3-D TERRAIN RECONSTRUCTION WITH AERIAL PHOTOGRAPHY

Bin-Yih Juang (莊斌鎰)\textsuperscript{1,2} and Chiou-Shann Fuh (傅楸善)\textsuperscript{3}

\textsuperscript{1}Ph. D candidate of Dept. of Mechanical Engineering
National Taiwan University, Taipei, Taiwan

\textsuperscript{2}Instructor of Dept. of Mechanical Engineering
St. John’s & St. Mary’s
Institute of Technology, Taipei, Taiwan

\textsuperscript{3}Prof. of Dept. of Computer Science and Information Engineering
National Taiwan University, Taipei, Taiwan

ABSTRACT
This paper presents the 3-D terrain reconstruction with aerial photography. The results of 3-D reconstruction using intensity-based matching and hybrid intensity-based and epipolar geometry approach matching are shown and compared. The intensity-based image matching uses maximum normalized cross correlation to find the displacement vector between two images. This method can overcome the effect of shadow. The hybrid method is divided into three steps: First, compute eight image-point matching using maximum normalized cross correlation. Second, epipolar line constraint is computed using the matching results of eight image points. Finally, the precise image matching is finished fast under the epipolar line constraint. In this paper, we add an artificial intelligence analysis procedure to improve the computing time for this kind of problem. The relative depth is computed through stereo vision analysis technique using two-images and presented in 3D.

1. INTRODUCTION

Three-dimensional (3D) terrain reconstruction with aerial photography is an important problem in computer vision application engineering. Three-dimensional terrain reconstruction data can be applied to a large construction project or national defense, especially, to achieve the maximum land resource utilization with dense population. In fact, the 3-D reconstruction problem has been focus of attention for a long time. However, It is restricted by the computer’s speed and capacity. Recently, high speed and high capacity computer is popular, such that the research and development of 3-D
reconstruction technologies have more possibilities and are more convenient and easy. There are many reconstruction researches based on feature-based matching. Vergauwen et al. [1] apply stereo vision reconstruction planetary surface. The position of real world point is unknown. Pedersini et al. [2] present a method for interpolating unstructured 3-D data to perform an accurate localization of the boundary of the object. Haralick [3, 4] and Nalwa [10] present many computer vision methods and techniques. Torkar [5] shows a feature-based stereo correspondence method to deal with a set of aerial stereo image of urban scenes. There are many stereo vision methods and techniques on Internet [7]. Forstner and Gulch [8] show the fast feature extracting such as corner detector and center of circular detector. Mendonca et al. [9] develop a novel method to achieve pose estimation and reconstruction from image profiles alone. Hartley [11] investigates projective reconstruction of geometric configurations and invariants from multiple images. Although feature-based matching has precise matching result, features are usually needed to deal with interest point or feature extraction. In this paper, first, we use maximum normalized cross correlation to measure the similarity between two images. Next, we combine epipolar geometry and maximum normalized cross correlation. Finally, we modify the epipolar line constraint to form a band to increase the precise matching.

2. MODELING

2.1 Normalized Cross-Correlation

Consider the typical aerial photography system in Fig. 1 to capture the terrain images. The airplane with a high-resolution camera flies over a target area and takes a picture per interval. There is nearly 60% overlap area between two images. The overlap area between two images is shown in Fig. 2.

![Illustration of aerial photography experiment.](image)

Through geometry relationship of perspective projection, the depth $Z$ from the object to camera can be found.

$$Z = \frac{f \cdot b_f}{(U_R - U_L)}$$  \hspace{1cm} (1)

The altitude of the object $h$ is the difference between the altitude of the camera to the depth $Z$.

$$h = T_h - Z$$  \hspace{1cm} (2)
The flight direction
Region1 Region2 Region3 Region4

Overlap area
The first image

The second image

Fig. 2 The overlap area between two images for reconstruction.

The overlap areas, Region2 and Region3, are 3D reconstruction areas. Points $A$, $B$, $C$, and $D$ are the locations of bird’s eye views. There is an assumption for reconstruction.

**Assumption 1:** The airplane with camera keeps fixed flight direction and altitude.

The procedure of the 3D reconstruction with aerial photography is divided into three steps. First, choose six points from the first image’s Region2 and Region3, to find the average offset. Second, using the offset, the global image matching problem is simplified to local image matching that can reduce much computing time. Finally, the maximum normalized cross correlation method is applied for image matching.

Using maximum normalized cross correlation to find the displacement vector $(dx, dy)$ such that the cost function is maximum.

\[
\max \frac{\sum(I_1 - \bar{I}_1)(I_2 - \bar{I}_2)}{\sqrt{\sum(I_1 - \bar{I}_1)^2 \sum(I_2 - \bar{I}_2)^2}}
\]  

where $I_1(x, y)$ is the intensity of the first image; $I_2(x + dx, y + dy)$ is the intensity of the second image; $\bar{I}_1(x, y)$ is the average intensity on $(x, y)$; and $I_2(x + dx, y + dy)$ is the average intensity on $(x + dx, y + dy)$.

The image matching process of the software is divided into three steps.

Step 1: The first image is divided into $n \times p$ patches, the size of each patch is $m \times m$ pixels. In our experiments, $m$ is 37 pixels; $n$ is 54 rows; $p$ is 28 columns. The horizontal and vertical steps are both 10 pixels. The START point is from offset $(50, 320)$.

Fig. 3 The first image divided for image matching

Step 2: From the second image, find a patch ($m \times m$), such that the normalized cross correlation is maximum. Record the relative displacement vector $(dx, dy)$ from the patch of the second image to the first image.

