

Design and Implementation of a Robust and Real-time Ultrasonic Motion-capture System

Tomohiko Sato Shigeki Nakamura Kotaro Terabayashi Masanori Sugimoto

Department of Electrical Engineering and Information Systems

University of Tokyo

7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656 Japan

Email: {tomo, shigeki, terabayashi, sugi}@itl.t.u-tokyo.ac.jp

<http://www.itl.t.u-tokyo.ac.jp/>

Hiromichi Hashizume

Information Systems Architecture

Research Division

National Institute of Informatics

Email: has@nii.ac.jp

<http://www.nii.ac.jp/>

Abstract—In this paper, we propose an innovative motion-capture system using ultrasonic communications. Compared with existing commercial motion-capture systems that use optical or magnetic sensing, the proposed system can provide a cost-effective solution for industrial and entertainment applications. To design and implement the system, a distance-estimation method called the Extended Phase Accordance Method (EPAM), which can measure the distance to a moving object with a standard deviation of less than 1 mm, was devised. To improve the capture rate of the proposed system, the EPAM algorithm was implemented in a field-programmable gate array (FPGA). The current version of the system conducts motion capture using five markers attached to a user. It can work at around 10 frames per second (fps), with an error of less than 55 mm and a standard deviation of 42 mm. This demonstrates a moderate level of accuracy, which will be useful for several applications.

I. INTRODUCTION

Recently, motion-capture systems have become widely used for tracking and representing human motion. Optical systems that employ passive or active markers [1] and magnetic systems that capture changes in electromagnetic fields [2] are available commercially. Using inertial sensors for tracking human motion has also been investigated in existing studies[3]. Although some of these systems demonstrate frame rates of a few thousand per second and better than one-millimeter tracking accuracy, their cost makes them out of reach for many users.

There are several studies related to ultrasonic motion-capture systems. Vlasic et al. [4] proposed a method for detecting errors in ultrasonic distance measurements by using inertial sensors. In addition, a system that integrates inertial and ultrasonic sensors was proposed by Stifmeier et al. [5]. The target of their system was to categorize users' activities into predefined classes. In this paper, we propose a motion-capture system that uses only ultrasonic signals. To the best of our knowledge, this is the first ultrasonic motion-capture system that achieves sufficient accuracy and update rates to be useful for several applications. The required performance of the proposed system was to have an error and standard deviation of less than 50 mm at more than 10 fps. These

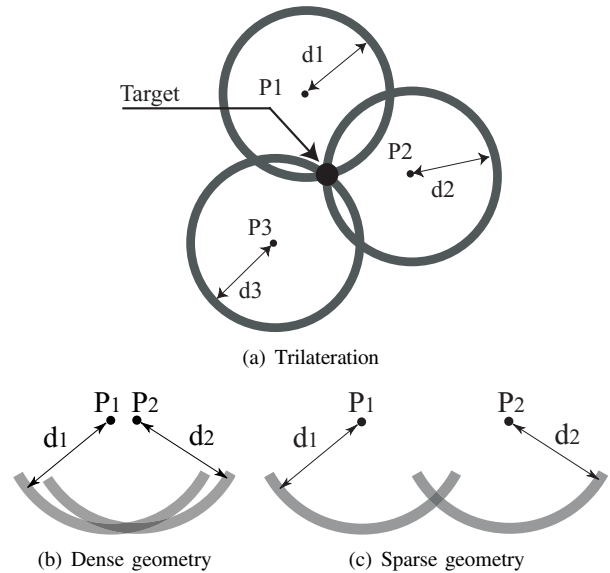


Fig. 1. Trilateration and the effects of the geometric arrangement of measuring nodes

requirements were specified to enable the system to identify and capture everyday human actions such as walking and exercising. We also investigated the software for several gesture-based games (e.g., 3D Monster Maze at about 6 fps) and found that these requirements were reasonable for such applications. Position estimation using ultrasonic signals requires lines of sight to prevent obstacles such as parts of the user's body from hindering the distance measurements between transmitters and receivers. The speed of sound being less than the speed of light is an advantage for ultrasonic signal processing, in that distance measurements require less computational complexity. On the other hand, the time needed for the propagation between transmitters and receivers determines the theoretical upper limit of the update rate for ultrasonic motion-capture systems. To cope with these problems, a robust and accurate tracking technique called the Extended Phase Accordance Method (EPAM), which can estimate accurately the distance to a moving object, was developed [9]. In addition,

