Proceedings of 2013 IEEE
International Conference on Mechatronics and Automation
August 4 - 7, Takamatsu, Japan

A Method of Decreasing Transmission Time of Visual Feedback for the Internet-based Surgical Training System

Jin Guo[1], Shuxiang Guo[2, 3], Nan Xiao[2]
[1]Graduate School of Engineering, Kagawa University, Takamatsu, Kagawa, Japan
[2]Intelligent Mechanical Systems Engineering Department, Kagawa University, Takamatsu, Kagawa, Japan
[3]Harbin Engineering University, Harbin, Heilongjiang Province, China
s12d503@stmail.eng.kagawa-u.ac.jp
guo@eng.kagawa-u.ac.jp

Abstract **- The Internet-based robotic catheter operating system for Vascular Interventional Surgery (VIS) provides a promising way to enhance emergency medical care and to train novice doctors to learn basic wire or catheter handling skills across the whole world. However, variable and unpredictable transmission time delay of visual feedback is always associated with the significant deterioration of the operability and may result in damage to organs in actual surgery. In this paper, we proposed a new method to reduce transmission time of visual feedback effectively based on image processing technique. At surgical site, the shape of active catheter is tracked in real time to provide coordinates of catheter in real vascular structure. And at operation site, position of catheter is reconstructed in 2-dimensional and 3-dimensional blood vessel model which has the same geometrical structure with that at surgical site according to the coordinate values received from surgical site. Transmission time of visual feedback is therefore decreased due to significant reduction of data volume. Finally, we present results from a visually guided tele-operative experiment where show that the proposed method makes a significant performance improvement for tele-operation with delays corresponding to inter-country distances.**

Index Terms – Internet-based robotic catheter operating system, image processing, transmission time,

I. INTRODUCTION

Vascular interventional surgery (VIS) represents one of the main evolutions of surgical techniques aimed at providing a greater benefit to the patient. The main advantage of this technique is to reduce trauma to healthy tissue since this trauma is the leading cause for patients' pain and scaring and prolonged hospital stay. However, minimally invasive surgery increases the operative difficulty since the depth perception is usually dramatically reduced. In addition, the surgeons could have prolonged exposure to radiation and be subjected to a high level of fatigue caused by poor ergonomics of the current procedure, which can pose danger or discomfort to the surgeons who perform the procedure over a prolonged period of time [1]. Therefore, a more accurate, safer, and more reliable approach that can reduce the radiation exposure to surgical doctors and disseminate new surgical knowledge, skills and techniques across the whole world should be developed.

With the capability of remote surgery, it is possible to realize participations of highly capable doctors to carry out an emergency medical care, which takes place in a distant place. And also, it is possible to facilitate a team work medical treatment between doctors from all around the world. In addition, doctors can be protected from longtime radiation. Marescaux et al. have conducted a remote surgery by using a surgical robot ZEUS through a dedicated fiber-optic line between New York and Strasbourg [1]. However, it is generally costly for laying dedicated fiber-optic lines and running them. Thus the use of conventional network infrastructures is preferable for the future expansion of remote surgery applications. There are a few research groups that have studied the robotic tele-operative surgical systems based on Internet [2]-[7]. In general, this kind of surgical system consists of a master system, a slave system and a communication link (**Fig. 1**). The surgeon operates the master system at operation site, and the actual operation is performed at surgery site. The two sites can be connected by IP networks for sending and receiving information such as image information, force data and control signals of manipulators.

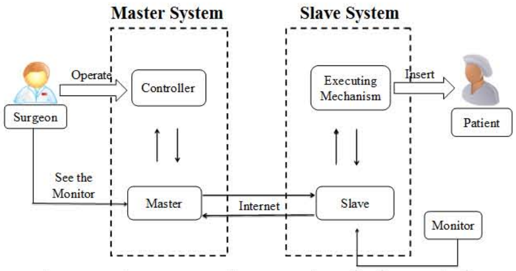


Fig.1 The general structure of Internet based tele-surgical system

Obviously, there are some unavoidable problems, referring to this kind of Internet-based remote surgical systems, time delay of image transmission and time difference between control signals and image information. The size of visual information is large and much more than control and force signals. The incremental time delay and time difference are always associated with the significant deterioration of the operability and may result in damage to training effect during remote surgical education. There are many research groups focused on delayed control and force signals through the Internet, nevertheless few academic institutions are engaged in the time delay of visual transmission. Traditionally, delays on visual feedback systems have been handled by predictive methods or through increased autonomy - i.e. higher level control, where the necessary degree of autonomy increases

