

3D Topological Reconstruction based on Hough Transform and Growing Neural Gas for Informationally Structured Space

Naoyuki Kubota, *Member, IEEE*, Tsubasa Narita, *Student Member, IEEE*,
Beom Hee Lee, *Fellow, IEEE*

Abstract—This paper proposes a method of 3D topological reconstruction for informationally structured space including sensor networks and robot partners for co-existing with people. The informationally structured space realizes the quick update and access of valuable and useful information for both people and robots on real and virtual environments. In this paper, we use distance information and color information measured by 3D distance image sensor and CMOS camera for 3D topological reconstruction. First, we propose an extraction method of objects from the background image based on Hough transform as preprocessing. Next, we propose a method of 3D topological reconstruction based on growing neural gas to construct informationally structured space. Finally, we show experimental results of the proposed method and discuss the effectiveness of the proposed method.

I. INTRODUCTION

Recently, various types of care systems have been developed as the increase of population of elderly people. One of the most popular care systems is a camera-based monitoring system, but the systems have problems of privacy and security in the network communication. Therefore, we should consider human-friendly, safe, and secure care systems for the aging society.

The emerging synthesis of information technology (IT), network technology (NT), and robot technology (RT) is one of the most promising approaches to realize a safe, secure, and comfortable society for the next generation [1-3,7,22,23]. NT can provide the robot with computational capabilities based on various types of information outside of robots. Actually the robot directly receives the environmental information through a local area network without the measurement by the robot itself. Wireless sensor networks [4-6] realize to gather the huge data on environments. However, it is very difficult to store all of huge data in real time. Furthermore, some features should be extracted from the gathered data to obtain the required information. Therefore, intelligent technology is required in wireless sensor networks. Intelligence technology and information technology have been discussed from various points of view. Information resources and the accessibility within an environment are essential for both people and robots.

Manuscript received March 9, 2010. N. Kubota is with Dept. of System Design, Tokyo Metropolitan University, 6-6 Asahigaoka, Hino, Tokyo, Japan (phone: +81-42-585-8441; fax: +81-42-585-8441; e-mail: kubota@tmu.ac.jp); T.Narita (e-mail: narita-tsubasa@sd.tmu.ac.jp); B.H.Lee is with School of Electrical Engineering and Computer Science, Seoul National University, Seoul, Korea (phone: +822-880-7311; bhlee@snu.ac.kr)

Therefore, the environment surrounding people and robots should have a structured platform for gathering, storing, transforming, and providing information. Such an environment is called informationally structured space [22,23] (Fig.1). The intelligent technology for the design and usage of the informationally structured space should be discussed from various points of view such as (1) data gathering of real environment and cyber space, (2) information extraction, (3) structuralization, (4) information visualization and display, (5) information search, and (6) information operations. The structuralization of informationally structured space realizes the quick update and access of valuable and useful information for both people and robots on real and virtual environments. The information is transformed into the useful form suitable to the features of robots and people. Furthermore, if the robot can share the environmental information with people, the communication with people might become very smooth and natural.

We proposed informationally structured space including sensor networks and robot partners for co-existing with people. The sensor network includes distance sensor, accelerometer, cameras, laser oscillators, pneumatic sensors, and illumination sensors. The measured data are converted into meaningful information based on the environmental information on layout and location of furniture and home appliances in a room. Furthermore, we proposed a 3D reconstruction method based on growing neural gas (GNG) for informationally structured space [7]. GNG is one of self-organizing neural network based on unsupervised learning [15-17]. The measured data are transformed into 3D colored topological model structure, and the spatial relationship of objects in the view is extracted from distance information. However, it is difficult to extract objects from the background environmental information such as wall and floor in this method. Therefore, we propose an extraction method of objects from the background image based on Hough transform as preprocessing. We can obtain 3D model by GNG after the separation based on Hough transform. Furthermore, the obtained map after removing the extracted objects can be used for topological roadmap for path planning of mobile robots.

