11

Detecting Human Activity by Location System and Stereo Vision

Yoshifumi Nishida, Koji Kitamura
National Institute of Advanced Industrial Science and Technology
Japan

1. Introduction

Information processing services centered around human activity in the real world has attracted increased attention recently (1). Human-centered applications require the facility to observe and recognize activities as a basis, and the present paper describes a method for quickly realizing a function for robustly detecting daily human activity events in the real world.

Generally, the problem of human activity recognition can be formulated as a kind of pattern recognition problem as follows.

$$P(\hat{W}|Y) = \max_{W_i} \frac{P(Y|W_i)P(W_i)}{P(Y)}, \tag{1}$$

where $P(W_i|Y)$ denotes the posterior probability that the meaning of an observed behavior pattern Y is W_i, $P(Y)$ denotes the probability that a behavior pattern Y will be observed, $P(W_i)$ denotes the probability that the behavior meaning W_i will occur, and $P(Y|W_i)$ denotes the conditional probability. Thus, the problem of human activity recognition becomes that of searching for the maximum posterior probability $P(\hat{W}|Y)$.

There are three problems in realizing and utilizing a function for recognizing human activity in the real world: the robust observation of a activity pattern Y, the efficient recognition of meaning W from the observed pattern, and quick implementation of a function for robustly observing and efficiently recognizing human activity. Without solving the first problem, equation (1) cannot be formed. Without tackling the second problem, guaranteeing a solution to the equation within the time frame demanded by the application is impossible. Without dealing with the third problem, it is difficult to utilize a funtion for observing and recognizing human activity for a basis of a real world application or various field researches.

As a method for efficient recognition of activites, the idea of object-based activity recognition has been proposed (2). In theory, the behavior of handling objects in an environment such as an office or home can be recognized based on the motion of the objects. However, when applying the method to real environments, it is difficult to even achieve an adequate level of object recognition, which is the basis of the method.

Separating the problems of object recognition and activity recognition is becoming increasingly realistic with the progress in pervasive computing technology such as microcomputers, sensor, and wireless networks technology. It has now become possible to resolve object recognition into the problems of sensorizing objects and tagging the objects

with identification codes (IDs), and to address activity recognition separately through the development of applied technology.
The present authors have developed a three-dimensional ultrasonic location and tagging system for the fundamental function of robustly tracking objects(3). This system enables a new approach of tag-based activity recognition. In terms of cost and robustness against environmental noise, the ultrasonic system is superior to other location techniques such as visual, tactile, and magnetic systems. Several types of ultrasonic location systems have been proposed. The Bat Ultrasonic Location System (4; 5; 6; 7) developed by AT&T, and the MIT Cricket Indoor Location System (8) are well known. Although a calibration method using a robot (9) has been proposed, the required calibration device is too large for use in a number of environments. An auto calibration method was considered in the DOLPHIN system (10), which can calibrate the positions of receivers/transmitters using a small number of reference receivers/transmitters having known positions. However, the system has only been tested in narrow areas having dimensions of approximately 2.5 m $\times$ 2 m. Bristol University proposed another auto calibration method, in which the positions of n transmitters and m receivers can be calculated given n$\times$m distance data among the transmitters and receivers and that the condition, $3(n+m)-6 < n \cdot m$, is satisfied(11). However, the scalability of this method is limited. In contrast, the present study proposes and examines a new calibration method, "global calibration based on local calibration," that requires a relatively small number of transmitters and is independent of room size. Using the proposed method, the calibration problem becomes a similar to a fitting problem in object modeling with multiple range images(12; 13) after local calibration. The present paper describes the method for global calibration based on local calibration and the constraints that are used in conjunction with the method for reducing the error of the calibrated receiver positions.
This paper focuses on a system for quickly realizing a function for robustly detecting daily human activity events in handling objects in the real world. This paper deals with a method for robustly measuring 3D positions of the objects handled by a person, a quick calibrating method for constructing a measuring system for 3D positions of the objects, and a quick registering method for target activity events. The next section describes the system for quick realization of the function for detecting human activity events. Section 3 shows algorithms for robustly measuring 3D positions of the objects handled by a person, and evaluates the algorithms. Section 4 describes a quick calibrating method, and Section 5 describes quick registration of human activity by a stereoscopic camera with ultrasonic 3D tags and interactive software for registering human activity events.

2. Quick realization of function for detecting human activity events

This section describes a system for quickly realizing a function for robustly observing and efficiently recognizing daily human activities.

2.1 System for quick realization for function of detecting human activity events

The configuration of the proposed system is shown in Fig. 1. The system consists of an ultrasonic 3D tag system, a calibrating device, a stereoscopic camera with ultrasonic 3D tags, and a host computer. The system has three functions: 1) robustly measuring 3D positions of the objects (Fig. 1(A)), 2) quickly calibrating a system for measuring 3D positions of the objects (Fig. 1(B)), 3) quickly registering target activity events (Fig. 1(C)), and 4) robustly detecting the registered events in real time (Fig. 1(D)).

As for 1), the system realizes robust measurement of 3D positions of the objects using an ultrasonic 3D tag system and robust estimation algorithm known as random sample consensus (RANSAC). As for 2), the system realizes quick calibration by a calibrating device having three or more ultrasonic transmitters. Quick calibration enables the system to be portable. As for 3), quick registration of target activity events is realized by a stereoscopic camera with ultrasonic 3D tags and interactive software for creating 3D shape model, creating virtual sensors based on the 3D shape model, and associating the virtual sensors with the target events.

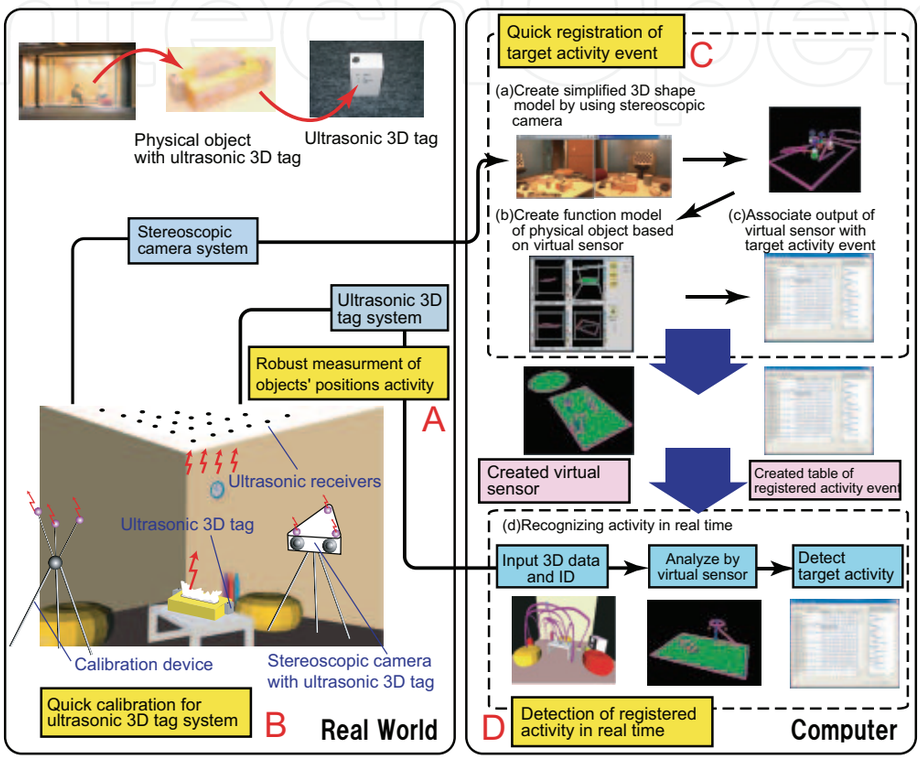


Fig. 1. Configuration of system for quick realization for function for detecting human activity events

2.2 Steps for quick realization for function of detecting human activity events

1. Install ultrasonic receivers in a target environment.
2. Calculate 3D positions of installed ultrasonic receivers using a calibration device. The details of a calibration method and a calibrating system are described in Section 4.
3. Register target activity events using a stereoscopic camera with ultrasonic 3D tags and interactive software. The details are described in Section 5.
4. Detect the registered target events using the ultrasonic 3D tags and the created virtual sensors.

2.3 Advantage of proposed system

Advantages of the proposed system are following points.

– **Utilization of user's knowledge** Since users know target activity to be detected, the system can make full use of knowledge of users familiar with target area by interactively registering target events.
– **Efficient processing** It is possible to create the minimum system by determining the number of ultrasonic receivers and the number of target events depending on the place where the users want to install and the activity events which the users want to target.
– **Inexpensive system** It is possible to utilize inexpensive sensors such as the ultrasonic 3D tag system (about $45 for a sensor and $200 for a tag), and the stereoscopic camera (about $200 in our system) to create the proposed system.
– **Robust system** It is easy to increase the number of ultrasonic receivers for robust estimation because they are inexpensive sensors. The details of an algorithm for robust estimation are described in Section 3.
– **Easy to improve** The function for quickly registration of target events enables to improve the constructed system by trial and error.

3. Robust observation of human activity in handling objects

3.1 System configuration of ultrasonic 3D tag system

Figure 2 shows the system configuration for the ultrasonic 3D tag system. The system consists of an ultrasonic receiving section, an ultrasonic transmitting section, a time-of-flight measuring section, a network section, and a personal computer. The ultrasonic receiving section receives ultrasonic pulses emitted from the ultrasonic transmitter and amplifies the received signal. The time-of-flight measuring section records the travel time of the signal from transmission to reception. The network section synchronizes the system and collects time-of-flight data from the ultrasonic receiving section. The positions of objects are calculated based on more than three time-of-flight results. The sampling frequency of the proposed system is 50 Hz.

