

Metropolitan Fixed Assets Change Judgment using Aerial Photographs

Hirokazu Koizumi, Hiroyuki Yagyu, Kazuaki Hashizume,
Toshiyuki Kamiya, Kazuo Kunieda, Hideo Shimazu

System Technology Laboratory, NEC System Technologies, Ltd.
8916-47, Takayama-cho, Ikoma-shi, Nara 630-0101, Japan
{koizumi-hxa, yagyu-hxa, hashizume-kxa, kamiya-txa}@necst.nec.co.jp,
k-kunieda@ak.jp.nec.com, shimazu-hxa@necst.nec.co.jp

Abstract

The Tokyo Metropolitan Government is the largest municipality in Japan and conducts building change identification work. Recently, Tokyo terminated its traditional visual identification work that has been in use for 20 years and shifted to a new automated system. This paper is intended to introduce the Fixed Assets Change Judgment (FACJ) system and its core tool, RealScape. RealScape automatically detects the changes in the height and color of buildings based on three-dimensional (3D) analysis of aerial photographs. It employs a unique pixel-by-pixel stereo processing method and enables the foot-level precision for each building. RealScape detects building changes more accurately than visual judgment operations by humans and reduces the labor costs to one third of the traditional approach and the required judgment duration to about two weeks per 100km².

Introduction

This paper describes the Fixed Assets Change Judgment (FACJ) system and its core tool, RealScape. RealScape automatically detects the changes in the height and color of buildings based on the 3D analysis of aerial photographs. The 3D analysis employs a pixel-by-pixel stereo processing method that calculates the height of each pixel in aerial photographs and thus enables precise difference detection between previous and current aerial photographs. The FACJ system reduces the labor costs to one third of the traditional approach and the required judgment duration to about two weeks per 100km². The FACJ system was first experimentally used by the Tokyo Metropolitan Government in 2005. Since then, it has been used at its Tax Bureau every year to calculate their fixed asset tax. After the success in Tokyo, other major city governments including Osaka and Sapporo have followed suit.

Problem Description

The Japanese fixed property tax is imposed by municipalities on the owners of land, buildings and depreciation assets (all hereinafter referred to as “fixed assets”) on January 1 of every year by calculating the tax sum according to current asset values. For this purpose, the municipalities take aerial photographs every year on January 1 and compare the photographs with those of the previous year to identify building change information (new construction, loss, enlargement, reform, reconstruction, work-in-progress, etc). The identification of changes is usually entrusted to survey companies who hire a large number of workers (Figure 1). However, reliance on human labor has led to problems as described below.

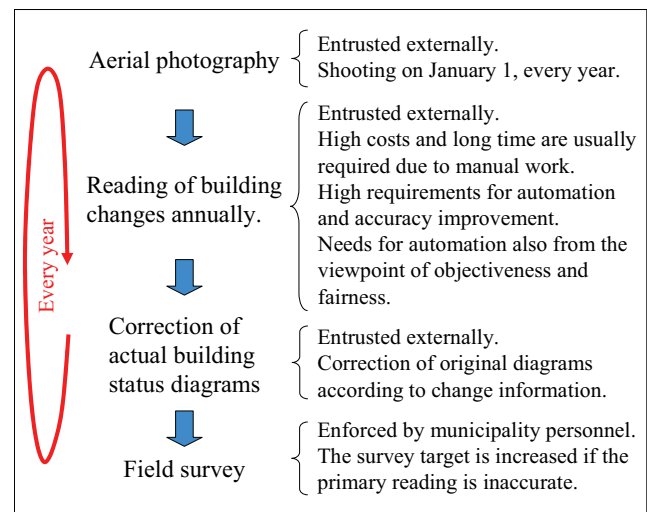


Figure 1: Flow of building change detection in local municipalities.

Huge Costs: It takes about ten hours to read and interpret a single photograph, and the average municipality must perform this work for several hundreds of photographs.

Impossibility of Eliminating Human Judgment Errors: Errors are not acceptable from the viewpoint of fair taxation, in particular, oversight of buildings subjected to actual change. Nevertheless, the current work done with the traditional system is dependent on the capabilities of individuals, so errors are unavoidable.

