Smart Structures and Systems, *Vol. 14, No. 5 (2014) 943-960*
DOI: http://dx.doi.org/10.12989/sss.2014.14.5.943

Pose-graph optimized displacement estimation for structural displacement monitoring

Donghwa Lee[1], Haemin Jeon[1] and Hyun Myung*[1,2]

[1]*Department of Civil and Environmental Engineering, KAIST, Daejeon 305-701, Republic of Korea*
[2]*Robotics Program, KAIST, Daejeon 305-701, Republic of Korea*

(Received September 3, 2013, Revised January 4, 2014, Accepted March 5, 2014)

Abstract. A visually servoed paired structured light system (ViSP) was recently proposed as a novel estimation method of the 6-DOF (Degree-Of-Freedom) relative displacement in civil structures. In order to apply the ViSP to massive structures, multiple ViSP modules should be installed in a cascaded manner. In this configuration, the estimation errors are propagated through the ViSP modules. In order to resolve this problem, a displacement estimation error back-propagation (DEEP) method was proposed. However, the DEEP method has some disadvantages: the displacement range of each ViSP module must be constrained and displacement errors are corrected sequentially, and thus the entire estimation errors are not considered concurrently. To address this problem, a pose-graph optimized displacement estimation (PODE) method is proposed in this paper. The PODE method is based on a graph-based optimization technique that considers entire errors at the same time. Moreover, this method does not require any constraints on the movement of the ViSP modules. Simulations and experiments are conducted to validate the performance of the proposed method. The results show that the PODE method reduces the propagation errors in comparison with a previous work.

Keywords: structural health monitoring (SHM); displacement measurement; pose-graph optimized displacement estimation (PODE); visually servoed paired light system (ViSP)

1. Introduction

Structural health monitoring (SHM) is an essential component in civil engineering for safety and integrity of civil structures such as buildings, bridges, and tunnels (Balageas *et al.* 2006). There are many kinds of disturbances in the structures from natural or artificial origin. Structural displacement measurement is one of the descriptors in evaluating deformation and variation from such disturbances effectively. Therefore, displacement monitoring is an important indicator for SHM (Ji and Chang 2008, Ni *et al.* 2011). Accordingly, many studies have been carried out in this area in conjunction with conventional sensors such as accelerometers, global positioning system (GPS), and Laser Doppler Vibrometers (LVDs). However, each of these sensors has shortcomings. Accelerometers measure the displacement indirectly and are neither stable nor accurate due to signal drift (Park *et al.* 2005). In GPS-based systems, costly RTK (Real Time Kinematics)-GPS (more than $20,000 USD) is needed for centimeter-level accuracy (Casciati and Fuggini 2011,

*Corresponding author, Professor, E-mail: hmyung@kaist.ac.kr

http://www.techno-press.org/?journal=sss&subpage=8 ISSN: 1738-1584 (Print), 1738-1991 (Online)

Psimoulis *et al.* 2008, Xu *et al*. 2009). LDVs are also very expensive and present difficulties when used in massive structures such as long-span bridges because the sensors should be installed in a fixed location such as on the ground (Nassif *et al*. 2005).

Due to these restrictions of previous approaches, vision-based methods that directly measure structural displacement have been studied as a possible alternative (Chang and Xiao 2009, Lee *et al.* 2012c, Lee and Shinozuka 2006, Leith *et al.* 1989, Marecos *et al.* 1969, Olaszek 1999, Park *et al.* 2010, Wahbeh *et al.* 2003). These vision-based systems use cameras and targets such as planar markers that have a simple black and white pattern. After capturing images of targets installed on a structure to be measured, the displacements are estimated in real-time through image processing algorithms. However, these systems are sensitive to environmental variations such as weather conditions since the cameras can see distant targets only in clean air.

In order to overcome the limitations of vision-based systems, Myung *et al.* proposed a paired structured light (SL) system based on lasers and cameras (Myung *et al.* 2011, Myung *et al.* 2012). The system consists of two sides facing each other; each side has a camera, a screen, and one or two lasers. The lasers are projected to the opposite side and the camera captures its own screen. The 6-DOF (Degree-Of-Freedom) relative displacement between the two sides is estimated from the positions of the projected laser beams. Since the distance between the cameras and the screens of this system is very short, this system robustly estimates the displacement. A visually servoed paired SL system (ViSP) was subsequently introduced to measure a wide range of displacement by using 2-DOF manipulators that control the pose of the lasers (Jeon *et al.* 2011). Therefore, in the ViSP, the projected laser beams can avoid the problem of travelling outside the opposite screen's boundary.

Since far distance displacement should be measured in massive structures, a single ViSP module is insufficient and hence multiple modules should be installed in a cascaded manner. In this case, the movement of the structure is estimated by combining the relative displacements of the multiple ViSPs, and hence each measurement error of the displacement is propagated along the ViSP modules. In order to solve this problem, a displacement estimation error back-propagation (DEEP) method was proposed (Jeon *et al.* 2013). Inspired by the error back-propagation algorithm used in neural networks, the DEEP minimizes estimation errors by using the Newton-Raphson or the gradient descent method. However, in the DEEP method, the displacement range of each ViSP module should be given to derive a solution. Furthermore, the DEEP corrects the displacement errors sequentially, and thus the entire estimation errors are not considered concurrently.

In this paper, we propose a pose-graph optimized displacement estimation (PODE) method for a multiple structural displacement monitoring system. The PODE method is based on a graph-based optimization technique. Graph-based optimization techniques are popularly used for solving SLAM (Simultaneous Localization And Mapping) problems in the robotics community (Dellaert and Kaess 2006, Grisetti *et al.* 2010, Kaess *et al.* 2008, Kaess *et al.* 2012, Lee *et al.* 2012a, Lee *et al*. 2012b, Lu and Milios 1997, Olson *et al.* 2006). The positions of the modules and the relative measurements build a graph structure. Next, the measurement noises are assumed as Gaussian distributions, and then the graph is iteratively optimized by a maximum likelihood method. This method considers the entire errors in each iteration step and does not require any constraints on the movement of the ViSP modules.

The remainder of this paper is organized as follows. In the second section, the ViSP and the DEEP method are briefly reviewed. Next, the PODE method is proposed in the third section. In the fourth section, the performance of the proposed method is validated with simulations and experiments. Finally, the last section offers concluding remarks.

2. Current displacement estimation approach using multiple ViSPs

2.1 Visually servoed paired structured light system (ViSP)

A visually servoed paired structured light system (ViSP) was proposed to estimate 6-DOF structural displacement with high accuracy regardless of environmental changes (Jeon *et al.* 2011). As shown in Fig. 1, the ViSP is composed of two sides facing each other, each with one or two lasers, a camera, a screen, and a 2-DOF manipulator. The manipulator controls laser beams such that they remain inside the screen all the time. In other words, the projected laser beams are forced to the center of the screen before moving outside the screen boundary. By calculating the rotation angles of the manipulators and positions of the projected laser beams, the 6-DOF displacement can be estimated.