978-1-4673-5558-2/13/$31.00 ©2013 IEEE

with the time delay [8]. Most prediction-based delay compensation methods relay on predicting the feedback from the remote environment, which may be difficult to model and predict, especially if the main feedback is a video signal [10]. Jumpei Arata et al. used a high speed image compression process in remote surgery experiment which reduces the negative effects caused by the time varying networks on the transmission of visual feedback data [9], however, the data processing time for compression and de-compression is time consuming.

The objective of this paper is to propose a new method based on image-processing technology to efficiently reduce the amount of communication data of image information in order to reduce transmission time of visual feedback and make easier to control the time difference between force, control signals and image information.

This paper is organized as follows. In Section II, the methods of reducing the transmission time of image information in both two-dimension and three-dimension are presented in a systematical way. Experiments to verify our method are described in section III. Finally, a brief conclusion and future work section is presented in Section IV.

II. METHOD AND IMPLEMENTATION

As many researchers point out, time delay of the transmission of image information from surgery site to operation site becomes a critical problem often encountered in tele-operative surgery. In terms of our novel Internet based robotic catheter operating system, the time delay is calculated as (Time delay = Transmission delay of control signal from the master site to the slave site + Response of slave + Transmission delay of image or force from slave to master). In this paper, we mainly focused on the method to solve the delay of image information by reducing the amount of data transmission. Although the length of the time delay depends on the configuration of the network used for the operation, it can reduce the time delay of image information effectively using our approaches.

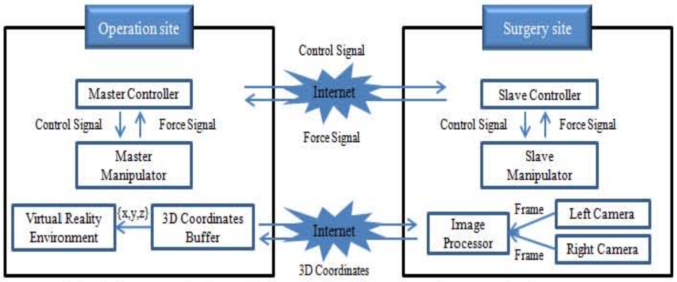


Fig.2 Transmission of signals in tele-operative surgical system

The overall structure can be described as **Fig. 2** shown. Two cameras are used to capture the images of catheter at the same time and two images are sent to image processor module to get the three dimensional coordinates of catheter. Next, these three dimensional coordinates are transmitted to operation site through the Internet. Then, at operation site, three dimensional coordinates received from surgery site are re-constructed in virtual reality environment in real time. According to the status of catheter model in virtual reality environment, doctors decide whether to insert catheter or not due to the fact that 3D visualization of surgery environment can be implemented to provide the doctor with a "live" virtual representation of the scene.

A. The Implementation Method of Two-dimensional Model

At the slave site, a digital camera capturing images at 30fps has been used to obtain images as the catheter is inserted into the EVE (Endo Vascular Evaluator Model) model. The images obtained from this camera are similar to X-ray fluoroscopic images in terms of contrast and frame capture rate. However, X-ray fluoroscopic images have a much higher resolution as compared to the images obtained from the camera. At the same time, the images were processed using a novel real-time algorithm to track the tip of the catheter to get the coordinate values in the Cartesian coordinates and we set the top-left point as the original point. Then the coordinates are sent to the master side. The size of coordinates is much less than that of image information. The real-time algorithm of motion tracking as showed in **Fig. 3** below.