In visual identification work, attempts to prevent oversight errors are made by performing several read operations per area, but this leads to a further increase in cost. Every photograph is taken over a scale that can cover an actual area of $800\text{m} \times 600\text{m}$ or $500\text{m} \times 600\text{m}$ (variable depending on the municipality), and every municipality has several hundreds of photographs to be read. As a result, it is not rare that the man-hours required for the photograph reading operation exceed 10,000.

Under these circumstances, the incentives among the municipalities to solve such problems by automating or systematizing the photograph-reading work are now higher than ever. The criteria for the identification of changes are based on guidelines given for the Research Center of Property Assessment System. Specifically, the criteria are required to detect the following types of changes without exception:

- Horizontal or vertical changes of 2 meters or more.
- Color changes in an area of about 2×2 meters.

As the above requirements set a high hurdle for achieving automatic processing, the attempts to automate have been limited to a partial use of certain tools and were far from a real systematization. The main judgment issues are the detection of height and color changes, the automation of which is accompanied by the following problems.

Height Change Detection: The height information may be obtained by means of aerial surveys by using a laser

profiler, but this method has high cost if required to detect height changes of around two meters. Moreover, aerial survey devices are too expensive and their availability is limited for many municipalities.

Color Change Detection: Precise overlapping of previous and new photographs is required for detecting color changes. However, two aerial photographs are difficult to overlap and match satisfactorily because the angles of the buildings are usually different due to variance of the shooting conditions (shooting position, altitude, focal distance, etc.)

We have solved the above two problems by adopting the method of pixel-by-pixel stereo processing, which makes it possible to obtain the height information of all pixels in aerial photographs and correct their positional information.

Application Description

The major component of the FACJ system is RealScape. Because RealScape is a general-purpose software package, the FACJ system is given some minor customization for the use of the tax bureau, such as data conversion and integration with other applications. In this paper, we focus on RealScape, which consists of two modules (Figure 2).

Stereo Processing Module: This module inputs two aerial photographs into the computer, converts all its pixels into 3D information and calculates the building height information (DSM: Digital Surface Model) within an accuracy of one meter. At the same time, it applies true orthorectification processing in order to correct for the inclinations of buildings in the photographs and to enable a precise overlapping of the previous and new photographs.

Change Detection Module: This module inputs the two, previous and new, orthorectification aerial photographs and the DSM information, and detects changes in the shapes and colors of buildings and land.

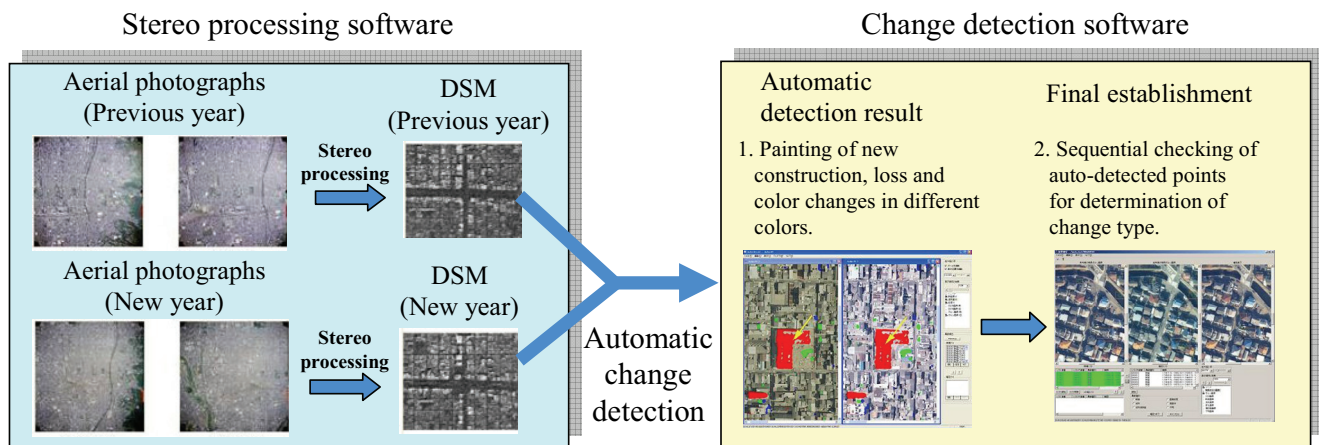


Figure 2: RealScape system configuration.

