OBJECT INFORMATION AUTOMATIC EXTRACTION FROM HIGH RESOLUTION STEREO PAIRS BY DENSE IMAGE MATCHING AND INFORMATION FUSION

Fang Yong, Hu haiyan, Chen Hong, Zhang Wu

The Xi'an Institute of Surveying and Mapping , No.1 YanTa Middle Road, Xi'an, ShaanXi, China，710054 - yong.fang@vip.sina.com.cn

WgS-PS: WG IV/9

KEY WORDS: Digital Surface Model, High Resolution, Stereo Image Pairs, Feature Matching, 3D Reconstruction, Image Classification, Object Information Extraction

ABSTRACT:

With the advent of digital sensor, the requirement for new Photogrammetric software is urgent to quick object information extraction from high-resolution stereo pairs. This paper describes an effective combined approach for digital surface reconstruction and thematic information extraction automatically. Digital surface reconstruction from stereo pairs is realized by multi-level feature matching, which use constraints based on the conditions of the point feature, edge feature, area feature, grey feature and epipolar line etc from low to high level. Then the highly reliable and accurate corresponding point pairs is obtained, and edge matching controlled with TIN constructed from initial points extends the number of requests to meet. Finally the stereo model is reconstruction accurately based on the aerial triangulation by iterative reconstruction technique. The dense image matching is used for Digital Surface Model (DSM) extraction, in which the adaptive match window is used to compensate for the blurring effect that occurs at object boundaries in high-resolution images. These adaptive techniques separate fore- from background information in a correlation window using image classification result for multispectral images. DSM refines the initial classification result, and the special layers for buildings and vegetation are generated.

1. INTRODUCTION

The rapid development of digital aerial-photo camera provides the effective means for the obtainment of remote sensing data, which not only improves the data attainment efficiency, but also enhances the quality of the attained data. At the same time, stereo digital images with large overlap degree can be obtained, which makes it possible to provide excellent data for auto extracting space information without any additional costs. The information extraction from high resolution stereo images is one of nowadays hot study fields in application of earth observation.
As high resolution images are concerned, the image characters which has obviously distinction to that of middle and low resolution ones. As the resolution improved, the rate that areas with poor texture or areas with inconspicuous characters increases gradually, which will lead the similar measure of corresponding window area to fail if the window match strategy is adopted and largen the matching results difference to the real instance. The traditional theory and method, which mainly concern point and linear characters suitable to deal mid and low resolution remote sensing images, are coming in for challenge with the improvement of resolution. The analyze method based on area or surface characters fusing will hold more important station.

The paper brings forward one technology to extract thematic objects based on high resolution stereo images depends on analyze of the techniques that can be obtained currently. The main characters of the technology is the integration of such techniques as dense image matching, remote sensing images' auto classifying and space analyzing and it bases on the ground surface reconstruct automatically to realize 3D information reconstruction and thematic information extraction such as buildings, vegetation and water system etc. The algorithm of dense image matching synthetically takes into account point, linear, spectrum characters and the restrict conditions between the corresponding character points, which realizes extracting the digital surface model automatic efficiently and provides basic 3D space information for further thematic information extraction.

2. BASIC THEORIES AND METHODS

A confederative method is presented to realize efficiently extracting digital surface model and ground thematic information automatically according to the characteristic of the current available high resolution stereo remote sensing images. The basic theories are described as following. Firstly, stereo reconstruction is realized through multi-level character matching, which using points, linear and area characters and epipolar condition restrict to get precise and reliable corresponding point pairs. At the same time, the corresponding point pairs are expanded to satisfy the requirement of high precise stereo directional by means of the edge matching controlled by irregular triangle net composed of initialize points. The stereo model then is reconstructed precisely through iterative stereo. Then, dense image matching is used to extract digital surface model. In this step self adaptive matching window is used to resolve the punch-drunk in edges during high resolution image matching. The selection of self adaptive window makes use of the results of high spectrum classify. The image in the window is divided to background and topic information so as to remove the affection of the discontinuous ground surface. The final digital surface model is refined by spectrum classify and thematic level as buildings and vegetation are generated. The basic theories reference to Figure1.

