# Multi-Image Matching for DTM Generation from SPOT-5 HRS/HRG and IRS-P5 Imagery - For the Project of West China Topographic Mapping at 1:50,000 Scale

ZHANG Li[1], ZHANG JiXian[1], Wang ShaoCheng[2]
{ zhangl, zhangjx }@casm.ac.cn
[1]Chinese Academy of Surveying and Mapping, Haidian, Beijing, 100039
[2]Beijing East Dawn Information Technology Inc., Haidian, Beijing, 100080

Theme Session ThS-3

**KEYWORDS:** The Project of Western China Topographic Mapping (WChTM) at 1:50,000 Scale, High-Resolution Satellite Imagery (HRSI), Multiply Primitive Multi-Image (MPM) Matching, Geometrical Constrained Cross-Correlation ($GC^3$), Digital Terrain Model (DTM)/Digital Surface Model (DSM)

**ABSTRACT:**

High-resolution satellite images (HRSI) at sub-5m footprint such as IKONOS, IRS-P5 (CartoSat-1) and SPOT-5 HRG/HRS images are becoming increasingly available to the earth observation community and their respective clients. HRSI is one of the main data sources for the Project of Western China Topographic Mapping (WChTM) which has been approved by the State Council of China in year 2006. New advanced technologies in field of photogrammetry and remote sensing, such as the block-adjustment with sparse GCPs, automatic DSM/DTM generation and 3D mapping with HRSI are very necessary to the project. In this paper, we present a matching approach for automatic DSM/DTM generation from HRSI. It can provide dense, precise and reliable results. The method matches multiple images simultaneously and it uses a coarse-to-fine hierarchical solution with an effective combination of several image matching algorithms and automatic quality control. The DSMs/DTMs are generated by combination of matching results of feature points, grid points and edges. The new characteristics provided by the high-resolution imaging systems, i.e. the multiple-view terrain coverage and the high quality image data, are also efficiently utilized in this approach.

The proposed approach has been applied to SPOT-5 HRS/HRG images over a testfield in Zone of headstream of Three rivers, eastern Tibet Plateau, China with variable terrain geomorphologic type. The accuracy study was based on the comparison between as many as 160 accurate GPS check points, more than 1400 manually measured check points and the automatically extracted DTMs. The RMS errors for the whole area are 2-7 m, while for flat/hilly areas the accuracy is about 2-3 m or even better. The proposed approach has also been applied to 23 IRS-P5 stereo pairs over Beijing city, the resulted 12.5 m DTM reproduced quite well not only the general features of the terrain relief but also small geomorphological and other features visible in the IRS-P5 images. From these experiments, it's shown that with the proposed automatic DTM generation approach, by using SPOT-5 HRS/HRG and IRS-P5 imagery, satisfactory 1:50000 DTMs can be completed with a better accuracy than those requirements from Chinese Surveying and Mapping regulations.

## 1. Introduction

In recent years, space-borne CCD linear array sensors have been widely used to acquire panchromatic and multi-spectral imagery in push-broom mode for photogrammetric and remote sensing applications. HRSI at sub-5m footprint such as IRS-P5, SPOT-5, IKONOS, and QuickBird/WorldView-I provide not only for high-resolution and multi-spectral data, but also for the capability of stereo mapping.

SPOT-5 HRS/HRG, IRS-P5 and other HRSI images are main data sources for the project of Western China Topographic Mapping (WChTP) at 1:50,000 scale. At the current time, in western part of China up to 2.02 million $km^2$ are not mapped at 1:50,000 scale, it includes Sorthern-XinJiang desert area, Qing-Tibet Plateau area and Heng-duan mountain ranges. This unmapped area covers about 20% of all areas of China; contains total number of about 5,032 sheets of 1:50,000 scale topographic maps. This situation greatly hiders the socio-economic development of this region, also posed potential threat to national security. The project has been approved by the State Council of China in year 2006, and is dedicated to complete 1:50,000 scale topographic map and construct the national geo-spatial database for the region within next 5 years through year 2006 to 2010. The unmapped area related to the project covers most unmans area of China, the very harsh natural conditions (average terrain elevation is more than 4000m) and difficult transportation conditions provide a very difficult situation for field works. In many areas, it's even impossible for accessing or collecting enough GCPs, therefore, the project of WChTP must across traditional surveying mode, utilize innovative aerial and satellite remote sensing techniques, such as DGPS/IMU assisted aerial photogrammetry, high resolution satellite imagery mapping technique and radar image mapping technique to achieve surveying for unmapping area in Western China.

The project of WChTP is the importance and complicated project which first uses HRSI cosmically for topographic map generation in China. Automatic DSM/DTM generation is one of the key techniques for 3D mapping from HRSI images. Although automatic DTM/DSM generation through image matching has gained much attention in the past years, a wide variety of approaches have been developed, and automatic DTM generation packages are in the meanwhile commercially available on several digital photogrammetric workstations, the accuracy performance and the problems encountered are very similar in the major systems and the performance of commercial image matchers does by far not live up to the standards set by manual measurements (Gruen et al., 2000). The main problems in DTM/DSM generation are encountered with:

(1) Little or no texture
(2) Distinct object discontinuities
(3) Local object patch is no planar face
(4) Repetitive objects
(5) Occlusions
(6) Moving objects, incl. shadows
(7) Radiometric artifacts like specular reflections and others

The key to successful and reliable matching is the matching of a dense pattern of features with an appropriate matching strategy, making use of all available and explicit knowledge, concerning sensor model, network structure and image content such as the epipolar geometry constraints and a piecewise smooth surface model. For an appropriate matching strategy, we have to consider combining the *area-based matching* (*ABM*) and the *feature-based matching* (*FBM*), matching parameter self-tuning, generation of more redundant matches and a coarse-to-fine hierarchical matching strategy. In particular, we also have to consider the fact that HRSI provides for some characteristics and possibilities for automatic image matching:

(1) Compared to the traditional scanned 8-bit images, images from these sensors have better radiometric performance. Most of the linear array sensors have the ability to provide more than 8-bit/pixel images. This results in a major improvement for image matching in terms of reducing the number of mismatches for "homogeneous" areas and especially for dark shadow areas.