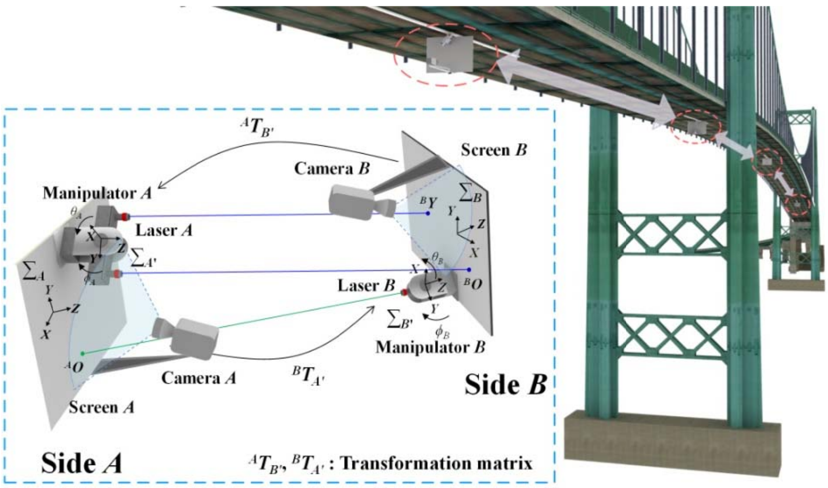


Fig. 1 A schematic diagram of the displacement measurement system using multiple visually servoed paired structured light systems (ViSPs). A schematic diagram of a single ViSP is shown in the box with the dashed line

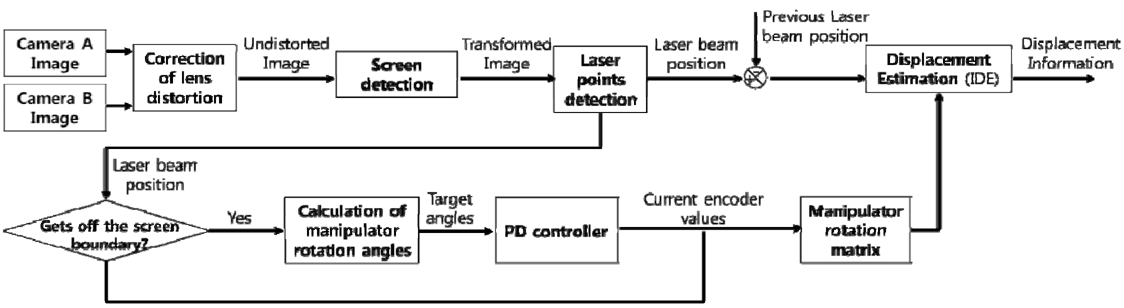


Fig. 2 Procedure of the 6-DOF displacement estimation using ViSP

The procedure of the displacement estimation is shown in Fig. 2 (Jeon *et al.* 2011). First, the camera on each side captures the image of its own screen from close proximity. Since the distance between the camera and the screen is short, such as less than 20 cm, it is robust to environmental changes such as weather or illumination. Therefore, if the lasers are successfully projected on each side, the ViSPs can be operated in the order of tens, or more, of meters without additional cost. Next, the lens distortion is corrected based on previously calculated distortion parameters. Afterwards, the screen boundary and the positions of the projected laser beams are calculated. If one of the laser beams travels outside the screen boundary, the manipulator forces the laser beam or the mid-point of the laser beam projected on side *A* or *B*, respectively, to remain inside the screen. The relative translational and rotational displacement between two sides is estimated by using the positions of the projected laser beams and rotation angles of the manipulators. In the estimation, an incremental displacement estimation (IDE) algorithm that updates a previously estimated displacement using the previous and the current observed data was proposed to reduce computation time. For the detailed explanation of the kinematics of the ViSP and the IDE algorithm, refer to Jeon *et al*. (2011, 2012).

2.2 Displacement estimation error back-propagation (DEEP)

To apply ViSP to massive civil structures, multiple ViSPs are placed in a cascaded manner and the dynamic displacement of the entire structure is estimated, as shown in Fig. 1. In other words, the relative displacement between adjoining ViSPs is combined with the next partition, and the entire movement of the structure can thereby be monitored. However, the performance of multiple ViSPs is impeded by the major problem that the displacement estimation error is propagated through the partitions.

To solve this problem Jeon *et al.* (2013), inspired by the error back-propagation algorithm used in neural networks, proposed a displacement estimation error back-propagation (DEEP) method that uses Newton-Raphson or gradient descent formulation. Fig. 3 shows a configuration of multiple ViSPs. $\mathbf{x}_i$ denotes a 6-DOF displacement, $[x_i, y_i, z_i, \theta_i, \phi_i, \psi_i]^T$, at the *i*-th ViSP module in the world coordinate. Relative poses $z_{i,i+1} = [x_{i,i+1}, y_{i,i+1}, z_{i,i+1}, \theta_{i,i+1}, \phi_{i,i+1}, \psi_{i,i+1}]^T$ $(i = 0,\ldots,N\text{-}1)$ denote measurement results through the ViSPs. A 6-DOF relative pose $\overline{\mathbf{z}}_{0,N}$ is obtained from the known data using a precise GPS or topographical surveying. The propagation error can be defined as the difference between the position of the last module given by the ground truth ($\overline{\mathbf{x}}_N$) and the estimated position from multiple ViSPs ($\mathbf{x}_N$), where *N* is the number of modules. $\overline{\mathbf{x}}_N$ and $\mathbf{x}_N$ are derived from $\overline{\mathbf{z}}_{0,N}$ and $z_{i,i+1}$, respectively. The error between the $\overline{\mathbf{x}}_N$ and the pose of the last module $\mathbf{x}_N$ is propagated sequentially for correcting the poses of the entire nodes. The DEEP method using the Newton-Raphson formulation with a learning rate α is formulated as follows

$$\mathbf{x}_i(k+1) = \mathbf{x}_i(k) - \alpha J^+_{\mathbf{x}_i} E_N \tag{1}$$

where $J_{\mathbf{x}_i} = \partial E_N / \partial \mathbf{x}_i$ is the Jacobian of the propagation error E_N, $J^+_{\mathbf{x}i}$ is the pseudo-inverse of $J_{\mathbf{x}i}$, and $E_N = \overline{\mathbf{x}}_N - \mathbf{x}_N$. Similar to the DEEP method using the Newton-Raphson formulation, the DEEP method with a gradient descent formulation can be depicted as follows

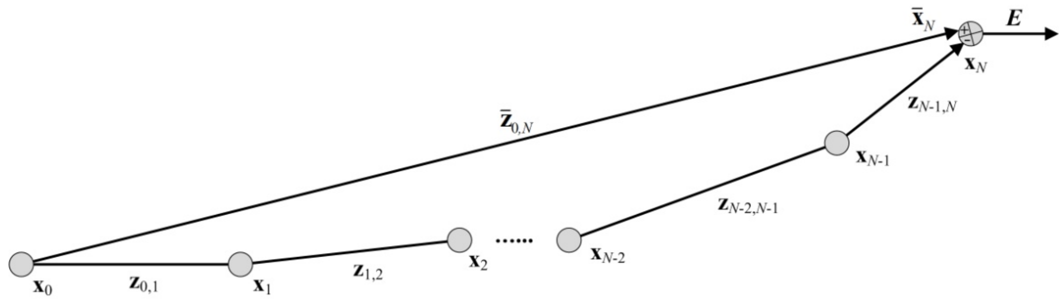


Fig. 3 The propagation error (E) of multiple ViSPs is defined as the difference of the position of the last module calculated by ground truth ($\bar{\mathbf{x}}_N$) and the ViSPs (x_N)

$$\mathbf{x}_i(k+1) = \mathbf{x}_i(k) - \alpha J_{\mathbf{x}_i} + \gamma \Delta \mathbf{x}_i(k) \tag{2}$$

Here, $J_{\mathbf{x}_i} = \partial E_G / \partial \mathbf{x}_i$ is the Jacobian of the propagation error E_G, where $E_G = \frac{1}{2}(\bar{\mathbf{x}}_N - \mathbf{x}_N)^2$, α is the learning rate, γ is a momentum parameter, and $\Delta\mathbf{x}_i(k)$= $\mathbf{x}_i(k)$- $\mathbf{x}_i(k$-1).

In the DEEP method, since the number of displacement variables is more than the number of constraints ($\bar{\mathbf{x}}_N$), there can be numerous solutions that minimize the propagation error. Therefore, we assume that the updated range of the displacement, $\Delta\mathbf{x}_i$, is limited by two constraint values, $\Delta\mathbf{x}_i^{\min}$ and $\Delta\mathbf{x}_i^{\max}$. The two constraint values are determined by considering physical limitations of the motors or error covariance of the ViSP. For the detailed explanation of the DEEP method, refer to Jeon *et al*. (2013).