The ultrasonic tag system calculates the 3D position of an object by trilateration using three distance measurements. Two methods of multilateration are investigated for use with the proposed system: multilateration based on a least-squares method using redundant distance data, and multilateration based on robust estimation.

The room used to conduct the experiments is shown in Fig. 3. The room was $3.5 \times 3.5 \times 2.7$ m in size, and was fitted with 307 ultrasonic receivers embedded in the wall and ceiling. Tags were attached to various objects, including a cup and a stapler as shown in and Fig. refsensor-room. Some objects were fitted with two transmitters. The purpose of the experimental room is to clarify the effect of the use of redundant sensors. More than 300 receivers do not mean that the algorithms described in the next section need such a large number of sensors. In actual usage, a smaller number of receivers can be used.

3.2 Multilateration method 1: linearization of the minimization problem

The receiver position (x,y,z) is calculated by a multilateration algorithm, such as that used in the Global Positioning System(14). Trilateration or multilateration algorithms have been proposed in the field of aerospace(15; 16). This paper presents the multilateration algorithms applicable to a more general case that multiple ultrasonic receivers are put on arbitrary positions. Using distance data l_i, l_j and the receiver positions $(x_i,y_i,z_i),(x_j,y_j,z_j)$, we obtain the following spherical equations for the possible position of the target.

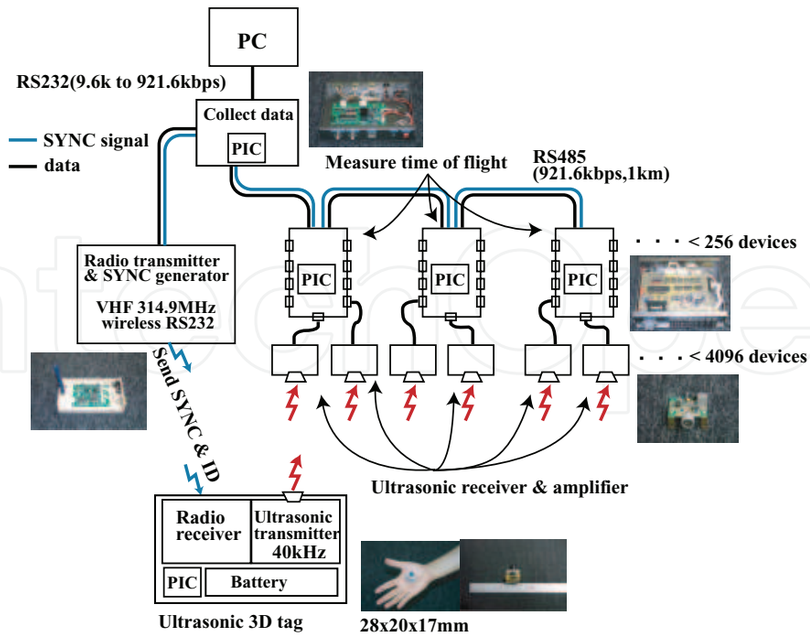


Fig. 2. System configuration of ultrasonic 3D tag system

$$(x_i - x)^2 + (y_i - y)^2 + (z_i - z)^2 = l_i^2, \tag{2}$$

$$(x_j - x)^2 + (y_j - y)^2 + (z_j - z)^2 = l_j^2. \tag{3}$$

By subtracting Eq. (3) from Eq. (2), we obtain an equation for intersecting planes between the spheres, as shown in Fig. 5.

$$\begin{aligned} &2(x_j - x_i)x + 2(y_j - y_i)y + 2(z_j - z_i)z = \\ &\quad l_i^2 - l_j^2 - x_i^2 - y_i^2 - z_i^2 + x_j^2 + y_j^2 + z_j^2 \end{aligned} \tag{4}$$

By inputting pairs of (i, j) into the above equation, we obtain simultaneous linear equations, as expressed by

$$\mathbf{AP} = \mathbf{B}, \tag{5}$$

$$\text{where} \quad \mathbf{P} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \tag{6}$$

$$\mathbf{A} = \begin{pmatrix} 2(x_0 - x_1) & 2(y_0 - y_1) & 2(z_0 - z_1) \\ 2(x_0 - x_2) & 2(y_0 - y_2) & 2(z_0 - z_2) \\ 2(x_0 - x_3) & 2(y_0 - y_3) & 2(z_0 - z_3) \end{pmatrix}, \tag{7}$$

Fig. 3. Experimental daily living space

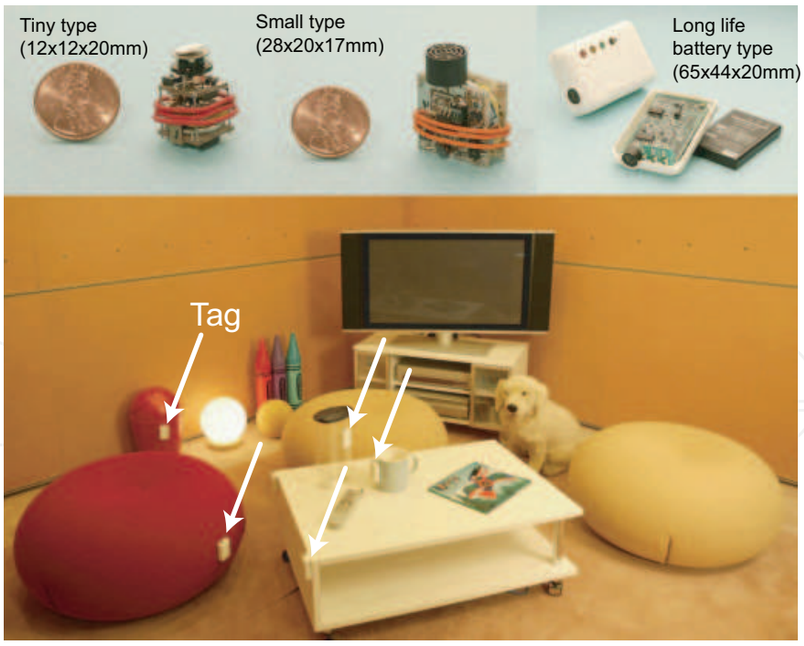


Fig. 4. Developed ultrasonic 3D tags and example of attaching tags to objects

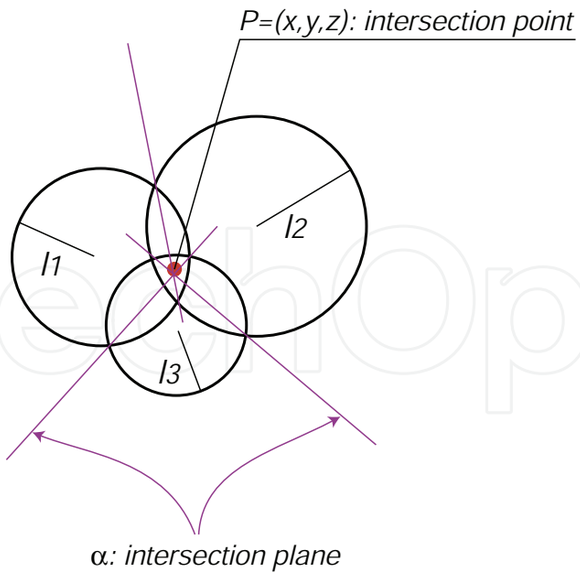


Fig. 5. Planes of intersection between spheres used to give the estimated position

$$\mathbf{B} = \begin{pmatrix} l_1^2 - l_0^2 - x_1^2 - y_1^2 - z_1^2 + x_0^2 + y_0^2 + z_0^2 \\ l_2^2 - l_0^2 - x_2^2 - y_2^2 - z_2^2 + x_0^2 + y_0^2 + z_0^2 \\ l_3^2 - l_0^2 - x_3^2 - y_3^2 - z_3^2 + x_0^2 + y_0^2 + z_0^2 \\ \vdots \end{pmatrix}. \quad (8)$$

The position $(\hat{x},\hat{y},\hat{z})$ can then be calculated by a least-squares method as follows.

$$\mathbf{P} = (\mathbf{A^T A})^{-1}\mathbf{A^T B}. \quad (9)$$

This method minimizes the square of the distance between the planes expressed by Eq. (4) and the estimated position. The algorithm is described in detail in Fig. 6. In actual usage, the rank of matrix **A** must be considered.

3.3 Multilateration method 2: robust estimation by RANSAC

Data sampled by the ultrasonic tagging system is easily contaminated by outliers due to reflections. Method 1 above is unable to estimate the 3D position with high accuracy if sampled data includes outliers deviating from a normal distribution. In the field of computer vision, robust estimation methods that are effective for sampled data including outliers have already been developed. In this work, the random sample consensus (RANSAC) (17; 18) estimator is adopted to eliminate the undesirable effects of outliers. The procedure is as follows.

1. Randomly select three distances measured by three receivers (jth trial).
2. Calculate the position (x_{cj},y_{cj},z_{cj}) by trilateration.
3. Calculate the error ε_{cji} for all receivers ($i = 0,1,...,n$) by Eq. (10), and find the median ε_{mj} of ε_{cji}.