This paper is organized as follows. Section 2 explains the data flow in the informationally structured space, mobile robots, and sensors used for 3D topological reconstruction. Section 3 proposes a method of integration of distance information and color information, the theory of 3D Hough transform to detect flat surface and curved surface, and the

algorithm of GNG. Section 4 shows experimental results of the proposed method. Section 5 summarizes the paper, and discusses the future works of this research.

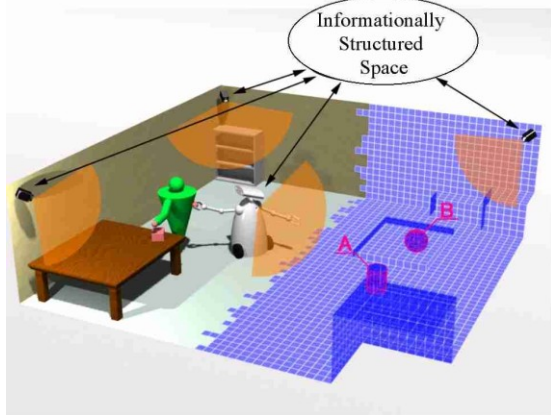


Fig.1. The concept of informationally structured space

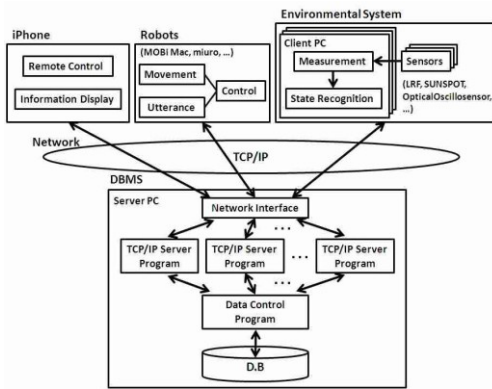


Fig.2. Data flow in informationally structured space

II. INFORMATIONALLY STRUCTURED SPACE

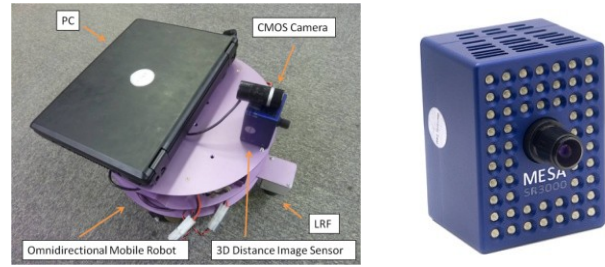
A. Data Flow in Informationally Structured Space

Figure 2 shows the data flow in the developed system based on informationally structured space. The developed system is divided into four components; (1) database management server, (2) robot systems, (3) environmental systems, and (4) human interface systems (iPhone). The environmental system is based on a ubiquitous wireless sensor network composed of sensors equipped with wall, floor, ceiling, furniture, and home appliances. These sensors measure the environmental data and human motions. The measured data are transmitted to the database server, and then feature extraction is performed. Each robot can receive the environmental information from the database server, and serve as a partner to people. Furthermore, the user interface device is used for the person to access the environmental information. In this paper, we focus on 3D topological model reconstruction using distance and color data measured by sensors in the environmental systems.

A. Reconstruction based on Distance Information

We use an omnidirectional mobile robot, MDT-R0-03, (Fig.3) developed by Mitsubishi Electric TOKKI Systems Corporation in order to perform 3D topological model

reconstruction. This robot is equipped with four independent omnidirectional wheels. The specification of this robot is shown in Table 1. We attach a computer for control, CMOS camera, and 3D distance image sensor (see Fig.3(a)).



