SYNTHESIS OF STEREO-MATE THROUGH THE FUSION OF A SINGLE AERIAL PHOTO AND LIDAR DATA

Howook Chang, Jaewan Choi, Hyejin Kim, Jaebin Lee, Kiyun Yu

School of Civil, Urban & Geo-System Engineering, Seoul National University, Korea
seeu942@snu.ac.kr, vicky2@snu.ac.kr, kiyun@snu.ac.kr

ABSTRACT :

Generally, stereo pair images are necessary for 3D viewing. In the absence of quality stereo-pair images, it is possible to synthesize a stereo-mate suitable for 3D viewing with a single image and a depth-map. In remote sensing, DEM is usually used as a depth-map. In this paper, LiDAR data was used instead of DEM to make a stereo pair from a single aerial photo. Each LiDAR point was assigned a brightness value from the original single image by registration of the image and LiDAR data. And then, imaginary exposure station and image plane were assumed. Finally, LiDAR points with already-assigned brightness values were back-projected to the imaginary plane for synthesis of a stereo-mate. The imaginary exposure station and image plane were determined to have only a horizontal shift from the original image's exposure station and plane. As a result, the stereo-mate synthesized in this paper fulfilled epipolar geometry and yielded easily-perceivable 3D viewing effect together with the original image. The 3D viewing effect was tested with anaglyph at the end.

KEY WORDS: LiDAR, Back projection, synthesis, epipolar, stereo

1. INTRODUCTION

Synthesizing new views is a frequently mentioned research topic in computer vision. In computer vision, researchers' concern is how to synthesize a multiple-view imagery of speakers in a teleconferencing system. To give physically plausible images of speakers regardless of the audience's view-angle, synthesizing new views from a given stereo pair is under intensive research. There have been many works that deal with the synthesis of new views for better teleconferencing. McMillan(1995), Sawhney(2000), Criminisi(2003), etc. employed IBR(image-based rendering). IBR refers to techniques that allow 3D scenes and objects to be visualized in a realistic way without full 3D model reconstruction. Scharstein(1996) and Isgro(2001) used disparity map together with stereo images. Chen(1993), Werner(1995), Katayama(1995), etc. exploited view interpolation. View interpolation approach works well when all of the partial sample images share a common gaze direction, and the synthesized viewpoints are restricted to stay within 90 degrees of this gaze angle. Those works mentioned above dealt with only indoor scenes, and it remains questionable whether those methods are appropriate for aerial photographs or not. An aerial photograph deals with much a larger scale and it cannot have sequential frames with a short interval between the frames contrary to indoor scene capturing. Consequently, this paper proposes a methodology to synthesize new views in aerial photography.

Synthesizing a stereo pair was also mentioned in anaglyph generation. According to Ideses, a stereo pair synthesized from a modified depth map yielded better anaglyph than the original pair. In remote sensing, synthesizing new views is also necessary. When a stereo pair was taken with a long base line and severely unsatisfactory epipolar geometry, stereo viewing was severely disturbed. In this case, synthesis of a new pair would help comfortable stereo viewing. In this paper, a new methodology is proposed to synthesize new views through the fusion of a single aerial photo and LiDAR data.

Synthesis of new views with an aerial photograph employed three techniques. They were registration between aerial photograph and LiDAR data, back-projection, and interpolation within the delauney triangles. The flow chart of our approach is as follows.

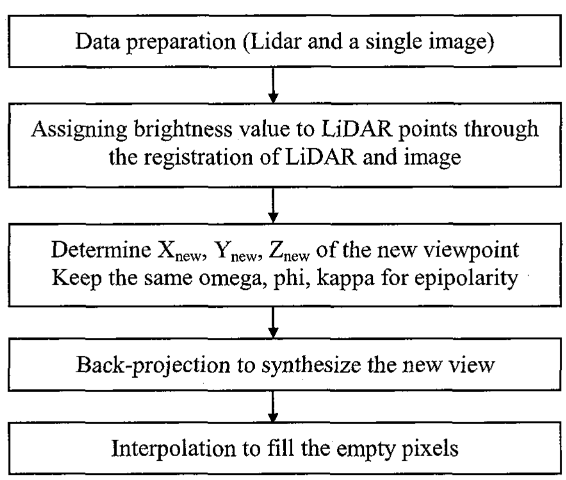

Figure 1. Flow chart

Registration of a photograph and LiDAR data can be carried out in several ways. Habib(2004) introduced a methodology for the fusion of LiDAR data and an aerial photograph using straight-line features. In his paper, straight line features were extracted from a photograph and LiDAR data, and LiDAR line features were used as control information. As a result the photograph was aligned relative to the LiDAR reference frame. On the other hand, if accurate interior and exterior orientation parameters are already known for the photograph, registration of aerial photograph and LiDAR data is performed quite easily by utilizing collinearity equation. It will be demonstrated in the following sections.