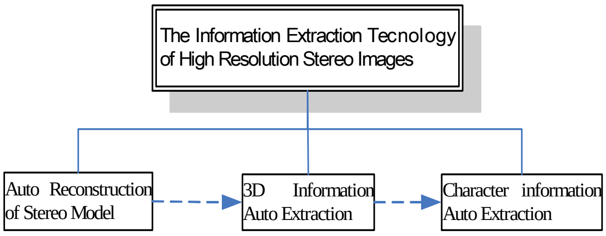


Figure 1, Basic Principles

3. STEREO MODEL AUTOMATIC RECONSTRUCTION FOR HIGH RESOLUTION IMAGES

To realize the auto reconstruction of stereo model, the main problem is to resolve the auto extraction of the corresponding image points on stereo image pairs, which is the key step to realize stereo auto reconstruction and is one of research content of computer vision. A new technique to realize stereo character auto extraction and high precision orientation is provided according to the character of the high resolution remote sensing images.

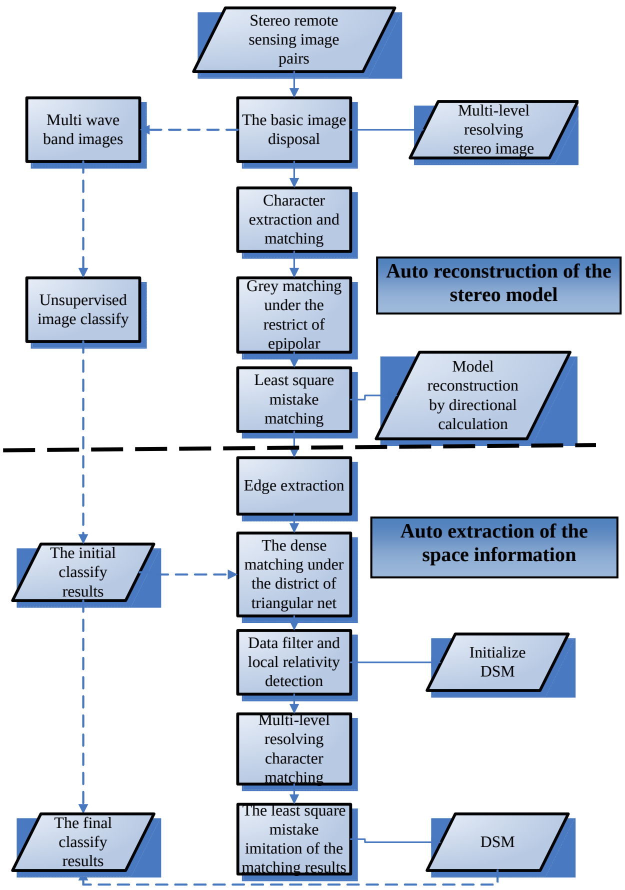


Figure 2, Technique flow of stereo model automatic reconstruction for high resolution images and the space information extraction

3.1 Character Points Extraction

Character extraction is the basis of the image analysis and image matching and it is also the most important task of single image processing. The character in the digital images can be divided into point characters, linear characters and area characters. Point characters refer to those distinct points such as acute point, circle points etc. The arithmetic operators used to extract point character is called interest operators or favorable operators. The goal of image processing is to extract as more as possible and precise orientation point characters from images. At present, many interest operators have been brought out and they can be divided into three classes according to their property. The first class is based on shape, that is to say the character point is positioned at location with the biggest curvature of the outline or at the point of intersection of two lines. The second class is based on signal. The digital image uses discrete grid to record continues information so that the image processing is one of application area of signal processing and the method used in signal processing can also be used into image information processing. The third class is based on template, but this kind of operator is designed for idiographic point character. Also the accuracy of template operator can achieve sub-pixel but it is inapplicable and its main application area is close photogrammetry.