(a) omnidirectional mobile robot (b) 3D distance image sensor

Fig.3. robot and sensor

Table.1 Specification of the omnidirectional mobile robot

Diameter	300 mm
Height	177 mm
Weight	8 kg (approximately)
Maximal Speed	1.5 km/h
Operating Time	1 hour
Maximal Payload Weight	15 kg
Communication Methods	Wi-Fi (2.4GHz)

Figure 3(b) shows a 3D distance image sensor, SR-3000. The SR-3000 is developed by MESA Corporation. SR-3000 is a range camera for measuring 3-dimensional distance up to 7.5 [m] based on time of flight principal by using infrared light source, and outputs the measuring data by a USB2.0 interface. The data size of SR-3000 is corresponding to distance information of 25344 directions because this sensor has a spatial QCIF resolution (176 * 144 pixels). The type of measured data is distance information transformed from the measurement time based on theory of the time-of-flight, but the distance information is further transformed into the position in the Cartesian space according to the measurement direction. Therefore, the output data is composed of position data in a Cartesian axis and luminance data (x, y, z, α) .

In computer vision, 3D reconstruction is used for the process of capturing the shape, appearance, and features of real objects [8-12]. Figure 4 shows 3D distance image of a person measured by SR-3000. The luminance image is represented by the data array on the camera image, but three other images are composed of points in the 3D Cartesian space. Therefore, the shape information should be extracted from the set of points in order to understand the feature of images. In the next section, we explain how to build up 3D topological models based on the measured data of distance information.

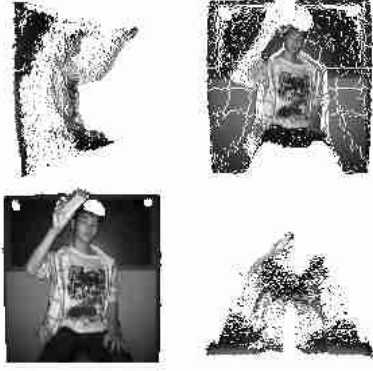


Fig.4. 3D distance image measured by SR-3000 (side view, front view, luminance image, and bottom view)

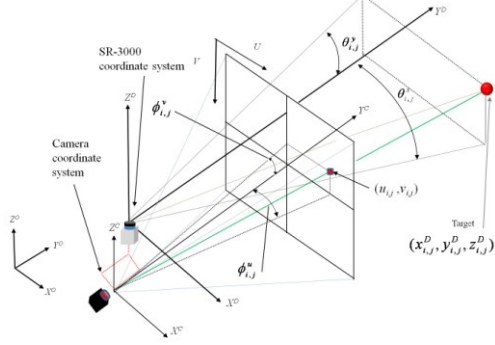


Fig.5. Synthesis of color and distance Information

III. 3D TOPOLOGICAL RECONSTRUCTION BASED ON HOUGH TRANSFORM AND GROWING NEURAL GAS

A. Synthesis Processing of Color and Distance Information

This section explains how to construct a 3D model using the color information measured by CMOS camera and the relative position (distance information) measured by 3D distance image sensor. Figure 5 shows the spatial relationship of coordinate systems used in the proposed method. The relative position measured by 3D distance image sensor is represented as $(x^D_{i,j}, y^D_{i,j}, z^D_{i,j})$ where the measurement point (pixel) on the camera image of the 3D distance image sensor is (i, j) ; the central position of lens is the origin. The position of the measurement point in the global (world) coordinate system is calculated by the following equation,

$$\begin{bmatrix} x_{i,j} \\ y_{i,j} \\ z_{i,j} \\ 1 \end{bmatrix} = \begin{bmatrix} \cos\theta & 0 & -\sin\theta & x^R \\ 0 & 1 & 0 & y^R \\ \sin\theta & 0 & \cos\theta & z^R \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x^D_{i,j} \\ y^D_{i,j} \\ z^D_{i,j} \\ 1 \end{bmatrix} \quad (1)$$

where the position of 3D distance image sensor equipped at the robot in the global coordinate system is (x^R, y^R, z^R) , and the rotation angle of robot is θ . We assume the position of the pixel on the camera image corresponding to the position of the measurement point (i, j) is $(u_{i,j}, v_{i,j})$. Furthermore, the position of the point

$$\begin{bmatrix} x^C_{i,j} \\ y^C_{i,j} \\ z^C_{i,j} \end{bmatrix} = \begin{bmatrix} x^D_{i,j} + x^C \\ y^D_{i,j} + y^C \\ z^D_{i,j} + z^C \end{bmatrix} \quad (2)$$