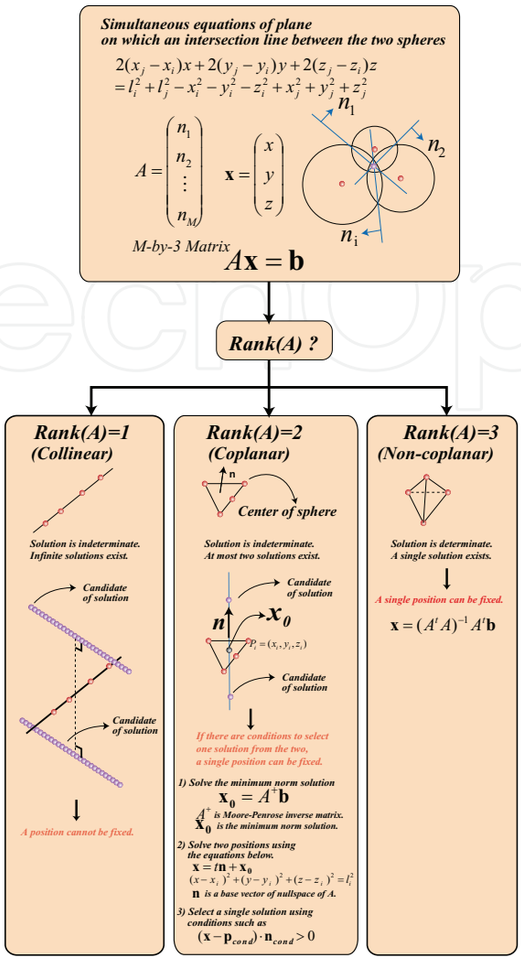


Fig. 6. Algorithm for estimating 3D position by a least-squares method considering the rank of **A**

4. Repeat steps 1 to 3 as necessary to find the combination of measurements giving the minimum error, and adopt the corresponding 3D position.

$$\varepsilon_{cji} = \left| l_i - \sqrt{(x_i - x_{mj})^2 + (y_i - y_{mj})^2 + (z_i - z_{mj})^2} \right| \tag{10}$$

$$\varepsilon_{mj} = med_j \left| \varepsilon_{cji} \right| \tag{11}$$

$$(\hat{x}, \hat{y}, \hat{z}) = min\ \varepsilon_{mj} \tag{12}$$

3.4 Resolution

Figure 7 shows the relationship between the number of receivers and the deviation of the estimated position for 4, 6, 9, 24, and 48 receivers in the ceiling. To compare the effect of the RANSAC method and that of the least-squares method, one receiver is selected randomly and 500[mm] is added to the distance data of the selected receiver as outlier. Each point was derived from 30 estimations of the position. The 5 lines in the figures represent estimation for 5 different locations of the transmitter. The resolution increases with the number of receivers,

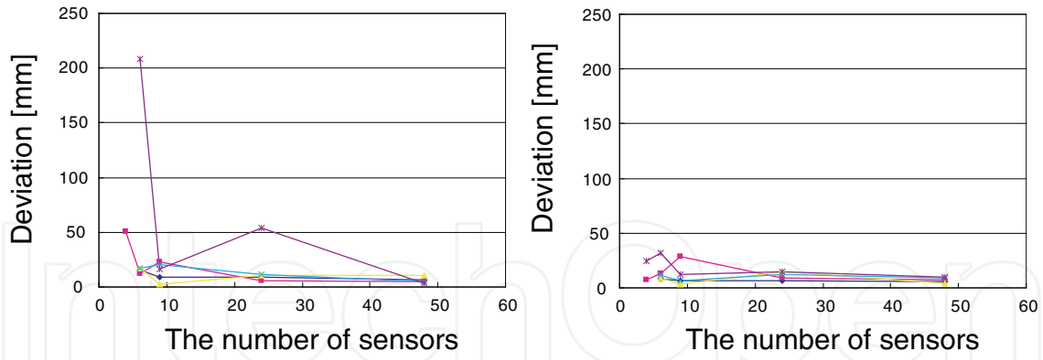


Fig. 7. Relationship between resolution and the number of sensors for the least-squares method (upper) and RANSAC (lower)

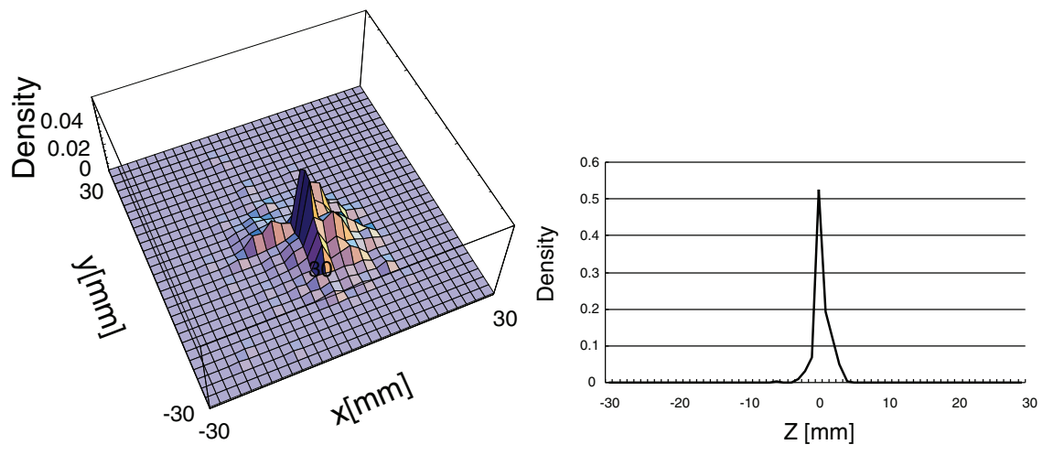


Fig. 8. Resolution in the x and y directions (upper) and z direction (lower) (grid size: 2 x 2 mm)

and the RANSAC method provides a more stable estimation with higher resolution compared to the least-squares method.

The resolution in the x, y, and z directions is illustrated in Fig. 8, which shows the probability density distribution for 1000 estimations using RANSAC. The resolution in x and y directions is about 15 mm, while that in the z direction is about 5 mm.

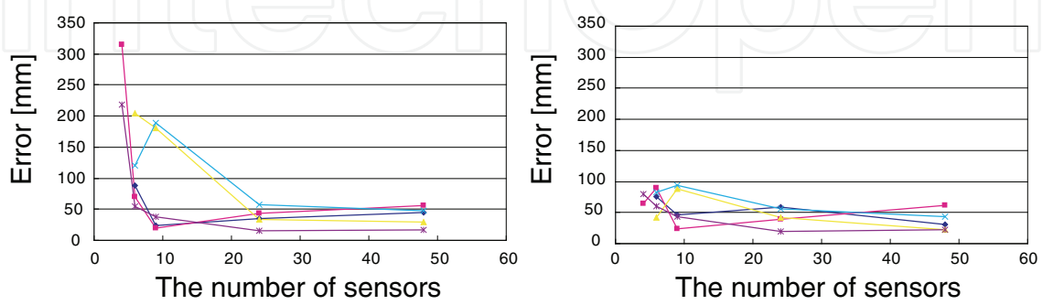


Fig. 9. Relationship between positioning accuracy and the number of receivers for the least-squares method (upper) and RANSAC (lower)

3.5 Positioning accuracy

Figure 9 shows the relationship between the number of receivers and the error of the estimated position for 4, 6, 9, 24, and 48 receivers. The error is taken as the distance from the position measured by a visual motion capture system. One receiver is selected randomly and 500[mm] is added to the distance data of the selected receiver as outlier. Each point was derived from 30 estimations of the position. The 5 lines in the figures represent estimation for 5 different locations of the transmitter. The error decreases as the number of receivers is increased, and the RANSAC method is appreciably more accurate with fewer receivers. It is considered that the least-squares method is easily affected by outliers, whereas the RANSAC method is not.
Figure 10 shows the 3D distribution of error for 1400 measured positions in the room. The figures show that the error is lowest (20–80 mm) immediately below the 48 receivers in the ceiling, increasing toward the edges of the room.
The results of experiments for evaluating accuracy and resolution demonstrate that it is possible to improve accuracy and resolution by increasing the number of receivers, and that the undesirable effect of outliers can be mitigated through the use of RANSAC estimation.

3.6 Robustness to occlusion

As in other measuring techniques such as vision-based methods, it is necessary to increase the number of sensors to solve the problem of sensor occlusion, where the line of sight to the target object is obstructed by other objects such as walls or room occupants. In the present tagging system, the problem of occlusion occurs often when a person moves or operates an object. These situations give rise to two separate problems; a decrease in the number of usable sensors for the target, and an increase in reflections due to obstruction and movement. As one of the most typical situations where occlusion occurs, this section focuses on occlusion due to a hand.
Figure 11 shows how the error increases and the number of usable sensor decreases as a hand approaches an object fitted with an ultrasonic transmitter for the least-squares and RANSAC methods. Although the error increases significantly by both methods when the hand approaches the object, the RANSAC method is much less affected than the least-squares method. This demonstrates that the proportion of outliers increases when occlusion occurs, and that RANSAC is more robust in this situation because it can mitigate the effect of such outliers.