(2) Ability to provide multiple-view terrain coverage in one flight mission or satellite orbit. This enables the multi-image matching approach, which leads to a reduction of problems caused by occlusions, multiple solutions, surface discontinuities and results in higher measurement accuracy through the intersection of more than two image rays. Also, along-track stereo images, from the same orbit within a very short time interval, have a distinct advantage to those across-track because they reduce the radiometric differences, and thus increase the correlation success rate.

In this paper, we present an advanced image matching approach and we will report about the key algorithms in details (Chapter 2). We give a DTM accuracy evaluation using SPOT-5 HRS/HRG triplets in a testfield in Zone of headstream of Three rivers, eastern Tibet Plateau, China. In another test, the proposed approach has been also applied to 23 IRS-P5 stereo pairs over Beijing city, the resulted 12.5 m DTM reproduced quite well not only the general features of the terrain relief but also small geomorphologic and other features visible in the IRS-P5 images. Through these experiments, we demonstrate that our approach leads to good results (Chpater 3).

## 2. The Automatic DSM/DTM Generation Approach

We have developed an advanced matching approach for automatic DSM/DTM generation from HRSI. It can provide dense, precise and reliable results. The approach uses a coarse-to-fine hierarchical solution with a combination of several image matching algorithms and automatic quality control. The new characteristics provided by HRSI imaging systems, i.e. the multiple-view terrain coverage and the high quality image data, are also efficiently utilized in this approach.

The approach essentially consists of 3 mutually connected components: the image pre-processing, the multiple primitive multi-image (MPM) matching and the geomorphologic refinement matching procedure. The overall data flow is shown schematically in Fig. 1. The images and the given or previously estimated orientation elements are used as input. After pre-processing of the original images and production of the image pyramids, the matches of three feature types (feature points, grid points and edges) in the original resolution images are found progressively starting from the low-density features in the lowest resolution level of the image pyramid. A TIN form DSM is reconstructed from the matched features at each pyramid level by using the constrained Delauney triangulation method. This TIN in turn is used in the subsequent pyramid level for derivation of approximations and adaptive computation of some matching parameters. Finally and optionally, least squares matching methods are used to achieve more precise results for all matched features and for the identification of some false matches.

In order to capture and model the detailed terrain features, our DSM/DTM generation approach not only generates a large number of mass points but also produces line features. Here we just give a detailed description about the core part of our approach, i.e. the Multiple Primitive Multi-Image Matching (MPM) matching procedure, for more details of this matching approach please refer to Zhang and Gruen, 2004,2006; Zhang, 2005; Baltsavias, et. al., 2006.

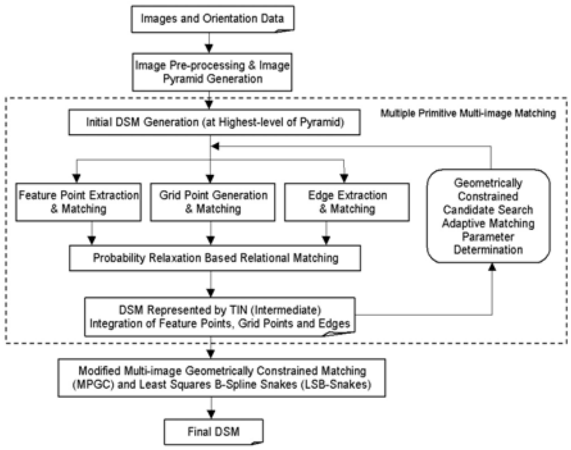



Fig. 1: Workflow of the proposed automated DTM/DSM generation approach.

### 2.1 The Multiple Primitive Multi-image (MPM) Matching Procedure

The Multiple Primitive Multi-Image (MPM) matching procedure is the core of our developed approach for accurate and robust DSM/DTM reconstruction. Results from this approach can be used as approximations for the refined matching procedure with least squares matching methods. In the MPM approach, the matching is performed with the aid of multiple images (two or more), incorporating multiple matching primitives – feature points, grid points and edges, integrating local and global image information and utilizing a coarse-to-fine hierarchical matching strategy. The MPM approach consists mainly of 3 integrated subsystems: the point extraction and matching procedure, the edge extraction and matching procedure and the relaxation based relational matching procedure.

In the MPM matching procedure, we do not aim at pure image-to-image matching. Instead we directly seek for image-to-object correspondences. We have developed a new flexible and robust matching algorithm – Geometrically Constrained Cross-Correlation ($GC^3$) method in order to take advantage of the multiple images. The algorithm is an extension of the standard Cross-Correlation technique and is based on the concept of multi-image matching guided from object space and allows reconstruction of 3D objects by matching all available images simultaneously, without having to match all individual stereo-pairs and merge the results.