where the position of CMOS camera on the robot in the SR-3000 coordinate system is (x^C, y^C, z^C) . In other words, (x^C, y^C, z^C) means translation vector of SR-3000 and CMOS camera. In Fig.6, the angle to the point from the origin against the x axis is calculated as

$$\phi^u_{i,j} = \tan^{-1} \left(\frac{2 \left(u_{i,j} - \frac{N}{2} \right) \cdot \tan \frac{\alpha}{2}}{N} \right) \quad (3)$$

where the image angle of the CMOS camera is α ; the number of pixels of the horizontal direction in the camera image is N . The angle to the point from the origin against the y axis is calculated as

$$\phi^v_{i,j} = \tan^{-1} \left(\frac{2 \left(v_{i,j} - \frac{M}{2} \right) \cdot \tan \frac{\beta}{2}}{M} \right) \quad (4)$$

where the image angle of the CMOS camera is β ; the number of pixels of the horizontal direction in the camera image is M . On the other hand, the spatial relationship is given in the following;

$$\theta^x_{i,j} = \tan^{-1} \left(\frac{x^C_{i,j}}{z^C_{i,j}} \right) \quad (5)$$

$$\theta^y_{i,j} = \tan^{-1} \left(\frac{y^C_{i,j}}{z^C_{i,j}} \right) \quad (6)$$

In the above formulation, because $\phi^u_{i,j} = \theta^x_{i,j}$ and $\phi^v_{i,j} = \theta^y_{i,j}$, we obtain the following relationship;

$$(u_{i,j}, v_{i,j}) = \left(\frac{N}{2 \tan \frac{\alpha}{2}} \cdot \frac{x^C_{i,j}}{z^C_{i,j}} + \frac{N}{2}, \frac{M}{2 \tan \frac{\beta}{2}} \cdot \frac{y^C_{i,j}}{z^C_{i,j}} + \frac{M}{2} \right) \quad (7)$$

Accordingly, we obtain the relationship between the position of the measurement point in 3D distance image sensor and that in camera image.

B. Hough Transform

Hough transform is often used for extracting flat surface, curved surface, and arbitrary surface of shapes [13, 14]. Hough transform can find surface candidates by a voting procedure. The voting procedure is performed in a parameter space by counting a set of parameters on discrete parameter space transformed sequentially from the point on the original space.

First, we explain how to detect a flat based on Hough transform (Fig.6). A point, (x, y, z) in the 3D distance image is transformed into the set of parameters, (θ, ϕ, ρ) in Fig.6 (a).

$$\rho = x \cos \theta \cos \phi + y \sin \theta \cos \phi + z \sin \phi \quad (8)$$

After the voting procedure is performed, we can obtain plain candidates with selecting the points of with the high counts step by step.

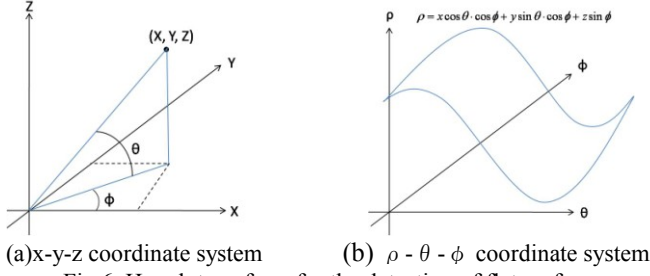


Fig.6. Hough transform for the detection of flat surfaces

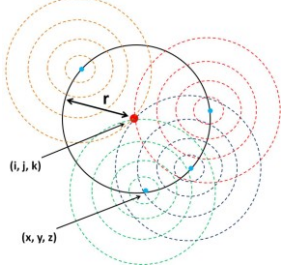


Fig.7. Hough transform for the detection of curved surfaces

Next, we explain how to detect a sphere where the center of sphere is (i, j, k) with the radius r . The relationship between a point (x, y, z) in the 3D Cartesian space and (i, j, k, r) is expressed as follows;

$$(i-x)^2 + (j-y)^2 + (k-z)^2 = r^2 \quad (9)$$

Therefore, a point, (x, y, z) in the 3D distance image is transformed into the set of parameters, (x, y, z, r) in Fig.7. After the voting procedure is performed, we can obtain the set of center and radius of sphere candidates with selecting the points of with the high counts step by step. In order to detect a cylinder, we can choose a continuous set of (x, y, z, r) where r is the constant on the parameter space of (x, y, z, r) . Furthermore, in order to detect a circular cone, we can choose a continuous set of (x, y, z, r) where r is increasing or decreasing on the parameter space of (x, y, z, r) .