3. PODE method

3.1 Pose-graph optimization

The PODE method proposed in this paper is based on a pose-graph optimization technique recently used in mobile robotics (Lee *et al.* 2012a, Lee *et al.* 2012b). A pose-graph consists of nodes and edges. Each node represents the robot pose. Relative measurements between the nodes are denoted as edges (Grisetti *et al.* 2010, Lu and Milios 1997, Olson *et al.* 2006). In the mobile robotics system, the measurements are obtained from dead reckoning and sensor measurement results, such as by use of an inertial measurement unit (IMU) or a variety of sensors such as a camera or a laser range finder (LRF). In the ViSP system, the nodes represent displacement of each module in the world coordinates. And the edges denote measurement results through the ViSPs and known data using precise GPS or surveying. Assuming the measurement noises follow Gaussian distributions, the maximum likelihood estimation (MLE) optimizes the graph structure. And then the corrected displacements of the ViSP modules are given from the nodes (Dellaert and Kaess 2006, Kaess *et al.* 2008, Kaess *et al.* 2012).

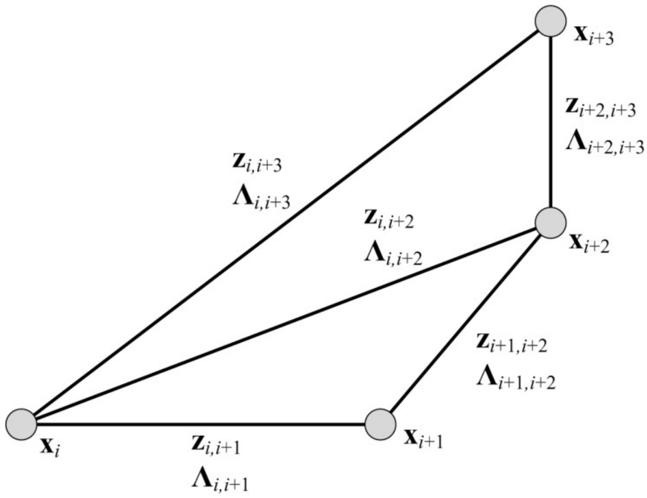


Fig. 4 Graphical model of pose-graph optimization. x_i is a robot pose vector, $z_{i,j}$ is the measurement value between i-th and j-th nodes from sensors, and $\Lambda_{i,j}$ denotes an information matrix of the measurement, which is the inverse covariance matrix

Fig. 4 shows a graphical model of pose-graph optimization, where $\mathbf{x}_i = [x_i, y_i, z_i, \theta_i, \phi_i, \psi_i]^T$ is a robot pose vector, $\mathbf{z}_{i,j} = [x_{i,j}, y_{i,j}, z_{i,j}, \theta_{i,j}, \phi_{i,j}, \psi_{i,j}]^T$ is the measurement value of the relative pose between i-th and j-th nodes from sensors, and $\mathbf{\Lambda}_{i,j}$ denotes an information matrix of the measurement, which is the inverse covariance matrix. The MLE of the pose-graph is obtained by minimizing the Mahalanobis distance of residuals as follows

$$\mathbf{x}^* = \arg\min_{\mathbf{x}} \frac{1}{2} \sum_{\langle i,j\rangle \in C} \mathbf{r}_{i,j}^T(\mathbf{x}) \mathbf{\Lambda}_{i,j} \mathbf{r}_{i,j}(\mathbf{x}) \tag{3}$$

where $\mathbf{r}_{i,j}$ is a residual between the prediction and the observation in the relative pose of i-th and j-th nodes and C strands for the set of edges connecting nodes. The residual $\mathbf{r}_{i,j}$ is represented as

$$\mathbf{r}_{i,j}(\mathbf{x}) = \mathbf{h}_{i,j}(\mathbf{x}) - \mathbf{z}_{i,j} \tag{4}$$

where $\mathbf{h}_{i,j}(\mathbf{x})$ is the prediction model between i-th and j-th nodes. Since $\mathbf{r}_{i,j}$ is generally a nonlinear function, the pose-graph optimization leads to an iterative method for solving a non-linear least square problem with respect to $\Delta\mathbf{x}$. A nonlinear cost function of the pose-graph is set as follows (Dellaert and Kaess 2006, Kaess *et al.* 2008, Kaess *et al.* 2012, Lu and Milios 1997)

$$\mathrm{F}(\mathbf{x}) = \frac{1}{2} \sum_{\langle i,j\rangle \in C} \mathbf{r}_{i,j}^T(\mathbf{x}) \mathbf{\Lambda}_{i,j} \mathbf{r}_{i,j}(\mathbf{x}) \tag{5}$$

$\mathbf{J}_{i,j}(\mathbf{x})$ is the Jacobian of $\mathbf{r}_{i,j}(\mathbf{x})$ with respect to $\mathbf{x}$, as delineated in Eq. (6).

$$\begin{aligned}\mathbf{r}_{i,j}(\mathbf{x}+\Delta\mathbf{x}) &\approx \mathbf{r}_{i,j}(\mathbf{x}) + \underbrace{\left.\frac{\partial \mathbf{r}_{i,j}(\breve{\mathbf{x}})}{\partial \breve{\mathbf{x}}}\right|_{\breve{\mathbf{x}}=\mathbf{x}}}_{\mathbf{J}_{i,j}(\mathbf{x})} \Delta\mathbf{x} \\ &= \mathbf{r}_{i,j}(\mathbf{x}) + \mathbf{J}_{i,j}(\mathbf{x})\Delta\mathbf{x}\end{aligned} \tag{6}$$

The cost function F($\mathbf{x}$) is approximated by its second-order Taylor series expansion as follows

$$\begin{aligned}\mathrm{F}(\mathrm{x}+\Delta\mathrm{x}) &= \frac{1}{2}\sum_{\langle i,j\rangle\in C} \mathrm{r}_{i,j}^T(\mathrm{x}+\Delta\mathrm{x})\Lambda_{i,j}\mathrm{r}_{i,j}(\mathrm{x}+\Delta\mathrm{x}) \\ &\approx \frac{1}{2}\sum_{\langle i,j\rangle\in C}\left(\mathrm{r}_{i,j}(\mathrm{x}) + \mathrm{J}_{i,j}(\mathrm{x})\Delta\mathrm{x}\right)^T \Lambda_{i,j}\left(\mathrm{r}_{i,j}(\mathrm{x}) + \mathrm{J}_{i,j}(\mathrm{x})\Delta\mathrm{x}\right) \\ &= \sum_{\langle i,j\rangle\in C}\left(\begin{aligned}&\frac{1}{2}\mathrm{r}_{i,j}(\mathrm{x})^T\Lambda_{i,j}\mathrm{r}_{i,j}(\mathrm{x}) + \Big(\underbrace{\mathrm{J}_{i,j}(\mathrm{x})^T\Lambda_{i,j}\mathrm{r}_{i,j}(\mathrm{x})}_{\mathrm{g}_{i,j}}\Big)^T\Delta\mathrm{x} \\ &+\frac{1}{2}\Delta\mathrm{x}^T\underbrace{\mathrm{J}_{i,j}(\mathrm{x})^T\Lambda_{i,j}\mathrm{J}_{i,j}(\mathrm{x})}_{\mathrm{H}_{i,j}}\Delta\mathrm{x}\end{aligned}\right)\end{aligned} \tag{7}$$

And then, by taking derivatives with respect to $\Delta\mathbf{x}$, the MLE problem of the pose-graph becomes a linear system as Eqs. (8) and (9).