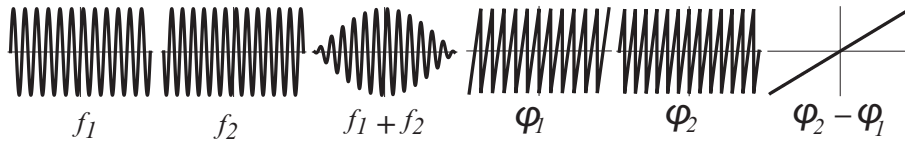


Fig. 2. Sync pattern and epoch

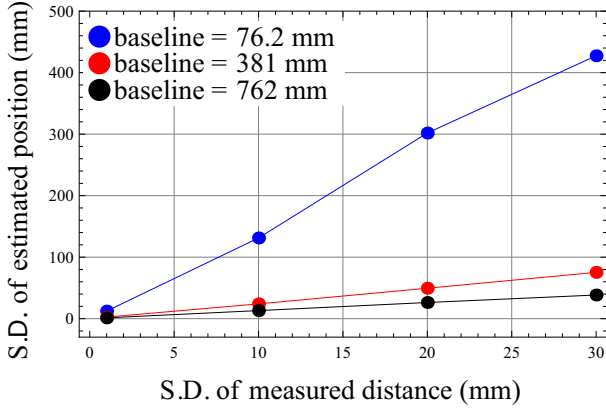


Fig. 3. How ranging accuracy affects 3D localization accuracy by changing baselines

a rapid calculation method on a field-programmable gate array (FPGA), which improved the update rate of the proposed motion-capture system to almost the theoretical limit, was implemented. Through experimentation in real environments, the system was found to perform motion-capture tasks at around 10 fps and within 54 mm accuracy. In this paper, the design and implementation of the proposed system and its performance evaluation are described.

II. ULTRASONIC MOTION-CAPTURE SYSTEM

A. Robust ultrasonic trilateration

To estimate the position of a target object in a 3D space using ultrasonic signals, a trilateration technique is used that requires at distance measurements from at least three different measuring nodes (transmitters or receivers) at known positions (P1, P2 and P3 in Fig. 1(a)) to the target. Trilateration has been implemented in ultrasonics-based localization systems such as Active Bat[6], Cricket[7] and Dolphin [8].

However, if there is even one obstacle between a target object and one of the three measuring nodes, with line of sight between them not being guaranteed, we cannot estimate the 3D position of the target successfully. To reduce the possibility of such a situation, we should consider arranging the measuring nodes close together, as shown in Fig. 1(b). This “dense geometry” arrangement has a short baseline (the distance between each measuring node), thereby degrading the accuracy of the localization, in contrast to the “sparse geometry” arrangement shown in Fig. 1(c).

Fig. 3 shows how distance measurements between transmitters and receivers affect the 3D localization accuracy. In this simulation, a 3D position of a transmitter was estimated by

using three receiver sensors with different baseline settings and its standard deviations were calculated through 1000 measurements for each setting. The simulation clarified that if the standard deviations of distance measurements between a transmitter and a receiver was less than 1 mm, it was possible to localize the transmitter with less than 2 centimeter accuracy, even with the 76.2 mm baseline. However, in many existing method using envelopes of received signals to detect the time of arrival, standard deviations of their distance measurements range from at least a few millimeters to several centimeters.

To implement a “dense geometry” system while retaining a moderate level of localization accuracy, a highly accurate distance-measurement method is indispensable. The EPAM algorithm, which is described in the next section, is a suitable distance-measurement technique for such requirements.

B. Multichannel ultrasonic communications

The proposed system in this paper requires multiple wireless communications between transmitters and receivers, for which time-division multiplexing (TDM) and frequency-division multiplexing (FDM) are investigated in this study. Each has pros and cons. TDM is easily implemented, but having a different time slot for each communication causes latency. On the other hand, FDM allows simultaneous multiple communications, but it needs broadband transducers that are more expensive than narrowband transducers. FDM requires more computational complexity in the receivers for detecting signals from different transmitters simultaneously. Therefore, for an initial version of the proposed system, TDM was chosen for its low cost and simple implementation.

Another issue to be examined was whether the system should be implemented as an active system or a passive system. In a passive system, a receiver is attached to the user’s body and receives signals from a transmitter. In a TDM-based system, because the receiver in motion cannot accept signals from multiple transmitters simultaneously, its 3D position cannot be estimated correctly via trilateration. Therefore, in this study, an active system was selected for implementation.