Accordingly, we can distinguish objects from the wall and floor by using the above Hough transform.

C. Growing Neural Gas

In this section, we explain the learning algorithm of GNG[15,16] for 3D topological reconstruction. The notation used in GNG is shown as follows;

- w_i : the n th dimensional vector of a node
- A : A set of nodes
- N_i : A set of nodes connected to the i th node
- c : A set of edges
- $a_{i,j}$: Age of the edge between the i th and j th nodes

Step 0. Generate two units at random position, w_1, w_2 in \mathbf{R}^n where n is the dimension of input data. Initialize the connection set.

Step 1. Generate at random an input data v . In this paper, we use the position (distance) information (x, y, z) of a point measured by the 3D distance image sensor.

Step 2. Select the nearest unit (winner) s_1 and the second-nearest unit s_2 from the set of nodes by

$$s_1 = \arg \min_{i \in A} \|v - w_i\| \quad (10)$$

$$s_2 = \arg \min_{i \in A \setminus \{s_1\}} \|v - w_i\| \quad (11)$$

Step 3. If a connection between s_1 and s_2 does not yet exist, create the connection ($c_{s_1, s_2} = 1$). Set the age of the connection between s_1 and s_2 at zero;

$$a_{s_1, s_2} = 0 \quad (12)$$

Step 4. Add the squared distance between the input data and the winner to a local error variable;

$$E_{s_1} \leftarrow E_{s_1} + \|v - w_{s_1}\|^2 \quad (13)$$

Step 5. Update the reference vectors of the winner and its direct topological neighbors by the learning rate η_1 and η_2 , respectively, of the total distance to the input data.

$$w_{s_1} \leftarrow w_{s_1} + \eta_1 \cdot (v - w_{s_1}) \quad (14)$$

$$w_j \leftarrow w_j + \eta_2 \cdot (v - w_j) \text{ if } c_{s_1, j} = 1 \quad (15)$$

Step 6. Increment the age of all edges emanating from s_1 .

$$a_{s_1, j} \leftarrow a_{s_1, j} + 1 \text{ if } c_{s_1, j} = 1 \quad (16)$$

Step 7. Remove edges with an age larger than a_{\max} . If this results in units having no more connecting edges, remove those units as well.

Step 8. If the number of input data generated so far is an integer multiple of a parameter λ , insert a new unit as follows.

i. Select the unit q with the maximal accumulated error.

$$q = \arg \max_{i \in A} E_i \quad (17)$$

ii. Select the unit f with the maximal accumulated error among the neighbors of q .

iii. Add a new unit r to the network and interpolate its reference vector from q and f .

$$w_r = 0.5 \cdot (w_q + w_f) \quad (18)$$

iv. Insert edges connecting the new unit r with units q and f , and remove the original edge between q and f .

v. Decrease the error variables of q and f by a temporal discounting rate α .

$$E_q \leftarrow E_q - \alpha E_q \quad (19)$$

$$E_f \leftarrow E_f - \alpha E_f \quad (20)$$

vi. Interpolate the local error variable of r from q and f

$$E_r = 0.5 \cdot (E_q + E_f) \quad (21)$$

Step 9. Decrease the local error variables of all units by a temporal discounting rate β .

$$E_i \leftarrow E_i - \beta E_i \quad (\forall i \in A) \quad (22)$$

Step 10. Continue with step 2 if a stopping criterion (e.g., the number of nodes or some performance measure) is not yet fulfilled.