### 2.2 Geometrically Constrained Cross-Correlation ($GC^3$) Algorithm

We developed a new flexible and robust matching algorithm –$GC^3$ method in order to take advantage of the multiple images. The algorithm is an extension of the standard *Cross-Correlation* technique and is based on the concept of multi-image matching guided from object space and allows reconstruction of 3D objects by matching all available images simultaneously, without having to match all individual stereo-pairs and merge the results.

Consider an IKONOS image triplet, as shown in Fig. 2. The middle image is chosen as the reference image and denoted as $I_0$, the other two images are search images and denoted as $I_i$, $i=1,2$.

In this case, we have two stereo pairs, that is, the pair $I_0$-$I_1$ and $I_0$-$I_2$. For a given point $p_0$ in the reference image, we can establish the image ray $Cp_0$ (here $C$ denotes the instant perspective center related to point $p_0$), on which the correspondence of $p_0$ in object space should lie, with the known image orientation parameters. By intersecting the image ray $Cp_0$ with a horizontal plane defined by a given approximate height $Z_0$, we obtain $P_0$ ($X_0$, $Y_0$, $Z_0$) in object space. The approximate height $Z_0$ may have an increment $\Delta Z$, such that the correct position of $P_0$ in object space should lie between $P_{min}$ and $P_{max}$, with height values of $Z_0$-$\Delta Z$ and $Z_0$+$\Delta Z$ respectively, along the image ray $Cp_0$. If we back-project the points between $P_{min}$ and $P_{max}$ onto the search images, the corresponding segments of the quasi-epipolar lines for the point $p_0$ can be easily defined. The correct matches $p_i$, $i=1,2$ in the search images $I_i$, $i=1,2$ must lie along its corresponding quasi-epipolar line segments.

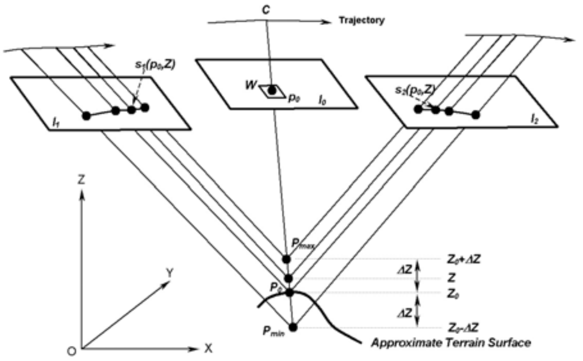



Fig. 2: Multiple image matching with the $GC^3$ algorithm. For details please refer to the text.

Let $I_0(p)$ and $I_i(p)$ be the image intensity values of the reference and the $i$th search image respectively. In the reference image, we define a correlation window $W$ around the point $p_0$. We assume that an approximate DSM is known either as a horizontal plane or from matching results at a higher level of the image pyramid. If we project this window onto the approximate DSM through the so-called mono-restitution procedure, we can obtain a piece of surface patch in object space. Then, we back-project this surface patch onto the search images, thus generating the corresponding image window in the search images. We named this process "*correlation window warping procedure*". Through this reshaping procedure, a square correlation window in the reference image can be correlated with windows of different size, shape and orientation in the search images. Therefore, multiple images with different image scale and orientation can be matched straightforward. The distortions caused by terrain relief and imaging geometry can be compensated (more details please refer to Zhang, 2005.

Now the *Normalized Correlation Coefficient* (*NCC*) value between the corresponding correlation windows in the reference image and the ***i***th search image can be defined, with respect to the height $Z$ for $p_0$, as:

$$NCC_i(p_0,Z)=\frac{\sum_{s\in W}(I_0(s)-\bar{I}_0)\times(I_i(s_i(Z))-\bar{I}_i)}{\sqrt{\sum_{s\in W}(I_0(s)-\bar{I}_0)^2}\sqrt{\sum_{s\in W}(I_i(s_i(Z))-\bar{I}_i)^2}} \qquad (1)$$

$$\text{Here,}\quad \bar{I}_0=\frac{1}{m\times n}\sum_{s\in W}I_0(s);\quad \bar{I}_i=\frac{1}{m\times n}\sum_{s\in W}I_i(s_i(Z))\qquad i=1,2$$

Where, $W$ and $s$ denote the correlation window in the reference image and a pixel in this window respectively; $m$ and $n$ denote the dimension of the correlation window $W$. $s_i(Z)$ denotes the corresponding point to $s$ in the $i$th search image. As mentioned before, $s_i(Z)$ can be computed through the correlation window warping procedure. The intensity values for point $s_i(Z)$ are interpolated from the $i$th search image by using the bilinear interpolation method.

As can be seen from equation (1), the value of $NCC_i$ is defined with respect to the height value Z, which could be any value between $Z_0$-$\Delta Z$ and $Z_0$+$\Delta Z$. Thus, given a point in the reference image, as well as its approximated height $Z_0$ and an increment $\Delta Z$ in object space, the *NCC* functions for all individual stereo pairs are defined within a unique framework. We then follow the procedure proposed by Okutomi and Kanade (1993), instead of computing the correct match of point $p_0$ by evaluating the individual *NCC* functions between the reference $I_0$ and search image $I_i$, $i=1,2$, we define the sum of *NCC* (*SNCC*) for point $p_0$, with respect to $Z$, as:

$$SNCC(p_0,Z)=\frac{1}{2}\sum_{i=1}^{2}NCC_i(p_0,Z) \qquad (2)$$

Therefore, by finding the value $Z$: $Z\in[Z_0-\Delta Z,\ Z_0+\Delta Z]$ which maximize the *SNCC* function, we can obtain the corresponding height value for point $p_0$. Here, the height increment $\Delta Z$ determines the search distance along the corresponding quasi-epipolar lines. Through the definition of the *SNCC* function, which simply accumulates the *NCC* functions of cross-correlation from all the stereo pairs, the correct match or correct height in object space for a given point in the reference image can be obtained. In general, the matching candidates show maxima in the *SNCC* function and each peak of the function *SNCC* corresponds to an object point with a certain height value. In the $GC^3$ algorithm, these object points are defined as the matching candidates for the given point. The method can be easily extended to a more general case, which is suitable for $n+1$ ($n \geq 1$) images.