3.7 Real-time position measurement

Figure 12 shows the measured trajectory for a person moving a cup to a chair, the floor, and a desk. The figure demonstrates that the system can robustly measure the positions of the objects in most places of the room regardless of occlusion by a hand or body.
In the current system, the sampling frequency is about 50 Hz. This frequency decreases to $50/n$ Hz when n objects are being monitored. However, it is possible to maintain a high sampling frequency by selecting which transmitters to track dynamically. For example, a transmitter can be attached to a person's wrist, and the system can select transmitters in the vicinity of the wrist to be tracked, thereby reducing the number of transmitters that need to be tracked at one time and maintaining the highest sampling frequency possible. Figure 13 shows the measured trajectory in a dynamic selection mode. The red sphere in the figure shows the position of the hand.

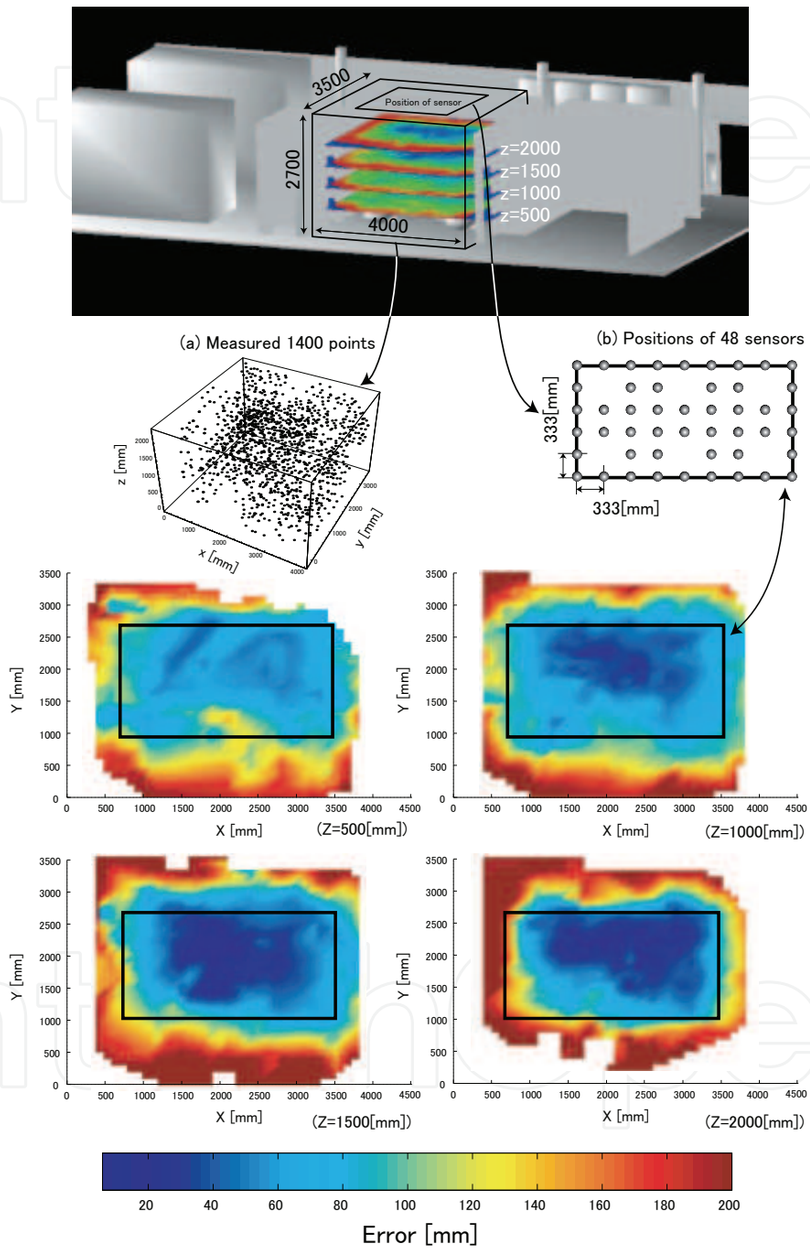


Fig. 10. 3D distribution of error in the experimental room

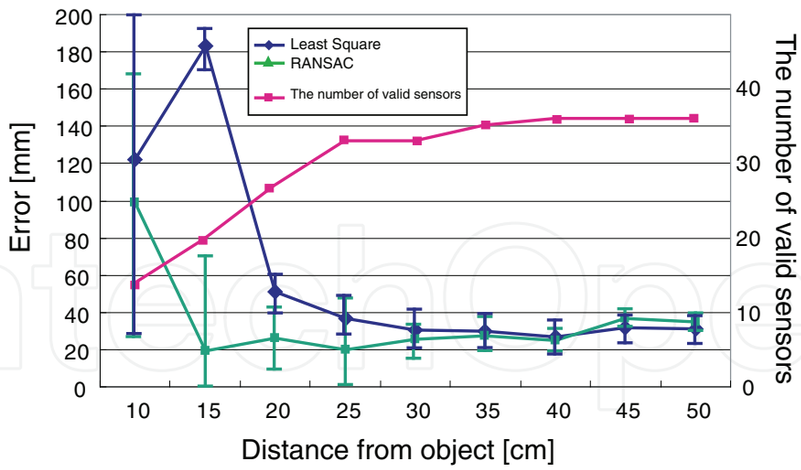


Fig. 11. Accuracy of the ultrasonic tagging system when occlusion due to a hand occurs

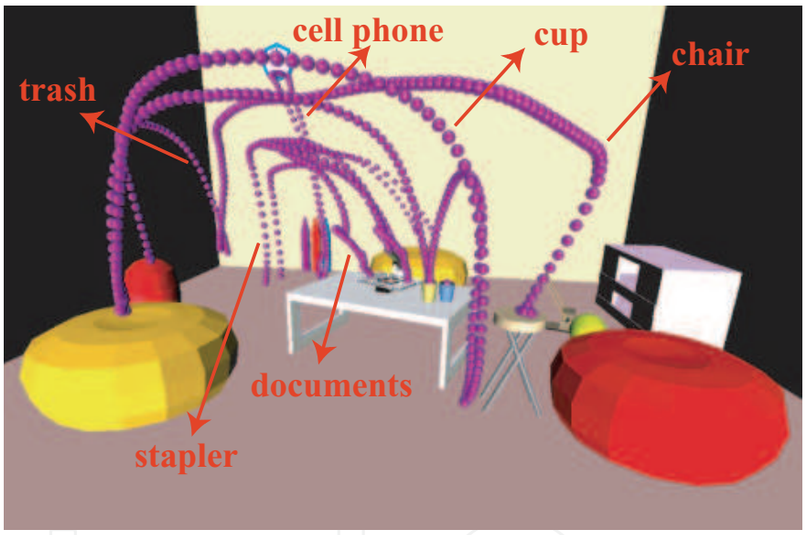


Fig. 12. Measured trajectory for movement of several objects one after another

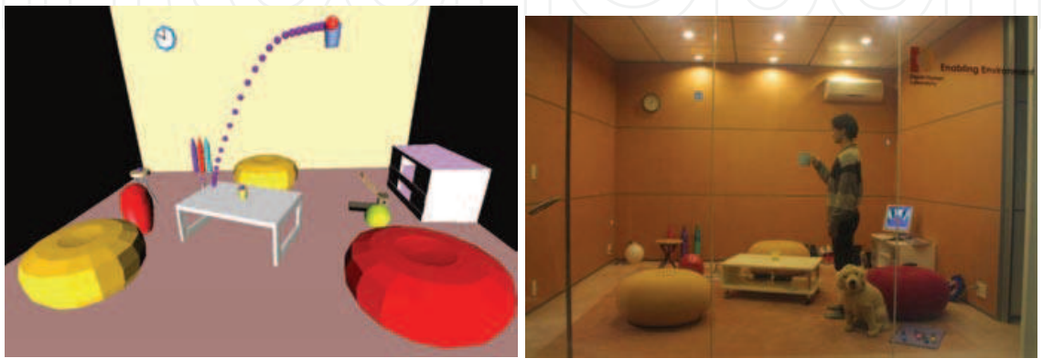


Fig. 13. Dynamic selection of transmitters

4. Quick calibration method for ultrasonic 3D tag system

4.1 Measurement and calibration

In the ultrasonic 3D tag system that the authors have developed, calibration means calculation of receivers' positions and measurement means calculation of transmitters' positions as shown in Fig. 14. Essentially, both problems are the same. As described in the previous section, the robustness of the ultrasonic 3D tag system can be improved by increasing the number of ultrasonic receivers. However, as the space where the receivers exist widens, it becomes more difficult to calibrate receivers' positions because a simple calibration method requires almost the same size of a calibration device which has multiple transmitters. This paper describes a calibration method which requires relatively small number of transmitters such as three or more and therefore doesn't require the same size of the calibration system as that of the space where the receivers exist.