The target we set before designing this system was to improve both the performance and accuracy (prevention of human errors and detection of height changes that are difficult to detect visually) of making judgments by automation. Even with the current technology level, it is difficult to achieve a perfectly error-free reading and judgment automatically. Therefore, we decided to apply judgment using automated software in the primary reading and to provide the final reading only with the tools for accurate judgment. According to actual data, the number of buildings shot per photograph is between 3,000 and 4,000, and the annual percentage of buildings subjected to changes is 3% to 5%. Therefore, the average number of changed buildings in a photograph is around 150. As the aim of primary reading in the traditional visual judgment work has been to select 500 to 600 candidate buildings among 3,000 buildings in a photograph, we set the target performance of our automatic judgment system, which is to be used in the primary reading, at a judgment capability level similar to that of the traditional primary reading level.

Stereo Processing of Aerial Photographs

AI technology is used in the stereo processing module. Specifically, various image processing techniques such as customized DP matching are adopted. Figure 3 shows the outline of the stereo processing module.

Input Data: The input images can be analog or digital aerial photographs similar to those used in the traditional change judgment system. Each of these photographs consists of a series of picture frames taken by overlapping

around a 60% area between the frames. This system executes stereo processing by assigning two adjacent picture frames as the left and right images and obtains the height information on all pixels in the overlapped area.

Relative Orientation: Unlike photographs in which images are taken using a stereo camera that can be calibrated every time before use, aerial photographs are taken using various camera attitudes. This procedure makes it necessary to perform image rectification, which is a method for parallelizing a pair of images. For this purpose, the system rectifies input images by using the external orientation parameter. Then, orientations of the input images are aligned.

Pixel-by-pixel Stereo Matching: The main task in stereo matching is to identify the corresponding points in the left and right images. When the left and right images are paralleled with relative orientation, all of the subjects are located on the same scanning line number in the left and right images. Therefore, the search for corresponding points can be limited in a single dimension. We adopted the DP (Dynamic Programming) matching method for the search that uses the cross-correlation as the evaluation value and outputs the DSM after the processing. In order to obtain high-quality DSM, it is necessary to select the cross-correlation parameters, such as the window sizes and threshold values, optimally according to the scales and types of input images. As their selection necessitates experience, we provided the system with parameter settings that are optimized according to the types of processed images, so that the user can perform optimum processing simply by selecting one of the parameter