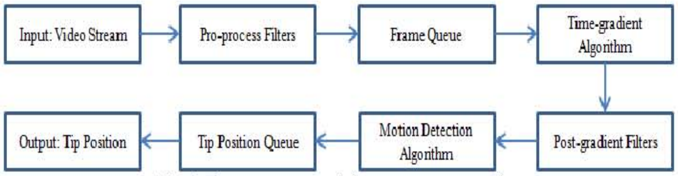


Fig.3 The structure of the surgeon console

The video stream consists of a stream of frames grabbed by the camera at 30fps. All frames are time-stamped. The pre-process filter block consists of a set of filters for suppressing noise, masking the image and improving the contrast of the image by adaptive thresholding. The frame queue is an image buffer structure managed by the queue manager block. This queue provides the appropriate input for the time-gradient algorithm. The time-gradient algorithm takes time gradients of the images in the frame queue. As the insertion speed is finite and limited, it concentrates on a neighborhood of the previous valid catheter tip position to limit the search area and decrease the processing time. The post-gradient filters suppress the noise in the gradient image and make it useful for the motion-detection algorithm. The motion-detection algorithm extracts the linear speed vector of the catheter (along the catheter axis) as well as the tip position. It also generates a true/false flag called motion flag which is true when any motion is detected. The signals generated by the motion detection algorithm are fed back to the master side.

At the master side, we used Canny edge detection algorithm to get the navigation chart, which is similar to the actual navigation chart in the real operation, from EVE model. The positions of the tip of catheter were reconstructed in the navigation chart according to received coordinate values from the slave site. The navigation chart was also defined in Cartesian coordinates. The canny edge detection algorithm was developed by John F. Canny as a means to detect edge lines and gradients for the purpose of image processing. This algorithm provides good detection and localization of real edges while providing minimal response in low noise environments. This algorithm is well known and explained in any introductory text on image processing. The main stages of the Canny Algorithm are as follows: Noise reduction by filtering with a Gaussian blurring filter; determining the gradients of an image to highlight regions with high spatial derivatives; relate the edge gradients to directions that can be traced; Tracing valid edges; and Hysteresis thresholding to eliminate breaking up of edge contours.

B. The Implementation Method of Three-dimensional Model

In this tele-operative surgical system, image processor module plays an important role, which consists of 5 steps to calculate the three-dimensional coordinates of catheter. **Fig. 4** shows the procedures of images processor module.

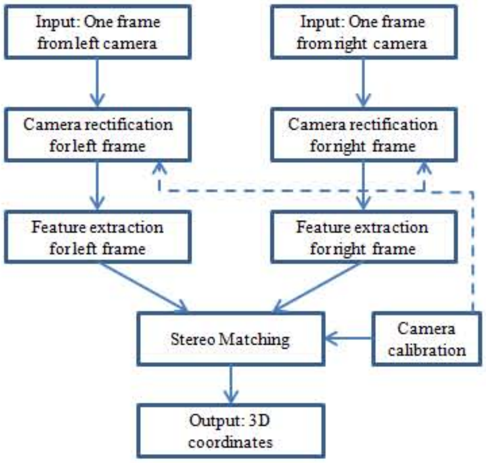


Fig.4 Procedures of images processor module

The purpose of camera calibration is to build a corresponding relationship between three-dimensional world coordinates and two-dimensional image coordinates. Two cameras, one on the left of catheter and another on the right of catheter, are used for capturing images. The two cameras are nonparallel.

To remove the barrel effect (due to lens distortion) and to find the perspective parameters, each of the cameras is calibrated using the Zhang's method. This method is simple and mainly based on different nonparallel views of a planar pattern (e.g., a chessboard). We have used a quadratic radial distortion model with four parameters. If the coordinates of a point in the undistorted image are (u, v), then the coordinates of the same point in the distorted image, $(\hat{u}, \hat{v})$ are obtained as follows:

$$\hat{u} = u(1 + k_1 r^2 + k_2 r^4 + 2p_1 v + 2p_2 u) + p_2 r^2$$
$$\hat{v} = v(1 + k_1 r^2 + k_2 r^4 + 2p_1 u + 2p_2 v) + p_2 r^2 \quad (1)$$

