Journal of the Korean Society of Surveying, Geodesy, Photogrammetry and Cartography
Vol. 35, No. 2, 75-80, 2017
https://doi.org/10.7848/ksgpc.2017.35.2.75

ISSN 1598-4850(Print)
ISSN 2288-260X(Online)
Original article

Analysis of the Increase of Matching Points for Accuracy Improvement in 3D Reconstruction Using Stereo CCTV Image Data

Moon, Kwang-il[1) · Pyeon, MuWook[2) · Eo, YangDam[3) · Kim, JongHwa[4) · Moon, Sujung[5)

Abstract

Recently, there has been growing interest in spatial data that combines information and communication technology with smart cities. The high-precision LiDAR (Light Dectection and Ranging) equipment is mainly used to collect three-dimensional spatial data, and the acquired data is also used to model geographic features and to manage plant construction and cultural heritages which require precision. The LiDAR equipment can collect precise data, but also has limitations because they are expensive and take long time to collect data. On the other hand, in the field of computer vision, research is being conducted on the methods of acquiring image data and performing 3D reconstruction based on image data without expensive equipment. Thus, precise 3D spatial data can be constructed efficiently by collecting and processing image data using CCTVs which are installed as infrastructure facilities in smart cities. However, this method can have an accuracy problem compared to the existing equipment. In this study, experiments were conducted and the results were analyzed to increase the number of extracted matching points by applying the feature-based method and the area-based method in order to improve the precision of 3D spatial data built with image data acquired from stereo CCTVs. For techniques to extract matching points, SIFT algorithm and PATCH algorithm were used. If precise 3D reconstruction is possible using the image data from stereo CCTVs, it will be possible to collect 3D spatial data with low-cost equipment and to collect and build data in real time because image data can be easily acquired through the Web from smart-phones and drones.

Keywords : 3D Reconstruction, Matching Point, SIFT Algorithm, Stereo CCTV

1. Introduction

The multidimensional geospatial information refers to varies forms of data collected during the construction of digital territory. This kind of geospatial information service is required to handle the non-geospatial region as well as geospatial areas (Lee, 2008). In terms of 3D geospatial information construction services, there has been the introduction of new sensors such as LiDAR, ToF (Time of Flight) and depth sensor (Lee *et al*., 2009). However, LiDAR has such problems as high price and long data update cycles although it guarantees high accuracy. The method using ToF is also expensive even though it provides high accuracy. Furthermore, it is difficult to acquire data outdoors due to the short range of sensors and is inappropriate for acquiring precise data due to relatively low resolutions. On the other hand, CCTVs are relatively low-priced and cutting-edge CCTV technologies are appearing owing to technical development. With these cutting-edge CCTVs, we can acquire high-quality image data and measure data even in nighttime and climate changes. This study aims to improve accuracy by increasing the number of matching points and

Received 2016. 04. 06, Revised 2016. 04. 25, Accepted 2017. 04. 30
1) Member, Dept. of Civil Engineering, Konkuk University (E-mail: rhkddlf1114@ konkuk.ac.kr)
2) Corresponding Author, Member, Dept. of Civil Engineering, Konkuk University (E-mail: Neptune@konkuk.ac.kr)
3) Member, Dept. of Advanced Technology Fusion, Konkuk University (E-mail: eoandrew@konkuk.ac.kr)
4) Dept. of Advanced Technology Fusion, Konkuk University (E-mail: fff307@konkuk.ac.kr)
5) Member, Dept. of Advanced Technology Fusion, Konkuk University (E-mail: msujung@konkuk.ac.kr)

minimizing the error range in image matching with the data acquired from CCTVs. For the experiment, a stereo CCTV module was fabricated and the number of extracted matching points was increased by applying the feature-based and area-based matching methods with the image data acquired from the stereo CCTV module. For the feature-based matching method, the SIFT matching results before and after image decomposition were compared and analyzed.

2. Methods for Stereo Image Matching and Image Decomposition

2.1 Methods for stereo image matching

The stereo vision matching method has been researched in the field of computer vision for a long time. Image matching methods can be largely divided into global matching method and local matching method (Koo, 2015). Among the stereo matching methods, the local matching methods are divided into feature-based matching method, which uses a certain range of surrounding information based on the pixel of the current position and extracts some features, and the area-based matching method based on window.

The feature-based matching method includes the SIFT keypoint detector for extracting feature points. The area-based matching method selects one of the left and right images and compares a random area of it with other images and determine similar corresponding points by using the changes in brightness, shapes of areas, average brightness, and the area size. The characteristic of the algorithm used in this experiment is that it only extracts the matching points whose SSD (sum of squared differences) and NCC (normalized cross correlation) values coincided and the epiplolar lines were aligned. Thus, matching points that were away by more than 3 pixels in the vertical direction were removed (Bae, 2011).

