CHANGE DETECTION OF TIME-SERIES 3D POINT CLOUDS
USING ROBUST PRINCIPAL COMPONENT ANALYSIS

T. Fuse 1, *, T. Yamano 1

1 Dept. of Civil Engineering, University of Tokyo, Hongo 7-3-1, Bunkyo-ku, Tokyo, 113-8656, Japan - fuse@civil.t.u-tokyo.ac.jp

Commission II, WG II/3

KEY WORDS: Change Detection, Time-Series, Robust PCA, DEM, Point Clouds, Depth Image

ABSTRACT:

The chances of acquiring three-dimensional (3D) point clouds have recently increased with the emergence of laser scanners. Hence, 3D monitoring of various objects through the accumulation of “time-series 3D point clouds,” which are point clouds of the same place at different times, is possible. Change detection is a task that is indispensable in 3D monitoring. One of the most common change detection method of 3D point clouds is simple subtraction between two data. However, this method is vulnerable to various errors. Therefore, change detection methods that are robust to errors are required. In this study, we developed robust principal component analysis, which has become popular in the background modeling of video images, to robustly recognize changes in time-series 3D point clouds. We first applied the proposed method to time-series depth images and confirmed its accuracy. We then applied the method to the digital elevation models of Mt. Unzen, which were acquired between 2003 and 2016, to recognize yearly elevation changes. The results show that the proposed method robustly recognizes elevation changes with a properly set parameter.

1. INTRODUCTION

The chances of acquiring three-dimensional (3D) point clouds have recently increased with the emergence of laser scanners. Hence, 3D monitoring of various objects through the accumulation of “time-series 3D point clouds,” which are point clouds of the same place at different times, is possible. Change detection is a task that is indispensable in 3D monitoring. One of the most common change detection method of 3D point clouds is simple subtraction between two data. However, this method is vulnerable to various errors. Errors are caused by the colour and material of objects, IMU and GNSS on platforms, time resolution of laser scanners, and overlaying of multiple data. Therefore, change detection methods that are robust to errors are required.

Many change detection methods and moving-object detection methods have been proposed for video images, such as frame subtraction, background subtraction, and optical flow detection. However, frame subtraction and background subtraction are at risk of errors. Furthermore, making a background model for background subtraction is sometimes challenging. Optical flow detection is also unsuitable in time-series 3D point clouds because it requires a high time resolution of data, which is difficult to acquire in time-series 3D point clouds.

Robust principal component analysis (RPCA), which is a method that can robustly detect moving objects in video images, can be applied to time-series 3D point clouds. RPCA is a method that separates input video images into a background and moving objects (Candes et al., 2011). In this method, a frame of video is converted into a vector. By converting all frames of video into vectors and merging them horizontally, one video is represented by a matrix. The input matrix is separated into two matrices: a matrix representing a background and a matrix representing moving objects. We minimized the rank of a matrix representing a background because a matrix representing moving objects is sparse. The accuracy of this method increases as the number of input frames increases. As more 3D point clouds accumulate, we can develop a robust change detection method through the application of RPCA to 3D point clouds. So far RPCA have not been applied to 3D data.

Once the robust change detection method enables 3D monitoring, it can be applied to various fields, such as data acquisition for the efficient and effective maintenance of infrastructure, efficient update of 3D street maps, and observation of volcanic activities and debris flows by recognizing elevation changes. In this study, we applied the method to depth images and digital elevation models (DEMs) and examined its accuracy.

The rest of this section describes related works on the change detection of 3D point clouds. There are some types of change detection methods of 3D point clouds, which depends on whether input data are two-dimensional (2D) depth image, 2D mesh or 3D data. The first type of change detection methods is the subtraction method, which is most applicable if input data are 2D depth image / mesh (DEMs) representing 3D information. In this method, depth image / mesh values in one data are subtracted from those acquired at different times. Change is recognized if subtracted values exceed the threshold. The subtraction method is often applied to DEMs for monitoring topographical changes, such as changes in channels (Carley et al., 2012) and landslides (Tsutsumi et al., 2007). Input data are not limited to DEMs. Murakami et al. (1999) applied the method to digital surface model (DSM) in urban areas to recognize shape changes in buildings. However, the subtraction method is vulnerable to measurement errors.