2. EXPERIMENTAL METHOD

To synthesize a new view (a new stereo-mate), a single aerial photograph and LiDAR data were used. LiDAR data presents quite accurate full 3D model and covers a wide area, so it can be adequately used in generating new views together with an aerial photograph.

2.1 Registration of a photograph and LiDAR data

For the synthesis of new stereo-mate, it is necessary to give every LiDAR point a brightness value in the first stage. Registration of a photograph and LiDAR data should be preceded for this purpose. When interior and exterior orientation parameters are not available, an elaborate process is required for the registration. However, quite accurate interior and exterior parameters were available in this experiment. Therefore, registration of the aerial photograph and LiDAR data was carried out easily through the collinearity equation. Every LiDAR point was transformed to the photograph's pixel coordinate through the collinearity equation, and acquired BGR information from the destination pixels.

2.2 Determination of new viewpoint's angle and location

After the registration of the photograph and LiDAR data, new viewpoint's attitude and location should be determined. They can be determined arbitrarily, but they were arranged to offer comfortable stereo viewing and to satisfy epipolar geometry. As it is known, comfortable stereo viewing requires that the length of base-line be 1/30 of the distance between the camera centre and the closest object from it. Thus, the new viewpoint's location was arranged to satisfy this requirement. To satisfy epipolar geometry, the new viewpoint's attitude was set to be the same as the original photograph. Supposing that the original photograph's rotations were ω , ϕ , and κ , the new viewpoint's rotations were also determined to be ω , ϕ , and κ . The following equations express the whole requirements all together.

$$M[(X_L, Y_L, Z_L) - (X_{new}, Y_{new}, Z_{new})] = (k_1, k_2, 0)$$

$$|(X_L, Y_L, Z_L) - (X_{new}, Y_{new}, Z_{new})| = \textit{distance to the closet object}/30 \quad (1)$$

$$\textit{where } L : \textit{the original photograph}$$

$$new : \textit{new viewpoint}$$

$$k_1, k_2 : \textit{arbitrary constant}$$

2.3 Back projection and interpolation in delauney triangles

After X_{new}, Y_{new}, and Z_{new} are determined, LiDAR points are back-projected to new viewpoint. Using X_{new}, Y_{new}, and Z_{new} together with ω , ϕ , and κ of the original photograph, the collinearity equation transforms each LiDAR point to its corresponding point that will comprise the new view. Because LiDAR points are already given brightness value from the original photograph, it is possible to generate a new image which is viewed from the new viewpoint. Even if all the LiDAR points are back-projected to the new viewpoint, the new image can not be filled completely. This phenomenon could be interpreted as follows. LiDAR points' density was not high enough to fill the new image completely and especially, there were not enough LiDAR points in the side wall of the tall buildings. To solve these problems, interpolation was performed based on delauney triangular network. Delauney triangular network was formed with the LiDAR points which had been back-projected to the new image. Then, the pixels inside each triangle were assigned brightness values through linear interpolation based on the brightness values of each triangle's vertices. This process will be demonstrated in detail in the following sections.

3. EXPERIMENTAL RESULT

The original aerial photograph and LiDAR data used in this experiment were as follows. The reference frames were identical for the photograph and LiDAR data.

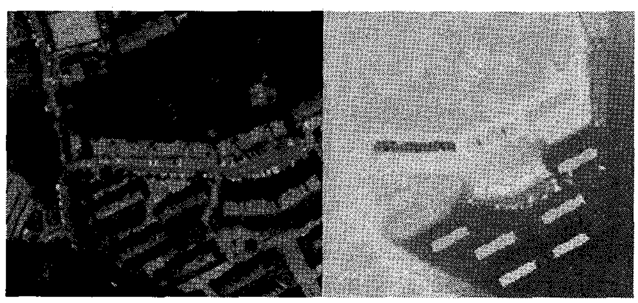

Figure 2. The original photo and LiDAR data

The aerial photograph was taken with the DSS SN0029 – 55mm lens, and the exterior orientation parameters were as in the following table.