Harris operator is a signal based operator for point extraction provided by C.Harris and M.J.Stephens in 1988 [1]. Harris has such characters: calculate simply, the acute point extracted distributed symmetrically, the characters can be extracted quantificationally and the operator has characteristic of stability. The operator is illuminated by the self correlation function used in information processing and given a matrix M which relates to self correlation function. The eigenvalue of the M is the first factorial curvature of the self correlation function. The point is considered to be a character point if the both curvature is high. The process is described as following:

$$M = G(\tilde{s}) \otimes \begin{bmatrix} g_x^2 & g_x g_y \\ g_x g_y & g_y^2 \end{bmatrix} \tag{1}$$

$$I = \det(M) - k \cdot tr^2(M), k = 0.04 \tag{2}$$

Where:

g_x is the grads of the x direction; g_y is the grads of the y direction; $G(\tilde{s})$ is Gauss template; det is the DET of the matrix; tr is the trace of the matrix; k is the default constant. The value of each element in the matrix I corresponds to the interest value of the initial image. The detection results using Harris operator refer to Figure 3. The corresponding acute points on the high resolution stereo images can be detected effectively as shown in the Figure.

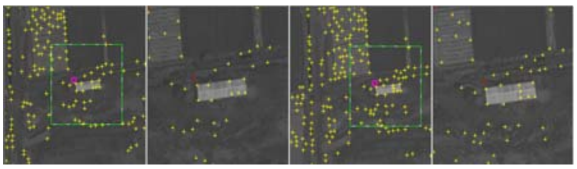


Figure 3, Stereo image and Harris operator extract results

3.2 Character Points Matching

Making use of the grey information round the acute point and the correlation method to build a local matching rule so as to divide the results detected by Harris operator into many-to-many matching pairs. The correlation coefficient is adopted to describe the measurement of similar degree of point p on image A and point q on image B. The correlation coefficient is defined as [2]:

$$Correl(p,q)=\frac{1}{\delta_1\delta_2}\sum_{j=-n}^{n}\sum_{i=-n}^{n}[A(u+i,v+j)-\mu_1]\cdot[B(u+i,v+j)-\mu_2] \qquad (3)$$

μ_1 （μ_2） and δ_1 （δ_2） express the mean and the square difference of point p (q) on image A（B）separately. The n is the neighbouring range of the acute point.

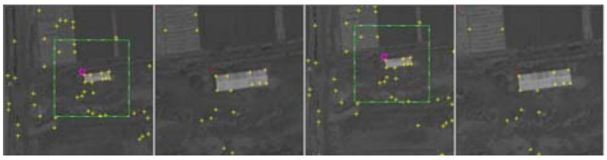


Figure 4, The matching results of the characters on stereo image pairs

An initial set of corresponding points is gained through character points matching, but if only grey character is relied as the correlation measurement there will exist some error matching points(refer to Figure 4, the matching results are not exactly corresponding), so that eliminating coarse error should be evolved.

3.3 Epipolar Detection

The corresponding points on the stereo image pairs should satisfy the coplanar condition according to the geometry relation of the imaging [3].

$$F=\begin{vmatrix} B_x & B_y & B_z \\ \bar{x}_1 & \bar{y}_1 & \bar{z}_1 \\ \bar{x}_2 & \bar{y}_2 & \bar{z}_2 \end{vmatrix} \qquad (4)$$

defining：

$$\begin{aligned}\frac{F_0}{Z_1X_2-X_1Z_2} &= \frac{\begin{vmatrix} B_x & B_y & B_z \\ X_1 & Y_1 & Z_1 \\ X_2 & Y_2 & Z_2 \end{vmatrix}}{Z_1X_2-X_1Z_2} \\ &= \frac{\begin{vmatrix} B_x & B_z \\ X_2 & Z_2 \end{vmatrix}}{Z_1X_2-X_1Z_2}Y_1-\frac{\begin{vmatrix} B_x & B_z \\ X_1 & Z_1 \end{vmatrix}}{Z_1X_2-X_1Z_2}Y_2-\frac{\begin{vmatrix} X_1 & Z_1 \\ X_2 & Z_2 \end{vmatrix}}{Z_1X_2-X_1Z_2}B_y \\ &= -N_1Y_1+N_2Y_2+B_y \\ &= -Q \end{aligned} \qquad (5)$$

Q is the model fluctuation parallax of the orientation points. When a stereo image pair have been done relative orientation, the value of the Q is 0, otherwise, Q≠0. That is to say the corresponding points must fall on the corresponding epipolar lines. This rule is used to eliminate the wrong matching point pairs in the initial corresponding points set.