$$SNCC(p_0,Z)=\frac{1}{n}\sum_{i=1}^{n}NCC_i(p_0,Z) \qquad (3)$$

In Fig. 3, an example of high-resolution airborne linear array image strips (ca. 5cm footprint) is shown in order to highlight the ability of the $GC^3$ algorithm to solve the multiple solution problem. Fig. 3 (b) shows that it is very difficult to determine the correct match by just evaluating each individual *NCC* value. However, the *SNCC* shows a sharp and clear maximum at the correct match, even within a large search distance.

### 2.3 Summary of the DSM/DTM Generation Approach

Our automatic DSM/DTM generation approach is characterized by the following items:

(1) **Multiple image matching**: We have developed a new flexible and robust matching algorithm –$GC^3$ method in order to take advantage of the multiple images. The algorithm is based on the concept of multi-image matching guided from object space and allows reconstruction of 3D objects by matching all available images simultaneously, without having to match all individual stereo-pairs separately and merge the results.

(2) **Matching with multiple primitives**: We have developed more robust hybrid image matching algorithms by taking advantage of both area-based matching and feature-based matching techniques and utilizing both local and global image information. In particular, we combine an edge matching method with a grid point matching method through a probability relaxation based relational matching process. The use of edges leads to the preservation of surface discontinuities, while grid points bridge areas with little or no texture.

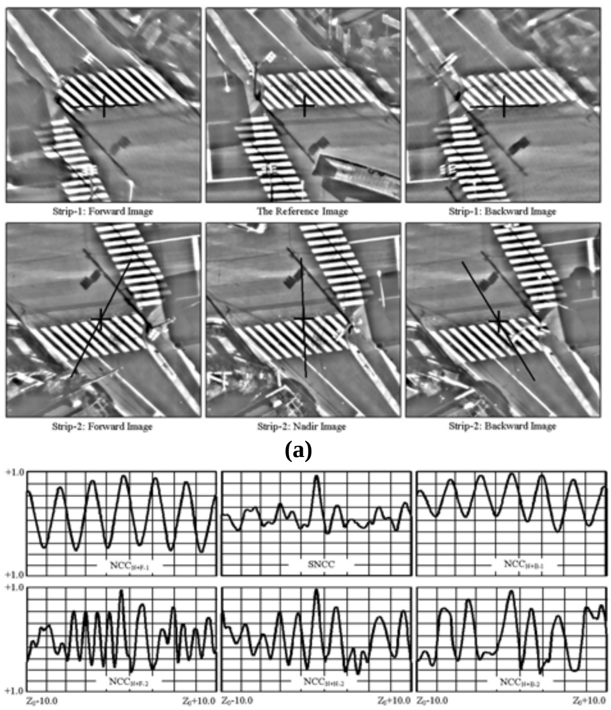
(3) **Self-tuning matching parameters**: The adaptive determination of the matching parameters results in higher

success rate and less mismatches. These parameters include the size of the correlation window, the search distance and the correlation threshold values. This is done by analyzing the matching results at the previous image pyramid level and using them at the current level.

(4) **High matching redundancy**: With our matching approach, highly redundant matching results, including points and edges can be generated. Highly redundant matching results are suitable for representing very steep and rough terrain and allow the terrain microstructures and surface discontinuities to be well preserved. Moreover, this high redundancy also allows for automatic blunder detection.

(5) **Efficient surface modeling**: The object surface is modeled by a triangular irregular network (TIN) generated by a constrained Delauney triangulation of the matched points and edges. A TIN is suitable for surface modeling because it integrates all the original matching results, including points and line features, without any interpolation. It is adapted to describe complex terrain types that contain many surface microstructures and discontinuities.

(6) **Coarse-to-fine hierarchical strategy**: The algorithm works in a coarse-to-fine multi-resolution image pyramid structure, and obtains intermediate DSMs at multiple resolutions. Matches on low-resolution images serve as approximations to restrict the search space and to adaptively compute the matching parameters.




**(b)**

Fig. 3: $GC^3$ matching with 6 high-resolution airborne linear array images (ca. 5cm footprint) from 2 strips with changing flight directions for solving multiple solution problems. The individual *NCC* functions and the *SNCC* function within the search range determined by height increment of ±10.0 meters is shown in (b)

## 3. Performance Evaluation

The height accuracy (or to be more precise the vertical accuracy) of DSMs/DTMs usually results from the quantitative and statistical evaluation of the DSMs/DTMs and it is determined by its root-mean-square error (RMSE), the square root of the average of the set of squared differences between height values of the DSM/DTM being evaluated and height values from an independent source with much higher accuracy. According to (McGlone, 2004), there are at least 3 major sources of error when DSMs/DTMs are generated by using the optical imaging systems, i.e. the Photogrammetric Modeling Error (**PME**), the Measurement Error (**ME**) and the Surface Modeling Error (**SME**).