2.2 Image decomposition and 3D reconstruction

A digital camera's color images can be improved and decomposed using diverse imaging techniques. In terms of an image-decomposing method, there is RGB (Red-Green-Blue) image decomposition (Yocky, 1995).

In terms of the bands used in the test, the RGB images were divided into gray band, red band, green band and blue band, and these separated images were histogram-equalized again, generating a total of 8 bands (Barsky, 2002). Then, each matching point was extracted and overlapped.

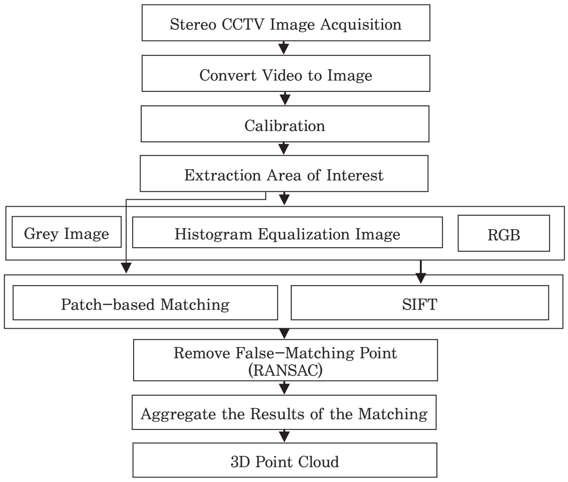

Fig. 1. Testing process to improve accuracy of 3D Point Cloud

Fig. 1 is the testing process that increase accuracy of 3D data. The videos acquired through stereo CCTVs are converted to images and the stereo camera calibration is carried out. In the stereo environment where two images are acquired simultaneously, stereo calibration and stereo rectification processes are required to obtain row-aligned images due to the distortion factors of external images (camera height, direction, etc.). Before matching is carried out, the regions of interest are extracted to reduce the computations and errors of unnecessary matching points and improve the matching accuracy.

The left and right images are decomposed into gray images and histogram equalization images, and each image is decomposed again into RGB bands, and the images are automatically matched using SIFT. In addition, PATCH-based matching is carried out after extracting the region of interest, and then 3D data is built using the DLT (Direct Linear Transformation) and the matching result values after removing mismatched points using RANSAC. For improve accuracy of 3D data (increasing mathing points), we progress patch technique and SIFT algorithm test to increase matching points.

3. Experiments

3.1 Stereo CCTV module & camera calibration

The network camera which was used for filming was AXIS-Q1755. After connecting the power, two IP ports, a router and two CCTV cameras, a system in which the images filmed by the two CCTV cameras can be confirmed and collected by a computer on a real-time basis, was developed (Jang *et al*., 2014)

Fig. 2 is Stereo CCTV module (AXIS' CCTV Q1755-E network cameras) that fixed to parallel camera arrangement with a 1.1m gap

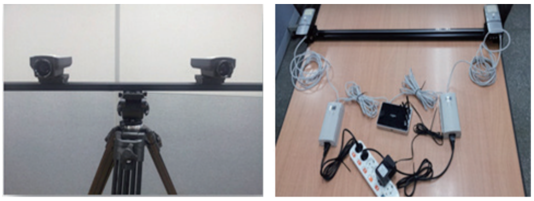

Fig. 2. Stereo CCTV module

And the photographed images that taken by Stereo CCTV module were applied Camera Calibration before matching. Fig. 3 is calibrated before matching and after camera calibration are as follows.

(a) (b)

Fig. 3. (a)Before and (b)After the Stereo Rectification

The resolution of the stereo CCTVs used in the experiment was 1440*1080. When the stereo calibration was carried out, the resolution was 1377*785.

3.2 Feature-based matching (SIFT)

The feature-based matching method was used a SIFT method. The matching points extracted through the SIFT method in each image are overlapped to increase their number. The decomposed images use the same coordinates, and the extracted matching points can be overlapped on the same position (Lowe, 2004).

In case of overlapped matching points, it is considered single-matched according to certain standards. Therefore, the rest should be removed. In this study, if the range of the overlapped matching points is within 1 pixel in both left and right images, they were deemed overlapped and eliminated.

Unless they are properly matched, in addition, a mismatching point may occur (problem of mismatching points). In terms of mismatching, if the matching point angle is 0.02 radian or larger, they were deemed mismatched and removed.