The second type of change detection methods is the use of 3D data as input. This type can be sub-classified into two common
methods. The first common method is to register point clouds at different times and calculate the distance between two point clouds. The distance used here is usually the Hausdorff distance (Girardeau-Montaut et al., 2005; Wawrzyniec et al., 2007) or a normal distance (Kromer et al., 2015), which is the distance between point clouds and a normal vector of the polygon made by point clouds of one data. However, this method is also vulnerable to measurement errors. The second common method for change detection is clustering (Palma et al., 2016; Aijazi et al., 2013). This method makes clusters from point clouds using distance and intensity information. By comparing the positions of corresponding clusters at different times, change is detected. However, this method is also vulnerable to errors, and the applicable situations are limited where clustering is possible.

Considering the above summary of related works, the challenges of this study are summarized as follows:

(a) Developing a method that can be applied to various situations
(b) Developing a change recognition method for point clouds that is robust to errors

In this study, a change detection method of time-series 3D point clouds based on RPCA is proposed against the above challenges.

2. DEVELOPMENT OF CHANGE DETECTION METHOD

2.1 Basic Concept

We first converted input 2D data representing 3D information into a matrix. A frame of data was converted into a vector. The values of element of the vector is distance information. By converting all frames into vectors and merging them horizontally, input data are represented by a matrix. Then, we separated the matrix into two matrices, each representing a background and changing objects.

When 3D point cloud data is used as is, voxels are used for matrix representation. The state in which a point exists in a voxel is defined as occupied state, and the state in which a point does not exist is defined as unoccupied state. The voxel value of the occupied voxel is 1, and the voxel value of the unoccupied voxel is 0. In this way, 3D point cloud data can be represented by describing the occupancy/non-occupancy of voxels. A column vector is created by arranging the voxel values at each time in the column direction, and a matrix is created by arranging the time series data in the row direction.

There may be a situation where there are a large number of unoccupied voxels in the target space. In that case, the matrix that is created will have most of the elements zero. In other words, the matrices cannot be separated because they are sparse and low-rank matrices. The proposed method acquires and stores the location information of voxels that remain unoccupied and unchanged during the target time, and excludes them from the calculation. After the calculation, the spatial data is reconstructed using the location information of the voxels that do not change their occupied state while remaining unoccupied at all times.

Assuming that the area of changing objects is small in the input data, a matrix that represents the changing objects becomes sparse. Furthermore, because the value in each pixel / mesh hardly changes between frames in a background, the columns of a matrix representing a background are similar to each other, i.e., a matrix representing a background will have a low rank. The input matrix can be separated into two matrices by minimizing the rank of a matrix representing a background on a condition that a matrix representing changing objects is sparse. When separating the input matrix, we represent the rank of a matrix by the nuclear norm (the sum of singular values) and the sparsity of a matrix by the L1 norm (the sum of absolute values of all elements of a matrix). Then, we solve Equation (1).

\[
\begin{align*}
\min_{A,E} & \|A\|_* + \lambda \|E\|_1, \quad \text{s.t.} \quad D = A + E,
\end{align*}
\]

where

- \(D\) is the input matrix
- \(A\) is the background matrix
- \(E\) is the changing objects matrix
- \(\lambda\) is the penalty parameter
- \(\|\cdot\|_*\) is the nuclear norm
- \(\|\cdot\|_1\) is the L1 norm

2.2 Optimization Algorithm

Equation (1) can be solved by some algorithms, such as iterative thresholding approach, accelerated proximal gradient approach, dual approach. We adopt inexact augmented Lagrange multiplier method (IALMM), which showed a low computational cost in a previous study (Lin et al., 2013). The algorithm is an extension of augmented Lagrange multiplier method using the iterative thresholding method. First, the input matrix is normalized to \(Y_0\) with the Equation (2).

