

Fast 2D-3D Registration for Navigation System of Surgical Robot

Yumi Iwashita, Ryo Kurazume
and Tsutomu Hasegawa

*Graduate School of Information Science and Electrical Engineering
Kyushu University
6-10-1, Hakozaki, Higashi-ku, Fukuoka, Japan
yumi@irvs.is.kyushu-u.ac.jp*

Kozo Konishi, Masahiko Nakamoto
and Makoto Hashizume

*Graduate School of Medical Sciences
Kyushu University
3-1-1, Maidashi, Higashi-ku, Fukuoka, Japan*

Abstract— This paper presents a new registration algorithm of 2D color images and 3D geometric models for navigation system of surgical robot. A 2D-3D registration procedure is used to superimpose a tumor model on an endoscopic image precisely, and is therefore indispensable for the surgical navigation system. And thus, the performance of the 2D-3D registration procedure influences directly the usability of the surgical robot operating system. One of the typical techniques developed so far is the use of external markers, but the accuracy of this method is getting worse due to the breath, the heart beat, or other unknown factors. For precise registration of 3D models and 2D images without external markers or special measurement devices, the new registration method is proposed which utilizes the 2D images and their distance maps created by the Fast Marching Method. We show some results of fundamental experiments using simulated models and actual images of the endoscopic operation.

I. INTRODUCTION

Development work on a practical surgical robot system is proceeding aiming at a minimally invasive surgical operation and a favourable postoperative process. Fig.1 shows the surgical robot system named Intuitive Surgical's "da Vinci".

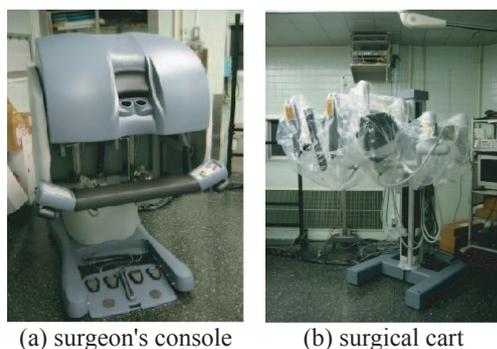


Fig. 1. The surgical robot system "da Vinci"

In general, the surgical robot system is composed of two main components, a manipulation system and a navigation system.

The manipulation system of the "da Vinci" is so-called the master-slave manipulator system. The slave manipulator consists of 3 arms; two 7 degrees-of-freedom manipulators and one 5 degrees-of-freedom camera holder. The manipulators are

used to hold and operate many types of forceps, which can be replaced according to an operation purpose. The camera holder is used for controlling the position of an endoscope, which is equipped with 2 CCD cameras. Captured stereo images of an internal organ are displayed on two LCDs in a surgeon's control console, and give stereognostic perception to the operator. The operator is able to control the forceps and the endoscope quite intuitively using the stereognostic perception.

Our navigation system for the "da Vinci" is designed to assist an operator by presenting a proper operation procedure to be taken and displaying patient's information related to the operation [1]. Another main function of the navigation system is to display an invisible internal structure of an organ, location of a tumor, or distribution of critical blood vessels on the endoscope image with high reality. The operator can obtain useful information from these realistic images intuitively so that the operation is performed smoothly.

The virtual images of the organ or the structure of blood vessels are created from 3D geometric models, which are obtained using medical imaging devices such as CT (Computerizing Tomography) scanner or MRI (Magnetic Resonance Imaging) beforehand. The obtained 3D geometric models are projected on the image plane of the endoscope and superimposed on 2D images of the target organs. The performance of this procedure influences directly the usability of the surgical robot operating system. Thus, the registration of 2D images and 3D geometric models has to be performed as precisely as possible.

Several registration techniques of 2D images and 3D geometric models have been proposed so far. One of the typical techniques is the use of external markers to calibrate the coordinate systems fixed on the patient's body and the endoscope. In this method, the position of several markers, pasted on the skin of the trunk and the endoscope respectively, are measured by an external optical device. Then, the calibration calculation is performed to determine the relative pose between these two coordinate systems. After the calibration, the 3D geometric models of organs are projected on the 2D images taken from the endoscope based on the standard position information of the organ in the patient's body. Therefore, the accuracy of this method is getting worse if the organ is away from the estimated position in the body due to the breath, the heart beat, or other

unknown factors.