These errors can be estimated empirically or estimated using error propagation. For instance, we could manually measure a sample of randomly spaced points using the stereo model with the same image orientation parameters and then compare them with their interpolated heights from the generated DTM. In this case, the height errors mainly come from ME, but also from SME in cases of very rough terrain. However, if we measure these points with a different method such as the traditional field surveying, the estimated errors may include all the errors mentioned above.

In order to evaluate the performance of our approach for DSM/DTM generation it has been verified extensively with several HRSI datasets, such as IRS-P5 and SPOT-5 HRS/HRG images, over different terrain types, which include hilly and rugged mountainous areas, rural, suburban and urban areas. In the following, we will report in detail about 2 experiments. The first involves the evaluation of SPOT-5 HRS/HRG triplet images over a testfield in Zone of headstream of Three rivers, eastern Tibet Plateau, China with accurate GCPs, more than 2500 m height range and variable land cover. The accuracy study was based on the comparison between as many as 160 accurate GPS check points, more than 1400 manually measured check points and the automatically extracted DTMs. In the second test, the proposed approach has been also applied to 23 IRS-P5 stereo pairs over Beijing city. Other processing and evaluation results of IKONOS and SPOT5 HRS/HRG can be found in Zhang and Gruen, 2004; Poli et al., 2004; Baltsavias et al., 2006 and Poon et al., 2005.

### 3.1 Automatic DTM generation from SPOT-5 HRS/HRG Images over Test-field in Zone of headstream of Three rivers, Eastern Tibet Plateau, China

The test area in Zone of headstream of Three rivers, Eastern Tibet Plateau, China covers 250 topographic maps at 1:50,000 scale with the area of about 12,000km$^2$, where contains large-area of seasonally and perennially frozen soil, mountain/valley glacier and perennial snowfield and large area of unman area. The test-field is the headstream of Yangtze River, Yellow River and Lancangjiang River, and the QingZang railway and national road cross the region from north-east to south-west. The study area consists of steep arid/semi-arid mountainous region in the northern part (transition zone between Kunlun Mountain and Tsaidam Basin), smooth hilly regions in the middle parts (plateau mountains and intermountain basins are well-developed) and high-plateau mountain ranges in the southern part (mountain/valley glacier, glacier canyon and knife-edge crest are well developed). The average elevation is 4000m in test-field. Main geological structures are in trend of nearly east-west direction (Fig. 4). The various landforms in study area provides better environment for DTM automatic generation.

Over the test area, totally 11 pairs of 10 × 5m SPOT-5 HRS and nearly twenty 5m HRG images were acquired. These images were used to generate DTM over the whole test area (Fig. 4). In particular, for DTM accuracy analysis, 2 SPOT-5 HRS satellite image pairs imaged in Nov 2003 and 6 HRG images, which can form the SPOT-5 stereo triplets, have been selected. The images have the fine quality and have no cloud coverage, which provide good data sources for DTM automatic generation and accuracy analysis. With the cooperation of several Institute of Surveying and Mapping in Shanxi, Heilongjiang and Sichuan, about 157 well-distributed GCPs were collected with differential GPS in 2006. The measurement accuracy was better than 1m in planimetry and 1.2m in height. With aids of these GCPs, block-adjustment was made and the orientation accuracy is about 5m in planimetry and 2m in height.

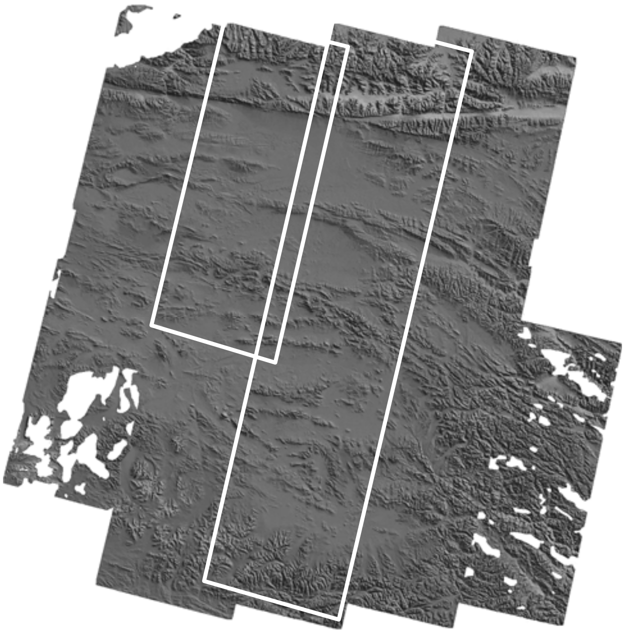


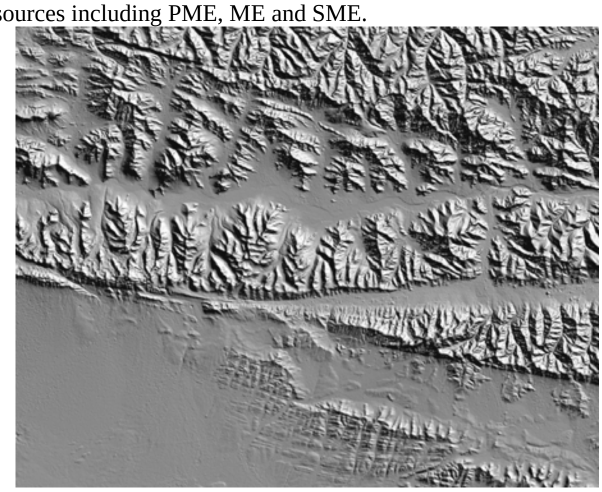
Fig. 4: Shaded DTM of the whole test-field in zone of headstream of Three Rivers, the DTM is automatically generated from 11 SPOT-5 HRS and nearly 20 HRG images. White areas (holes in DTM) are cloud-covered areas. White boxes show that coverage of 2 SPOT-5 HRS images, which used for DTM accuracy study.