\[
Y_0 = D / J(D), \quad J(D) = \max \left( \|D\|_2, \lambda^{-1} \|D\|_1 \right),
\]

where

- \(\|\cdot\|_2\) is the Frobenius norm
- \(\|\cdot\|_\infty\) is the infinity norm (the maximum absolute value of the matrix entries)

Next, we apply the augmented Lagrangian to Equation (1). The extended Lagrangian and the update rule are as follows:

\[
L(A, E, Y, \mu) = \|A\|_* + \lambda \|E\|_1 + \langle Y, D - A - E \rangle + \frac{\mu}{2} \|D - A - E - Y - Y_0\|_2^2,
\]

where

- \(L\) is the augmented Lagrange function
- \(\mu\) is the positive scalar as weight
- \(\langle A, B \rangle\) is the product of matrices \(A\) and \(B\)
- \(\|\cdot\|_F\) is the Frobenius norm.

The optimization problem of \(A\), Equation (5), is then solved.

\[
A_{k+1} = \arg \min_A L(A, E, Y, \mu)
\]

In this case, the singular values are thresholded using the iterative thresholding method. The method reduces the dimensionality of matrix \(A\) by thresholding the singular values, resulting in a lower rank. Accordingly, Equations (6) and (7) are to be solved.
\[
(U, S, V) = \text{svd} \left( D - E + \mu^{-1}Y_s \right); \quad (6)
\]

\[
A_{k+1} = US_k^\top S_k V^\top
\]

\[
S[x] = \begin{cases} 
  x - \varepsilon & \text{if } x > \varepsilon \\
  x + \varepsilon & \text{if } x < -\varepsilon \\
  0 & \text{otherwise} 
\end{cases}
\]

where \( U, V = \) matrix by arranging singular value vectors for diagonal components \( S = \) Matrix with singular values \( \mu = \text{the singular value decomposition} \)

\[ \varepsilon = \text{positive threshold.} \]

After the optimization of \( A \), we solve Equation (9) to optimize \( E \) shown in equation (8).

\[
E = \arg \min_E L (E, A, Y, \mu); \quad (8)
\]

\[
E_{k+1} = S_{\mu_k} \left[ D - A_{k+1} + \mu_k^{-1}Y_s \right] \quad (9)
\]

The above calculation is repeated until the value of \( D-A-E \) converges below a certain value. We set the initial values and parameters as follows: \( \varepsilon_0 = 0, \mu_0 = 10^{-3}, \varepsilon = 10^{-8} \). The setting is according to Lin et al. (2013), and confirmed that there is no major sensitivity.

### 2.3 Setting Parameter and Dividing Input Data

The \( \lambda \) in Equation (1) is an important parameter that has an impact on the results of change detection. The \( \lambda \) is set to Equation (10) in this study according to Candès et al. (2011) to effectively detect the change.

\[
\lambda = \left\{ \max \left( n_1, n_2 \right) \right\}^{-1/2}, \quad (10)
\]

where \( n_1 = \) the number of pixels/meshes in one data \( n_2 = \) the number of frames in the input data

Only a few sets of point clouds acquired at different times may be available. Hence, a pre-processing of dividing the area of input data is required because an extremely high or low ratio of the number of columns to rows of the input matrix leads to a poor performance of the method. Considering the results of a previous study (Candès et al., 2011) and various experiments in this study, we assume that the ratio of the number of columns to rows of the input matrix should be more than 1/100. Therefore, we divide the area of input data into several pieces so that each piece meets the condition.

### 3. EXPERIMENTAL RESULTS

#### 3.1 Application to Depth Images

We applied the proposed method to depth images (128 x 126 pixels), which were taken by a time-of-flight distance camera (DISTANZA P-401D by Brainvision Inc., Table 1). A total of 14 background depth images of the target space were taken without any change in the space (Figure 1(a)). Then, one changed depth image was taken with a cubic object in the target space (Figure 1(b)).

We compared the accuracy of the change detection of the proposed method and subtraction method. In the proposed method, we divided the input images into small patches of 32 x 42. After the calculation, we applied a threshold of 100 mm to eliminate change that is below the threshold based on the accuracy of the data. In the subtraction method, we chose a background depth image and subtracted each pixel value from that of the changed depth image. We performed the same procedure to all 14 background depth images, applied the same threshold as the proposed method, and calculated the average of the index in the confusion matrix.