This paper presents a new registration algorithm of 2D color images and 3D geometric models without external markers or special measurement devices. This method utilizes the 2D images and their distance maps created by the Fast Marching Method, and is able to determine a precise relative pose between 2D images and 3D models. Since the registration process can be executed at a quite high speed using actual images, the registration accuracy is expected to be higher than the one of the conventional method using the external markers, even if the position of the organ is disturbed by the breath, the heart beat, or other factors.

After an overview of previous works in Section II, we will give a short introduction of the Fast Marching Method in Section III, which is utilized for the fast construction of a distance map. A detail description of the new 2D-3D registration algorithm is presented in Section IV. In Section V, we show some results of fundamental experiments using actual images of the endoscopic operation.

II. RELATED WORKS

For the alignment of a 2D color image and a 3D geometric model, Viola proposed a technique based on a statistical method [2]. This method evaluated the mutual information between the 2D image and the image of the 3D model based on the distribution of intensity. Allen et al. [3] also proposed a method using the intersection lines of several planes extracted from the range data and the color edges. These methods work well on the surface with little albedo variance. But these are easily trapped in a local minimum on surfaces with rich textures.

Range sensors often provide reflectance images as side products of range images. This 2D reflectance image is aligned with the 3D range image because both images are obtained through the same receiving optical device. Therefore, we can utilize the 2D reflectance image for solving the 2D-3D registration problem instead of the original 3D model. By using the 2D reflectance image, the 2D-3D problem is modified to the simple 2D-2D problem. Kurazume et al. [4] proposed the registration algorithm using the reflectance image and the 2D texture image. In his method, a number of photometric edges extracted from both images were registered and the relative pose is determined using a robust M-estimator. Elstrom et al. [5] and Umeda et al. [6] also proposed the methods using the reflectance image and the texture image. However, an organ doesn't have a rich texture on its surface and there are also differences of textures between individuals. Therefore, the method using texture is quite difficult for the 2D-3D registration problem of the surgical navigation system.

On the other hand, some registration techniques using a silhouette image or a contour line have been proposed. Lensch [7],[8] proposed a silhouette-based approach. He used XOR value of silhouette images of a 2D image and a 3D model for evaluating a registration error. The Downhill Simplex Method was used for the convergent calculation of the 2D-3D registration. In contour-based approach, the error is computed

as the sum of distances between points on a contour line in a 2D image and on a projected contour line of a 3D model [9], [10]. Since a number of point correspondences between both contours has to be determined for calculating the registration error, these algorithms are computationally expensive.

Our 2D-3D registration algorithm belongs to the contour-based approach. However, we adopt a distance map for evaluating a registration error instead of the point correspondence. Therefore, once the distance map is created, our algorithm runs faster than the conventional point-based approach[9], [10]. In addition, since the distance map can be created quite rapidly using the Level Set Method named the Fast Marching Method, our method is able to track an object continuously even if the object moves.

III. FAST MARCHING METHOD FOR RAPID CONSTRUCTION OF DISTANCE MAP

Before describing our registration algorithm, the Fast Marching Method proposed by Sethian [11],[12] is briefly introduced, which is utilized for the fast construction of a distance map in our method.

The Fast Marching Method [11],[12] was initially proposed as a fast numerical solution of the Eikonal equations ($|\nabla T(p)| F = 1$), where $T(p)$ is arrival time of a front at a point p and F is a speed function. In general, since this equation is solved using a convergent calculation, it takes a long time to obtain a proper solution. However, by adding restriction such that a sign of the speed function F is invariant, the Fast Marching Method solves the Eikonal equation straightforwardly, and thus, quite rapidly. In this method, the arrival time of a front at each point is determined in the order from the old to the new.

At first, the Fast Marching Method modifies the Eikonal equation into the next difference equation.

$$\begin{aligned} & (\max(D_{ij}^{-x}T, -D_{ij}^{+x}T, 0))^2 + \\ & \max(D_{ij}^{-y}T, -D_{ij}^{+y}T, 0))^{\frac{1}{2}} = 1/F_{ij} \end{aligned} \quad (1)$$

Next, since the arrival time is propagated in one direction from the old to the new, a point which holds the oldest arrival time in a whole region is chosen, and the arrival time of the boundary at this point is determined from Eq.(1).