After the block-adjustment, the proposed DTM generation approach was applied to 2 SPOT HRS stereo image strips and 6 HRG images simultaneously. As a result, 25m grid-spacing DTM of the accuracy study area has been generated automatically after about 26 hour. As results, about 90 millions feature points and 5 millions feature lines have been matched which is equivalent to match a $4 \times 4$ density grid on the original image, and it is good enough for generation of DTM with 25m interval. Since our DTM generation approach not only generates a large number of mass points but also produces line features, which are necessary for the modeling the rugged mountainous terrain. Parts of DTM results are shown in Fig 5. It can be seen that the resulted DTM reproduced quite well not only the general features of the terrain relief but also small geomorphologic and other features visible in the SPOT-5 images.

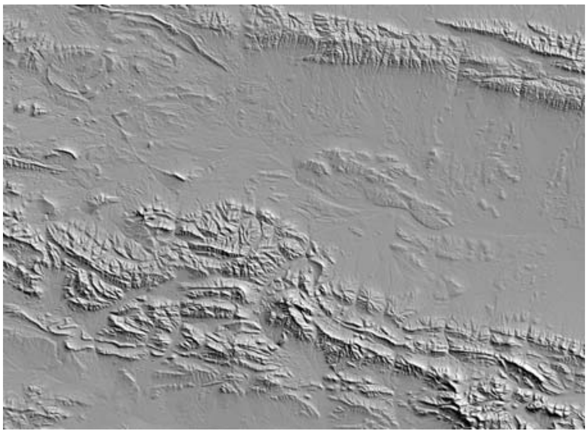
Since the study area is within the WChTM project covered area, there are not good enough reference data. Therefore we apply the following three accuracy evaluation methods for DTM accuracy evaluation:

(1) Overlay the automated generated DTM onto stereo image pairs to apply the manual visual checking under stereoscopic display device. Checking results show a good enough match between DTM and stereo images expect some small blunders within shadow areas, which have to apply necessary manual editing.

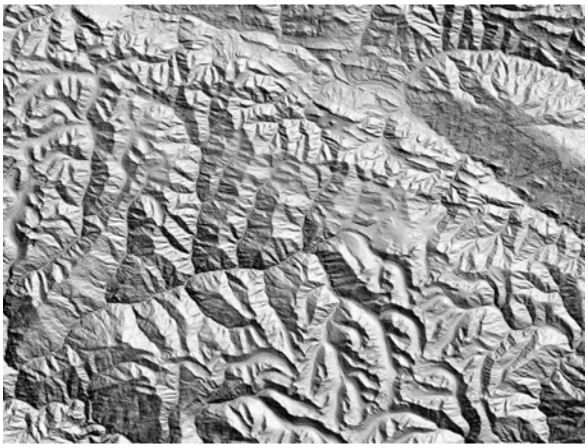
(2) Accuracy checking by using GCPs and check points measured by differential GPS. Upside of Table 1 give the DTM accuracy evaluation results by using this method. We computed the differences as GCPs minus the interpolated heights from our generated DTM. The accuracy of the DTM is between 1.7 – 6.7 m depending on the terrain relief and land cover. With this method, the errors probably include influence of all three error sources including PME, ME and SME.



(a): Shaded DTM of arid/semi-arid mountainous terrain.



(b): Shaded DTM of hilly region



(c): Shaded DTM of high-plateau mountain ranges

Fig. 5: The shaded terrain models of 25m over 3 sub-areas in Zone of headstream of Three rivers, Tibet Plateau, China. The resulting DTM reproduced quite well not only the general features of the terrain relief but also small geomorphologic and other features visible in the images. Results show that it very necessary to use the multiply image matching and edge matching with SPOT-5 HRS/HRG images over rough and steep mountainous areas.

(3) Accuracy checking by using manually measured checking points which were acquired on digital photogrammetric

workstation. Operators of the Photogrammetry and Remote Sensing Department of Shanxi Bureau of Surveying and Mapping have manually collected 1489 well-distributed checking points in study area. We computed the differences as these points minus the interpolated heights from our generated DTM. The accuracy of the DTM is between 3.6 – 4.5m depending on the terrain relief and landform. Since the DTM generation and point measurement procedure share the same image orientation results, the errors of the DTM would be from ME derived from stereo image matching, which is directly relate to the matching accuracy. Detailed accuracies are shown in Table 1 downside.

According to the experiment results, the following conclusions can be made: (a) DEM at 1:50,000 scale can be generated automatically from SPOT-5 HRS/HRG stereo/triplet-stereo images by using the automatic DTM generation approach proposed in this paper. The elevation accuracy can roughly be 50% better than requirements of the Chinese national surveying regulation; (b) simultaneously matching SPOT-5 HRS/HRG triplet stereo strips and matching both feature points and feature-lines are quite necessary for deeply incensed mountainous area and rugged terrain, such as arid and semi-arid broken mountains in northern part and rugged mountain ranges in southern part of the study area; (c) the proposed automatic DTM generation approach can largely increase the working efficiency, however, careful manually stereoscopic checking and editing are still necessary to remove some small matching blunders, which are in most cases within shadow and gully/rill/steep-valley well-developed areas.