$$\left.\frac{\partial \mathrm{F}(\mathbf{x}+\Delta\mathbf{x})}{\partial \Delta\mathbf{x}}\right|_{\Delta\mathbf{x}=0} = 0 \;\rightarrow\; \mathbf{g} + \mathbf{H}\Delta\mathbf{x} = 0 \tag{8}$$

$$\mathbf{H}\Delta\mathbf{x} = -\mathbf{g} \tag{9}$$

where $\mathbf{H}$ and $\mathbf{g}$ are represented by Eqs. (10) and (11), respectively.

$$\begin{aligned}\mathbf{H} &= \sum_{\langle i,j\rangle\in C}\mathbf{H}_{i,j} \\ \mathbf{H}_{i,j} &= \mathbf{J}_{i,j}^T(\mathbf{x})\mathbf{\Lambda}_{i,j}\mathbf{J}_{i,j}(\mathbf{x})\end{aligned} \tag{10}$$

$$\begin{aligned}\mathbf{g} &= \sum_{\langle i,j\rangle\in C}\mathbf{g}_{i,j} \\ \mathbf{g}_{i,j} &= \mathbf{J}_{i,j}^T(\mathbf{x})\mathbf{\Lambda}_{i,j}\mathbf{r}_{i,j}(\mathbf{x})\end{aligned} \tag{11}$$

All robot poses $\mathbf{x}$ is updated from $\Delta\mathbf{x}$ as

$$\mathbf{x} \leftarrow \mathbf{x} + \Delta\mathbf{x} \tag{12}$$

Using the updated $\mathbf{x}$, the cost function is recalculated to obtain a new incremental value $\Delta\mathbf{x}$ iteratively. These calculations are iterated until it converges or meets some termination criteria.

A variety of pose-graph optimization algorithms have been developed for computational efficiency in robotics communities. In this paper, the iSAM (Incremental Smoothing and

Mapping) algorithm is used to optimize the poses of the multiple ViSP modules. iSAM provides a graph optimization solution using sparse linear algebra, and hence computational speed is dramatically increased (Kaess *et al.* 2008, Kaess *et al.* 2012).

3.2 Pose-graph optimized displacement estimation (PODE)

In order to estimate the displacement of massive civil structures, ViSPs have been applied in a cascaded manner, and then each pair of ViSP modules monitors the relative 6-DOF movement between the modules. However, in this configuration, measurement error of each ViSP is propagated, and then the last module has a large displacement error. In order to overcome this drawback, we assume that the ground truth at the last module is known *a priori* using precise sensors such as a high-priced GPS or topographical surveying. It is then necessary to estimate 6-DOF poses of all modules by considering the propagated error. In this section, we propose a PODE method to obtain correct results for multiple ViSPs in civil infrastructures while overcoming the shortcomings of the previous work, the DEEP method: the DEEP needs constraints on the movement of each ViSP module and corrects the displacement errors sequentially, not concurrently.

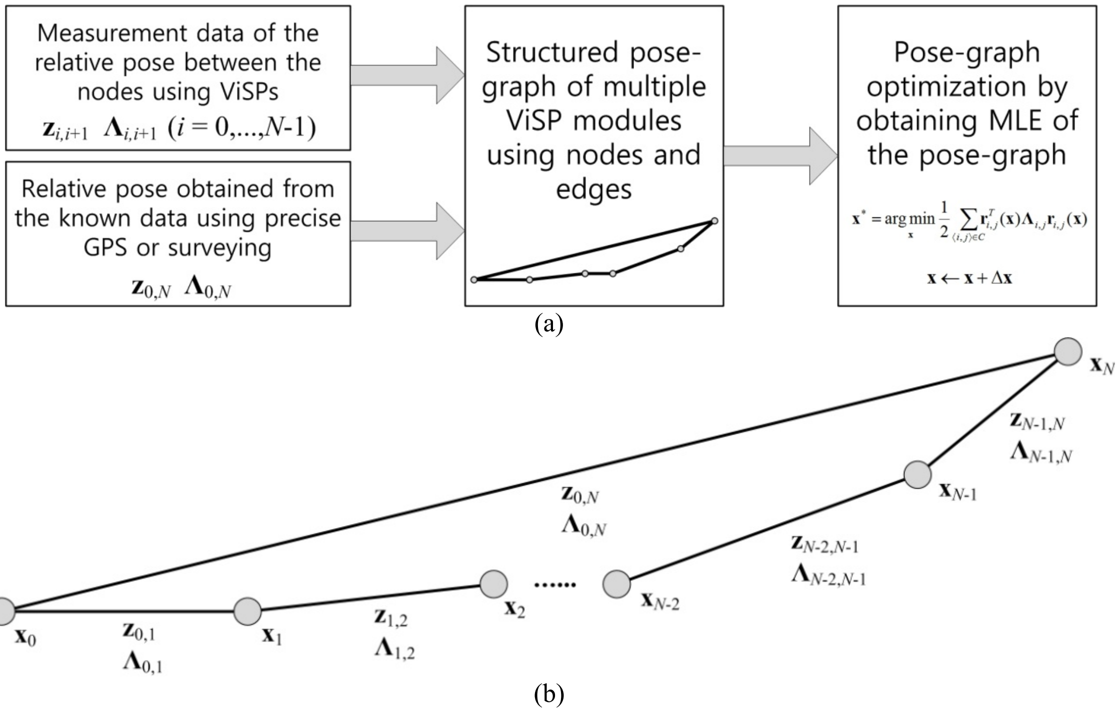


Fig. 5 (a) Procedure of the PODE method. (b) Pose-graph representation of multiple ViSP modules using nodes and edges

The PODE method is based on pose-graph optimization. Fig. 5 shows the procedure of the proposed PODE method and a pose-graph representation of multiple ViSP modules using nodes and edges. Pose vectors $\mathbf{x}_i$ (i = 0,…,N) represent 6-DOF poses of each module in the world coordinates, where N is the number of modules. Measurement data of the relative pose between the nodes are assumed to have Gaussian noise. Relative poses $\mathbf{z}_{i,i+1}$ and information matrices $\mathbf{\Lambda}_{i,i+1}$ (i = 0,…,N-1) denote measurement results and their uncertainties through the ViSPs, respectively. A relative pose $\mathbf{z}_{0,N}$ and an information matrix $\mathbf{\Lambda}_{0,N}$ are obtained from the known data using precise GPS or surveying, and hence the measurement $\mathbf{z}_{0,N}$ has lower uncertainty than the other measurements. Moreover, since the information matrix is the inverse covariance matrix of the Gaussian noise, the information matrix $\mathbf{\Lambda}_{0,N}$ has larger values than the other information matrices. The structured pose-graph with all initial poses $\mathbf{x}_{0\ldots N}$ and all measurement data $\mathbf{z}_{0\ldots N\text{-}1,1\ldots N}$, $\mathbf{z}_{0,N}$, $\mathbf{\Lambda}_{0\ldots N\text{-}1,1\ldots N}$, and $\mathbf{\Lambda}_{0,N}$, is optimized by the procedure shown in Section 3.1 above. The initial poses of the nodes are set using only measurement data of the ViSPs. Therefore, in our method, only the measurement information of the ViSPs and the prior pose of the last module are used to obtain correct poses of the entire modules. Contrary to the DEEP method, the PODE method does not need displacement constraints to find a unique solution, because it simply optimizes the graph structure under the maximum likelihood criterion.

4. Simulations and experiments

4.1 Simulation results

In order to verify the performance of the PODE method, simulation studies have been conducted and the results were compared to those of the DEEP method. Three configuration sets, with four, six, and ten ViSP modules, respectively, are used in the simulation. The estimated 6-DOF poses ($\mathbf{x}_i$) of the modules are evaluated against the ground truth for the poses ($\bar{\mathbf{x}}_i$). In these simulations, we focused on numerical aspects of displacement estimation rather than the real displacement order. This scheme also had been applied to the simulation studies of the DEEP method (Jeon *et al.* 2013). Therefore, we performed the same procedures and analyzed the simulation results for comparison purpose.