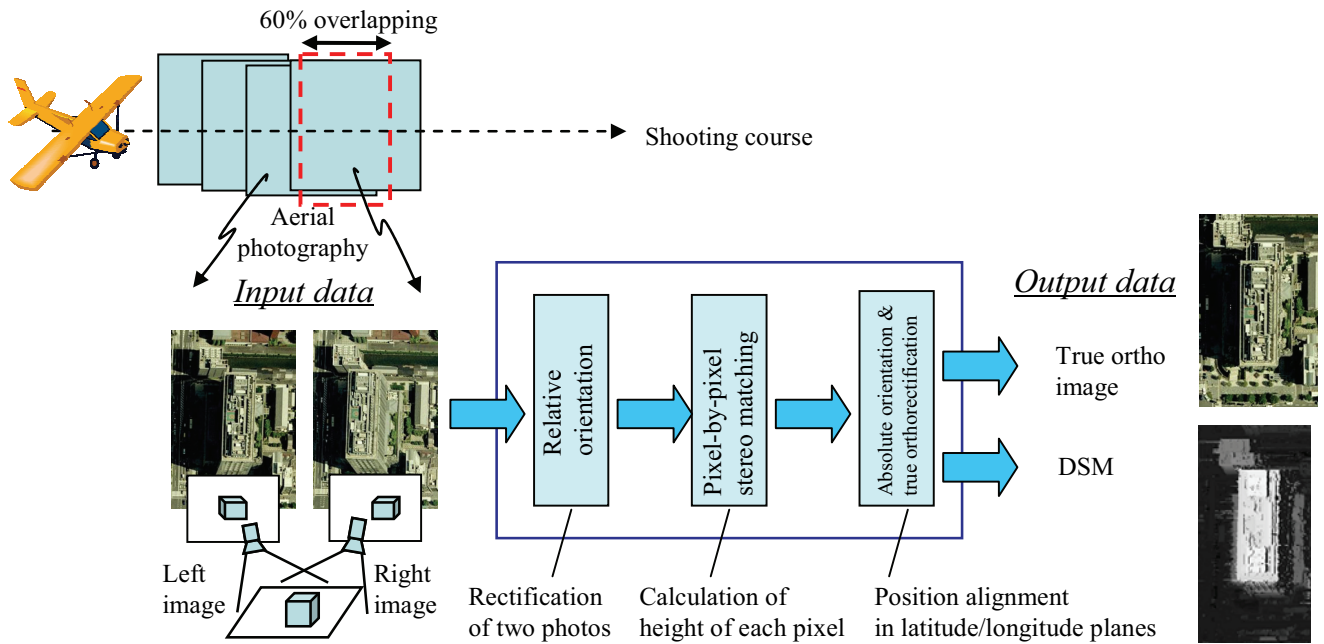


Figure 3: Outline of the stereo processing module.

settings.

Absolute Orientation: Since processing for the above is performed in the image coordinate space, it eventually becomes necessary to compare the matched correspondence of the photograph and the DSM with the latitude and longitude of the land location in this procedure. At the same time, a conversion of the parallax values obtained by the stereo matching of the altitude value is also performed.

True Orthorectification: Since stereo processing requires a large amount of calculations, traditional aerial survey software generally obtains the altitude information only for the characteristic points and the contour lines of topography and buildings, and applies interpolation to other points. This has resulted in problems such as dealing with buildings that lack contours and are undistinguishable from the ground and consequently remain inclined in/with the images. On the other hand, RealScape can determine the absolute positions of all pixels because the stereo processing applied by it offers the height information of all pixels without a need for contour information. We call this processing method the “True Orthorectification” method. True ortho images show the roof surfaces of all buildings in their real positions without tilting (Figure 4). This makes it possible to overlap a photograph precisely onto a map or to overlap two photographs taken under different shooting conditions. Figure 5 shows the 3D view example of downtown Tokyo generated by RealScape without any manual editing.

Hardware Configuration: RealScape runs on multiple PCs (Figure 6). It includes a server-and-client system for concurrent execution of stereo processing using a grid computing architecture, and software for use in a series of operations. The system can perform work for generating DSM from aerial photographs. The clients are composed of 1 to 48 PCs and can join or leave the stereo processing dynamically, enabling a flexible system configuration.

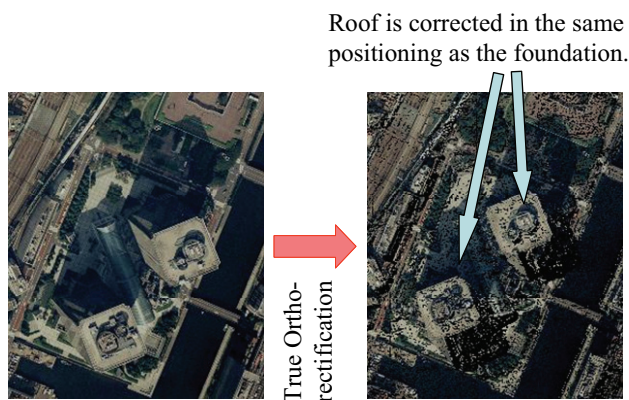


Figure 4: True orthorectification.

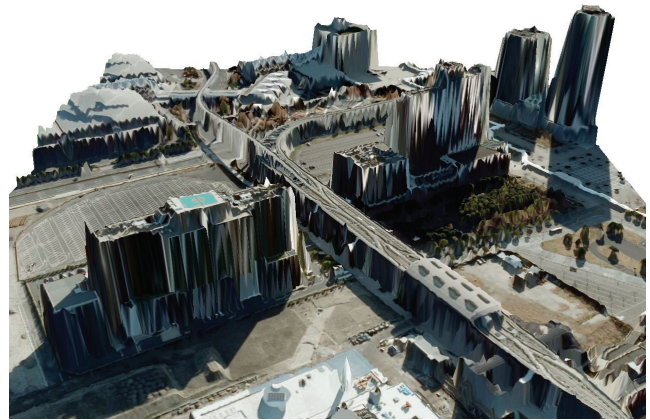


Figure 5: 3D view example by RealScape.