The concrete procedure is as follows:

Step.1 (Initialization) The whole region is divided into a number of grid points with a proper grid width. Before starting the calculation, all the grid points are categorized into three categories (*known*, *trial*, *far*) according to the following procedure.

- 1) The grid points belonging to the initial front (denote the boundary hereafter) are added to the category of *known* and the arrival time of these grid points is set to 0 ($T = 0$).
- 2) Among the 4 neighboring grid points of a grid point belonging to the category of *known*, the one which doesn't belong to the category of *known* is categorized into to the category of

trial, and the arrival time of this grid point is calculated temporarily from the equation, $T_{ij} = 1/F_{ij}$. In addition, the arrival time of these grid points is stored in heap data structure using the heap sort algorithm in ascending order of T .

- 3) The grid points except for the above grid points are all categorized into the category of *far*, and the arrival time is set to infinity ($T = \infty$).

Step.2 Choose the grid point (i_{min}, j_{min}) which is placed at the top of the heap structure. This grid point has the smallest arrival time in the category of *trial*. Remove this grid point from the category of *trial* and the heap structure, and categorize it as the category of *known*. Run “downheap” algorithm to reconstruct the heap structure.

Step.3 Among the 4 neighbouring grid points $((i_{min} - 1, j_{min}), (i_{min} + 1, j_{min}), (i_{min}, j_{min} - 1), (i_{min}, j_{min} + 1))$ of the selected grid point (i_{min}, j_{min}) , the one which belongs to the category of *far* is changed to the category of *trial*.

Step.4 Among the 4 neighbouring grid points of the selected grid point (i_{min}, j_{min}) , the arrival time of the one which belongs to the category of *trial* is calculated using Eq.(1). Then, run “upheap” algorithm to reconstruct the heap area.

Step.5 If there is a grid point which belongs to the category of *trial*, then go to Step 2. Otherwise, the process is terminated.

This Fast Marching Method constructs a distance map of a whole region quite rapidly, which indicates the distance from a boundary (an initial position of the front) to a certain point. To construct the distance map, the speed function F_{ij} in Eq.(1) is set to 1 at first. Next, the arrival time of the front T is determined using the above procedure. Since the speed function is 1, T suggests the distance from the boundary to the point. Fig.2 shows an example of the calculated distance map.

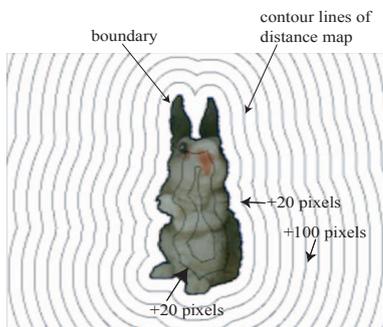


Fig. 2. An example of the distance map calculation using the Fast Marching Method

IV. A NEW 2D-3D REGISTRATION ALGORITHM BASED ON THE DISTANCE MAP

This section describes our fast 2D-3D registration algorithm using the distance map in detail. We assume that a 3D geometric model of the target, such as a structure of an organ or a tumor, has been constructed by CT or MRI images beforehand, and is represented by a number of triangular patches. A 2D color image is also taken from a camera attached to an endoscope.

The proposed method is summarized as follows:

- 1) At first, a boundary of a target on the 2D color image is detected using an active contour model (ex. Snakes or the Level Set Method[13]).
- 2) Next, the distance map from the detected boundary on the 2D image plane is constructed using the Fast Marching Method.
- 3) The 3D geometric model of the target is placed at an arbitrary position and projected on the 2D image plane.
- 4) Contour points of the projected image and their corresponding patches on the 3D model are identified.
- 5) Force is applied to the selected patch of the 3D model in 3D space according to the distance value obtained from the distance map at each contour point.
- 6) Total force and moment around the COG (center of gravity) of the 3D model is determined using the robust M-estimator.
- 7) The pose of the 3D model is changed according to the total force and moment.
- 8) Repeat from step 1 to 7 until the projected image of the 3D model and the 2D image coincide each other.