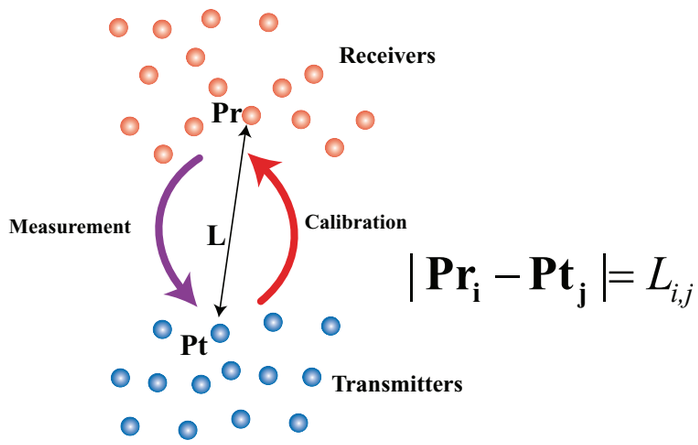


Fig. 14. Calibration and measurement

4.2 Quick calibration method

In the present paper, we describes "a global calibration based on local calibration (GCLC)" method and two constraints that can be used in conjunction with the GCLC method.
The procedure for GCLC is described below.
1. Move the calibration device arbitrarily to multiple positions (A, B, and C in Fig. 15).
2. Calculate the positions of the receivers in a local coordinate system, with the local origin set at the position of the calibration system. The calculation method was described in the previous section.
3. Select receivers for which the positions can be calculated from more than two calibration system positions.
4. Select a global coordinate system from among the local coordinate systems and calculate the positions of the calibration device in the global coordinate system using the receivers selected in Step 3. Then, calculate transformation matrices ($\mathbf{M_1}$ and $\mathbf{M_2}$ in Fig. 15).
5. Calculate the receiver positions using the receiver positions calculated in Step 2 and the transformation matrices calculated in Step 4.
Steps 4 are described in detail in the following.

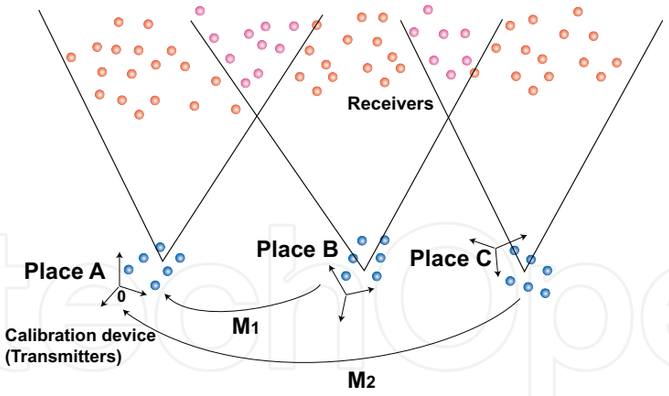


Fig. 15. Quick calibration method

4.3 Details of quick calibration

4.3.1 Calculating the positions of the calibration device in the global coordinate system (Step 4)

The error function E can be defined as follows:

$$E = \sum_{i=0}^{n} \sum_{j=i+1}^{n} ||\mathbf{M}_i \mathbf{P}_i^{(i,j)} - \mathbf{M}_j \mathbf{P}_j^{(i,j)}||^2, \tag{13}$$

where $\mathbf{M_i}$ is the transformation matrix from the local coordination system i to the global coordination system, and $\mathbf{P}_j^{(i,j)}$ denotes points in the local coordination system j for the case in which the points can be calculated in both local coordination systems i and j.

$$\begin{aligned}
&\frac{\partial E}{\partial \mathbf{M_i}} \\
&= \frac{\partial}{\partial \mathbf{M_i}} \sum_{\substack{j=0 \\ (i \neq j)}}^{n} Tr\left\{ \left(\mathbf{M}_i \mathbf{P}_i^{(i,j)} - \mathbf{M}_j \mathbf{P}_j^{(i,j)} \right)^T \left(\mathbf{M}_i \mathbf{P}_i^{(i,j)} - \mathbf{M}_j \mathbf{P}_j^{(i,j)} \right) \right\} \\
&= \frac{\partial}{\partial \mathbf{M_i}} \sum_{\substack{j=0 \\ (i \neq j)}}^{n} Tr\left\{ -(\mathbf{M}_j \mathbf{P}_j^{(i,j)})^T \mathbf{M}_i \mathbf{P}_i^{(i,j)} - (\mathbf{M}_i \mathbf{P}_i^{(i,j)})^T \mathbf{M}_j \mathbf{P}_j^{(i,j)} \right. \\
&\quad \left. + (\mathbf{M}_i \mathbf{P}_i^{(i,j)})^T \mathbf{M}_i \mathbf{P}_i^{(i,j)} + (\mathbf{M}_j \mathbf{P}_j^{(i,j)})^T \mathbf{M}_j \mathbf{P}_j^{(i,j)} \right\} \\
&= -2\mathbf{M}_0 \mathbf{P}_0^{(i,n)} (\mathbf{P}_i^{(i,n)})^T - \cdots - 2\mathbf{M}_{i-1} \mathbf{P}_{i-1}^{(i,i-1)} (\mathbf{P}_{i-1}^{(i,i-1)})^T \\
&\quad + 2\mathbf{M}_i \sum_{\substack{j=0 \\ (i \neq j)}}^{n} \mathbf{P}_i^{(i,j)} (\mathbf{P}_i^{(i,j)})^T \\
&\quad - 2\mathbf{M}_{i+1} \mathbf{P}_{i+1}^{(i,i+1)} (\mathbf{P}_i^{(i,i+1)})^T - \cdots - 2\mathbf{M}_n \mathbf{P}_n^{(i,n)} (\mathbf{P}_i^{(i,n)})^T.
\end{aligned} \tag{14}$$

If we select the local coordinate system 0 as the global coordinate system, $\mathbf{M_0}$ becomes an identity matrix. From Eq. (14), we can obtain simultaneous linear equations and calculate $\mathbf{M_i}$ using Eq. (15),

$$\begin{aligned}
&\begin{pmatrix} \mathbf{M_1} & \mathbf{M_2} & \cdots & \mathbf{M}_n \end{pmatrix} = \\
&\begin{pmatrix} \mathbf{P}_0^{(0,1)} (\mathbf{P}_1^{(0,1)})^T & \mathbf{P}_0^{(0,2)} (\mathbf{P}_2^{(0,2)})^T & \cdots & \mathbf{P}_0^{(0,n)} (\mathbf{P}_n^{(0,n)})^T \end{pmatrix} \times \\
&\begin{pmatrix}
\sum_{i=0}^{n} \mathbf{P}_1^{(1,i)} (\mathbf{P}_1^{(1,i)})^T & -\mathbf{P}_1^{(1,2)} (\mathbf{P}_2^{(1,2)})^T & \cdots & -\mathbf{P}_1^{(1,n)} (\mathbf{P}_n^{(1,n)})^T \\
-\mathbf{P}_2^{(1,2)} (\mathbf{P}_1^{(1,2)})^T & \sum_{i=0}^{n} \mathbf{P}_2^{(2,i)} (\mathbf{P}_2^{(2,i)})^T & \cdots & -\mathbf{P}_2^{(2,n)} (\mathbf{P}_n^{(2,n)})^T \\
\vdots & \vdots & \ddots & \vdots \\
-\mathbf{P}_n^{(1,n)} (\mathbf{P}_1^{(1,n)})^T & -\mathbf{P}_n^{(2,n)} (\mathbf{P}_2^{(2,n)})^T & \cdots & \sum_{i=0}^{n} \mathbf{P}_n^{(n,i)} (\mathbf{P}_n^{(n,i)})^T
\end{pmatrix}^{-1}.
\end{aligned} \tag{15}$$

4.4 Considering the environment boundary condition

Regarding the GCLC method as presented above, the error of calibration will accumulate as the space in which the ultrasonic receivers are placed becomes larger. Therefore, the number of moving calibrating devices becomes larger. For example, if we place receivers on the ceiling of a corridor of size 2 x 30 m, the accumulated error may be large. This section describes the boundary constraint with which we can reduce the error accumulation.

In most cases, the ultrasonic location system will be placed in a building or on the components of a building, such as on a wall or ceiling. If we can obtain CAD data of the building or its components or if we can measure the size of a room inside the building to a high degree of accuracy, then we can use the size data as a boundary condition for calibrating the receiver positions.

Here, let us consider the boundary constraint shown in Fig. 16. We can formulate this problem using the Lagrange's undecided multiplier method as follows:

$$E' = \sum_{i=0}^{3}\sum_{j=i+1}^{3}\left\|M_i P_i^{(i,j)} - M_j P_j^{(i,j)}\right\|^2 + \lambda F(M_3), \tag{16}$$

$$F(M_3) = (M_3 P_{b1} - P_{b0})\cdot n + l_0 - l_1 = 0 \tag{17}$$

where λ denotes a Lagrange's undecided multiplier. By solving this equation, we can obtain the following equations:

$$\begin{pmatrix} \mathbf{M_1} & \mathbf{M_2} & \mathbf{M_3} \end{pmatrix} = \begin{pmatrix} \mathbf{P}_0^{(0,1)}(\mathbf{P}_1^{(0,1)})^T & 0 & -1/2\lambda \mathbf{n}\mathbf{P}_{b1}^T \end{pmatrix} \times \begin{pmatrix} \mathbf{P}_1^{(0,1)}(\mathbf{P}_1^{(0,1)})^T + \mathbf{P}_1^{(1,2)}(\mathbf{P}_1^{(1,2)})^T & -\mathbf{P}_1^{(1,2)}(\mathbf{P}_2^{(1,2)})^T & 0 \\ -\mathbf{P}_2^{(1,2)}(\mathbf{P}_1^{(2,1)})^T & \mathbf{P}_2^{(1,2)}(\mathbf{P}_2^{(1,2)})^T + \mathbf{P}_2^{(2,3)}(\mathbf{P}_2^{(2,3)})^T & -\mathbf{P}_2^{(2,3)}(\mathbf{P}_3^{(2,3)})^T \\ 0 & -\mathbf{P}_3^{(2,3)}(\mathbf{P}_2^{(2,3)})^T & \mathbf{P}_3^{(2,3)}(\mathbf{P}_3^{(2,3)})^T \end{pmatrix}^{-1}. \tag{18}$$

By substituting $\mathbf{M_3}$ into Eq. (17), we can solve λ and eliminate it from Eq. (18).