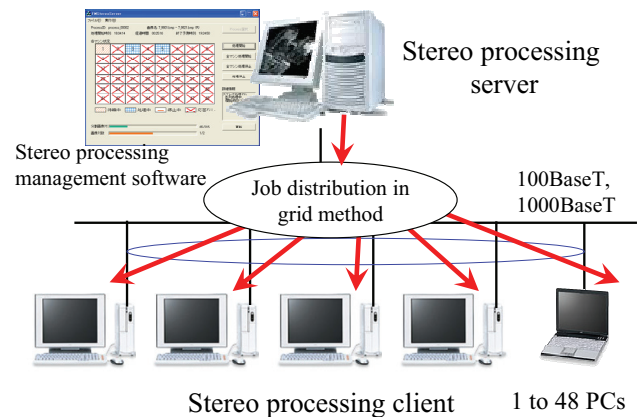


Figure 6: RealScape hardware configuration.

Change Detection of Buildings

The change-detection software inputs both previous and new color images and DSMs that are obtained as a result of stereo processing, and detects any changes in them. This software is composed of an automatic detection tool for the primary reading operation and the final establishment tool for supporting the final judgments by experts (Figure 7).

Automatic Detection Tool: The automatic detection tool was designed to specifically analyze and detect changes in buildings such as new construction, loss or color changes. For the altitude changes, this software refers to both previous and new DSM and then detects increase in height as new construction and decrease in height as loss. For the color changes, it refers to the true ortho color images and detects changes above threshold values as color changes. Since it is difficult for users to specify the threshold value for color change, the system at first presents the threshold calculated automatically in consideration of the illumination difference between two images. Here, all the

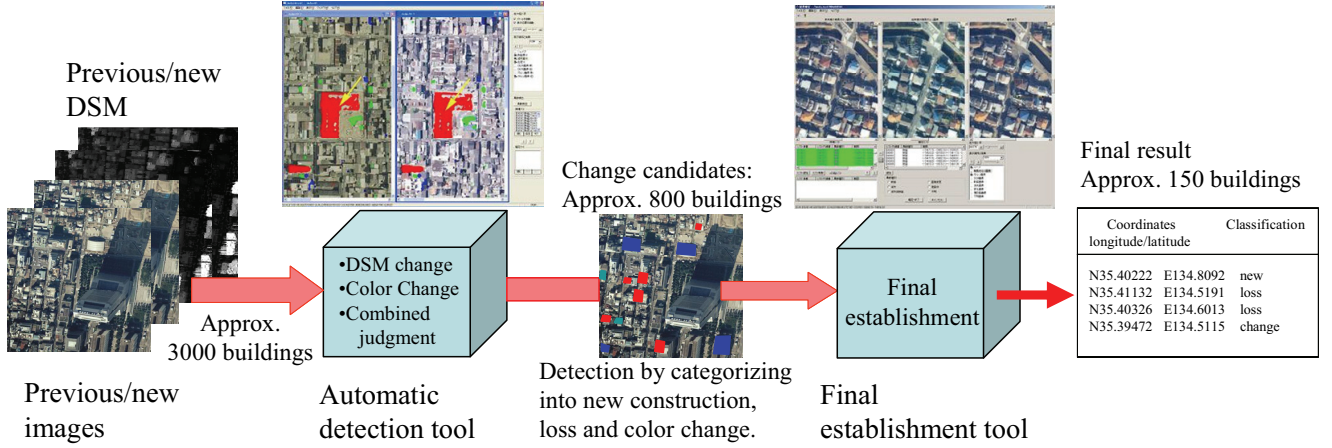


Figure 7: Outline of the change detection module.

areas that were detected with color changes cannot be interpreted as actual changes because objects of certain size such as automobiles could have been falsely detected as areas with color changes. Therefore, our system also refers to the DSM in order to eliminate color changes in areas with a height of below two meters from the ground. The results obtained with this tool are output as a map image showing new construction, losses and color changes by particular painted pixels, and the center coordinates (latitude and longitude) of each change area detected as a mass by means of labeling.

Final Establishment Tool: The final establishment tool is designed to assist experts performing the final judgments. The colored map and center coordinate list obtained from the automatic detection tool are taken as input for this tool. While the results of automatic detection indicate three types of changes including new construction, loss and color change, the final establishment tool is also capable of more detailed categorization of change types (new construction after loss, works in progress, etc.) according to the requirements of municipalities.