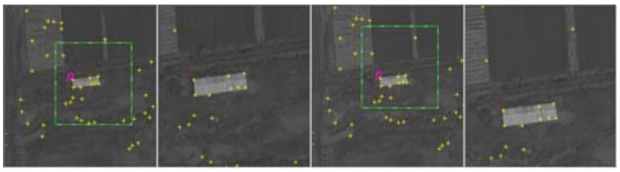


Figure 5 Detect Results under Epipolar Line Condition

3.4 Using Least Squares Image Matching to refine the matching results

There are two shortages exist in the obtained corresponding points: one is that the coordination of the character points detected by Harris operator is pixel and its accuracy is not high; the other is that during character matching, the geometry and radiate difference between matching windows have not been taken into account, so the matching precision is not high. In order to realize orientating the corresponding points precisely, the least squares image matching is used to eliminate the coarse error and orientate accurately.

The basic principle of the least squares image matching can be described as following [2]: the grey and geometry aberration parameter are imported when doing image matching, and these parameters are calculated through least squares image matching principle to improve the precision of image matching.
The simple aberrance is considered in practice because the size of the matching windows is often small.

$$\begin{aligned} x_2 &= a_0 + a_1x + a_2y \\ y_2 &= b_0 + b_1x + b_2y \end{aligned} \qquad (6)$$

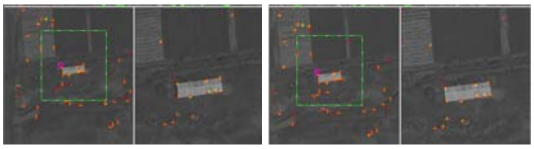
If the linear grey aberrance of the right image relative to the left one are taken into account, then the following formula can be acquired:

$$g_1(x,y)+n_1(x,y)=h_0+h_1g_2(a_0+a_1x+a_2y,b_0+b_1x+b_2y)+n_2(x,y) \qquad (7)$$

The error formula of least squares image matching can be acquired by establishing error formula and do linearization to the formula.

$$\upsilon=c_1dh_0+c_2dh_1+c_3da_0+c_4da_1+c_5da_2+c_6db_0+c_7db_1+c_8db_2-\Delta g \qquad (8)$$

Establishing error formula pixel by pixel in the object area and solving the grey and geometry aberrance by normalizing. Then geometry and grey transformation is applied to right window to get new image, calculating iteratively the correlation coefficients of the left and the right window, the object place is obtained when the value of correlation coefficients do not rise and the loop course to be stopped. During the loop course, the threshold is set as a condition to eliminate coarse error and eliminate the unreliable matching points.



(Left Image) (Right Image)

Figure 6, Superior Results by Using Least Squares Image Matching(The Orange Points are the Original Position of the Character，The Red Points are the Position of the Characters be Disposed Superiorly)

3.5 Iterative Stereo Orientation

The corresponding coordinates (x_1 , y_1)and(x_2 , y_2)are obtained by accurate measurement, then resolving the relative orientation elements ($B_x, B_y, B_z, \Delta\alpha_x, \Delta\omega, \Delta\kappa$), the model coordinates are obtained by front intersection formula and then the stereo model is reconstructed.

During the program from character point detection to orientation, a series of criterion is adopted as conditions to insure the correct of process. The criterion includes point character value, correlation coefficient and fluctuate parallax. Because the process following is a sequence one, the criterion used in each phase is localization which results in the quantity and distributing of the results can't satisfy the requirement of high precise orientation. The iterative strategy is adopted to settle the localization problem, which repeat step two to four on basis of the initial orientation. If integrated criterion is used in processing the effect of the stereo image auto-orientation can be improved farther.

4. DENSE IMAGE MATCHING TO AUTO-EXTRACT DSM

Digital surface model (DSM) is extracted by dense image matching. The initial irregular triangular web constituted with corresponding points gotten from the stereo orientation is used as geometry restriction conditions. The self-adaptive window is used in the matching course to settle the edge fuzzy in the high resolution image matching. The selection of self-adaptive window makes use of the classify results of multi-spectrum images. The image in the matching window is divided into background and thematic information to eliminate the effects of the discrete ground surface. The final digital surface model is the refined of the spectrum classify and the thematic information levels such as building and vegetation are generated. The basic technique following refers to the down part of Figure 3.