The general case of the GCLC method with multiple boundary constraints is as follows:

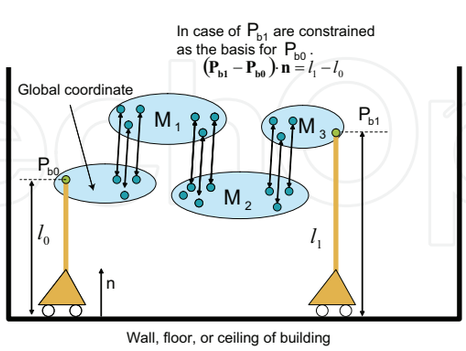


Fig. 16. Example of a boundary condition as the basis for the building

$$
\begin{pmatrix} \mathbf{M_1} & \mathbf{M_2} & \cdots & \mathbf{M}_n \end{pmatrix} =
\begin{pmatrix}
\mathbf{P}_0^{(0,1)}(\mathbf{P}_1^{(0,1)})^T & & & \mathbf{P}_0^{(0,n)}(\mathbf{P}_n^{(0,n)})^T \\
-1/2\sum_{i=0}^{n_i}\lambda_{1,i}\mathbf{n}_{1,i}\mathbf{P}_{1,i}^T & \cdots & \cdots & -1/2\sum_{i=0}^{n_n}\lambda_{n,i}\mathbf{n}_{n,i}\mathbf{P}_{n,i}^T
\end{pmatrix} \times
\begin{pmatrix}
\sum_{\substack{i=0\\ i\neq 1}}^{n}\mathbf{P}_1^{(1,i)}(\mathbf{P}_1^{(1,i)})^T & -\mathbf{P}_1^{(1,2)}(\mathbf{P}_2^{(1,2)})^T & \cdots & -\mathbf{P}_1^{(1,n)}(\mathbf{P}_n^{(1,n)})^T \\
-\mathbf{P}_2^{(1,2)}(\mathbf{P}_1^{(1,2)})^T & \sum_{\substack{i=0\\ i\neq 2}}^{n}\mathbf{P}_2^{(2,i)}(\mathbf{P}_2^{(2,i)})^T & \cdots & -\mathbf{P}_2^{(2,n)}(\mathbf{P}_n^{(2,n)})^T \\
\vdots & \vdots & \ddots & \vdots \\
-\mathbf{P}_n^{(1,n)}(\mathbf{P}_1^{(1,n)})^T & -\mathbf{P}_n^{(2,n)}(\mathbf{P}_2^{(2,n)})^T & \cdots & \sum_{\substack{i=0\\ i\neq n}}^{n}\mathbf{P}_n^{(n,i)}(\mathbf{P}_n^{(n,i)})^T
\end{pmatrix}^{-1}, \tag{19}
$$

where $\lambda_{i,j}$, $\mathbf{n_{i,j}}$, and $\mathbf{P_{i,j}}$ denote the j-th undecided multiplier, the j-th constraint vector, and the j-th constrained point in the i-th local coordinate system, respectively. In this case, the boundary constraints are as follows:

$$F_{i,j} = \left(\mathbf{M_i P_{i,j}} - \mathbf{P_{b0}}\right) \cdot \mathbf{n_{i,j}} - \Delta l_{i,j} = 0, \tag{20}$$

where $\Delta l_{i,j}$ denotes a distance constraint. The above GCLC method with boundary constraints is applicable to, for example, the case in which more complex boundary conditions exist, as shown in Fig. 17.

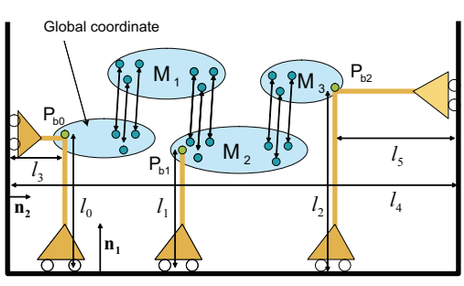


Fig. 17. Example of a greater number of boundary conditions as the basis of the building

4.5 Experimental results of GCLC

4.5.1 Method for error evaluation

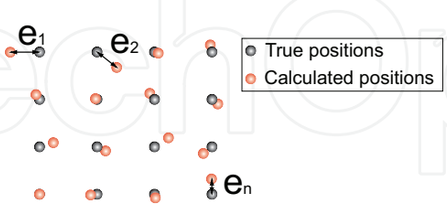


Fig. 18. Method for calculating error

Figure 18 shows the method used to calculate error. The distances between the calculated receiver positions and the true receiver positions are denoted by $e_1, e_2, \cdots, e_n$. The average error is defined by

$$E = \frac{1}{n}\sum_{i=1}^{n} e_i. \tag{21}$$

4.5.2 Accuracy evaluation

Calibration was performed in a room (4.0×4.0×2.5 m) having 80 ultrasonic receivers embedded in the ceiling. Figure 19 shows the experimental results obtained using the GCLC method without any constraints. The authors performed calibration at 16 points in the room. Seventy-six receivers were calculated. In the figure, the red spheres indicate calculated receiver positions, the black crosses indicate the true receiver positions, and the blue spheres indicate the positions of the calibration device. Figure 20 shows the experimental results for the GCLC method considering directivities. Seventy-six receivers were calculated. Table 1 shows the average error E, maximum error, and minimum error for these methods. The above results show that using the GCLC method we can calibrate the position of receivers placed in a space of average room size and that the error can be reduced significantly by considering directivity.

Another calibration was performed in a rectangular space (1.0×4.5) having a longitudinal length that is much longer than its lateral length. Seventy-six ultrasonic receivers are embedded in the space. Figure 21 shows the experimental results obtained using the GCLC method without any constraints. Seventy-five receivers were calculated. Figure 22 shows the experimental results obtained using the GCLC method with directivity consideration and a boundary constraint. Table 2 shows the average error E, maximum error, and minimum error for these methods. The above results show that with the GCLC method with directivity consideration and boundary constraint has a significantly reduced error.

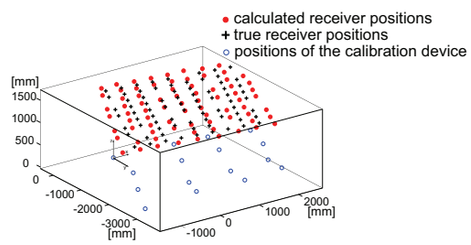


Fig. 19. Experimental result obtained by the GCLC method

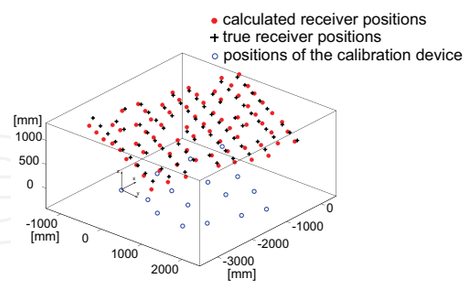


Fig. 20. Experimental result obtained by the GCLC method considering directivity

4.6 Advantages of the GCLC method

The advantages of the GCLC method are listed below.

– The method requires a relatively small number of transmitters, at least three transmitters, so that the user can calibrate the ultrasonic location system using a small calibrating device having at least three transmitters.

– The method can calibrate the positions of the receivers independent of room size.

	Ave. error	Max. error	Min. error
GCLC	195 mm	399 mm	66 mm
GCLC with directivity consideration	75 mm	276 mm	9 mm

Table 1. Errors (mm) of the proposed method for the case of a square-like space

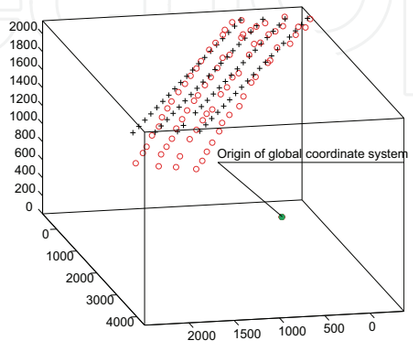


Fig. 21. Experimental results obtained by the GCLC method

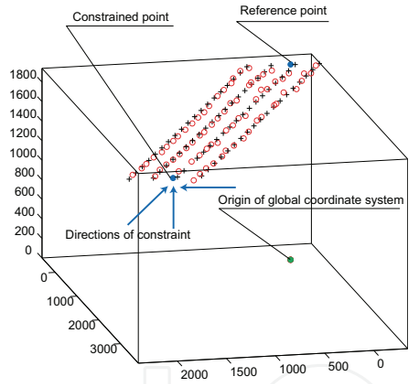


Fig. 22. Experimental results obtained by the GCLC method with directivity consideration and a boundary constraint

	Ave. error	Max. error	Min. error
GCLC	236 mm	689 mm	17 mm
GCLC with directivity consideration and boundary constraint	51 mm	121 mm	10 mm

Table 2. Errors (mm) of the proposed method for the case of a rectangular space having a longitudinal length that is much longer than its lateral length

– The error can be reduced by considering the directivity constraint. The constraint is useful for cases in which the ultrasonic location system adopts a method in which the time-of-fight is detected by thresholding ultrasonic pulse.
– The error can be reduced by considering the boundary constraint. The constraint is useful for cases in which the receivers to be calibrated are placed in a rectangular space having a longitudinal length that is much greater than the lateral length, such as a long corridor.