where $r = \sqrt{u^2 + v^2}$ and k_1, k_2, p_1 and p_2 are the coefficients of radial distortion. The perspective transformation matrix is assumed as:

$$A = \begin{bmatrix} f_u & \alpha & c_u \\ 0 & f_v & c_v \\ 0 & 0 & 1 \end{bmatrix}$$

where (c_u, c_v) is the principal point of the image (near the center of the image), α is the parameter describing the skewness of the two image axes, and f_u and f_v are the focal lengths along the horizontal and vertical axes of the image, respectively. The algorithm that is used for computing intrinsic and distortion parameters is as follows: 1) capture different nonparallel views of a 6 × 8 black and white planar chessboard pattern; 2) detect the corners of the pattern in each of the views; 3) find a homography for all points in the set of images, where a homography is a matrix of perspective transforms between the calibration pattern plane and the camera view plane; 4) initialize intrinsic parameters and set the distortion parameters to zero; 5) find extrinsic parameters for each image of the pattern; and 6) minimize the error of the projection points with all the parameters using a maximum likelihood algorithm.

TABLE I
Distortion Parameters For Left and Right Cameras

	k_1	k_2	p_1	p_2
Left Camera	-0.197	0.628	0	0
Right Camera	-0.2443	0.8192	0	0

TABLE II
Intrinsic Parameters For Left and Right Cameras

	f_u	f_v	α	c_u	c_v
Left Camera	991.74	991.74	0	319.5	239.5
Right Camera	1028	1028	0	319.5	239.5

The distortion parameters k_1, k_2, p_1, p_2 and the intrinsic parameters f_u, f_v, c_u, c_v, α, resulting from this algorithm for each of the cameras are listed in Table I and Table II, respectively.

The accuracy of this calibration method has been reported to be 0.7197 pixel.

After camera calibration process, a mapping relationship is built between two-dimensional space and three-dimensional space. Next is the method to obtain three-dimensional information according to two calibrated cameras.

We use epipolar geometric constraints to shorten stereo matching time and improve matching precision by reducing the searching scope of stereo matching algorithm from a two-dimensional plane to a one-dimensional line which is called epipolar line. **Fig. 5** shows the epipolar geometry of two cameras system, where the line between C1 and C2 is baseline and e1, e2 are epipolar points for left and right camera respectively. Also, *l1* and *l2* are epipolar lines. According to the epipolar geometric constraints of two-camera system, for any pixel pl in the left image, its correspondence must lie on the epipolar line in the right image. In addition, we can calculate the corresponding point according to intrinsic camera matrix computed in last step.

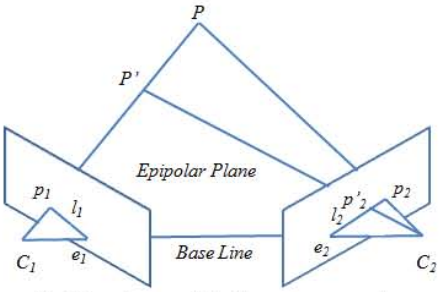


Fig.5 Imaging model of two cameras system

Furthermore, camera rectification comes to improve the epipolar geometric constraints. The purpose of camera rectification is to normalize the distribution of epipolar lines in left and right views. After camera rectification, the epipolar lines become parallel to the horizontal image axis and collinear (the same scan lines in both images) and for such a configuration, to find the correspondence of pixel (x, y) in the right image, only pixels (*, y) are considered.

The process of image rectification is to make two-dimensional transform for left and right images as equation 2 shown, where p1 and p2 are the coordinates of uncorrected images and $\overline{p_1}$, $\overline{p_2}$ are rectified points corresponding to p1 and p2. U1 and U2 are the two-dimensional transforms.

$$\overline{p_1} = U_1 p_1$$

$$\overline{p_2} = U_2 p_2$$

U1 and U2 consist of three kinds of sub-transforms. Assumed that:

$$U_1 = U = \begin{bmatrix} u_1^T \\ u_2^T \\ u_3^T \end{bmatrix} = \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & 1 \end{bmatrix}$$