The above procedure is explained in more details with some examples in the following sections.

A. Construction of distance map

Fig.3 shows the calculated distance map. At first, a boundary is extracted from the color image using the Level Set Method[13]. Next, the distance map from this boundary is calculated using the Fast Marching Method explained in Section III.

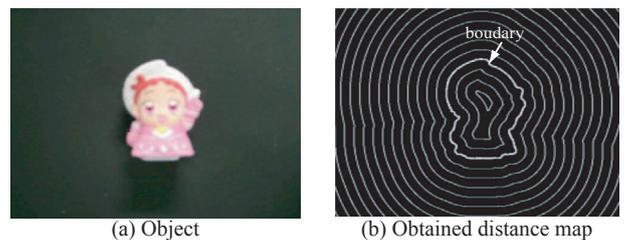


Fig. 3. Detected boundary and distance map

B. Fast detection of triangular patches of contour points in 3D model

Fig.4 shows an example of the contour detection of the projected 3D geometric model. The contour detection and identifying triangular patches on the 3D model corresponding

to points on the contour line are computationally expensive and time consuming. In our implementation, we utilize the high-speed rendering function of the OpenGL hardware accelerator and execute these procedures quite rapidly.

The detailed algorithm is as follows: Initially, we assign different colors for all the triangular patches in the 3D model and draw the projected image of the 3D model on the image buffer using the OpenGL hardware accelerator. The contour points of the 3D model are detected by raster scanning of the image buffer. By reading colors of the detected contour points, we can identify the corresponding triangular patches on the 3D geometric model. Fig.5 shows the color image of the projected 3D model with the patches of different colors.

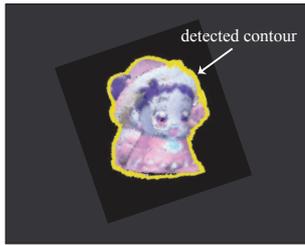


Fig. 4. Contour detection of 3D geometric model

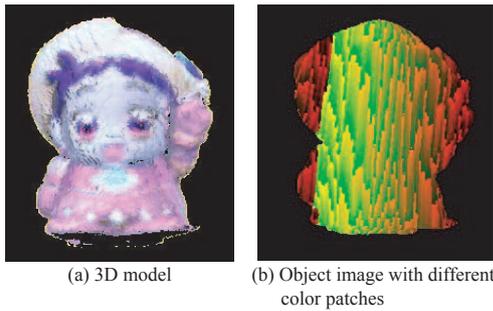


Fig. 5. Detection of triangular patches of the contour points

C. Force and moment calculation using the M-estimator and determination of relative pose

After obtaining the distance map on the 2D image and the list of the triangular patches of the 3D model corresponding to the contour points, the force f_i is applied to all the triangular patches of the contour points (Fig. 6). The force f_i is the vector perpendicular to the line of sight, and the projection of the f_i onto the 2D image plane coincides with the f_{DM_i} , which is the vector toward the direction of the steepest descent of the distance map (Fig. 7(a)). We assume that the magnitude of the f_{DM_i} is proportional to the value on the distance map at the projected contour point of the 3D model. Then, total force and moment around the COG are calculated using the following equations as shown in Fig.7.

$$F = \sum_i \rho(f_i) \quad (2)$$

$$M = \sum_i \rho(r_i \times f_i) \quad (3)$$

where r_i is a vector from the COG to the triangular patch i and $\rho(z)$ is a particular estimate function.

In practice, a part of the target organ is occasionally hidden by the other organs in the endoscopic image. In this case, the obtained boundary does not coincide with the projected contour of the 3D model and the correct distance map cannot be obtained. Therefore, the robust M-estimator is used in our system to ignore contour points with a large amount of errors. Let's consider the force f_i and the moment $r_i \times f_i$ as an error z_i . Then, we modify Eqs.(2) and (3) as

$$E(P) = \begin{pmatrix} F \\ M \end{pmatrix} = \sum_i \rho(z_i) \quad (4)$$

where P is the pose of the 3D geometric model.