IV. EXPERIMENTAL RESULTS

This section shows experimental results of the proposed method. The number of maximal neurons of GNG is 300, and the number of the maximal iterations of GNG is 100000. The parameters used for the learning in GNG are as follows; λ is 300; $\eta_1=0.1$; $\eta_2=0.02$; $\alpha=0.5$; $\beta=0.005$

Figure 8 shows an experimental result of the synthesis of color and distance information. Because 3D distance image sensor measures only the distance to the surface, the synthesis result of color and distance information is not complete (Fig.9 (b)). In this example, the 3D model of the backside of a ball and the backward of the ball is not obtained. Therefore, we use the omnidirectional mobile robot in order to control the position and posture of the robot easily. Figures 8 (c) and (d) show a synthesis result of measured data from three different view angles.

Figure 9 shows an experimental result of detection of sphere. This experimental result shows that the proposed method can distinguish target objects from the background image according to the posture of the robot.

Figure 10 shows an experimental result of the proposed method. GNG can perform both clustering and topological mapping according to the distribution of input data. Figure 10 (a) shows an experimental result using the rest of data after Hough transform (Case 1). Blue and Green area shows detected flat surface. In the result, GNG generates two clusters of objects, but the partial data of the floor is also included in the result, because Hough transform cannot completely divide original data into several regions owing to the measurement noise. Therefore, we use the selection process of the obtained object models after the clustering GNG. Figure 10 (b) shows an experimental result of the selection process after Hough transform and GNG (Case 2). Furthermore, because GNG makes polygons, we can show the target object by a textured model (Figs.10 (c) and (d)). In Figure 9 (c), colored target have about 4700 points. However, the colored target in Fig.10 is represented by only 300 nodes. In this way, we can reduce the number of data dramatically by using GNG.

Figure 11 shows an experimental result of the synthesis of color and distance information after Hough transform for the object extraction. In order to build a 3D model, the robot obtains color and distance information from both frontward and backward directions. The data corresponding the floor are

removed after Hough transform. As a result, we obtain a 3D model of the target object.

Finally, we discuss the applicability of the proposed method. Figure 12 shows an application example of the proposed method. The obtained topological model includes the edge connections among the nodes. Therefore, we can use the topological map for the path planning of a mobile robot. Figures 12 (b) and (c) show a topological environmental map and a result of path planning, respectively, where the position and target of the robot are depicted as yellow and blue, respectively. This is obtained by GNG after Hough transform. In this way, the proposed method can be applied to the visualization and map building of environment, and the path planning for mobile robots.

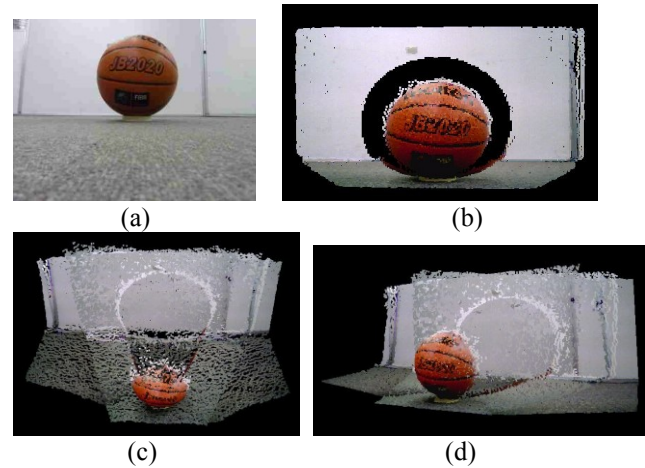


Fig.8. An example of synthesis of color and distance information; (a) original color image taken by CMOS camera; (b) a synthesis result from a single view; (c) a synthesis result from three different viewpoints (overview); (d) a synthesis result from three different viewpoints (side view)