U1 can be decomposed to Us, Ur and Up. Up is perspective transformation in which epipolar point is moved to infinite distance to make epipolar lines become to group of parallel lines.

$$U_p = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ c_a & c_b & 1 \end{bmatrix}$$

Ur is similarity transformation in which parallel epipolar lines are transformed to parallel to horizontal axis of image coordinates.

$$U_r = \begin{bmatrix} b_2 - b_3 c_2 & b_3 c_1 - b_1 & 0 \\ b_1 - b_3 c_1 & b_2 - b_3 c_2 & b_3 \\ 0 & 0 & 1 \end{bmatrix}$$

In order to reduce distortion, Us, shear transformation, is performed in left and right images in which image distortion in horizontal direction reach the minimum value.

$$U_s = \begin{bmatrix} s_1 & s_2 & s_3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

In our case, we used Canny edge detection algorithm to get contour of catheter as the feature regions of interested. First step is to separate catheter from other objects in left and right images such as blood vessel model. Then, Canny detection algorithm is applied to calculate the feature points. In order to extract contours of catheter efficiently, dilate algorithm is used.

In terms of stereo matching between left and right images, we applied a robust and accurate algorithm, block matching algorithm, to compute stereo correspondence. The block matching stereo correspondence algorithm is very fast single-pass stereo matching algorithm that uses sliding sums of absolute differences between pixels in the left image and the pixels in the right image, shifted by some varying amount of pixels. In order to improve quality and readability of the disparity map, the algorithm includes pre-filtering and post-filtering procedures

As for the rectified images captured from left and right camera, there is a fast algorithm to rebuild three-dimensional coordinates of catheter due to the fact that the pair of left and right images can be considered as two images captured by two parallel cameras.

Assumed that p1= (u1 v1 1)^T and p2= (u2 v1 1)^T are a pair of matched points. Also, the coordinates value of point P in world coordinates defined by p1 and p2 from left camera perspective view is: P = (X, Y, Z, 1)^T and P = (X-b, Y, Z, 1)^T from right camera perspective view, where b is the baseline between two cameras.

Thus, three dimensional coordinates can be computed as:

$$u_1 - u_0 = \frac{fX}{kZ} - \frac{fY}{kZ} \cot\theta$$

$$v_1 - v_0 = \frac{fY}{lZ \sin\theta}$$

$$u_2 - u_0 = \frac{f(X-b)}{kZ} - \frac{fY}{kZ} \cot\theta$$

Due to d = u1 - u2 as the disparity value, (X, Y, Z) can be calculated according to the three equations above:

$$\begin{cases} X = \dfrac{b}{d}\left[u_1 - u_0 + \dfrac{(v_1 - v_0) l \cot\theta}{k}\right] \\ Y = \dfrac{(v_1 - v_0) l b \sin\theta}{kd} \\ Z = \dfrac{fb}{kd} \end{cases}$$

III. EXPERIMENTAL RESULTS

To testify our image-processing method proposed in this paper, we designed a series of experiments to compare the time delay of transmission of visual feedback from surgical site to operation site before and after applying the proposed architecture. In addition, we carried out a train of experiments to compare the shape of real catheter in world space to virtual catheter in virtual reality environment.

A. Internet-based Catheter Operating System

The conceptual principle of our novel master-slave robotic catheter operating system has been shown in **Fig.6-8** [12]-[16]. The whole system concludes two parts, operation site and surgery site. Surgeons operate master system at operation site, and actual operation is performed at surgery site. And the two sites are connected by Internet.