4.1 Irregular Triangular Web Construction by Corresponding Points

An increment style random dynamic Delaunay triangular web construction algorithm is promoted in the paper according to how to make use of discrete corresponding image points to quickly construct Delaunay triangular web with optimize space structure and how to do effective edition (such as add or delete vertex) on the already exist web [4]. The algorithm has such characters: adaptability, randomicity and high efficiency dynamic edition for great capacity of data. The capability of the algorithm is analyzed thoroughly: the space complication degree of the data structure involved in the algorithm is "linearity factorial" $O(n)$; the time complication degree of the web construction algorithm is "power factorial" $O(n \log n)$; the complication degree of search time of the triangle cell or web vertex is "logarithm factorial" $O(\log n)$ which is the most optimized resolution for such kind of question as query and search. The application effect refers to Figure 7.

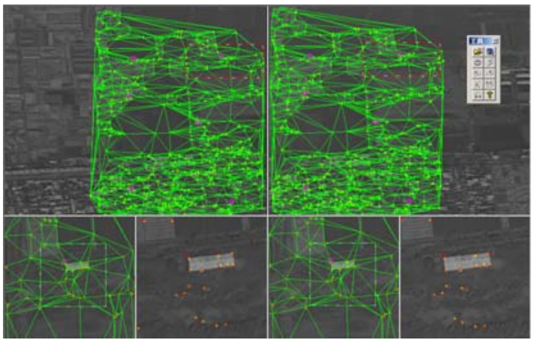


Figure 7, The Irregular Triangular Web Constructed by Corresponding Character Points

4.2 Iterative Edge Character Matching under Space Restriction Condition

The initial irregular triangular web obtained by image character points matching is rather coarse because the infection of adverse factors such as occlusion, noise, shadow and aberrance etc. The space restriction condition needs to be extended and perfected for dense image matching. In order to avoid resulting in wrong results, the edge character matching is adopted to obtain typical space character lines. The lines are used to restrict the error in the DSM and improve the accuracy of DSM.

There are abundant research on edge extraction and matching. The Canny operator is adopted in the paper to detect edges [5], tracking the edges to extract the vector character of the edge after burr having been wiped off. The inflexion of the edge character can be confirmed when the grad value of the point is greater than a threshold. After inflexion obtained, the inflexions on the character lines can be matched with the epipolar condition to obtain corresponding points. Creating web with inflexions obtained to establish triangular web. Based on the relative connection of these grids to restrict the search range in line character matching and make edge matching quick and exact. Character matching methods refer to reference [6].

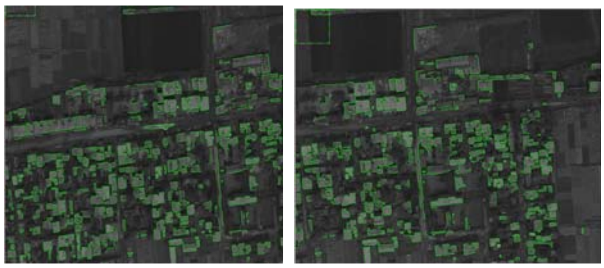


Figure 8, Vector Edges Extraction from Stereo Image Pairs

4.3 Dense Grid Matching

4.3.1 Ground Matching: The purpose of image matching is to extract object's geometry information and confirm its space position. Space front intersection should be employed to calculate the space coordination of the corresponding points after obtaining parallax of left and right photo through image matching. The digital surface model is created after the ground coordinates of corresponding points have been obtained. Some inter methods have been employed in DSM generation, so the precision willreduce more or less. The image matching based on ground only needs calculating Z value because the planar coordinates (X, Y) are known.