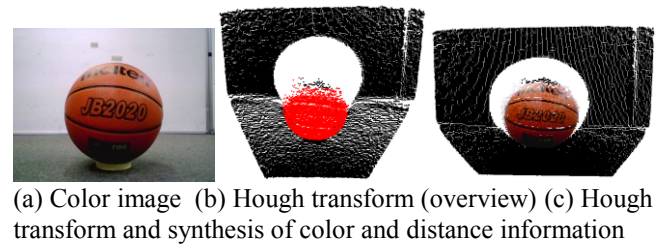
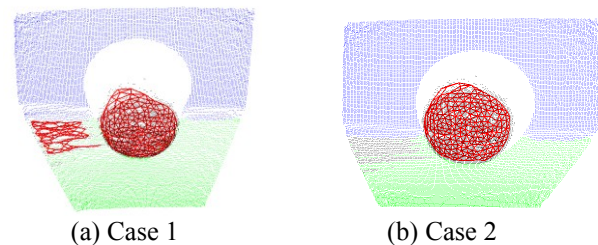
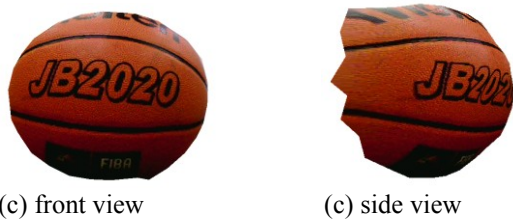


Fig.9. An example of Hough transform





(c) front view (c) side view
Fig.10. Visualization of target object by GNG

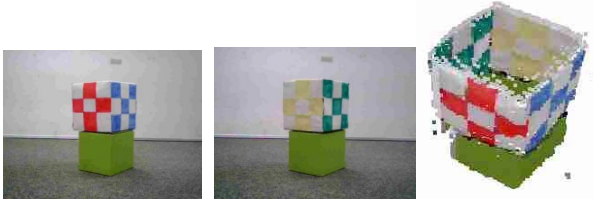
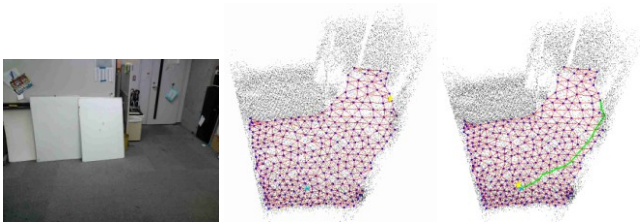


Fig.11. 3D Modeling based on the synthesis of color and distance information after Hough transform



(a) Original color image (b) Topological map (c) Path planning

Fig.12. 3D topological environmental Map

V. SUMMARY

This paper proposed an extraction method of objects from the background image based on Hough transform, and 3D topological reconstruction method based on growing neural gas (GNG). The experimental results show that the proposed method can sequentially conduct the extraction of objects and 3D topological reconstruction of the extracted objects as surface models. However, the proposed method can extract objects from the background and walls, but cannot separate the adjacent objects each other because the unsupervised learning of GNG is done according to only the distribution of distance information. Therefore, we should incorporate the object information through the communication and interaction between robots and co-existing people.

As future work, in order to improve the performance of 3D topological reconstruction, we intend to use simultaneous localization and mapping [18,19]. Furthermore, we integrate the path planning method [20,21] to the proposed method in order to automatically perform 3D topological reconstruction of objects in a room for informationally structured space.