The pose P which minimizes the $E(P)$ is obtained as the following equation.

$$\frac{\partial E}{\partial P} = \sum_i \frac{\partial \rho(z_i)}{\partial z_i} \frac{\partial z_i}{\partial P} = 0 \quad (5)$$

Here, we define the weight function $w(z)$ as the following equation in order to evaluate the error term.

$$w(z) = \frac{1}{z} \frac{\partial \rho}{\partial z} \quad (6)$$

From the above equation, we obtain the following weighted least squares method.

$$\frac{\partial E}{\partial P} = \sum_i w(z_i) z_i \frac{\partial z_i}{\partial P} = 0 \quad (7)$$

In our implementation, we adopt the Lorentzian function for the estimation function $\rho(z)$, and gradually minimize the error E in Eq.(4) by the steepest descent method.

$$\rho(z) = \frac{\sigma^2}{2} \log(1 + (z/\sigma)^2) \quad (8)$$

$$w(z) = \frac{1}{1 + (z/\sigma)^2} \quad (9)$$

The P which minimizes the error in Eq.(4) is the estimated relative pose between the 2D image and the 3D geometric model.

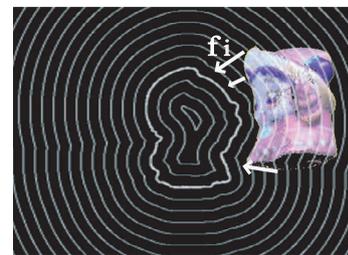


Fig. 6. Applying the force f to all the triangular patches of the contour points.

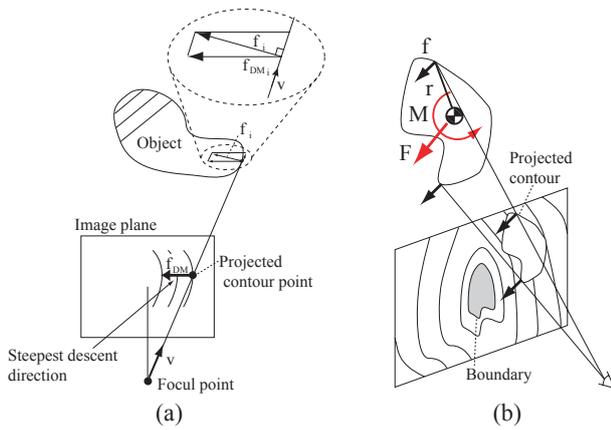


Fig. 7. Force and moment around COG

D. Coarse-to-fine strategy and the “distance band”

Though the 2D distance map can be constructed efficiently by the Fast Marching Method, computing the distance value in whole region of the color image is wasteful. Therefore we adopt coarse-to-fine strategy and a band-shaped distance map named a distance band. The distance band is a narrow band along the image boundary and the precise distance map is constructed in this region using the Fast Marching Method. In other region except the distance band, the distance value is roughly estimated as the distance from the center of the image boundary. In the beginning of the registration process, we use a coarse image and a wide distance band in order to align the 2D image and the 3D model roughly. After several iterations, we change the image to the fine one and compute the distance band with narrower width. In experiments, we used the distance band with 10 pixels and three images with difference resolution, that is, 160x120, 320x240, 640x480. Computation time of the distance band in each image is shown in Table I.

TABLE I
COMPUTATION TIME OF THE DISTANCE BAND BY THE FAST MARCHING METHOD.

Image size	Whole region [ms]	Distance band (10 pixels) [ms]
160x120	6.9	0.48
320x240	32.8	1.1
640x480	189.1	2.7

V. EXPERIMENTS

In this section, we show some fundamental results of the registration and tracking experiments using simulated images and actual images of the endoscopic operation.

A. 2D-3D registration of simulated images

Firstly, Fig.8 shows how the proposed 2D-3D registration algorithm works with a model of a small doll (a height of 5cm). The boundary of the object on the 2D color image

is detected by applying the Level Set Method[13] to the background subtraction image.

The computation time of the force f_i for one patch on the 3D model is 0.30 [μs] for Pentium IV processor, 3.06 GHz. The total processing time is 9.6 [ms] for one update period including projected contour detection, calculation of force and moment, and the steepest descent method. The average and the variance of the registration error between the boundary and the projected contour are 1.19[pixels] and 0.90[pixels²] respectively after 60 iterative calculations. Here, the number of points on the boundary is 652, image size is 640 × 480 pixels, and the 3D model is composed of 73,192 meshes with 219,576 vertices. On the other hand, the computation time of the force f_i for the conventional point-based method [4] is 1.15 [μs], which uses the k-D tree structure to search the nearest points between the contour points of the 2D image and the projected 3D model.