Performance Evaluations

Accuracy of DSM (Analog photographs)

We evaluated horizontal accuracy and vertical accuracy of DSM created from analog aerial photographs at 40 points in the mountains, hills and urban areas by comparing DSM with field surveys. Table 1 shows specifications of aerial photographs and field surveys. Table 2 shows the result: 119cm horizontal accuracy and 49cm vertical accuracy. This satisfies the requirement defined in the public survey regulations.

Table 1: Specifications of aerial photographs and field surveys

Aerial photographs	Scale	1/10,000
	Spatial resolution	21cm
	Overlap rate	OL: 60%, SL: 30%
	Focal length	152.96mm
	Scan resolution	1,200dpi
Field surveys	Survey method	Network-based RTK-GPS
	Number of points	40 points

Table 2: Accuracy of DSM created from analog aerial photographs

	Horizontal accuracy	Vertical accuracy
DSM	119cm	49cm
Public survey regulation	Within 175cm	Within 67cm

Accuracy of DSM (Digital Photographs)

We evaluated the vertical accuracy of DSM created from digital aerial photographs by comparing DSM with field surveys and laser profiler data. Table 3 shows specifications and Table 4 shows the result, which were 15.9cm vertical accuracy for DSM and 12.4cm vertical accuracy for laser profiler data. DSM was nearly equal to laser profiler data in vertical accuracy. On the other hand, the resolution of laser profiler data was only about 2m, but that of DSM was 7.2cm, which was equal to the resolution of an aerial photograph. Hence, DSM created by RealScape had much higher resolution than laser profiler data.

Table 3: Specification of aerial photograph, field survey and laser profiler data

Aerial photographs	Camera	UltraCamD
	Scale	1/8,000
	Spatial resolution	7.2cm
	Overlap rate	OL: 60%, SL: 30%

	Focal length	105.2mm
Field surveys	Survey method	Mobile receiver by FKP
	Number of points	100 points
Laser profiler	Spatial resolution	About 2m

Table 4: Accuracy of DSM created from digital aerial photographs

	Vertical accuracy	Resolution
DSM	15.9cm	7.2cm
Laser profiler data	12.4cm	About 2m

Accuracy of Change Detection

In the 660m x 1200m flat land area, which included 1,636 buildings in a certain city, we evaluated change-detection accuracy of RealScape by comparing the result of RealScape with that of human visual identification. Table 5 shows specifications of the photographs.

Table 5: Specifications of photographs

Year	2006	2004
Camera	UCD (digital)	RC30 (analog)
Scale	1/8,000	1/8,000
Scanning	-----	21 μ m
Spatial resolution	7.2cm	16.8cm
Overlap rate	80%	60%

Table 6: Evaluation result of change detection

	Human visual identification	RealScape
Number of buildings	1636	1636
New construction	66	73
Loss	35	41
New construction after loss	42	42
Color change	33	40

Table 6 shows the results. Compared with visual identification, in detection of new construction, loss and new construction after loss, RealScape detected all changes detected by visual identification. Furthermore, RealScape additionally detected seven new construction changes and six loss changes not detected by the human visual identification. Figure 8 shows the example of change that RealScape detected as loss and that the human visual identification failed to detect. In the human visual identification, the oversight of a building occurred if the color of a roof in one photograph was very similar to that of the ground color in another photograph. Even in such a case, RealScape correctly judged the change as loss by their height difference.

Figure 9 shows an example of change that RealScape detected as a newly constructed building but the human visual identification judged it incorrectly as none. It seemed in the visual identification that the area was judged as a parking space because the color of the roof in a photograph was similar to that of asphalt. RealScape

judged it correctly as new construction. As for color change detection, RealScape detected ten color-changed buildings not detected by the human visual identification and detected 30 color-changed buildings out of 33 color-changed buildings detected by the human visual identification. The three additional changes detected by the human visual identification were mistakes because the humans concluded that the heights were different when they were the same.

This evaluation reveals that change detection by RealScape has almost equal accuracy to that of visual identification. Furthermore, RealScape is able to detect height change, which is difficult to judge by human visual identification.