The frame rate of the proposed system is limited by the number of time slots assigned to each transmitter. To achieve 10 fps or more when using five transmitters, the time slot for each transmitter is limited to 20 ms. Assuming a sound velocity of 340 m/s in air (at room temperature), the theoretical maximum distance between a transmitter attached to a user and a receiver at a static location must not exceed 6.8 m. This distance provides a user with sufficient working space to enable a motion-capture system to be used. Therefore,

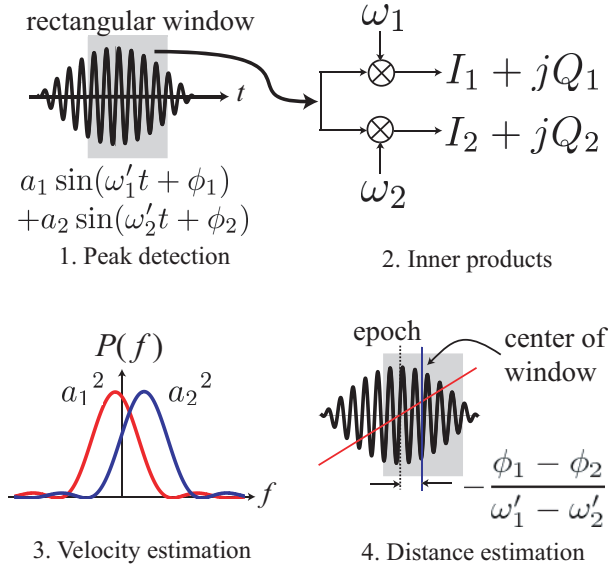


Fig. 4. Overview of the EPAM algorithm

the challenge is to make the calculations for the distance measurements and localization as rapidly as possible. To satisfy this requirement, we implemented the signal-processing calculations in hardware, using an FPGA.

III. THE EXTENDED PHASE ACCORDANCE METHOD

The EPAM can estimate both the distance to a moving object and the object's velocity simultaneously and in real time. In this method, an ultrasonic signal burst, called a sync pattern, which is composed of two or more frequency signals, is transmitted from a transmitter, as shown in Fig. 2. The phase of each signal coincides at only one point in time, called the epoch. The algorithm for detecting the distance and velocity of a moving object is described in [9].

In the current implementation, a sync pattern comprising two sinusoidal waves of frequencies 39.75 kHz and 40.25 kHz, and lasting 2 ms is used. By processing the waveform of 1 ms retrieved from the received sync pattern at a receiver, the EPAM can measure the distance and velocity very accurately, with standard deviations of 1 mm, and 15 mm/s, respectively.

Because the EPAM uses a short-burst ultrasonic signal (approximately 680 mm, at room temperature), multipath interference is unlikely to occur, which is another advantage of the motion-capture system described in this paper.

IV. SYSTEM IMPLEMENTATION

A. Calculations on the FPGA

The core algorithm of the EPAM involves finding the phases of the received signal via a quadrature-detection method that includes inner-product calculations (Fig. 4). In an earlier system [9], the EPAM was implemented on a microcontroller (SH2/7145, 48MHz, by Renesas Technology) and its update rate was around 10 Hz. When five transmitters were used as active markers attached to the user, the frame rate reduced to 2 fps, which was too slow for tracking the user's motion.

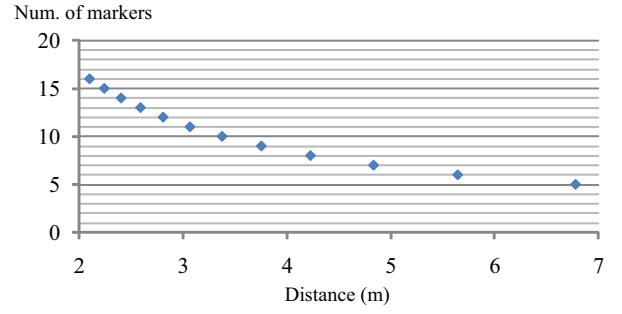


Fig. 5. Maximum numbers of markers by changing distances to receivers

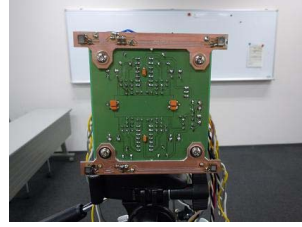


Fig. 6. Ultrasonic receiver unit with four sensors

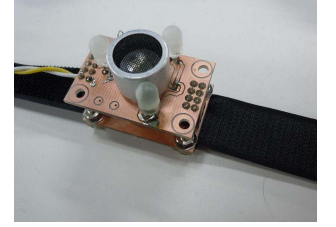


Fig. 7. A marker to be attached to the user (an ultrasonic receiver and three IR-LEDs are mounted)