Table II show the comparison of the processing time for several cases in which different numbers of points on the boundary were used. From these results, it is clear that the processing time of the proposed method is almost constant even if the number of points on the boundary increases. On the other hand, the point-based approach requires 3 to 7 times longer processing time than the proposed method. Moreover, the larger the number of points becomes, the longer processing time is required for the point-based approach. Therefore, it is verified that the proposed 2D-3D registration algorithm has an advantage comparing with the conventional point based approach in views of execution time, especially when there are a large number of points on the boundary.

Next example is the tracking experiment of the moving object. Fig.9 shows the tracking results of the doll which moves on the image plane. As seen in this example, the 2D-3D registration can be performed even if the object moves up to 20.5 pixels/second. This result suggests that it is possible to overcome the registration error due to unknown disturbance such as the breath or the heart beat in endoscopic images.

TABLE II
COMPARISON OF PROCESSING TIME FOR ONE PATCH

Number of points on boundary	Proposed method [μs]	Point-based method [4] [μs]
628	0.30	1.15
1265	0.30	1.50
1868	0.30	1.70
2490	0.30	2.22

B. 2D-3D registration using actual images of the endoscopic operation

We carried out fundamental experiments for a navigation system of the endoscopic operation.

Firstly, we tried the 2D-3D registration using actual images of the endoscopic operation of the cirrhosis of the liver. Fig.10 shows the initial results of the experiments. Fig.10(a) shows the actual image of the endoscopic operation of the cirrhosis of the liver. In this image, the liver is partially hidden by the other

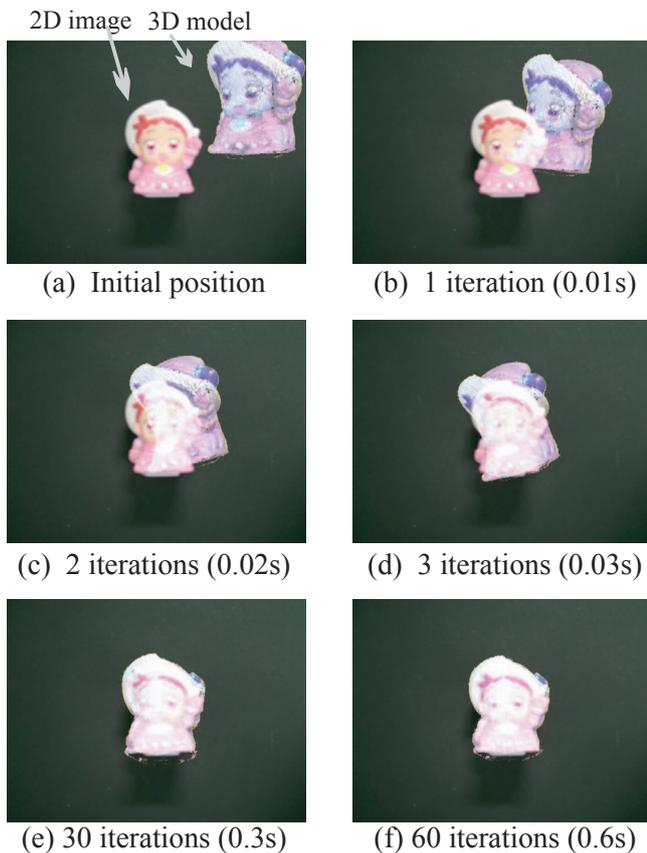


Fig. 8. 2D-3D registration of simulation images

organs. Fig.10(b) and Fig.10(c) show the 3D model of the liver obtained from the CT images and some experimental results for the 2D-3D registration, respectively. The registration is successfully executed if the initial position of the 3D model is close enough to the image of the cirrhosis of the liver.

Fig. 11 shows the tracking results of the gallbladder in video images taken by the endoscope. Boundaries of the gallbladder on the every 2D video images are detected using by Snakes. As seen in this result, our algorithm is able to track the gallbladder even if the gallbladder image moves due to the change of the position of the endoscope.