The simulation data sets are the same with the data used in the previous work for the evaluation of the DEEP (Jeon *et al.* 2013). For the data it was assumed that each ViSP module has uniform random noise of [-0.1 0.1] cm in laser point measurements. The random noise corresponds to approximately ten image pixel errors of a camera with a screen size of 0.15 m×0.1 m and image resolution of 640×480 pixels. Initial poses are obtained from the measurement of ViSP and then the errors are propagated along the modules. And the relative pose between the first and the last modules is given; that is, it is assumed that the pose is measured by a precise GPS or topographical surveying. Then, for the PODE method, a graph structure is built by using the measurement data from the ViSP and the precise relative pose between the front and the end modules. In this simulation, an information matrix $\mathbf{\Lambda}$ is formed such as

$$\Lambda = \begin{bmatrix} \sigma_x^2 & 0 & 0 & 0 & 0 & 0 \\ 0 & \sigma_y^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma_z^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_\theta^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sigma_\phi^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & \sigma_\psi^2 \end{bmatrix}^{-1} \tag{13}$$

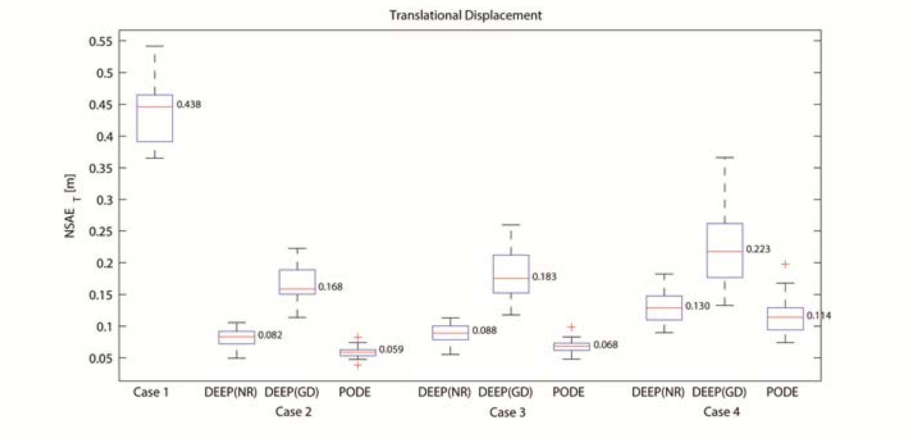
The six elements (σ_x, σ_y, σ_z, σ_θ, σ_ϕ, σ_ψ) of the information matrix represent the standard deviations of the Gaussian noise in measuring the 6-DOF relative pose. In the ViSP and the precise relative pose measurements, the standard deviations are set to (0.055 m, 0.055 m, 0.085 m, 0.4°, 0.4°, 0.4°) and (0.0055 m, 0.0055 m, 0.0085 m, 0.04°, 0.04°, 0.04°), respectively. The pose of each node is updated by a graph optimization algorithm, iSAM (Kaess *et al.* 2008, Kaess *et al.* 2012). The simulations with the DEEP are performed with the same parameters as reported in Jeon *et al.* (2013). In DEEP method, the motion limits and the error threshold are set to $[\pm 0.01\text{ m}, \pm 0.01\text{ m}, \pm 0.15\text{ m}, \pm 1^\circ, \pm 1^\circ, \pm 1^\circ]^T$ and 1.0×10^{-4}, respectively. The learning rate of the DEEP method with Newton–Raphson formulation is set to 0.1, and the learning rate and momentum parameter of the DEEP method with gradient descent formulation are set to (0.1, 0.3), (0.05, 0.1) and (0.003, 0.005) for the simulations with four, six and ten ViSP modules, respectively.

The performance of the PODE method is analyzed statistically with a Monte Carlo simulation (MCS). A multi-module simulation set is randomly generated twenty times with four, six, and ten modules, respectively. In particular, with the four modules, additional simulations are performed by adding random errors to the precise measurement data of the last modules. The updated displacement is then evaluated against the ground truth based on the normalized sum of absolute errors (NSAE_T and NSAE_R) at each module, which are defined by the following equations

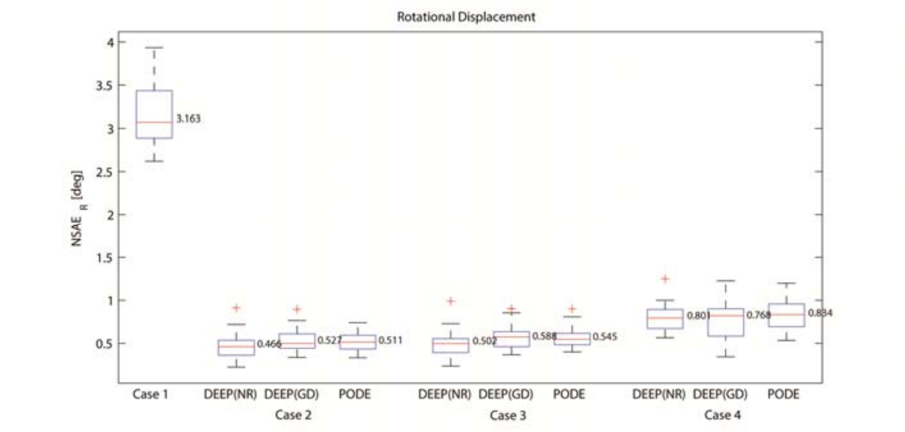
$$\text{NSAE}_T = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{3}\left|\bar{\mathbf{x}}_i^j - \mathbf{x}_i^j\right| \tag{14}$$

$$\text{NSAE}_R = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=4}^{6}\left|\bar{\mathbf{x}}_i^j - \mathbf{x}_i^j\right| \tag{15}$$

where n is the total number of modules, the subscript i denotes the i-th module, the superscript j denotes the j-th component of the pose vector, and $\bar{\mathbf{x}}$ indicates the ground truth. Boxplots of the simulations with four modules are used to analyze the results, as shown in Fig. 6. Cases 1 and 2 are estimated displacements by using only ViSP and ViSP with DEEP or PODE, respectively. In Cases 3 and 4, uniform random noises of $[\pm 0.02\text{ m}, \pm 0.02\text{ m}, \pm 0.02\text{ m}, \pm 0.1^\circ, \pm 0.1^\circ, \pm 0.1^\circ]^T$ and $[\pm 0.1\text{ m}, \pm 0.1\text{ m}, \pm 0.1\text{ m}, \pm 0.5^\circ, \pm 0.5^\circ, \pm 0.5^\circ]^T$ are added to the pose vector of the last modules, respectively. The simulation results with three different numbers of modules are shown in Fig. 7 and Table 1. In Fig. 7, the three dimensional axes X, Y, and Z represent coordinate system of the first ViSP module as shown in Fig. 1. The proposed PODE method greatly reduces the estimation errors of the ViSP, and has strength in reducing translational errors with respect to the other methods. In addition, using a t-test, the difference between the PODE and the others is verified, as shown in Table 2. The null and alternative hypotheses in these tests are that two distributions are equal or not equal, respectively. p_T and p_R represent p-values of the NSAE_T and NSAE_R, respectively. All p-values of $NSAE_T$ (p_T) are less than 0.05. In $NSAE_R$, the comparison to the result with ten modules and the DEEP(GD) reveals good performance ($p_R < 0.05$), while the other four cases are similar to the PODE ($p_R > 0.05$). In particular, comparing the result with six modules and that of the DEEP(GD) reveals that the PODE has poor performance with $p_R < 0.05$. However, since the p-value of this case is 0.024, the result of the PODE is not far from those of the other methods.