Matching points that extracted by automatically match contains Outlier which is false-matching points. Removing the Outlier efficiently is important part of extracting 3D data which has high accuracy. We use RANSAC algorithm to remove Outlier (Kim *et al*., 2015).

3.3 Region-based matching (patch technique)

The region-based matching used a patch technique. Because of the characteristics of stereo images, disparity takes place for patch-based matching on a similar position, therefore, the right image was moved by 37 pixels from the center, and matching research was conducted. After creating a patch in 200×60 pixels and another patch in 160×50 pixels on the left and right images respectively, the most similar patch was found by allowing the small patch on the right image to retrieve the big patch on the left image. Then, the patch's midpoint was matched. Each patch's search region was set to research most regions where left and right images are overlapped. Then, algorithm was configured in a manner to find matching points on all pixels in the search region based on the right image. The matching points where SSD and NCC are matched were only used. Since the epiline was aligned, the matching points which were off more than 3 pixels vertically were eliminated. The matching points which were off more than 13 pixels horizontally were removed as well. When more than two points were matched on a single spot based on the right image, they were removed (Chen *et al*., 1999).

4. Results and Analysis

4.1 Results of the feature-based matching

As shown in Table 1 and Fig. 4, interest point and matching point counts were extracted, using gray band and HE gray band. In terms of gray band, left and right images were 3982 and 3777 each while the matching results revealed 2087. In HE gray band, in contrast, left and right images were 5306 and 5609 respectively while matching pairs were 2859. When the two matched images were overlapped, the number of matching points was 4938, eliminating 8 overlapped points.

Table 1. SIFT matched points before image decomposition

Image Type	Interest Points Image(Left)	Interest Points Image (Right)	Matching Points	overlapping matching points	Removal of False Matching Points Points
gray	3982	3777	2087	8	4938
HE gray	5306	5609	2859		

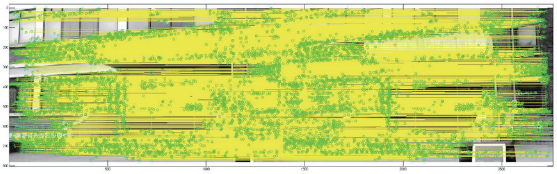

Fig. 4. SIFT matching results before image decomposition

The results of the estimation of matching points using the same images under the image decomposition method are shown in Table 2 and Fig. 5 below. When the CCTV color images were converted into gray and HE gray bands, matching results were the same with the one before the image decomposition. After converting the color images into gray image and decomposing them to the RGB band, they were individually transformed into HE. Then, eight matching points were overlapped. As a result, a total of 18,682 matching points were acquired. When the overlapped points were eliminated, the figures were reduced to 8635, higher than the count before RGB image decomposition (4938) by 3697.

Table 2. Number of SIFT matched points after image decomposition

Image Type	Interest Points Image (Left)	Interest Points Image (Right)	Matching Points	overlapping matching points	Removal of False Matching Points Points
gray	3982	3777	2087	10047	8635
HE_gray	5306	5609	2859		
rgb_1band	3864	3926	2006		
rgb_2band	3797	3844	2063		
rgb_3band	3768	3853	2029		
HE_rgb_1band	5166	5391	2465		
HE_rgb_2band	5403	5782	2565		
HE_rgb_3band	5621	5885	2608		

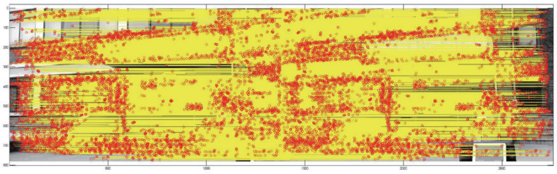

Fig. 5. SIFT matching results after image decomposition

4.2 Region-based matching results

The number of region-based matching points ranged from 1,000 to 500,000. Therefore, there are limitations in measuring them with the naked eye. Therefore, the accuracy was assessed using the slope and distance of the line connected matching pairs. Table 3 and Fig. 6 below reveal the results of matching of left and right images photographed with stereo CCTV cameras using SIFT and their largest deviation and standard deviation.