<table>
<thead>
<tr>
<th>Number of Pixels</th>
<th>128 x 126</th>
</tr>
</thead>
<tbody>
<tr>
<td>External Messy Light Removal Ratio</td>
<td>Approx. 20:1</td>
</tr>
<tr>
<td>Maximum Frame Rate</td>
<td>40 fps</td>
</tr>
<tr>
<td>Use of Environmental Illumination</td>
<td>200,000 lx</td>
</tr>
<tr>
<td>Angle of View (Horizontal/Vertical)</td>
<td>27/27, 0.5 m - 15 m</td>
</tr>
<tr>
<td>Measure Distance</td>
<td>52/52, 0.5 m - 5 m</td>
</tr>
<tr>
<td>92/92, 0.5 m - 3 m</td>
<td></td>
</tr>
<tr>
<td>LED Emission Wavelength</td>
<td>850 nm</td>
</tr>
<tr>
<td>LED Emission Frequency</td>
<td>10 MHz</td>
</tr>
<tr>
<td>Output Image</td>
<td>Depth Image, Modulated Light Intensity Image, Monochrome Brightness Image</td>
</tr>
</tbody>
</table>

Table 1. Distance camera specifications.

(a) background depth image

(b) changed depth image

Figure 1. Input depth images.
The results of the proposed method are summarized in Table 2 and shown in Figure 2. Figure 2 (a) and (b) show the generated background and the detected change by the proposed method. The findings show that the method successfully detected the change.

<table>
<thead>
<tr>
<th>Proposed Method</th>
<th>Detected Class</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual Class</td>
<td>Change</td>
<td>Unchange</td>
</tr>
<tr>
<td>Change</td>
<td>0.5 %</td>
<td>0.0 %</td>
</tr>
<tr>
<td>Unchange</td>
<td>24.2 %</td>
<td>75.3 %</td>
</tr>
</tbody>
</table>

Table 2. Results of the proposed method (depth image).

Figure 2. Results of the proposed method (depth image).

The results of the subtraction method are summarized in Table 3. The results is lower than that of the proposed method.

<table>
<thead>
<tr>
<th>Proposed Method</th>
<th>Detected Class</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual Class</td>
<td>Change</td>
<td>Unchange</td>
</tr>
<tr>
<td>Change</td>
<td>0.5 %</td>
<td>0.0 %</td>
</tr>
<tr>
<td>Unchange</td>
<td>35.8 %</td>
<td>63.7 %</td>
</tr>
</tbody>
</table>

Table 3. Results of the subtraction (depth image).

3.2 Application to DEM

We applied the proposed method to the time-series DEMs of Mt. Unzen in Nagasaki, Japan. The DEMs were derived by converting point clouds acquired by aerial laser scanners into 1m x 1m square meshes. The target area has a rectangular shape area of 5075 m × 3046 m. As time-series DEMs, we prepared nine DEMs that were acquired in different years between 2003 and 2016. The proposed method was applied to examine if elevation changes in the target area can be detected. In the experiment, we evaluated the accuracy of the proposed method by comparing the results with the visual observation of orthophotographs. We also compared the results of the proposed method with those of the subtraction method. In the proposed method, we divided input DEMs into small patches of 108 × 11. After the calculation, we applied a threshold of 1 m to eliminate the change that is below the threshold based on the changes. In the subtraction method, we used the DEMs of 2003 and 2016. The same threshold as the proposed method was applied.

First, we set parameter \( \lambda \), as shown in Equation (10). However, a lot of misdetection occurred. After many trials with different values of parameter \( \lambda \), we modified it as shown in Equation (11) according to the results.

\[
\lambda = 2 \left( \max \left( n_1, n_2 \right) \right)^{1/2} ,
\]

The results of the change detection are shown in Figure 3. The results of three target areas (a), (b), and (c) are shown. The first column shows orthophotographs in 2003, and the second column shows orthophotographs in 2016 (white dashed line rectangles show detailed area of the change). The third and fourth columns show the results of the proposed method and the subtraction method.