The vertical line locus (VLL) is adopted in the paper. If there is a vertical line locus on ground, its projection on the photo will also be a line (refer to fig 9). That is to say the projection of the intersection point A of VLL must be on the parallax line on the photo. To solute the Z value of point A on ground through making use of VLL to search the corresponding image points a_1 and a_2.

$$Z_i = Z_{\min} + i \cdot \Delta Z \qquad (i = 1,2,\cdots)$$

$$\begin{cases} x_i = -f \dfrac{a_1(X - X_S) + b_1(Y - Y_S) + c_1(Z_i - Z_S)}{a_3(X - X_S) + b_3(Y - Y_S) + c_3(Z_i - Z_S)} \\ y_i = -f \dfrac{a_2(X - X_S) + b_2(Y - Y_S) + c_2(Z_i - Z_S)}{a_3(X - X_S) + b_3(Y - Y_S) + c_3(Z_i - Z_S)} \end{cases}$$

$$\begin{cases} x_i' = -f \dfrac{a_1'(X - X_S) + b_1'(Y - Y_S) + c_1'(Z_i - Z_S)}{a_3'(X - X_S) + b_3'(Y - Y_S) + c_3'(Z_i - Z_S)} \\ y_i' = -f \dfrac{a_2'(X - X_S) + b_2'(Y - Y_S) + c_2'(Z_i - Z_S)}{a_3'(X - X_S) + b_3'(Y - Y_S) + c_3'(Z_i - Z_S)} \end{cases}$$

$$(x_i, y_i) \qquad (x_i', y_i') \Rightarrow \rho_i$$

$$if \qquad \rho_k \geq \rho_i \qquad then \qquad Z = Z_k$$

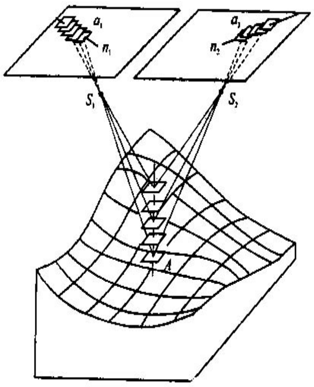


Figure 9, Sketch Map for Ground Matching

The ground matching is a linear searching along Z with △Z in [Zmin，Zmax] because the planar coordinate (X，Y) is known. The search strategy can improve the space stereo matching success ratio.

4.3.2 Multi-level Matching: According to the characters of the remote sensing image, multi-level matching strategy is adopted to overcome partial image distortion and facular infection and increase the matching speed. Image matching is a "sick" program, that is to say there would exist no corresponding point on one image for a given point on the other image because of occlusion or there are more than one corresponding points because of overlap or transparency of the object. In order to overcome the matter and make it an "order" program, additive condition and restrict should be imported. In order to control the complication degree and the reliability of the matching, it is very important to take into account the multi-level data structure and geometry restrict.

From low to high level strategy is adopted to generate remote sensing pyramid image and do image matching frequency divisionally. On the level 0 using single point grey matching which has a wider drag in range. The initial Z value of each point in the DEM is given the roughly average value of the area. Smooth is done to the DEM after each level matching is finished and the result is regarded as approximation value for the next level. The search rang includes all height distributing extent of the area. The higher level uses the results from the former level as its initial value and using shift spherical surface approach to forecast the result. The correlation coefficient is set a threshold and using statistical probability model to eliminate coarse error.

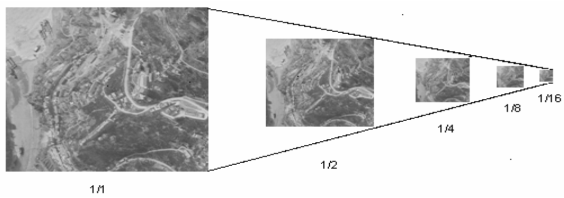
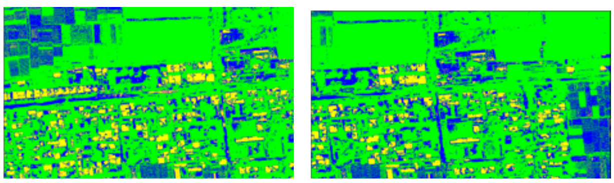


Figure 10, The Image's Pyramid Structure

4.3.3 Matching window adjustment: It can refer to the reference [7] as matching windows self adjustment strategy is considered for high resolution stereo image matching. According to the multi-spectrum info character that remote sensing image has, the initial result which is classified from original image can be treated as the basis of window adjustment for the reconstruction of thematic element such as building.