Table 3. SIFT matching results

Type		Value
Matching Points		1,865 pairs
The Largest Deviation	Slope	0.0114
	Length	1340.80 pixels
Standard Deviation	Slope	0.0011
	Length	31.06 pixels

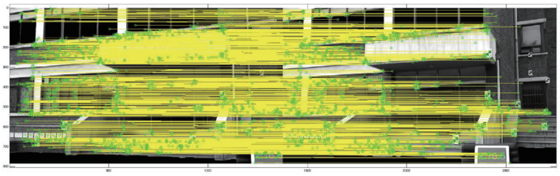

Fig. 6. SIFT matching results

The results of the patch-based matching using the same images are shown in Table 4 and Fig. 7. When the SIFT-based matching results were compared to the patch-based matching results, the latter was greater than the former by 286 times in terms of the number of matching points. In terms of deviation, in contrast, the SIFT algorithm was higher than patch-based algorithm.

Table 4. Patch-based matching results

Type		Value
Matching Points		532,582 pairs
The Largest Deviation	Slope	0.0038
	Length	15.08 pixels
Standard Deviation	Slope	0.0007
	Length	6.06 pixels

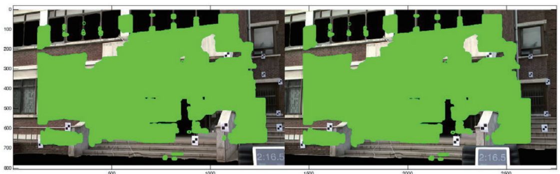

Fig. 7. Patch-based matching results

5. Conclusion

In this study on RDM (Realtime Digital Map), experiments were conducted to increase the number of matching points so as to enhance the precision of 3D reconstruction. For this experiment, the color images captured with stereo CCTVs were converted to gray images and decomposed into RGB colors. Then the number of matching points was increased using the method of overlapping the extracted matching points and the PATCH-based method.

In the feature-based matching results, the number of SIFT matching points increased by approximately 40% after image decomposition. In the area-based matching results, the number of PATCH-based matching points was higher than the number of SIFT matching points, and the standard deviation also decreased.

Therefore, it is more efficient to construct 3D data using both the SIFT- and PATCH-based matching method rather than using the SIFT matching method alone. It appears that it would be able to extract high-precision 3D information in an effective manner by applying them to the stereo images or those photographed with fixed devices such as CCTV cameras. And although this experiment used stereo CCTV, it is expected that real-time 3D data can be constructed by using image data (smart phone, drones, action cam, etc.) that can be easily obtained on the web.

Acknowledgment

This work is financially supported by Korea Ministry of Land, Infrastructure and Transport (MOLIT) as 「U-City Master and Doctor Course Grant Program」. & This research was supported by a grant (15CHUD-C061156-05) from National Spatial Information Research Program funded by Ministry of Land, Infrastructure and Transport of Korean government.

Reference

Bae, K.R. (2011), *A Hardware Architecture Based on the NCC Algorithm for Fast Disparity Estimation*, Master's thesis, Kyungpook National University, Daegu, Korea, 52p.

Barsky, B.A. (2002), *Light, color, perception, and color space theory*, University of California, Berkeley, California.

Chen, C.S., Hung, Y.P., and Cheng, J.B. (1999), RANSAC-based DARCES: a new approach to fast automatic registration of partially overlapping range images, *IEEE Computer Society*, IEEE, Vol. 21, No. 11, pp. 1229–1234.

Jang, I.W., Pyeon, M.W., Hong, K.H., Kim, J.H., Moon, K.I., and Kim, K.S. (2014), Realtime digital map by using multi source video image, *International Conference on Innovative Engineering Technologies (ICIET)* , 28-29 Dec., Bangkok, Thailand, pp. 169-171.

Kim, K.S., Pyeon, M.W., Kim, J.H., and Moon, K.I. (2015), Time series image stereo matching experiment using the overlap method, *Journal of Korean Society for Geospatial Information System*, Vol. 23, No. 1, pp. 123-128.

Koo, J.M. (2015), *A Study on the Robust Stereo Matching Algorithm under Radiometric Change based on Weighted Local Descriptor*, Master's thesis, Chung-Ang University, 65p.

Lee, H.J. (2008), Producing true orthophoto using multi-dimensional spatial information, *Journal of the Korean Society of Surveying*, Vol. 26, No. 3, pp. 241–253.

Lee, H.J., You, J.H., and Koh, Y.C. (2009), Analysis for practical use as KOMPSAT-2 imagery for product of geo-spatial information, *Journal of Korean Society for Geospatial Information System*, Vol. 17, No. 1, pp. 21-35.

Lowe, D.G. (2004), Distinctive image features from scale-invariant keypoints, *International Journal of Computer Vision*, Vol. 60, No. 2, pp. 91–110.

Yocky, D.A. (1995), Image merging and data fusion by means of the discrete two-dimensional wavelet transform, *Journal of the Optical Society of America A*, Vol. 12, No. 9, pp. 1834p.