VI. CONCLUSIONS

This paper described a new registration algorithm of 2D color images and 3D geometric models for navigation system of surgical robot. This method utilizes the 2D images and their distance map created by the Fast Marching Method, and determines a precise relative pose between 2D images and 3D models. The efficiency of the proposed algorithm was verified through the fundamental experiments using simulation images and actual images of the endoscopic operation of liver.

We have conducted experiments with various images and 3D models and obtained almost good results. Occasionally the performance of the proposed algorithm becomes worse especially in case that only the partial edge of an organ is visible from the endoscope. However, once the registration

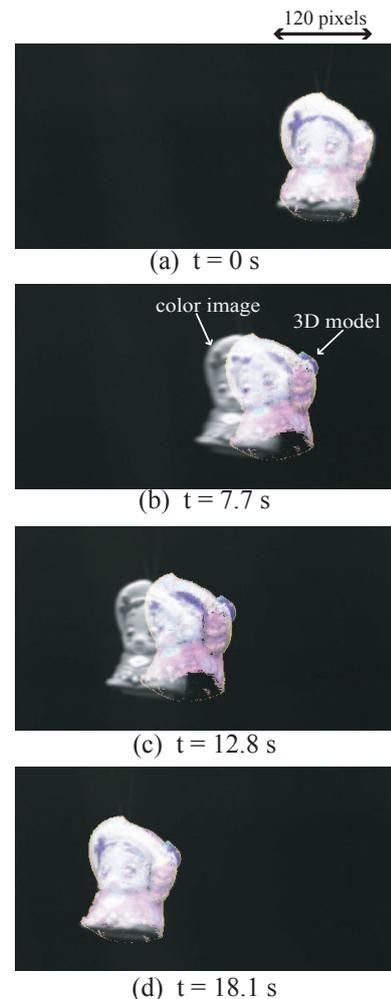


Fig. 9. An example of moving object tracking

is successfully performed, then the algorithm can track a moving object stably and continuously. Therefore the proposed algorithm is quite useful for developing a practical navigation system since it overcomes the registration error caused by organ's movement due to unknown disturbance such as the breath or the heart beat.

ACKNOWLEDGMENT

This research was partly supported by the 21st Century COE Program "Reconstruction of Social Infrastructure Related to Information Science and Electrical Engineering", and the Ministry of Public Management, Home Affairs, Posts and Telecommunications in Japan under Strategic Information and Communications R&D Promotion Programme (SCOPE).