(a)



(b)

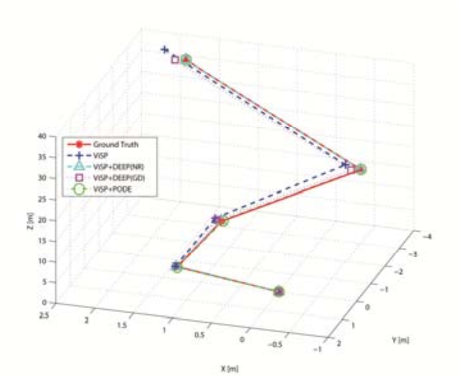
Fig. 6 Boxplots of 20 Monte Carlo simulation results with four modules. Cases 1, 2, 3, and 4 are estimated displacements by using ViSP, ViSP with DEEP or PODE, and ViSP with DEEP or PODE with uniform random noises of $[\pm 0.02\ \mathrm{m}, \pm 0.02\ \mathrm{m}, \pm 0.02\ \mathrm{m}, \pm 0.1^{\circ}, \pm 0.1^{\circ}, \pm 0.1^{\circ}]^T$, and $[\pm 0.1\ \mathrm{m}, \pm 0.1\ \mathrm{m}, \pm 0.1\ \mathrm{m}, \pm 0.5^{\circ}, \pm 0.5^{\circ}, \pm 0.5^{\circ}]^T$, respectively. The DEEP method is performed with the Newton-Raphson (NR) and gradient descent (GD) formulation. The random noises are applied to the ground truth positions at the last modules. (a) Normalized sum of absolute translational errors ($NSAE_T$). (b) Normalized sum of absolute rotational errors ($NSAE_R$)

Table 1 Normalized sum of absolute errors ($NSAE_T$ and $NSAE_R$) with different numbers of modules. The results of the simulations are the medians of 20 cases. Compared to the ViSP, the results of the ViSP+PODE exhibit significantly reduced errors, which are indicated by the values in parenthesis. Compared to the DEEP(NR) and DEEP(GD), the PODE method especially has strength in reducing translational errors

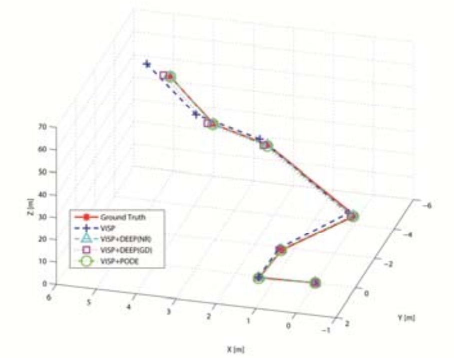
# of modules	Algorithms	$NSAE_T$ (m)	$NSAE_R$ (deg)
4	ViSP	0.438	3.163
	ViSP+DEEP(NR)	0.082 (81% ↓)	0.466 (85% ↓)
	ViSP+DEEP(GD)	0.168 (62% ↓)	0.527 (83% ↓)
	ViSP+PODE	0.059 (87% ↓)	0.511 (84% ↓)
6	ViSP	0.870	3.535
	ViSP+DEEP(NR)	0.124 (86% ↓)	0.818 (77% ↓)
	ViSP+DEEP(GD)	0.331 (62% ↓)	0.782 (78% ↓)
	ViSP+PODE	0.104 (88% ↓)	0.884 (75% ↓)
10	ViSP	1.598	3.847
	ViSP+DEEP(NR)	0.301 (81% ↓)	1.558 (59% ↓)
	ViSP+DEEP(GD)	0.902 (44% ↓)	1.774 (54% ↓)
	ViSP+PODE	0.163 (90% ↓)	1.537 (60% ↓)

Table 2 Welch's *t*-test results of normalized sum of absolute errors ($NSAE_T$ and $NSAE_R$) with different numbers of modules. The null and alternative hypotheses in these tests are that two distributions are equal or not equal, respectively. The PODE is compared to each previous method for the difference between two means of each 20 cases. p_T and p_R represent *p*-values of the $NSAE_T$ and $NSAE_R$, respectively. The PODE method especially has strength in reducing translational errors ($p_T < 0.05$), while some rotational errors have similar distributions with the DEEP(NR) and DEEP(GD) ($p_R > 0.05$)

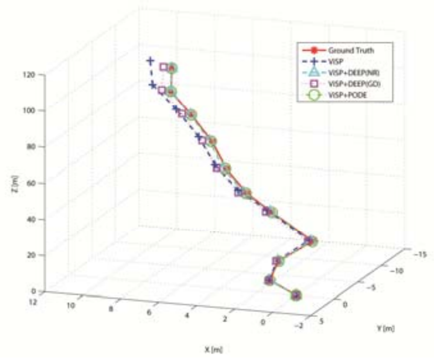
# of modules	Algorithm	p_T	p_R
4	ViSP	3.036×10^{-19}	2.835×10^{-20}
	ViSP+DEEP(NR)	1.988×10^{-6}	0.311
	ViSP+DEEP(GD)	1.322×10^{-13}	0.665
6	ViSP	1.531×10^{-21}	2.050×10^{-21}
	ViSP+DEEP(NR)	0.042×10^{-2}	0.091
	ViSP+DEEP(GD)	1.373×10^{-15}	0.024
10	ViSP	5.766×10^{-22}	3.564×10^{-15}
	ViSP+DEEP(NR)	2.128×10^{-17}	0.832
	ViSP+DEEP(GD)	3.224×10^{-9}	0.002



(a)



(b)



(c)

Fig. 7 Simulation results with (a) four, (b) six, and (c) ten modules. Solid lines with asterisks: ground truths. Dashed lines with crosses: estimated results from the ViSP. Dashed lines with triangles: updated positions by using the DEEP method with Newton-Raphson formulation (DEEP(NR)). Dotted lines with rectangles: updated position by using the DEEP method with gradient descent formulation (DEEP(GD)). Dot-dashed lines with circles: updated positions by using the PODE method. The three dimensional axes X, Y, and Z represent coordinate system of the first ViSP module as shown in Fig. 1

4.2 Experimental results

The PODE method has been applied to real experiments with a ViSP module. The experimental setup of the ViSP system is the same as in previous studies (Jeon *et al.* 2011, Jeon *et al.* 2013). Fig. 8 shows the ViSP system and the experimental setup. As shown in Fig. 8(a), each side has a screen, a camera, and one or two lasers. One side is installed on a motion stage that generates rotational and/or translational variations. The lasers are manipulated by the electrical components, as shown in Fig. 8(b). Experimental results are compared with the motion stage outputs, which can be considered as the ground truth. The motion stage outputs consist of data from the motorized rotation and translation stages, as shown in Fig. 8(c). The estimated displacement results using the ViSP module have been applied four times to simulate a multi-module system. Simulated positions of four ViSP modules are represented in Fig. 8(d).