Table 1: DTM accuracy evaluation results with SPOT-5 HRS/HRG imagery

| Reference data | Map sheets | Num. of check points | Height accuracy (RMSE) | Accuracy requirements of 1:50,000 scale 1st/2nd level DTM | Type of terrain |
|---|---|---|---|---|---|
| GPS points | 3 | 99 | 6.7 m | 14.0 m/19.0 m | Mountainous area |
| | 1 | 26 | 3.6 m | 8.0 m/11.0 m | Mountainous area |
| | 2 | 64 | 1.7 m | 5.0 m/ 7.0 m | Hilly area |
| | 3 | 127 | 2.8 m | 3.0 m/ 4.0 m | Flat/hilly area |
| Manually measured points | 47 | 937 | 3.6 m | 5.0 m/ 7.0 m | Hilly area |
| | 21 | 367 | 4.5 m | 8.0 m/11.0 m | Mountainous area |
| | 8 | 185 | 4.4 m | 14.0 m/19.0 m | Mountainous area |

### 3.2 DSM Generation from 23 IRS-P5 Stereo Images over Beijing Test-field

The test-field is an area around the city of Beijing, China. It consists of a steep mountainous region in the north-western part and flat regions in the middle and southern parts. The city of Beijing is located in the lower part of the study area. The whole area is about 70 × 210 km$^2$. The site has an average terrain height of 300m and an elevation range of more than 1100 m.

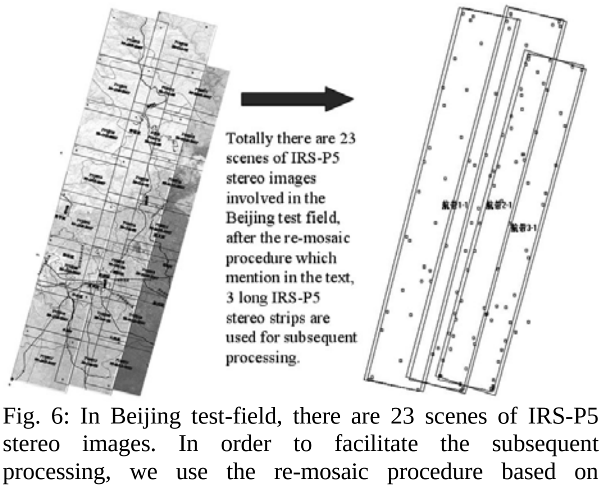



Fig. 6: In Beijing test-field, there are 23 scenes of IRS-P5 stereo images. In order to facilitate the subsequent processing, we use the re-mosaic procedure based on metadata to combine these 23 stereo images into 3 long stereo strips.

Over the test area, totally 23 scenes of IRS-P5 stereo images were acquired. The ground resolution of IRS-P5 images is 2.5m, and each scene covers about 30 × 30km$^2$. In order to precisely georeference these images, with the cooperation of the 1$^{rd}$ Institute of Surveying and Mapping of Heilongjiang SBSM, about 66 well-distributed GCPs were collected with differential GPS in 2006. The measurement accuracy was better than 0.5m in planimetry and 1m in height. The GCPs are well-distributed in the test area, their ground intervals are about 30km, and half of them are located at the center of road intersection which can be precisely measurement both in image and object space.

Normally, we could commercially get so-called IRS-P5 "standard scene" of images, which have 12000 × 12000 pixels and cover 30 × 30km$^2$. Basically the standard scenes could be sub-images of a long IRS-P5 strip. They are just subdivided from a long IRS-P5 strip for commercial reason and normally they have 5%-15% overlap between adjacent scenes. This situation will cause difficulties or complicate the sub-sequent processing. To avoid these problems, we develop a procedure to re-mosaic adjacent IRS-P5 standard scenes into a long stereo strip (Fig. 6) in condition that these scenes are sub-images of the original long IRS-P5 strip. After this re-mosaic procedure, 23 IRS-P5 scenes of the Beijing test-field are re-mosaic into only 3 IRS-P5 long stereo strips.

With aids of 66 GCPs, block-adjustment was made and the orientation accuracy is about 1.7m in planimetry and 2m in height. After the block-adjustment, the proposed DSM/DTM generation approach was applied to 3 re-mosaic IRS-P5 stereo image strips simultaneously. As a result, 12.5m grid-spacing DSM of the study area has been generated automatically. Since our DSM/DTM generation approach not only generates a large number of mass points but also produces line features, which are necessary for the modeling the rugged mountainous terrain. Parts of DSM results are shown in Fig. 7, the resulted 12.5 m DSM reproduced quite well not only the general features of the terrain relief but also small geomorphological and other features visible in the IRS-P5 images. The DSM shows many topographic details and features such as small valleys in the mountains, detailed patterns related to streets and buildings in suburban and urban areas, linear features related to highways and main road networks, sparse trees, small clusters of houses and forest areas. Checking by GCPs measured by differential GPS shows that the accuracy of the DSM is 3.7 m and it meet the requirements from the Chinese Surveying and Mapping regulations for 1:50000 topographic maps.