Low resolution multi-spectrum (RGB+NIR) image is used and spectrum character extraction is applied firstly, then max similar classification is adopted to realize image auto classifying to get original classification result [8]. The original classification result is used as covering template to distinguish foreground and background information. The image character information is integrated in the matching window to determine the image matching result and improve the matching reliability of character indistinctive area.



(Left Image) (Right Image)

Figure 11, The Spectrum Classify Results of the Stereo Images

5. EXPERIMENT AND RESULTS

The experiment data is the digital aerial stereo image gotten from DMC with scale 1:4000 and 5cm ground resolving power, 60% overlap, corresponding area is urban and suburban combination of plain with abundant terrestrial object. There would be 5000~6000 points detected for a single photo when the window size is 7X7. About 1025 corresponding points will be obtained after character point matching or 794 points after epipolar condition detection. The result of relative orientation shows that the accuracy of measurement is ±21.6μm that is 18 pixels size. After least squares matching 774 accurately orientation points can be obtained and the accuracy will be improved to ±2.3μm, the coarse error is partially eliminated, point orientation accuracy has been improved obviously and can reach 1/5 pixel. Dense image matching fusing spectrum classification result realize the DSM auto extraction on stereo model. The buildings and terraces in corresponding area can be all denoted as shown in fig 13. Terrain info combined with thematic character info obtained through spectrum classification can realize the expression and extraction of terrain objects effectively.

Figure 12, DMC Stereo Images (Scale 1:4000)

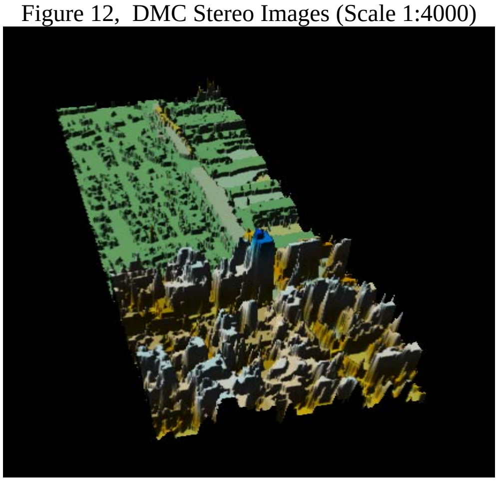


Figure 13, The 3D Play for Auto Extracted DSM

6. CONCLUSIONS

An effective method is put forward in the paper for resolving the difficulty be confronted with high resolution stereo image auto processing. The main target is to resolve the problem that the stereo matching success ratio falls as the resolution improving. The main cause is that with the resolution improved the ratio that character indistinct rise, which leads to big difference between matching result and factual situation. The paper introduces a strategy fusing multi characters which integrates point, linear and area character and spectrum classify technique to improve the accuracy and reliability of the analyze result. The experiment results show that the method introduced in paper is effective, but there need further research in such fields: improving the holistic geometry restriction in linear character and the integrality of multi-images matching.

REFERENCES

[1] C. Harris and M. J. Stephens, 1988. A Combined Corner and Edge Detector, In Alvey88, pp 147-152.

[2] Shen Bangle, 1995. Computer Image Processing, PLA publishing company

[3] Wang Zhizhuo, 1979. Theory of Photogrammetry， Mapping and Surveying Publishing Company.

[4] Hu Haiyan, etc. 2007(1). Design of Randomized Incremental Algorithm for Terrain Modelling, Geomatic Science and Engineering,

[5] F. Canny, 1986. A Computational Approach to Edge Detection, IEEE Trans. Pattern Analysis and Machine Intelligence, 8(6), pp. 679-698.

[6] Fang Yong, Chang Benyi, 2000(1). The Study of SAR Image Map Production Based on the fuse of Multi-Sensors Remote Data, Geomatic Science and Engineering.

[7] Zhang, L., 2005. Automatic Digital Surface Model (DSM) Generation from Linear Array Images, Ph. D. Dissertation, Report No. 88, Institute of Geodesy and Photogrammetry, ETH Zurich, Switzerland. pp.206

[8] Lukas Z., Andreas K., Barbara G., Konrad K., 2006. Towards 3D map generation from digital aerial images. ISPRS Journal of Photogrammetry & Remote Sensing， 60(6), pp.413-427.