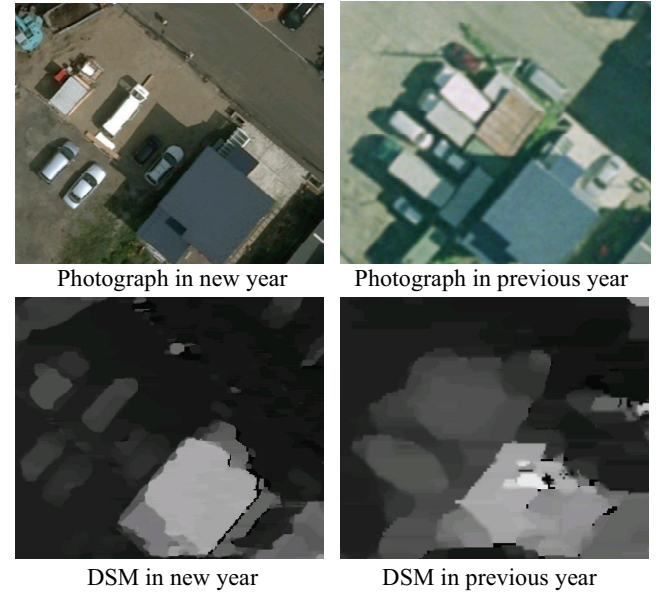


Figure 8: Example of loss. RealScape detected it as loss but visual identification missed it.

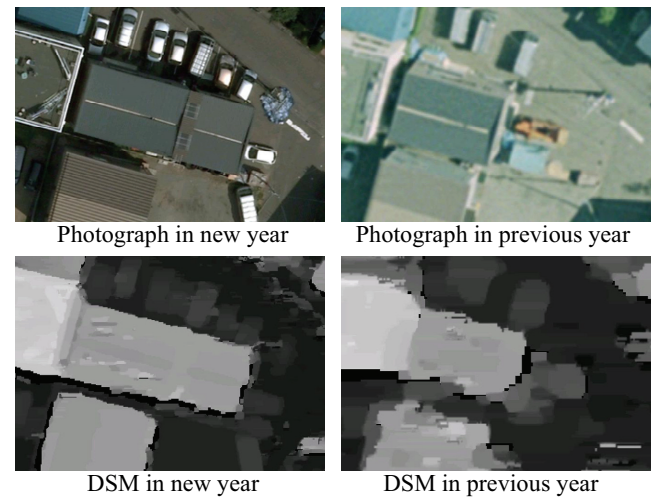


Figure 9: Example of new construction. RealScape detected it as new construction but visual identification judged it as loss.

Application Development, Use and Pay off

It took two years for four of us to develop RealScape since 2003. We developed the prototype in the first year and the software package in the second year. In fiscal 2005, we experimentally deployed the fixed assets change detection work with the support of the Tokyo Metropolitan Government. We created the full-scale DSM of 2,800 photographs (on-ground resolution 12.5cm) of the Tokyo Metropolitan area, and obtained confirmation from the Japan Association of Surveyors that the results were within the standard error tolerance of 50cm horizontal accuracy and 100cm vertical accuracy.

For detecting the change, we conducted comparisons for 10 photographed districts in the Shinagawa District, which is the central part of downtown Tokyo, and confirmed it by comparing the results with separately conducted visual judgment operations and found that the following changes were detected without exception.

- Height changes: Horizontal or vertical change by 2 meters or more, in areas of 1m × 2m or wider.
- Color changes: In areas of 2 m× 2m or wider.

As it is estimated that the time taken for the manual final establishment operation is 2 to 3 hours per pair of images, our system can offer an operation time reduction effect of about 60% compared to the time taken for the traditional judgment process by one person as shown in Figure 10. The time-saving effect may be even larger if our system is compared to a multi-operator system.

Maintenance

The FACJ systems are running internally in different survey companies that use the RealScape software package. Since RealScape is a self-contained software package, many traditional maintenance tasks, such as data backup, are provided in self-service mode. For the stereo processing part, no real maintenance is needed since the rules and algorithms are static. However, since the time for

the stereo processing is still large, we are striving to improve its algorithm to shorten the time.

Comparison with Previous Works

Change detection of buildings is an important task and is addressed in many studies. Lillestrand compared pixel intensity of images taken at different times (Lillestrand, 1972). Comparing intensity values is not very effective because such changes don't necessarily reflect actual changes in shape, but could be caused by changes in viewing, in illumination conditions and in seasonal variations. Carlotto first predicted pixel intensities by extrapolating the changes that were seen in a sequence of satellite images taken from the same position (Carlotto, 1997). Pixels whose intensities do not match the predicted value are judged as changes. However, this method does not distinguish among building changes, parallax, and shadow. Huertas and Nevatia created 3-D wireframes of buildings from previous stereo image pairs, and then projected the 3-D wireframes onto the current image (Huertas and Nevatia, 1998). Building changes were detected by analyzing the overlap of the 3-D wireframes and the edges of current image objects.