X_L	233763m

Y_L	320843.619m
Z_L	1550.510m
ω	4.03625°
Φ	-2.44085°
κ	2.37510°

Table 1. EOP of the original aerial photograph

Registration of the original photograph and LiDAR data was carried out with the following equation.

$$x_L = -f\frac{m_{11}(X_{Lidar}-X_L)+m_{12}(Y_{Lidar}-Y_L)+m_{13}(Z_{Lidar}-Z_L)}{m_{31}(X_{Lidar}-X_L)+m_{32}(Y_{Lidar}-Y_L)+m_{33}(Z_{Lidar}-Z_L)}$$

$$y_L = -f\frac{m_{21}(X_{Lidar}-X_L)+m_{22}(Y_{Lidar}-Y_L)+m_{23}(Z_{Lidar}-Z_L)}{m_{31}(X_{Lidar}-X_L)+m_{32}(Y_{Lidar}-Y_L)+m_{33}(Z_{Lidar}-Z_L)} \quad (2)$$

where L : *the original photograph*
x, y : *The original photograph 's coordinate*
f : *focal length of DSS SN - 0029*
X, Y, Z : *Lidar point's coordinate in reference frame*
m : *rotation matrix according to* ω, ϕ, κ

New viewpoint's angle and location were decided according to equation (1). X_L, Y_L, and Z_L of the original photograph were put into the equation (1) and k_1 and k_2 were set arbitrarily. As a result, two equations with three unknowns were formed (As a result, multiple-solution was obtained). After calculating X_{new}, Y_{new}, and Z_{new}, back-projection to the new viewpoint was carried out with the equation similar to equation(2). In carrying out back-projection, equation(2) was reused with X_L, Y_L, and Z_L replaced with X_{new}, Y_{new}, and Z_{new} respectively. The new image synthesized by back-projection was as follows.

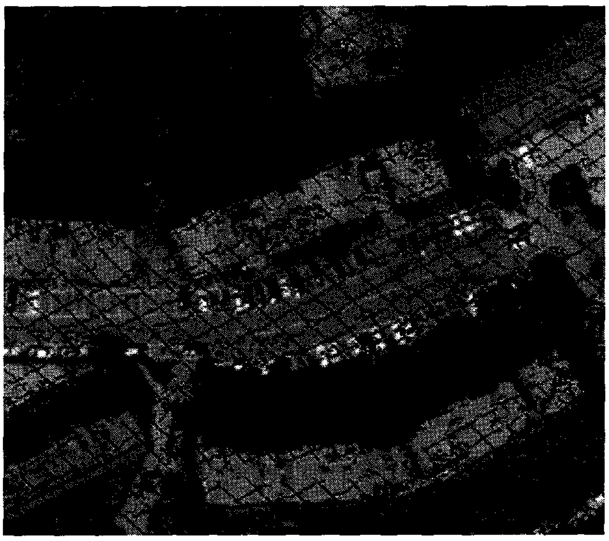

Figure 3. The image synthesized for the new viewpoint (only partial image was displayed)

As shown in figure 3, there were many unfilled pixels in the synthesized image. Only part of the image was displayed to clearly show unfilled pixels in detail. You can figure out that tall buildings' side walls were not filled at all, so stereo viewing was severely disturbed when using the synthesized image. In addition, unfilled pixels were distributed all around the image in the net shape. The reason that tall buildings' side walls remained empty was that LiDAR points were rare in tall side walls. Net-shaped empty pixels were also due to insufficient LiDAR points' density. To solve these problems, empty pixels were filled out by interpolation. Delauney triangular network was formed with the points that already had brightness values. Within each triangle, linear interpolation was carried out to fill the empty pixels within the triangle. The image synthesized through the interpolation process was as follows.

Figure 4. The synthesized image after interpolation

As you can see, unfilled pixels in tall buildings' side walls were filled out completely, and net-shaped empty pixels were also removed.

4. ACCURACY EVALUATION

Accuracy was evaluated in two ways. First, we tested whether epipolar lines coincided horizontally or not. We sampled about 175 matched point-pairs between the original image and the synthesized image, and observed that every pair existed on the same horizontal line within the error range of plus or minus 0.38 pixel. The statistics about epipolarity is shown in Table2. Then, the original image and the synthesized image were combined into an anaglyph to test if the anaglyph gave us enough 3D viewing effect or not. The quality of the anaglyph proved that the synthesized image was completely valid. The resulting anaglyph is presented on the next page. Although the anaglyph did not yield enough 3D viewing effect around the tall buildings as we had expected, it still yielded satisfactory 3D viewing effect. Especially, it described the highs and lows of the forest area quite well.