As shown in the results of (a), the proposed method successfully detected the change. The rectangular artifacts are seen in the result. It is effect of the patch division. The proposed method detected the shape of the change in (b) and (c), whereas the subtraction method detected measurement errors as changes in (b); the upper right part of (b) is detected as change although it is unlikely that the part changed because we do not see any change in the orthophotographs.

3.3 Application to MMS Data

We applied the proposed method to MMS point cloud. It was acquired using Nikon-Trimble’s MX-5, which is equipped with GNSS and IMU measuring devices, POS LV 520 (Applanix), and one LASER scanner, the Riegl VQ 450 (Table 4). The current vehicles do not yet capture sequential data. Therefore, for the purpose of this study, sequential data are synthetically created from the above data (Figure 4).

<table>
<thead>
<tr>
<th>Riegl VQ 450</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Irradiation</td>
<td>Max. 550,000 /s</td>
</tr>
<tr>
<td>Reaching Distance</td>
<td>Max. 800 m</td>
</tr>
<tr>
<td>Shortest Distance</td>
<td>1.5 m</td>
</tr>
<tr>
<td>Distance Measurement Accuracy</td>
<td>± 0.008 m</td>
</tr>
<tr>
<td>POS LV 520</td>
<td></td>
</tr>
<tr>
<td>Error (RMS)</td>
<td>X, Y: 0.035 m</td>
</tr>
<tr>
<td>Z: 0.05 m</td>
<td></td>
</tr>
</tbody>
</table>

Table 4. MMS specifications.

Two types of changes, constant and temporary, were assigned to the data as changes in the data, which were used as time series data. The constant changes include changes in the shape of buildings and changes in shape, additions, and removals of structures on roads. In this study, four types of constant changes are assigned to the data. Specifically, these are the removal of electric lights, and the addition of signs, electric lights, and guardrails. The temporal changes are changes caused by pedestrians, automobiles, and so on (Figure 5). In the figure, temporal changes are shown in blue.
We applied the proposed method to the data of 10 time points with temporary changes (referred to as “background”) and one time point with constant and temporary changes (referred to as “change”). The length of one side of a voxel was set to 1 m. The number of voxels was 150 × 120 × 95. According to the voxel exclusion in section 2.1, 1,682,549 voxels of the total 1,710,000 voxels were excluded from the analysis because they did not remain unoccupied at all 10 time points. The results are shown in Table 5. Figure 6 shows an example of change detection. In the figure, actual changes and detect changes are shown in red. The line-like artifacts in Figure 6 (b) is effect of voxel exclusion.

<table>
<thead>
<tr>
<th>Proposed Method</th>
<th>Detected Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual Class</td>
<td>Change</td>
</tr>
<tr>
<td>Change</td>
<td>0.4 %</td>
</tr>
<tr>
<td>Unchange</td>
<td>4.5 %</td>
</tr>
</tbody>
</table>

Table 5. Results of the proposed method (MMS data).

Figure 3. Comparison of the results of the proposed method and subtraction method (DEM).

Figure 4. MMS Point Cloud.

Figure 5. Example of Temporal Changes.

Figure 6. Example of Change Detection (MMS data).
4. DISCUSSION

4.1 Application to Depth Images

We compared the proposed method with the subtraction method. The subtraction method is vulnerable to measurement errors in the background depth images and changed depth image. Moreover, the proposed method is robust to measurement errors in the background depth images because a background is generated by reducing the measurement errors of each background image in the proposed method.

Some unchanged pixels were detected as changes, which can be attributed to measurement errors in the changed depth image. The RMSE of the depth images in the experiment was ±91.7 mm. Considering that the measurement errors follow a normal distribution, 27.6% of the unchanged pixels contain errors that exceed the threshold. Because the ratio of unchanged pixels is 99.5% in all pixels, the ratio of pixels that were misrecognized because of measurement errors is 27.5% in all pixels. Hence, most of the misdetection in the proposed method, which is 24.2%, was caused by the measurement errors in the changed depth image. Accordingly, the accuracy can be improved by reducing the measurement errors in the changed depth image.