REFERENCES

- [1] N.Kubota, Y.Shimomura, Human-Friendly Networked Partner Robots toward Sophisticated Services for A Community, *Proc. of SICE-ICCAS 2006*, pp. 4861-4866, 2006.
- [2] N.Kubota, K.Nishida, Cooperative Perceptual Systems for Partner Robots Based on Sensor Network, *International Journal of Computer Science and Network Security*, Vol.6, No.11, pp.19-28, 2006.
- [3] N.Kubota, T.Obo, T.Fukuda, An Intelligent Monitoring System based on Emotional Model in Sensor Networks, *The 18th IEEE International Symposium on Robot and Human Interactive Communication*, pp.346-351, 2009.
- [4] C. Corderio, D. Agrawal, *Ad Hoc & Sensor Networks*, World Scientific, 2006.
- [5] I. Khemapech, I. Duncan and A. Miller, A Survey of Wireless Sensor Networks Technology, *Proc. of The 6th Annual PostGraduate Symposium on The Convergence of Telecommunications, Networking and Broadcasting*, 2005.
- [6] J-H. Kim, Y-D. Kim, and K-H. Lee, The Third Generation of Robotics: Ubiquitous Robot, *Proc. of 2nd International Conference on Autonomous Robots and Agents*, 2004, pp.1-12.
- [7] M.Satomi, H.Masuta, N.Kubota, Hierarchical Growing Neural Gas for Information Structured Space, *Proc. of IEEE Symposium Series on Computational Intelligence 2009*, pp.54-59, 2009
- [8] Sanjeev J. Koppal, Srinivasa G.Narasimhan, Appearance Derivatives for Isonormal Clustering of Scenes, *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol.31, No.8, 2009.
- [9] A. Saxena, M. Sun, and A.Y.Ng, Make 3D:Learning 3D Scene Structure from a Single Still Image, *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol.31, No.5, pp.824-40, 2009.
- [10] Y. Moses, I. Shimshoni, 3D Shape Recovery of Smooth Surface:Dropping the Fixed-Viewpoint Assumption, *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol.7, No.7, pp.1310-24, 2009.
- [11] S. Liu, K. Kang, J.-P. Tarel, D.B. Cooper, "Free-Form Object Reconstruction from Silhouettes, Occluding Edge, and Texture Edge:A unified and Robust Operator Based on Duality, *IEEE Transaction on Pattern Analysis and Machine Intelligence* Vol.30, No.30, pp.131-146, 2008.
- [12] T. Nishida, T. Kitamori, A. Eguchi, Y. Fuchikawa and S. Kurogi, Integration of Color and Range Image for Plant Recognition, *Proc. of ISAPCS2003*, pp. 219-224, 2003.
- [13] K.Okada, S.Kagami, M.Inaba, H.Inoue, Plane Segment Finder: Algorithm, Implementation and Applications, *Proc. of the 2001 IEEE International Conference on Robotics & Automaton*, Seoul, Korea, May 21-26, 2001, Vol.2, pp.2120-2125.
- [14] H. Wu, G. Yoshikawa, T. Shioyama, S. Lao, M. Kawade. Glasses Frame Detection with 3D Hough Transform. In *Proceedings of ICPR (2)2002*. pp.346~349.
- [15] B. Fritzke: A growing neural gas network learns topologies; *Advances in Neural Information Processing Systems*, Vol. 7, pp. 625-632, 1995.
- [16] B. Fritzke: Growing self-organizing networks - why?; *European Symposium on Artificial Neural Networks*, pp. 61-72, 1996.
- [17] B. Fritzke: Unsupervised clustering with growing cell structures; *Neural Networks*, Vol. 2, pp 531-536, 1991.
- [18] S.Thrun, W.Burgard, D.Fox, *Probabilistic Robotics*, MIT Press, 2005.
- [19] H.Sasaki, N.Kubota, K.Taniguchi, Topological Map and Cell Space Map for SLAM of A Mobile Robot, *GESTS International Transactions on Computer and Engineering*, Vol.45, No.3, 2008.
- [20] K.S.Fu., R.C.Gonzalez., and C.S.G.Lee, *Robotics*, McGraw-Hill Book Company, 1987.
- [21] T. Fukuda, N.Kubota, An Intelligent Robotic System Based on A Fuzzy Approach, *Proceedings of The IEEE*, Vol.87, No.9, pp.1448-1470, 1999.
- [22] N.Kubota, D.Tang, T.Obo, S.Wakisaka, Localization of Human Based on Fuzzy Spiking Neural Network in Informationally Structured Space, *IEEE World Congress on Computational Intelligence 2010*, 2010 (submitted).
- [23] N.Kubota, T.Obo, Localization of Human Based on Fuzzy Spiking Neural Network in Informationally Structured Space, *IEEE World Congress on Computational Intelligence 2010*, 2010 (submitted).