4.7 Development of Ultrasonic Portable 3D Tag System

The GCLC method enables a portable ultrasonic 3D tag system. Figure 23 shows a portable ultrasonic 3D tag system, which consists of a case, tags, receivers, and a calibration device. The portable system enables measurement of human activities by quickly installing and calibrating the system on-site, at the location where the activities actually occur.

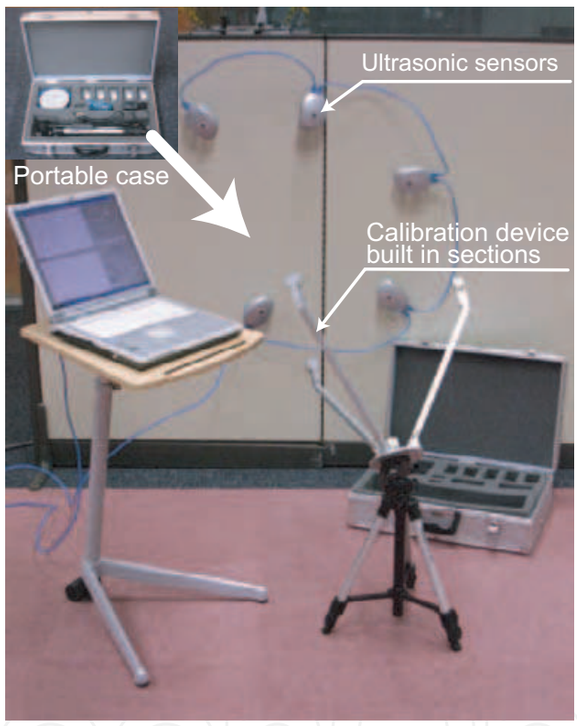


Fig. 23. Developed portable ultrasonic 3D tag system

5. Quick registration of human activity events to be detected

This section describes quick registration of target human activity events. Quick registration is performed using a stereoscopic camera with ultrasonic 3D tags as shown in Fig. 24 and interactive software. The features of this function lie in simplification of 3D shape, and simplification of physical phenomena relating to target events. The software abstracts the shapes of objects in real world as simple 3D shape such as lines, circles, or polygons. In order to describe the real world events when a person handles the objects, the software abstracts the function of objects as simple phenomena such as touch, detouch, or rotation. The software adopts the concept of virtual sensors and effectors to enable for a user to define the function of the objects easily by mouse operations.

For example, if a person wants to define the activity of ”put a cup on the desk”, firstly, the person simplifies the cup and the desk as a circle and a rectangle respectively using a photo-modeling function of the software. Second, using a function for editting virtual sensors, the person adds a touch type virtual sensor to the rectangle model of the desk, and adds a bar type effector to the circle model of the cup.

5.1 Software for quick registration of human activity events to be detected

5.1.1 Creating simplified 3D shape model

Figure 26 shows examples of simplified 3D shape models of objects such as a Kleenex, a cup, a desk and stapler. The cup is expressed as a circle and the desk is a rectangle. The simplification is performed using a stereoscopic camera with the ultrasonic 3D tags and a photo-modeling function of the software. Since the camera has multiple ultrasonic 3D tags, the system can track its position and posture. Therefore, it is possible to move the camera freely when the user creates simplified 3D shape models and the system can integrate the created 3D shape models in a world coordinate system.

5.1.2 Creating model of physical object's function using virtual sensors/effectors

The software creates the model of a object's function by attaching virtual sensors/effectors which are prepared in advance in the software to the 3D shape model created in step (a). Virtual sensors and effectors work as sensors and ones affecting the sensors on computer. The current system has ”angle sensor” for detecting rotation, ”bar effector” for causing phenomenon of touch, ”touch sensor” for detecting phenomenon of touch. In the right part of Fig. 27, red bars indicate a virtual bar effector, and green area indicates a virtual touch sensor. By mouse operations, it is possible to add virtual sensors/effectors to the created 3D shape model.

5.1.3 Associating output of model of physical object's function with activity event

Human activity can be described using output of the virtual sensors which are created in Step (b). In Fig. 28, red bar indicates that the cup touches with the desk and blue bar indicates that the cup doesn't touch with the desk. By creating the table describing relation between the output of the virtual sensors and the target events, the system can output symbolic information such as ”put a cup on the desk” when the states of virtual sensors change.

5.1.4 Detecting human activity event in real time

When the software inputs position data of ultrasonic 3D tag, the software can detect the target events using the virtual sensors and the table defined in Step (a) to (c) as shown in Fig. 29

6. Conclusion

This paper described a system for quickly realizing a function for robustly detecting daily human activity events in handling objects in the real world. The system has three functions: 1) robustly measuring 3D positions of the objects, 2) quickly calibrating a system for measuring 3D positions of the objects, 3) quickly registering target activity events, and 4) robustly detecting the registered events in real time.

As for 1), In order to estimate the 3D position with high accuracy, high resolution, and robustness to occlusion, the authors propose two estimation methods, one based on a least-squares approach and one based on RANSAC.

Fig. 24. UltraVision (a stereoscopic camera with the ultrasonic 3D tags) for creating simplified 3D shape model

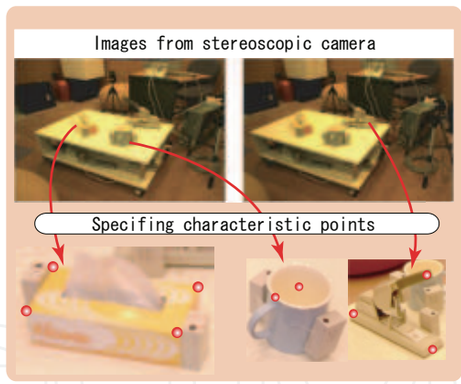


Fig. 25. Photo-modeling by stereoscopic camera system

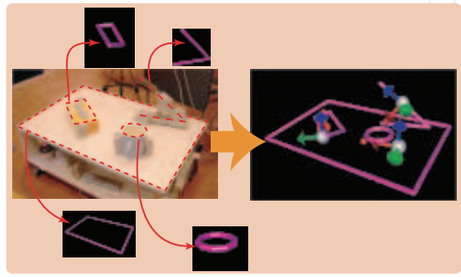


Fig. 26. Create simplified shape model

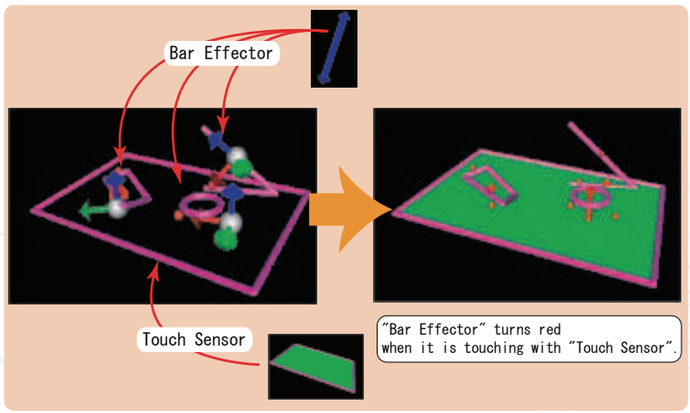


Fig. 27. Create model of physical object's function using virtual sensors/effectors

The system was tested in an experimental room fitted with 307 ultrasonic receivers; 209 in the walls and 98 in the ceiling. The results of experiments conducted using 48 receivers in the ceiling for a room with dimensions of $3.5 \times 3.5 \times 2.7$ m show that it is possible to improve the accuracy, resolution, and robustness to occlusion by increasing the number of ultrasonic receivers and adopting a robust estimator such as RANSAC to estimate the 3D position based on redundant distance data. The resolution of the system is 15 mm horizontally and 5 mm vertically using sensors in the ceiling, and the total spatially varying position error is 20–80 mm. It was also confirmed that the system can track moving objects in real time, regardless of obstructions.

As for 2), this paper described a new method for quick calibration. The method uses a calibration device with three or more ultrasonic transmitters. By arbitrarily placing the device at multiple positions and measuring distance data at their positions, the positions of receivers can be calculated. The experimental results showed that with the method, the positions of 80 receivers were calculated by 4 transmitters of the calibration device and the position error is 103 mm.