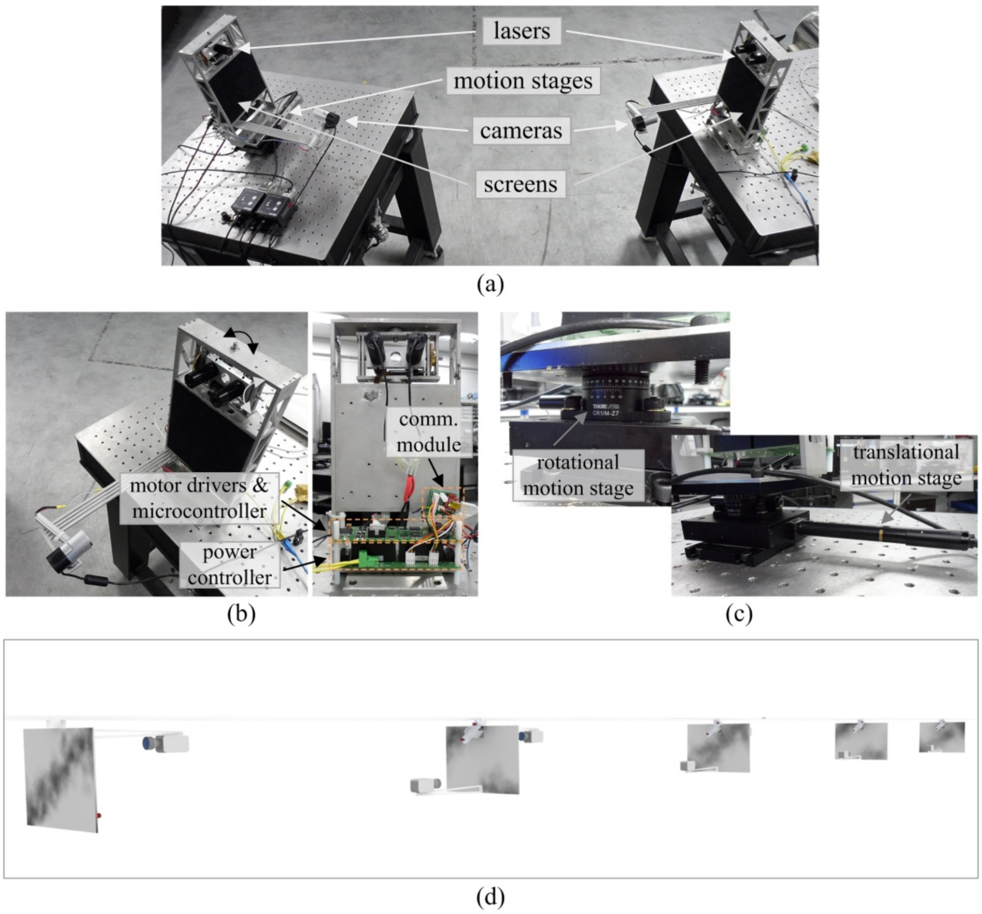
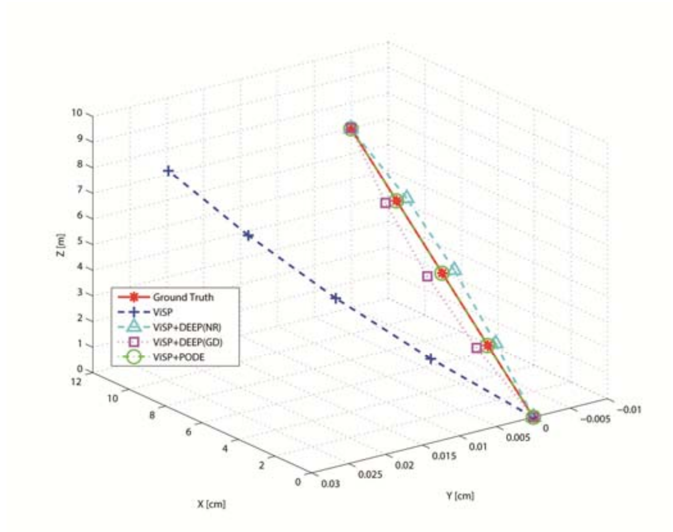
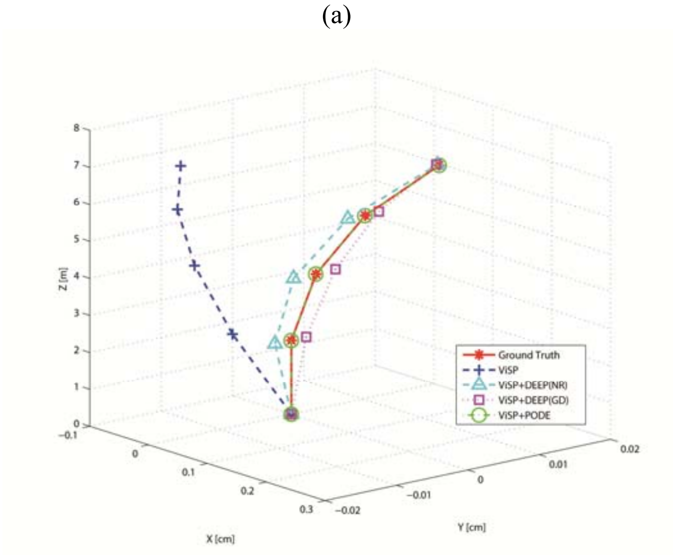


Fig. 8 Experimental setup. (a) Overall experimental setup, (b) front and rear view of side A, (c) motorized rotational and translational motion stages, and (d) simulated positions of four ViSP modules



(a)



(b)

Fig. 9 Experimental results of the proposed method PODE in two cases: (a) translational displacement along *X* axis and (b) rotational displacement about *Y* axis. Solid lines with asterisks: ground truths. Dashed lines with crosses: estimated results from the ViSP. Dashed lines with triangles: updated positions by using the DEEP method with Newton-Raphson formulation (DEEP(NR)). Dotted lines with rectangles: updated position by using the DEEP method with gradient descent formulation (DEEP(GD)). Dot-dashed lines with circles: updated position by using the PODE method. The three dimensional axes X, Y, and Z represent coordinate system of the first ViSP module as shown in Fig. 1

Table 3 Normalized sum of absolute errors ($NSAE_T$ and $NSAE_R$) from real experiments with four modules. Compared to the ViSP, the results of the ViSP+PODE exhibit significantly reduced errors, which are indicated by the values in parentheses. Compared to the DEEP(NR) and DEEP(GD), the PODE method especially has strength in reducing translational errors

Case	Algorithm	$NSAE_T$ (m)	$NSAE_R$ (deg)
Translation	ViSP	0.149	0.621
	ViSP+DEEP(NR)	0.011 (93% ↓)	0.010 (98% ↓)
	ViSP+DEEP(GD)	0.009 (94% ↓)	0.001 (99% ↓)
	ViSP+PODE	0.001 (99% ↓)	0.002 (99% ↓)
Rotation	ViSP	0.139	1.046
	ViSP+DEEP(NR)	0.031 (78% ↓)	0.024 (98% ↓)
	ViSP+DEEP(GD)	0.033 (76% ↓)	0.004 (99% ↓)
	ViSP+PODE	0.012 (91% ↓)	0.009 (99% ↓)

The updated displacements with the proposed PODE method are compared to the ground truths and the results using the DEEP method with the Newton-Raphson (NR) and gradient descent (GD) formulation (Fig. 9 and Table 3). Information matrix values of the PODE are the same as those used in the previous simulation experiments. The DEEP is performed with the same parameters as used in Jeon *et al.* (2013). In DEEP method, the motion limits and the error threshold are set to $[\pm 0.05 \text{ cm}, \pm 0.05 \text{ cm}, \pm 0.01 \text{ m}, \pm 0.4^{\circ}, \pm 0.4^{\circ}, \pm 0.4^{\circ}]^T$ and 1.0×10^{-4}, respectively. As shown in Fig. 9, the errors of the ViSP increase as the number of modules increases. By applying the proposed PODE method, the entire propagation errors in the estimated displacement are minimized. In particular, the results of the PODE are in good agreement with the ground truth at every module, while the DEEP agrees only at the last module. The three dimensional axes X, Y, and Z represent coordinate system of the first ViSP module as shown in Fig. 1. As shown in Table 3, the PODE method reduced errors drastically, especially in the estimation of translation. The rotational errors were already corrected significantly in the DEEP(NR) and DEEP(GD) methods (above 98%). Therefore, the $NSAE_R$ of the PODE also shows very small errors compared to the other methods. Also, these results conform to those from the simulation studies: the rotational errors of the PODE have had similar distribution to the DEEP(NR) and DEEP(GD).