The ATOMI project (Eidenbenz et al., 2000, Niederost, 2000) aims at detecting changes, and enhancing planimetric accuracy from the 5 meter level to the 1 meter level based on colored aerial photos, a high resolution digital elevation model (DEM) and a photogrammetrically generated digital surface model (DSM). The ATOMI update work uses the DSM as the primary data source, and uses the image information primarily in order to discern man made objects from natural objects. Vogtle's approach uses airborne laser scanning data (Vogtle et al., 2004). Changes are detected by comparison of DSMs, which were acquired at two different times. To eliminate non-building objects, segmentation procedure is used before the comparison. Though it was necessary to get 50 cm horizontal and 1 meter vertical accuracy for our specific application, these above approaches including laser scanning methods had not achieved the required accuracies.

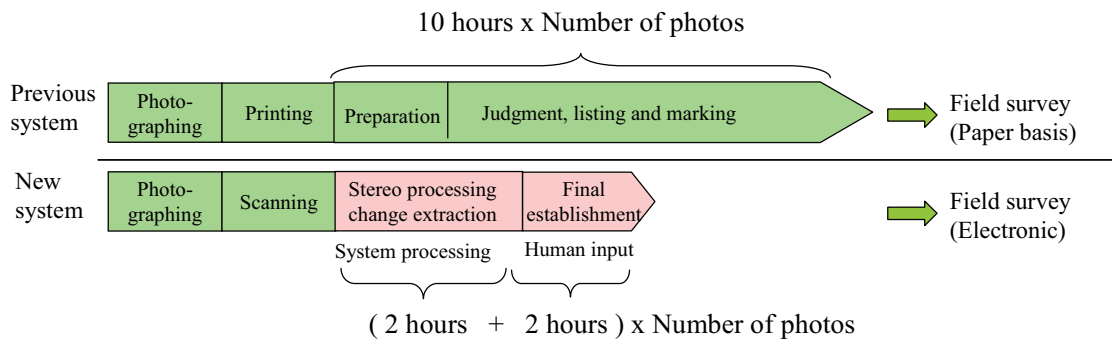


Figure 10: Time reduction effect of automation.

Conclusion

This paper described the Fixed Assets Change Judgment (FACJ) system and its core tool, RealScape. Recently, Tokyo terminated its traditional visual identification work that has been in use for 20 years and shifted to a new automated system. Because this systemization had an impact on both the municipality community and the survey industries, after Tokyo's success, Osaka (Japan's third largest city with a population of 2.6 million), Sapporo (Japan's fifth largest city with a population of 1.8 million) and other city governments followed suit. The challenge of this application was to achieve precise height calculation in aerial photographs. For this, we developed a unique pixel-by-pixel height calculation algorithm. Because it requires huge computation, AI technologies were adopted, including customized DP matching to optimize the stereo processing algorithm. The algorithm also required a grid computing architecture. The precision RealScape achieved is now approved by the Japan Association of Surveyors.

References

- Lillestrand, R. (1972) Techniques for Change Detection. *IEEE Transactions on Computers*, No.7, 654-659.
- Carlotto, M. (1997). Detection and analysis of change in remotely sensed imagery with application to wide area surveillance. *IEEE Transaction on Image Processing*, Vol.6, No.1, 189-202.
- Huertas, A. and Nevatia, R. (1998). Detecting changes in aerial views of man-made structures. *In International Conference on Computer Vision, IEEE*, 73-80.
- Eidenbenz, C., Kaeser, C. and Baltsavias, E. (2000). ATOMI-automated reconstruction of topographic objects from aerial images using vectorized map information. *International Archives of Photogrammetric and Remote Sensing XXIII*, 8.
- Niederost, M. (2001). Reliable reconstruction of buildings for digital map revision. *International Archives of Photogrammetric and Remote Sensing XXIII*, 8.
- Vogtle, T. and Steinem, E. (2004). Detection and recognition of changes in building geometry derived from multi temporal laser scanning data. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol.34, Part XXX.
- Ohta, Y and Yamada, H. (1989). Pattern matching using dynamic programming. In *Jour. IPSJ*, 1058-1066.