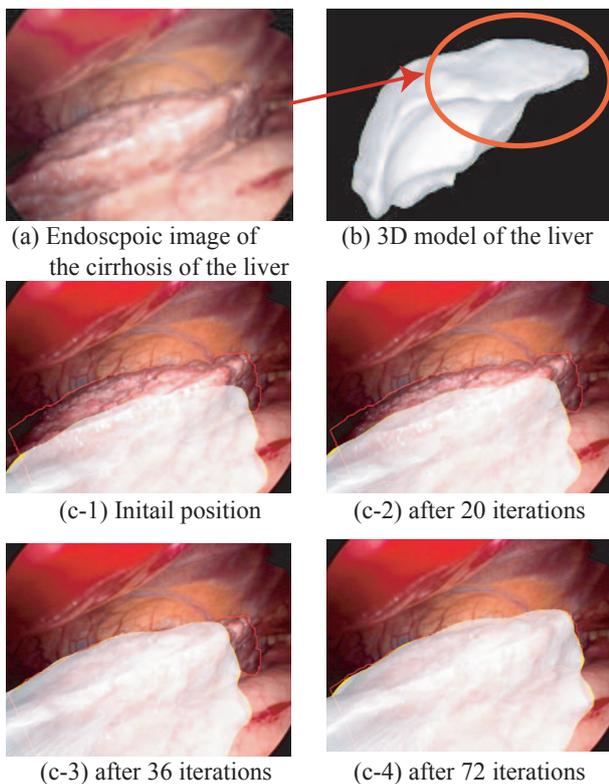


Fig. 10. 2D-3D registration using actual images of the cirrhosis of the liver

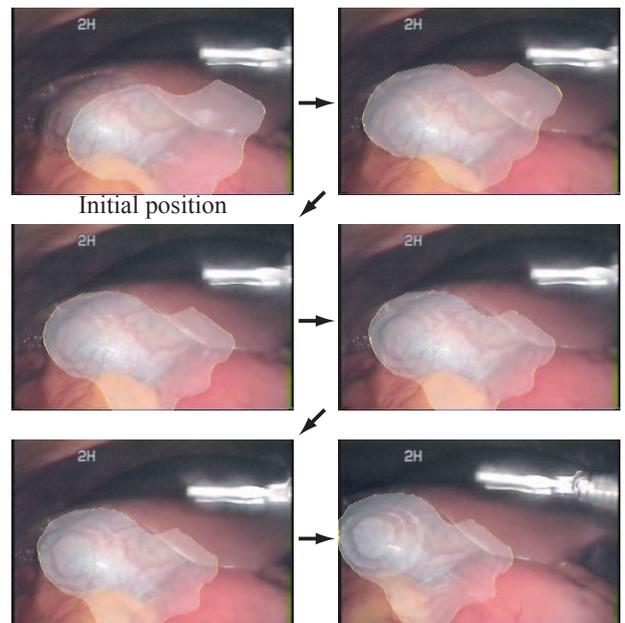


Fig. 11. 2D-3D registration in endoscopic video images of the gallbladder.

REFERENCES

- [1] M. Hashizume, M. Shimada, M. Tomikawa, Y. Ikeda, I. Takahashi, R. Abe, F. Koga, N. Gotoh, K. Konishi, S. Maehara, and K. Sugimachi, "Early experiences of endoscopic procedures in general surgery assisted by a computer-enhanced surgical system," in *Surg Endosc*, 2002, pp. 1187–1191.
- [2] P. Viola and W. W. III, "Alignment by maximization of mutual information," *Int. J. of Computer Vision*, vol. 24, no. 2, pp. 137–154, 1997.
- [3] I. Stamos and P. K. Allen, "Integration of range and image sensing for photorealistic 3d modeling," in *Proc. of the 2000 IEEE International Conference on Robotics and Automation*, 2000, pp. 1435–1440.
- [4] R. Kurazume, K. Noshino, Z. Zhang, and K. Ikeuchi, "Simultaneous 2d images and 3d geometric model registration for texture mapping utilizing reflectance attribute," in *Proc. of Fifth Asian Conference on Computer Vision (ACCV)*, 2002, pp. 99–106.
- [5] M. D. Elstrom and P. W. Smith, "Stereo-based registration of multi-sensor imagery for enhanced visualization of remote environments," in *Proc. of the 1999 IEEE International Conference on Robotics and Automation*, 1999, pp. 1948–1953.
- [6] K. Umeda, G. Godin, and M. Rioux, "Registration of range and color images using gradient constraints and range intensity images," in *Proc. of 17th International Conference on Pattern Recognition*, 2004, pp. 12–15.
- [7] H. Lensch, W. Heidrich, and H.-P. Seidel, "Automated texture registration and stitching for real world models," in *In Pacific Graphics '00*, 2000, pp. 317–326.
- [8] H. Lensch, W. Heidrich, and H. P. Seidel, "Hardware-accelerated silhouette matching," in *In SIGGRAPH Sketches*, 2000.
- [9] K. Matsushita and T. Kaneko, "Efficient and handy texture mapping on 3d surfaces," in *Comput. Graphics Forum 18*, 1999, pp. 349–358.
- [10] P. J. Neugebauer and K. Klein, "Texturing 3d models of real world objects from multiple unregistered photographic views," in *Computer Graphics Forum 18*, 1999, pp. 245–256.
- [11] J. Sethian, "A fast marching level set method for monotonically advancing fronts," in *Proceedings of the National Academy of Science*, vol. 93, 1996, pp. 1591–1595.
- [12] ———, *Level Set Methods and Fast Marching Methods, second edition*. UK: Cambridge University Press, 1999.
- [13] Y. Iwashita, R. Kurazume, T. Tsuji, K. Hara, and T. Hasegawa, "Fast implementation of level set method and its realtime applications," in *IEEE International Conference on Systems, Man and Cybernetics 2004*, 2004.