Step 3: Use each point’s displacement vector to compute the range data for each point.
2.2 Epipolar Geometry Approach

Epipolar geometry is the fundamental geometric relationship between two images [9]. Consider an epipolar geometry system in Fig. 4.

**Assumption 2:** Neglect the distortion effect of camera lens.

![Fig. 4 The epipolar geometry between two cameras.](image)

where \( P \) is the object in real world; \( O \) is the projection center of Camera 1; and \( O' \) is the projection center of Camera 2. The plane of \( OPO' \) is called epipolar plane; \( b_x \) is base line; and the intersection of image plane and epipolar plane is called epipolar line. There are two epipolar lines in Fig. 4. Point \( p = (u,v,1) \) is the perspective projection of image point \( P \) on Camera 1. Point \( p' = (u',v',1) \) is the perspective projection of image point \( P \) on Camera 2. The epipolar geometry describes the point \( p \) and point \( p' \) satisfying

\[
pFp' = 0
\]

where \( F \) is called fundamental matrix.

\[
F = \begin{bmatrix}
    f_{11} & f_{12} & f_{13} \\
    f_{21} & f_{22} & f_{23} \\
    f_{31} & f_{32} & f_{33}
\end{bmatrix}
\]

Extend Eq. (4) we can obtain

\[
\sum_{i=1}^{3} \sum_{j=1}^{3} f_{ij} x^i y^j = 0
\]

choose eight points to solve the elements of the fundamental matrix \( F \). In this paper, we choose eight points near the center of the lens to have larger feature image area. Make sure rank(\( F \))=2, and \( \sum f_{ij}^2 = 1 \). Then we can solve the epipolar line equation. The epipolar line constraint is the strongest constraint in image matching. It is independent of the shape of the object. It reduces the two-dimensional search to one-dimensional search [4]. The epipolar line constraint is usually applied to stereo vision correspondence.

3. EXPERIMENT AND RESULTS

3.1 The Reconstruction Results Using Normalized Cross Correlation

The experimental equipments and data [6] are the following. The camera is Rmktop15/23. Camera’s focal length is 15.28cm. The film size is 23cm by 23cm. The scanner resolution for each pixel is 21um. The scanner converts analog image to digital image. The baseline is 1350m. Each reconstruction patch size is 63.42m by 63.42m.
Case 1

In this case, the image pairs in Fig. 5 are located at middle altitude (400–700 m) from sea level. Fig. 6 presents the top view of 3D reconstruction result and defines the four corner points \(A, B, C,\) and \(D\). Fig. 7a–7d presents the bird’s eye views of 3D reconstruction result from points \(A, B, C,\) and \(D\) for Case 1.

Three dimensional terrain reconstruction using normalized cross correlation measures the similarity between two images.

Case 2

In this case, the image pairs in Fig. 8 are located at low altitude (90–200 m) from sea level. Fig. 9 presents the top view of 3D reconstruction result and defines the four corner points \(A, B, C,\) and \(D\) for Case 2. Fig. 10a–10d presents the bird’s eye views of 3D reconstruction result from points \(A, B, C,\) and \(D\) for Case 2.
3.2 The Reconstruction Results Using Hybrid Normalized Cross Correlation and Epipolar Geometry

In order to find the parameters of fundamental matrix, we choose eight points near the center of the lens. These points all locate at the image areas with more features.

Fig. 11 The eight point locations and the epipolar lines for Case 1.

Fig. 10a~10d The bird’s eye views of 3D reconstruction result at positions A, B, C, and D for Case 2.

Fig. 12 (a) The top view of 3D reconstruction result with hybrid maximum normalized cross correlation and epipolar line constraint for Case 1. (b) the result using a band around epipolar line constraint.

Fig. 13a~13d The bird’s eye views of 3D reconstruction result with hybrid maximum normalized cross correlation and epipolar line constraint at positions A, B, C, and D for Case 1.
Fig. 14a–14d The bird’s eye views of 3D reconstruction result with hybrid maximum normalized cross correlation and a band around epipolar line constraint at positions A, B, C, and D for Case 1.

Fig. 15 The eight point locations and the epipolar lines for Case 2.

Fig. 16 (a) The top view of 3D reconstruction result with hybrid maximum normalized cross correlation and epipolar line constraint for Case 2. (b) the result using a band around line constraint.

Fig. 17a–17d The bird’s eye views of 3D reconstruction result with hybrid maximum normalized cross correlation and epipolar line constraint at positions A, B, C, and D for Case 2.

Fig. 18a–18d The bird’s eye views of 3D reconstruction result with hybrid maximum normalized cross correlation and a band around epipolar line constraint at positions A, B, C, and D for Case 2.

4. DISCUSSION
Comparing the results of Cases 1 and 2, low-texture image area has high mismatching rate. Therefore, high-texture image area has high successful matching rate. However low-texture area has low successful matching rate. Modifying the mask size affects matching
results. Small mask size gives more precise altitude but increases mismatching rate. Large mask size decreases mismatching rate but gives less precise altitude and wastes computing time.

5. CONCLUSION
The main idea of the maximum normalized cross correlation tries to achieve the maximum similarity degree between two images. The epipolar geometry reduces the two-dimensional search to one-dimensional search. In this paper, we present three kinds of 3D terrain reconstruction results using aerial photograph image pairs. The result of hybrid maximum normalized cross correlation and epipolar line constraint is the fastest. However the result of hybrid maximum normalized cross correlation and a band around epipolar line constraint is more exact.

6. REFERENCES