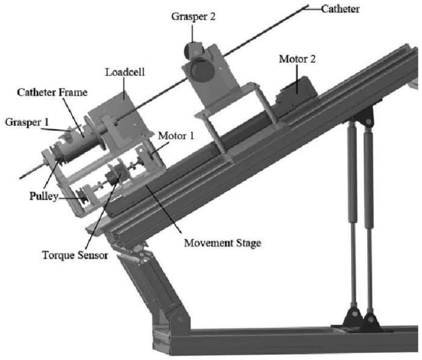


Fig.6 The catheter manipulator

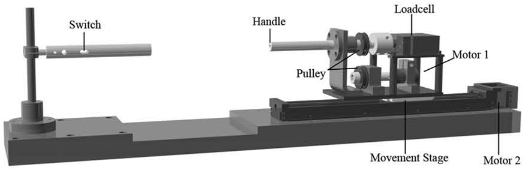


Fig.7 The structure of the surgeon console

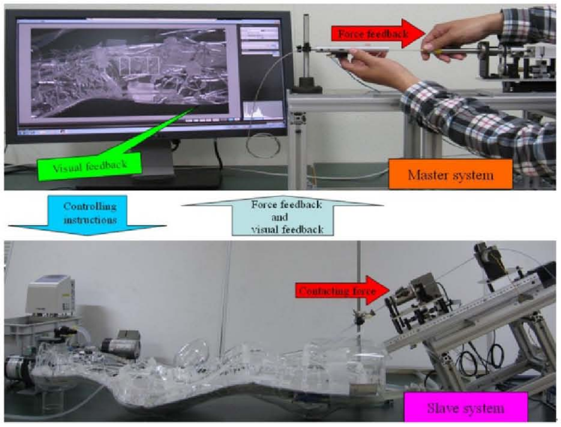


Fig.8 The surgical procedure

Fig.8 shows surgical procedure using our catheter operating system. At operation site, surgeons see the monitoring image and operate the handle of master manipulator to insert or rotate catheter, as operate the catheter directly, at the same time, controlling instructions are transmitted to surgery site. According to the controlling signals from master system, slave manipulator can insert or rotate catheter as if surgeons operate catheter directly. Based on force feedback and monitoring image information, surgeons carry out the surgical operations.

B. *Experiments on Two-dimensional Reconstruction*

To testify the accuracy of our method, we operated the handle of master manipulator to insert or rotate catheter, throughout the aortas located in the slave side of the master-slave catheter operating system, as showed in figure 7, using Endo Vascular Evaluator Model (EVE) to simulate blood vessels and load cell, as shown in figure 7. EVE model is made of a special silicone that recreates the elasticity and friction of human vasculature, simulating the sensation and behavior of catheter manipulation.

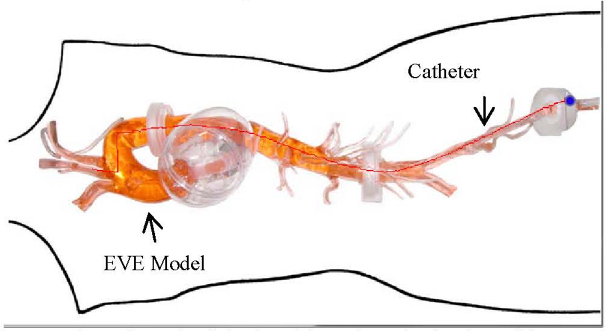


Fig.9 The path of the tip of the catheter in the slave side

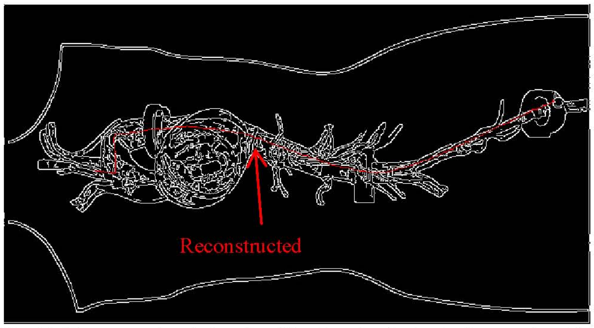
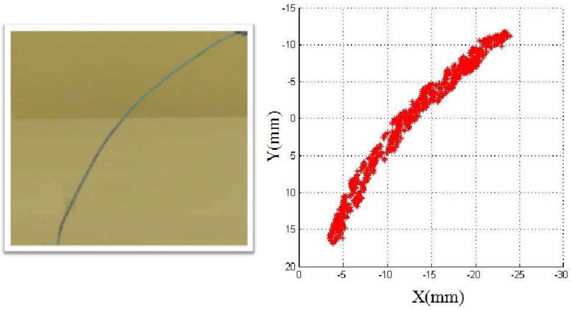