4.2 Application to DEM

With $\lambda = \left( \max \left( n_1, n_2 \right) \right)^{1/2}$, the proposed method misrecognized many meshes. One reason for this error is the ratio of change amount to mesh values. The change amount in each mesh is about several meters, whereas the value in each mesh is several hundreds of meters or more than a thousand meters. The misdetection is caused by the application of the same parameter as video or depth images, even though the ratio of change amount to mesh values is much smaller in time-series DEMs compared to that in video or depth images.

We compared the proposed method with the subtraction method, as shown in Figure 3(b). We do not visually recognize any difference between the two orthophotographs at the upper-right corner although that area is misrecognized as a change in the subtraction method. In the proposed method, this area is correctly detected as an unchanged area. Moreover, the proposed method misrecognizes the shape of the change in some areas, such as in Figure 3(c). Because the proposed method generates a background from multiple input data, the exact shape of the change is hard to detect sometimes. For example, if the change occurs in the early stage of the observation, the exact shape of the change in the newest data is hard to detect. Therefore, the proposed method needs to be carefully applied to data acquired over a long period. The proposed method is suitable for a situation where robustly locating changing area is a higher priority than detecting the exact shape of the change.

4.3 Application to MMS Data

The change detection by the proposed method showed a high detection accuracy of about 95% for both the “change” and the “background”. On the other hand, there was some misdetection mainly in areas where the point density was decreasing. Specifically, the boundary area and the window glass of a building. One possible reason for this is that the number of points contained in a voxel is small in such regions, so the occupancy of the voxel is likely to change. The misdetection caused by the decrease in point density can be improved by changing the way of the voxel representation. In this study, when a point is contained inside a voxel, the voxel value was set to 1 as the occupied state. However, such a representation method may cause misdetection of voxels with low point density. Accordingly, setting a threshold for determining the occupancy status may reduce false positives. The introduction of voxels of variable size, such as oct-tree representation, may also be effective.

The automatic generation of the background in the proposed method enables us to grasp the shape of the structure. In this study, the length of one side of a voxel is set to 1 m. It is expected that the detailed shape of a structure can be grasped by shortening the length of one side of a voxel. However, care should be taken because reducing the size of the voxel is expected to increase misdetection in areas with low point density.

We also applied the proposed method to various values of the parameter $\lambda$. In this example, we did not find significant differences in the results depending on the parameter values, and confirmed the robustness of the proposed method to the parameter values.

In this study, only relatively small-scale constant changes, such as the addition of road signs and lights, were given. However, in the real road space, large-scale constant changes such as the addition or removal of lanes or the construction or demolition of buildings can occur. Since the proposed method assumes the sparsity of the matrix represented by the change voxels, the scale of change is expected to affect the accuracy of change detection.

5. CONCLUSIONS

In this paper, we proposed a method to detect changes in time-series 3D point clouds for the future object monitoring through laser scanning. The accuracy of the proposed method was shown by its application in depth images, DEMs, and MMS data. We quantitatively showed the robustness of the proposed method against measurement errors by comparing it with the subtraction method. By applying the proposed method to real/synthesis data and examining its accuracy, we showed the future possibility of 3D monitoring via laser scanning. We obtained knowledge on the possible applications of the proposed method. Most of the misdetection errors was caused by recognizing changes in unchanged parts. Such measurement errors are attributed to the changed data, which can be reduced by preprocessing. We also identified some limitations of the proposed method, including the lack of preprocessing and the ratio of the changing amount to pixel/mesh/voxel values. These limitations signify the need to further improve the detection method in the future.

ACKNOWLEDGEMENTS

The DEM used in this study was provided by the Unzen Restoration Work Office (Nagasaki, Japan). The study was conducted with the cooperation of the Asia Air Survey Corporation (Tokyo, Japan). We would like to express our gratitude for their cooperation. We also express our thanks to anonymous reviewers and editors for their constructive and useful comments. This research was funded by the Japan Society for the Promotion of Science (JSPS), Grant-in-Aid for Scientific Research (B) 18H01554.
REFERENCES


