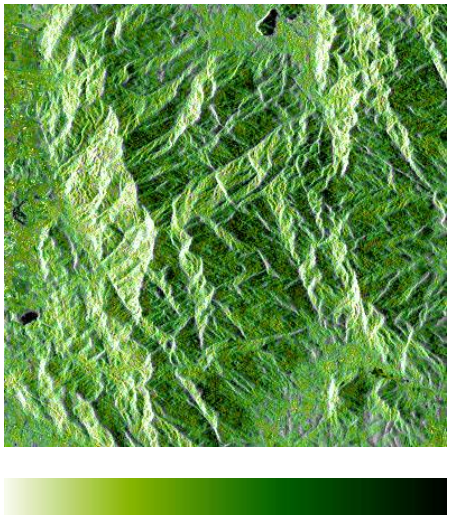


Figure 8. Overlay of correlation image and amplitude image

5. CONCLUSION

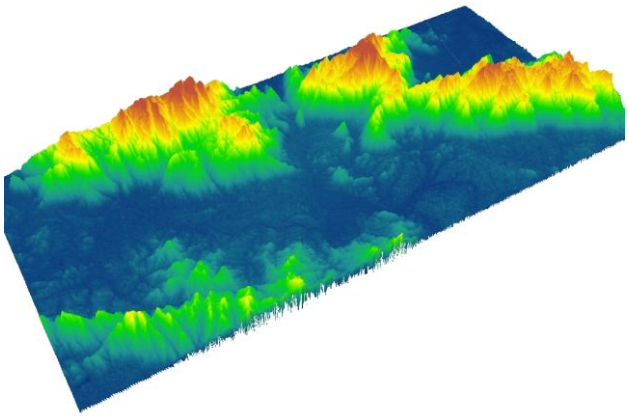


Figure 9. 3D model of the DSM

Stereo radargrammetry with TerraSAR-X can provide precise DSMs even in mountainous and strongly vegetated areas as shown above and in Figure 9. We achieved a low mean error in the DSM, but the standard deviation of the error is comparably high. Compared with our previous work (Balz *et al.* 2009), we improved the accuracy and strongly reduced the number of outliers in the data by using the SRTM data and by improving the processing. We tested different matching criteria but found that the two-dimensional normalized correlation delivers the best results for our dataset.
The RPC based geo-coding improved our processing speed and the overall geo-accuracy of our DSM. Still, the DSM created by Infoterra's Pixel Factory is more precise and less noisy. There is therefore still a lot of room for improvement.

ACKNOWLEDGEMENT

The authors thank the Infoterra GmbH for providing the test data and the reference data. The work was supported by the Research Fellowship for International Young Scientists of the National Natural Science Foundation of China (Grant No. 60950110351).

REFERENCES

Balz, T., He, X.Y., Zhang, L., Liao, M.S., 2009. TerraSAR-X stereo radargrammetry for precise DEM generation in South-East Asia. In: *Proc. Asian Conference on Remote Sensing*, Beijing.

Fayard, F., Méric, S., and Pottier, E., 2007. Matching stereoscopic SAR images for radargrammetric applications. In: *Proc. Geoscience and Remote Sensing Symposium (IGARSS 2007)*, Barcelona, Spain.

Fraser, C.S. and Hanley, H.B. 2005. Bias-compensad RPCs for Sensor Orientation of High-resolution Satellite Imagery. *Photogrammetric Engineering & Remote Sensing*, 71, pp. 909-915.

Infoterra, 2009. Pixel Factory™, the Next Generation Solution for Industrial Geo-Production. Available online: http://www.infoterra.fr/Infoterra/Pixel_Factory.php [last accessed on May 26, 2010]

Infoterra, 2010. Free TerraSAR-X data. Available online: http://www.infoterra.de/tsx/freedata/start.php [last accessed on June 3, 2010]

Leberl, F., 1990. Radargrammetric Image Processing. Artech House, Boston, MA.

Raggam, H., Gutjahr, K., Perko, R., and Schardt, M., 2010. Assessment of the stereo-radargrammetric mapping potential of TerraSAR-X multibeam spotlight data, *IEEE Transactions on Geoscience and Remote Sensing*, 48, pp. 971-977.

Tao, C.V. and Hu, Y., 2001. A comprehensive study of the rational function model for photogrammetric processing. *Photogrammetric Engineering & Remote Sensing*, 67, pp. 1347-1357.

Toutin, T., 2000. Evaluation of radargrammetric DEM from RADARSAT images in high relief areas. *IEEE Transactions on Geoscience and Remote Sensing*, 38, pp. 782–789.

Tupin, F. and Nicolas, J.M., 2002. Matching criteria for radargrammetry. In: *Proc. IGARSS 2002*, Toronto.

Volpe, F., 2004. RPC based processing of QuickBird high resolution satellite data. In: *Proc. Polish Remote sensing and GIS Conference 2004*, Bialobrzegi, Poland.

Wei, X.H., He, X.Y., Zhang, L., Balz, T., Liao, M.S. 2010. RPC modelling for spaceborne SAR and its application in radar image geocoding. In: *Proc. 2010 Dragon-2 Mid Term Results Symposium*, Guilin, China.