Fig.10 The reconstructed path of the tip of catheter in the master side

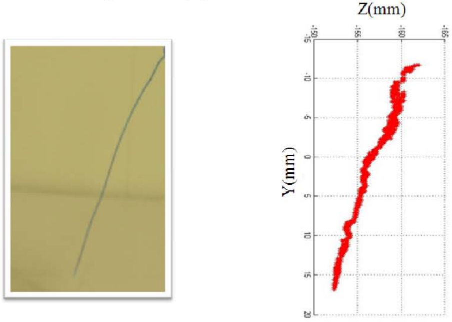
C. *Experiments on Three-dimensional Reconstruction*

The reconstruction algorithm has been introduced in part II of this paper. Here are two experiments to show the results using our proposed methods.

Fig. 11(a) and **Fig. 11(c)** show the extracted real catheter in world space from front view and lateral view and **Fig. 11(b)** and **Fig. 11(d)** show the corresponding coordinates of front view and lateral view.



(a) Real catheter in X-Y plane (b) Reconstructed coordinates in X-Y plane



(c) Real catheter in Y-Z plane (d) Reconstructed coordinates in Y-Z plane

Fig. 11 Results of experiment I:(a)front view (c) lateral view

D. Remote Transmission Experiments between China and Japan

Based on our method, the transmission of visual feedback is changed from image transmission to coordinates extraction and transmission. We conducted five experiments to compare the transmission time before and after applying the proposed architecture. We delivered 100 pieces of images captured in our remote expriments and their corresponding coordinates rebuit based on our method from Takamatsu, Japan to Beijing, China. The sizes of each image and its coordinates files are approximately 90,939 Bytes and 17,038 Bytes, respectively. The left bar in **Fig.12** shows transmission time of images and the other bar shows the sum time of coordinates extraction and coordinates transmission. The transmission time can be reduced efficiently.

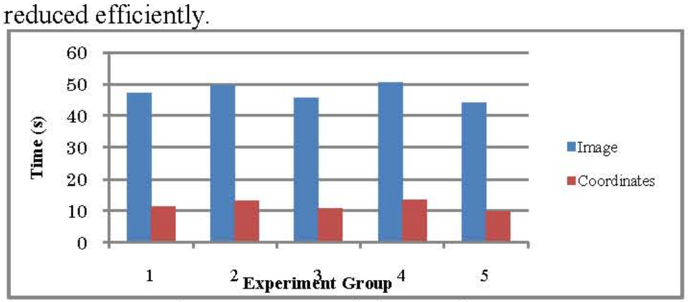


Fig.12 Remote transmission experiments

IV. CONCLUSION AND FUTURE WORK

Due to the image-processing algorithm and virtual reality technology, a new method is designed and implemented to reduce the time delay of image transmission through the Internet. In addition, it is easier to synchronize visual information with other signals such as force signals and control instructions due to significant reduction of visual data volume converting visual information to the coordinate's values. Experimental results show that the shape of virtual catheter in virtual reality environment is similar to that in real world space.

It is not realistic to lay dedicated fiber-optic lines all around the world for remote surgery in terms of cost. Therefore, our research approach with conventional network infrastructures is important for the future wide use of the remote surgery applications.

The experimental result positively shows the feasibility of remote surgery using conventional network infrastructures. However, shorter time-delay and stabilized network (no time variation of quality) are desirable for the patient's safety and the surgeon's operability. The stabilization of the image of our catheter operating system is a subject for a further study. Further improvements of our Internet based catheter operating system can allow the remote surgery applications to be more realistic and practical.

ACKNOWLEDGMENT

This research is supported by Kagawa University Characteristic Prior Research fund 2013.

REFERENCES

[1] A. R. Abreu,M. A. Campos, and B. P. Krieger, "Pulmonary artery rupture induced by a pulmonary artery catheter: A case report and review of the literature," *Journal of Intensive Care Medicine.*, vol. 19, pp. 291–296, 2004.