5. Conclusions

In order to estimate displacement of massive structures, the ViSP modules should be placed in a cascaded manner. Estimation error is then propagated through the multiple modules, and the error of the last module thereby becomes large as the number of modules increases. In a previous work it was shown that the DEEP method minimizes the estimation errors by using the Newton-Raphson or the gradient descent method. However, in the DEEP method, the displacement range of each

ViSP module must be constrained and the entire estimation errors are not considered concurrently. To overcome these limitations, the PODE method based on a pose-graph optimization technique has been proposed in this paper. In the PODE method, a graph structure is built using the positions of the modules and the relative measurements. The graph is then iteratively optimized by a maximum likelihood method. The PODE method considers entire errors simultaneously and does not need any constraints on the ViSP modules. In order to validate the performance of the PODE method, simulations with different numbers of modules and real experiments have been performed and the obtained results were compared with those of the DEEP method. The results show that the displacement errors are reduced significantly, and the PODE especially has strength in reducing translational errors. In the future, multiple modules of the ViSP with the PODE method will be applied to a variety of massive structures.

Acknowledgments

This research was supported by a grant (13SCIPA01) from Smart Civil Infrastructure Research Program funded by Ministry of Land, Infrastructure and Transport (MOLIT) of Korea government and Korea Agency for Infrastructure Technology Advancement (KAIA). It was also supported by the R&D program of the Korea Ministry of Trade, Industry and Energy (MOTIE) and the Korea Evaluation Institute of Industrial Technology (KEIT). (The Development of Low-Cost Autonomous Navigation Systems for a Robot Vehicle in Urban Environment, 10035354). The students are partially supported by Ministry of Land, Infrastructure and Transport (MOLIT) as U-City Master and Doctor Course Grant Program.

References

Balageas, D., Fritzen, C.P. and Guemes, A. (eds) (2006), *Structural Health Monitoring*, New Jersey: John Wiley & Sons, Inc.

Casciati, F. and Fuggini, C. (2011), “Monitoring a steel building using GPS sensors”, *Smart Struct. Syst*., **7**(5), 349-363.

Chang, C.C. and Xiao, X.H. (2009), “Fusion of vision-based displacement and acceleration using Kalman filter”, *Proceedings of the 5th Int. Workshop on Advanced Smart Materials and Smart Structures Technology*, Boston.

Dellaert, F. and Kaess, M. (2006), “Square root SAM: simultaneous localization and mapping via square root information smoothing”, *Int. J. Robot. Res.*, **25**(12), 1181-1203.

Grisetti, G., Kummerle, R., Stachniss, C. and Burgard, W. (2010), “A tutorial on graph-based SLAM”, *IEEE Intell. Transport. Syst. Mag.*, **2**(4), 31-43.

Jeon, H., Bang, Y. and Myung, H. (2011), “A paired visual servoing system for 6-DOF displacement measurement of structures”, *Smart Mater. Struct.*, **20**(4), 45019-45034.

Jeon, H., Shin, J.U. and Myung, H. (2012), “Incremental displacement estimation of structures using paired structured light”, *Smart Struct. Syst.*, **9**(3), 273-286.

Jeon, H., Shin, J.U. and Myung, H. (2013), “The displacement estimation error back-propagation (DEEP) method for a multiple structural displacement monitoring system”, *Meas. Sci. Technol.*, **24**(4), 045104.

Ji, Y.F. and Chang, C.C. (2008), “Nontarget stereo vision technique for spatiotemporal response measurement of line-like structures”, *J. Eng. Mech. – ASCE*, **134**(6), 466-474.

Kaess, M., Johannsson, H., Roberts, R., Ila, V., Leonard, J.J. and Dellaert, F. (2012), “iSAM2: Incremental smoothing and mapping using the Bayes tree”, *Int. J. Robot. Res.*, **31**(2), 216-235.

Kaess, M., Ranganathan, A. and Dellaert, F. (2008), "iSAM: Incremental smoothing and mapping", *IEEE T. Robot.*, **24**(6), 1365-1378.
Lee, D., Kim, H. and Myung, H. (2012a), "2D image feature-based real-time RGB-D 3D SLAM", *Proc. of Robot Intelligence Technology and Applications 2012*, 485-492.
Lee, D., Kim, H. and Myung, H. (2012b), "GPU-based real-time RGB-D 3D", *Proc. of 2012 9th Int. Conf. on Ubiquitous Robots and Ambient Intelligence (URAI)*, 26-28.
Lee, J.J., Ho, H.N. and Lee, J.H. (2012c), "A vision-based dynamic rotational angle measurement system for large civil structures", *Sensors*, **12**(6), 7326-7336.
Lee, J.J. and Shinozuka, M. (2006), "Real-time displacement measurement of a flexible bridge using digital image processing techniques", *Exp. Mech.*, **46**(1), 105-114.
Leith, J.G., Thompson, A. and Sloan, T.D. (1989), "A novel dynamic deflection measurement system for large structure", *Proceedings of the 4th Int. Conf. on Civil and Structural Engineering Computing*, London.
Lu, F. and Milios, E. (1997), "Globally consistent range scan alignment for environment mapping", *Auton. Robot.*, **4**(4), 333-349.
Marecos, J., Castanheira, M. and Trigo, J. (1969), "Field observation of Tagus river suspension bridge", *J. Struct. Div.- ASCE*, **95**(4), 555-583.
Myung, H., Jung, J. and Jeon, H. (2012), "Robotic SHM and model-based positioning system for monitoring and construction automation", *Adv. Struct. Eng.*, **15**(6), 943-954.
Myung, H., Lee, S.M. and Lee, B.J. (2011), "Paired structured light for structural health monitoring robot system", *Struct. Health Monit.*, **10**(1), 49-64.
Nassif, H.H., Gindy, M. and Davis, J. (2005), "Comparison of laser doppler vibrometer with contact sensors for monitoring bridge deflection and vibration", *NDT & E Int.*, **38**(3), 213-218.
Ni, Y.Q., Wong, K.Y. and Xia, Y. (2011), "Health checks through land mark bridges to sky-high structures", *Adv. Struct. Eng.*, **14**(1), 103-119.
Olaszek, P. (1999), "Investigation of the dynamic characteristic of bridge structures using a computer vision method", *Measurement*, **25**(3), 227-236.
Olson, E., Leonard, J. and Teller, S. (2006), "Fast iterative alignment of pose graphs with poor initial estimates", *Proceedings of the 2006 IEEE Int. Conf. on Robotics and Automation (ICRA)*.
Park, J.W., Lee, J.J., Jung, H.J. and Myung, H. (2010), "Vision-based displacement measurement method for high-rise building structures using partitioning approach", *NDT & E Int.*, **43**(7), 642-647.
Park, K.T., Kim, S.H., Park, H.S. and Lee, K.W. (2005), "The determination of bridge displacement using measured acceleration", *Eng. Struct.*, **27**(3), 371-378.
Psimoulis, P., Pytharouli, S., Karambalis, D. and Stiros, S. (2008), "Potential of global positioning system (GPS) to measure frequencies of oscillations of engineering structures", *J. Sound Vib.*, **318**(3), 606-623.
Sciavicco L. and Siciliano B. (1988), "A solution algorithm to the inverse kinematic problem for redundant manipulator", *IEEE J. Robot. Autom*., **4**(4), 403-410.
Wahbeh, A.M., Caffrey, J.P. and Masri, S.F. (2003), "A vision-based approach for the direct measurement of displacements in vibrating systems", *Smart Mater. Struct.*, **12**(5), 785-794.
Xu, Y.L. and Chan, W.S. (2009), "Wind structural monitoring of long span cable-supported bridges with GPS", *Proceedings of the 7th Asia-Pacific Conf. on Wind Engineering*, Taipei.

CY