As for 3), this paper described a quick registration of target human activity events in handling objects. To verify the effectiveness of the function, using a stereoscopic camera with ultrasonic 3D tags and interactive software, the authors registered activities such as "put a cup on the

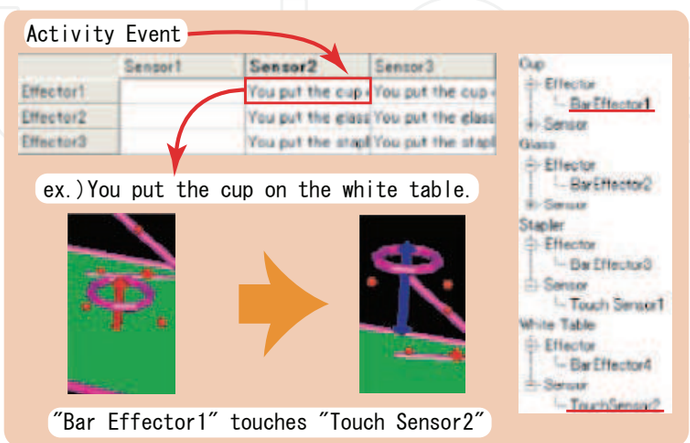


Fig. 28. Associate output of virtual sensors with target activity event

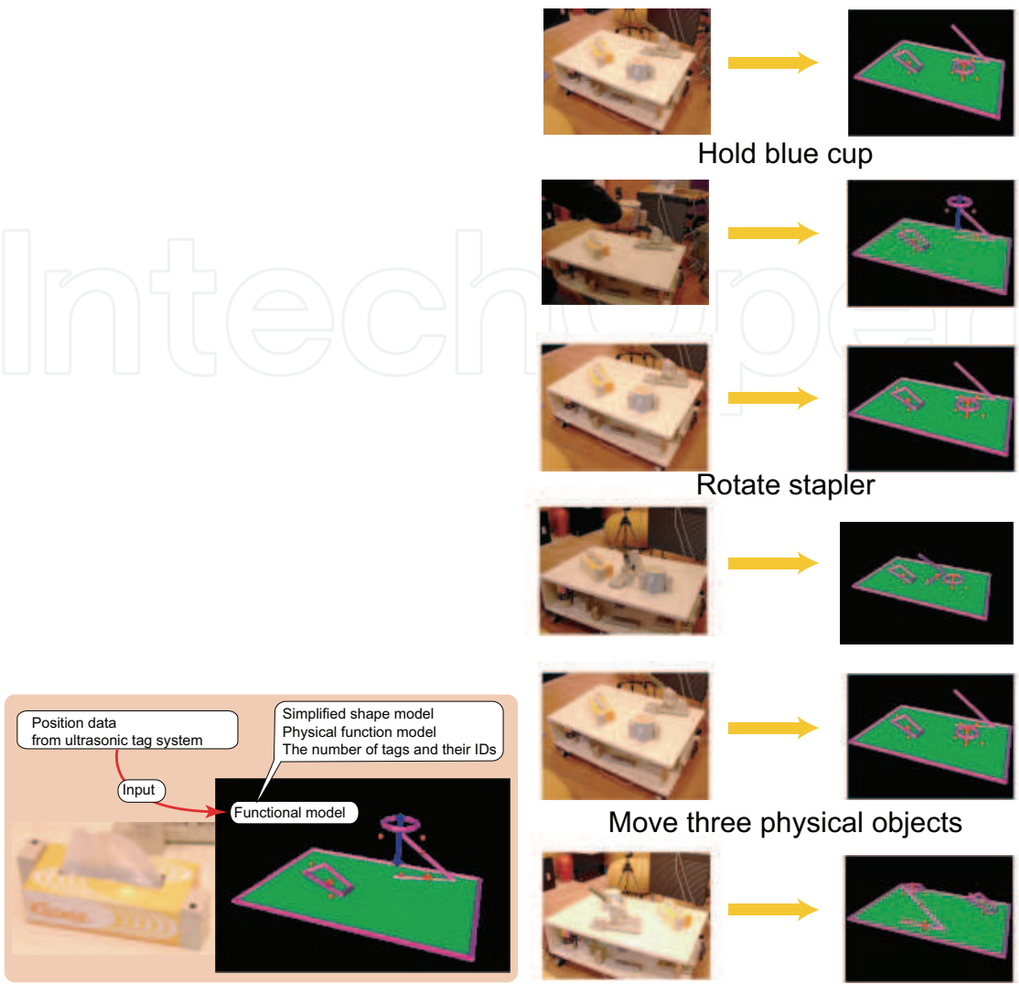


Fig. 29. Recognize human activity in real time by function's model

desk" and "staple document" through creating the simplified 3D shape models of ten objects such as a TV, a desk, a cup, a chair, a box, and a stapler.
Further development of the system will include refinement of the method for measuring the 3D position with higher accuracy and resolution, miniaturization of the ultrasonic transmitters, development of a systematic method for defining and recognizing human activities based on the tagging data and data from other sensor systems, and development of new applications based on human activity data.

7. References

[1] T. Hori. Overview of Digital Human Modeling. *Proceedings of 2000 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS2000), Workshop Tutorial Note,* pp. 1–14, 2000

[2] H. Mizoguchi, T. Sato, and T. Ishikawa. Robotic Office Room to Support Office Work by Human Behavior Understanding Function with Networked Machines. *IEEE/ASME Transactions on Mechatronics*, Vol. 1, No. 3, pp. 237–244, September 1996

[3] Y. Nishida, H. Aizawa, T. Hori, N.H. Hoffman, T. Kanade, M. Kakikura, "3D Ultrasonic

Tagging System for Observing Human Activity, " in *Proceedings of IEEE International Conference on Intelligent Robots and Systems* (*IROS2003*), pp. 785-791, October 2003.

[4] A. Ward, A. Jones, A. Hopper, "A New Location Technique for the Active Office, " *IEEE Personal Communications*, Vol. 4, No. 5, pp. 42-47, October 1997.

[5] A. Harter, A. Hopper, P. Steggles, A. Ward, P. Webster, "The Anatomy of a Context-Aware Application, " in *Proceedings of the ACM/IEEE MobiCom*, August 1999.

[6] M. Addlesee, R. Curwen, S. Hodges, J. Newman, P. Steggles, A. Ward, A. Hopper, "Implementing a sentient computing system, " *IEEE Computer*, Vol. 34, No. 8, pp. 50-56, August 2001.

[7] M. Hazas and A. Ward, "A Novel Broadband Ultrasonic Location System, " in *Proceedings of UbiComp 2002*, pp. 264-280, September 2002.

[8] N.B. Priyantha, A. Chakraborty, H. Balakrishnan, "The Cricket Location-Support system, " in *Proceedings of the 6th International Conference on Mobile Computing and Networking* (*ACM MobiCom2000*), pp. 32-43, August 2000

[9] A. Mahajan and F. Figueroa, "An Automatic Self Installation and Calibration Method for a 3D Position Sensing System using Ultrasonics," *Robotics and Autonomous Systems*, Vol. 28, No. 4, pp. 281-294, September 1999.

[10] Y. Fukuju, M. Minami, H. Morikawa, and T. Aoyama, "DOLPHIN: An Autonomous Indoor Positioning System in Ubiquitous Computing Environment, " in *Proceedings of IEEE Workshop on Software Technologies for Future Embedded Systems* (*WSTFES2003*), pp. 53-56, May 2003.

[11] P. Duff, H. Muller, "Autocalibration Algorithm for Ultrasonic Location Systems," in *Proceedings of 7th IEEE International Symposium on Wearable Computer*, pp. 62-68, October 2003.

[12] Y. Chen, G. Medioni, "Object Modeling by registration of multiple range images," *Image and Vision Computing,* Vol. 10, No. 3, pp. 145-155, April 1992.

[13] P.J. Neugebauer, "Geometrical Cloning of 3D Objects via Simultaneous Registration of Multiple Range Images," in *Proceedings of the 1997 International Conference on Shape Modeling and Application* (*SMA'97*), pp. 130-139, 1997.

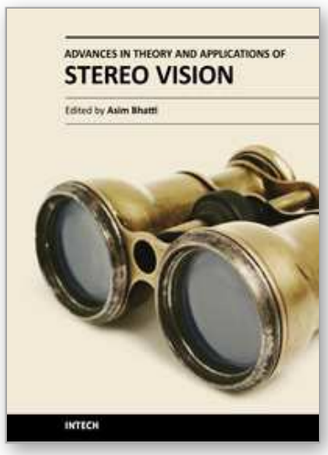
[14] B.W. Parkinson, J.J. Spilker, P. Axelrad, P. Enge, *The Global Positioning System: Theory and Applications,* American Institute of Aeronautics and Astronautics, 1996.

[15] K.C. Ho. Solution and Performance Analysis of Geolocation by TDOA. *IEEE Transaction on Aerospace and Electronic Systems*, Vol. 29, No. 4, pp. 1311–1322, October 1993.

[16] D.E. Manolakis, "Efficient Solution and Performance Analysis of 3-D Position Estimation by Trilateration, " *IEEE Trans. on Aerospace and Electronic Systems*, Vol. 32, No. 4, pp. 1239–1248, October 1996

[17] P. J. Rousseeuw, and A. M. Leroy. *Robust Regression and Outlier Detection.* Wiley, New York, 1987.

[18] M.A. Fishler, and R.C. Bolles. Random Sample Consensus: A Paradigm for Model Fitting with Application to Image Analysis and Automated Cartography. *Communication of the ACM*, Vol. 24, No. 6, pp. 381–395, June 1981.



Advances in Theory and Applications of Stereo Vision
Edited by Dr Asim Bhatti

ISBN 978-953-307-516-7
Hard cover, 352 pages
Publisher InTech
Published online 08, January, 2011
Published in print edition January, 2011

The book presents a wide range of innovative research ideas and current trends in stereo vision. The topics covered in this book encapsulate research trends from fundamental theoretical aspects of robust stereo correspondence estimation to the establishment of novel and robust algorithms as well as applications in a wide range of disciplines. Particularly interesting theoretical trends presented in this book involve the exploitation of the evolutionary approach, wavelets and multiwavelet theories, Markov random fields and fuzzy sets in addressing the correspondence estimation problem. Novel algorithms utilizing inspiration from biological systems (such as the silicon retina imager and fish eye) and nature (through the exploitation of the refractive index of liquids) make this book an interesting compilation of current research ideas.

INTECH
open science | open minds

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

IntechOpen