[2] J. Marescaux, J. Leroy, M. Gagner, F. Rubino, D. Mutter, M. Vix, S.E. Butner and M. K. Smith, "Transatlantic robot-assisted telesurgery," NATURE, vol.413, no.27, pp.379–380, Septemper, 2001.

[3] F. Araf, R. Fuji, T. Fukuda, and M. Negoro, "New catheter driving method using linear stepping mechanism for intravascular neurosurgery," *Proceedings of IEEE International Conference on Robots and Automation*, Washington, DC, pp. 2944–2949, May 2002.

[4] F. Arai, M. Tanimoto, T. Fukuda, I. Shimojma, H. Matsuurq, and M. Negoro, "Multimedia tele-surgery using high speed optical fiber network and its application to intravascular neurosurgery," *Proceedings of IEEE International Conference on Robots and Automation*, pp. 878–883, 1996.

[5] R. Beyar, T. Wenderow, D. Lindner, G. Kumar, and R. Shofti, "Concept, design and pre-clinical studies for remote control percutaneous coronary interventions," *Eurointervention*, vol. 1, pp. 340–345, 2005.

[6] P. Arcara, C. Melchiorri, Control schemes for teleoperation with time delay: A comparative study, Robotics and Autonomous Systems 38 (2002) 49–64.

[7] P. F. Hokayem, M. W. Spong, Bilateral teleoperation: An historical survey, Automatica 42 (2006) 2035–2057.

[8] Christian Smith, Patric Jensfelt, A Predictor for Operator Input for Time-Delayed Teleoperation, Elsevier Mechatronics Journal, 2010.

[9] Jumpei Arata, Hiroki Takahashi et al, A remote surgery experiment between Japan and Thailand over Internet using a low latency CODEC system, IEEE International Conference on Robotics and Automation, Roma, Italy, 10-14 April 2007.

[10]H. Hawkeye King, Kotaro Tadano, Regina Donlin, Diana Friedman, Mitchell J.H. Lum, Victoria Asch, Cong Wang, Kenji Kawashima, Blake Hannaford, "Preliminary Protocol for Interoperable Telesurgery", Proceedings of the 2009 International Conference on Advanced Robotics, *pp.1-6* , June 22-16, 2009.

[11]Kazushi Onda, Takayuki Osa, Naohiko Sugita, Makoto Hashizume and Mamoru Mitsuishi, "Asynchronous Force and Visual Feedback in Teleoperative Laparoscopic Surgical System", The 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems, Taipei, Taiwan, pp844-849, October 18-22, 2010.

[12]Nan Xiao, Jian Guo, Shuxiang Guo, and Takashi Tamiya, A Robotic Catheter System with Real-time Force Feedback and Monitor, Journal of Australasian Physical and Engineering Sciences in Medicine, Vol. 35, No. 3, pp. 283-289, 2012.

[13]Jian Guo, Shuxiang Guo, Nan Xiao, Xu Ma, Shunichi Yoshida, Takashi Tamiya and Masahiko Kawanishi, "A Novel Robotic Catheter System with Force and Visual Feedback for Vascular Interventional Surgery", International Journal of Mechatronics and Automation, Vol.2, No.PP., 2012.

[14]Jin Guo, Shuxiang Guo, Nan Xiao, and Baofeng Gao, Virtual Reality Simulators based on a Novel Robotic Catheter Operating system for Training in Minimally Invasive Surgery, Journal of Robotics and Mechatronics, Vol. 24, No. 4, pp. 649-655, 2012.

[15]Xu Ma, Shuxiang Guo, Nan Xiao, Jian Guo, Shunichi Yoshida, Takashi Tamiya, and Masahiko Kawanishi, Development of a Novel Robotic Catheter Manipulating Systemwith Fuzzy PID Control, International Journal of Intelligent Mechatronics and Robotics (IJIMR), Vol. 2, No. 2, pp. 58-77, 2012.

[16]Yili Fu, Anzhu Gao, Hao Liu, and Shuxiang Guo, The master-slave catheterisation system for positioning the steerable catheter, Int. J. Mechatronics and Automation, Vol. 1, Nos. 3/4, pp. 143-152, 2011.