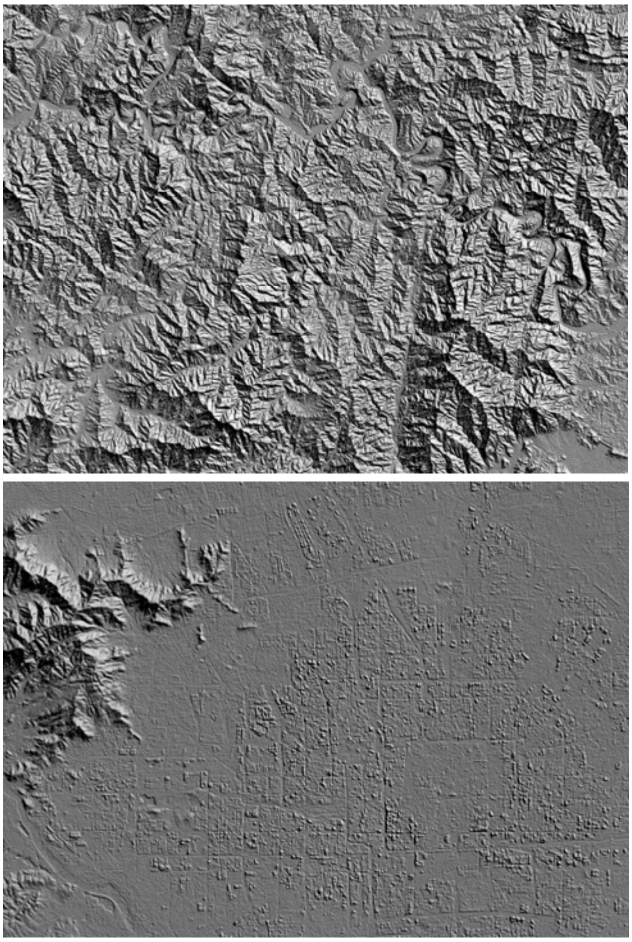


Fig. 7: Shaded terrain models (12.5m grid spacing) of 2 sub-areas in Beijing test-field, where the upper one shows the mountain area and the lower one shows the area around the city of Beijing.

**4. Conclusion**

In this paper we have reported about an advanced matching approach for automatic DSM/DTM generation from high-resolution satellite images. It can provide dense, precise and reliable results. The method uses a coarse-to-fine hierarchical solution with an effective combination of several image matching algorithms and automatic quality control. We have developed a matching strategy combining local point matching of feature and grid points, robust edge matching and a relaxation based global relational matching. The DSMs/DTMs are generated by a combination of matching results of feature points, grid points and edges.

The proposed approach has been applied to SPOT-5 HRS/HRG images over a testfield in Zone of headstream of Three rivers, Tibet Plateau, China with variable terrain geomorphologic type. The accuracy tests were based on the comparison between as many as 160 accurate GPS check points, more than 1400 manually measured check points and the automatically extracted DEMs. The RMS errors for the whole area are 2-7 m, while for flat/hilly areas the accuracy is about 2-3 m or even better. The proposed approach has been also applied to 23 IRS-P5 stereo pairs over Beijing area, the resulting 12.5 m DEM reproduced quite well not only the general features of the terrain relief but also small geomorphological and other features visible in the IRS-P5 images. The DEM shows many topographic details and features like small valleys in the mountains, detailed patterns related to streets and buildings in suburban and urban areas, linear features related to highways and main road networks, sparse trees, small clusters of houses and forest areas. From these experiments, it's shown that with the proposed automatic DEM generation approach, by using SPOT-5 HRS/HRG and IRS-P5 imagery, satisfactory 1:50000 DEMs can be completed with a better accuracy than those requirements from Chinese Surveying and Mapping regulations.

A major problem of our automatic DSM generation approach is the automated detection of small blunders, which still may exist in the results and need carefully manually editing. This constitutes a relevant topic for further research and refinement of our approach. We also have to collect good enough reference DSM/DTM for extensive accuracy test and analysis of our approach.

**References**

Baltsavias, E., Zhang, L., Eisenbeiss, H., 2006. DSM Generation and Interior Orientation Determination of IKONOS Images Using a Testfield in Switzerland. Journal of Photogrammetrie, Fernerkundung, Geoinformation, (1), pp. 41-54.

Gruen, A., Bär, S., Bührer, Th., 2000. DTMs Derived Automatically From DIPS - Where Do We Stand? Geoinformatics 3(5), 36-39

McGlone, J. C. (Editor). 2004. Manual of Photogrammetry (fifth edition). American Society for Photogrammetry and Remote Sensing, 5410 Grosvenor Lane, Suite 210, Bethesda, Maryland 20814, ISBN 1-57083-071-1. 1151 pp

Okutomi, M., Kanade, T., 1993. A Multiple-baseline Stereo. IEEE Transactions on Pattern Analysis and Machine Intelligence 15(4), 353-363.

Poli, D., Zhang, L., Gruen, A., 2004. SPOT-5/HRS Stereo Image Orientation and Automatic DSM Generation. IAPRS 35 (Part B1), 421-432.

Poon, J., C. Fraser, C. Zhang, L. Zhang, A. Gruen, 2005. Quality Assessment of Digital Surface Models Generated from IKONOS Imagery, Photogrammetric Record 20(110), 162-171.

Zhang, L., Gruen, A., 2004. Automatic DSM Generation from Linear Array Imagery Data. IAPRS 35 (Part B3), 128-133.

Zhang, L., Gruen, A., 2006. Multi-image matching for DSM generation from IKONOS imagery. ISPRS Journal of Photogrammetry and Remote Sensing, 60(2006), pp 195-211

Zhang, L., 2005. Automatic Digital Surface Model (DSM) Generation from Linear Array Images. Ph. D. Dissertation, No. 88, Institute of Geodesy and Photogrammetry, ETH Zurich, Switzerland. URL: http://e-collection.ethbib.ethz.ch/ecol-pool/diss/fulltext/eth16078.pdf