Max discrepancy along y-axis	0.38 pixel (175 matched point-pairs were sampled)
RMSE of discrepancy	0.11 pixel

Table 2. Accuracy test about epipolar geometry

5. CONCLUSION

In this paper, synthesis of new views with a single aerial photograph and LiDAR data was presented. Even if there are already many researches related to the synthesis of new views with indoor scenes for better teleconferencing, aerial photographs should be treated in a different way from those for indoor image frames. It has a much larger scale and long interval between image frames. For these properties of aerial photographs, LiDAR data that can cover wide area was used with them. The synthesized image through this paper's process yielded quite good result. Epipolar lines coincided within 0.38 pixel discrepancy between the original image and the synthesized image, and the resultant anaglyph also yielded good 3D viewing effects.

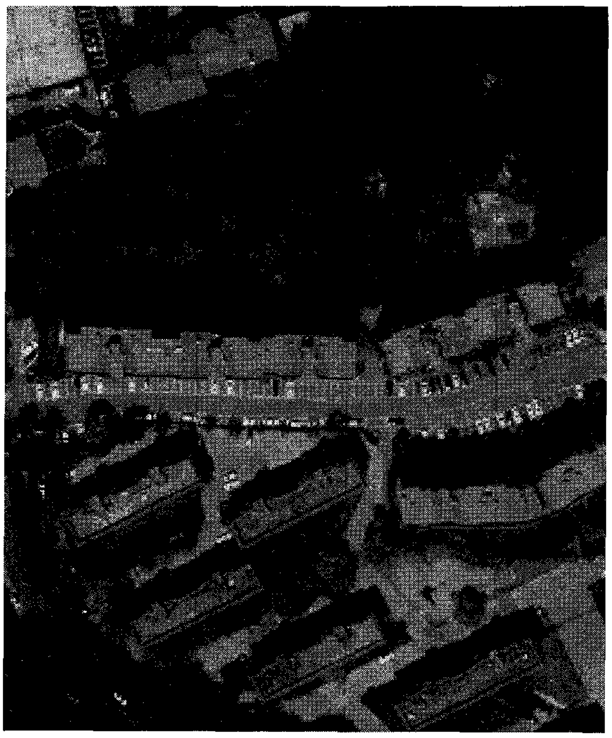

Figure 5. Anaglyph generated from the original image and synthesized image

Acknowledgments

Authors are gratefully acknowledging the financial support by Seoul Research & Business Development (10540).

References

Chen S.E. and Williams L., 1993, View interpolation for image synthesis, SIGGRAPH, pp273~288

Criminisi A.et al, 2003, Efficient dense-stereo and novel-view synthesis for gaze manipulation in one-to-one teleconferencing. Microsoft research LTD, UK

Habib A. et al, 2002, Alternative approaches for utilizing LiDAR data as a source of control information for photogrammetric models. Department of Geomatics Engineering in University of Calgary, Canada

Habib A. et al, 2004, Co-registration of photogrammetric and LiDAR data: Methodology and case study. Department of Geomatics Engineering in University of Calgary, Canada

Ideses I. and Yaroslavsky L., 2005, methods that improve the visual quality of colour anaglyphs, Journal of Optics A: Pure and applied optics 7, pp755~762

Isgro F. et al, 2001, Towards teleconferencing by view synthesis and large baseline stereo, 11th Intl. conference on image analysis and processing

Katayama et al. 1995, A view point independent stereoscopic display using interpolation of multi-view point images, SPIE stereoscopic displays and virtual reality systems II, pp11~20,

Mark W.R. et al, 1997, Post-rendering 3D warping, In Symposium on interactive 3D graphic, pp7~16

McMillan L. and Bishop G., 1995, Plenoptic modelling: An image based rendering system, SIGGRAPH, pp39~46

Sawhney H.S. et al, 2000, Hybrid stereo camera: An IBR approach for synthesis of very high resolution stereoscopic image sequences, Sarnoff Corp

Scharstein D., 1996, IEEE computer society conference on computer vision and pattern recognition, pp852~858

Werner T.et al, 1995, Rendering real world objects using view interpolation, In Intl conference of computer vision, pp957~962