In the proposed system, we implemented the EPAM algorithm on an FPGA (Spartan3 XC3S200, by Xilinx) for rapid parallel processing. Through experimentation, we found that an EPAM calculation on the FPGA requires only 83.5 μ sec. Therefore, the proposed system can work at 9.96 fps, which is almost at the theoretical upper limit of the update rate. To make the system work at 10 fps, the maximum distance between its transmitters and receivers was set to 6.77 m, which is slightly below the maximum sound propagation distance (6.8 m). By reducing the maximum distance between transmitter and receiver, we can increase either the update rate (to more than 10 fps) or the number of markers attached to the user (Fig. 5).

B. Ultrasonic receivers and transmitters

Because the EPAM can measure the distance to a moving object accurately, it is possible to set a short baseline between the receiver sensors for the trilateration. In the current implementation, four ultrasonic receiver sensors (SPM0404UD5, by Knowles Acoustics) are arranged to form a 76.2 mm square on one receiver unit, as shown in Fig. 6.

Fig. 7 shows a marker that includes an ultrasonic transmitter (UT1612MPR, by SPL) and three infrared light-emitting diodes (IR-LEDs), which is captured by multiple cameras (Firefly MV, 1320x1048 pixels, by PointGrey) to evaluate the performance of the proposed motion-capture system. Five markers are attached to the user's body, with each being activated interchangeably to transmit ultrasonic signals at 20 ms intervals. The ultrasonic transmitters and receivers are synchronized by a trigger pulse via a wireless sensor-network module (MICAz, by Crossbow Technology).

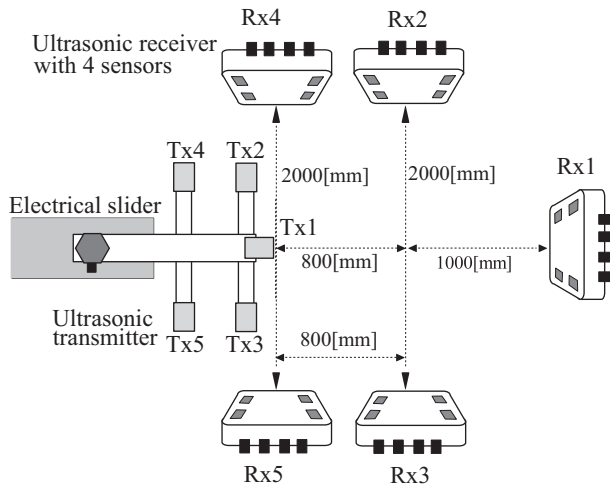


Fig. 8. Experimental setup using an electrical slider

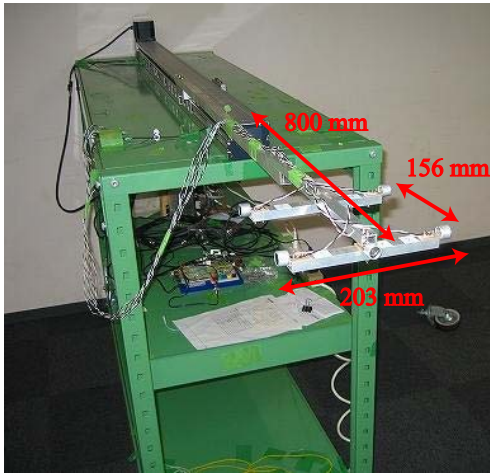


Fig. 9. Five transmitters are moved by the electrical slider

V. EXPERIMENT

Two experiments were conducted. Although the proposed system can identify positions of transmitters about 6.8 m away from receivers, its localization accuracy becomes worse due to the signal attenuation from transmitters. By examining a sufficient size of a motion-capture area that enabled user's fullbody actions, shorter distances than the theoretical limit between transmitters and receivers were set in the experiments.

A. Motion tracking using electrical slider

In the first experiment, five transmitters were moved forwards and backwards at a constant velocity by an electrical slider, as shown in Fig. 8 and Fig. 9. The purpose of the experiment was to reveal any problems with the proposed system in a controlled setting, where the electrical slider could measure the true position and velocity of the moving transmitters.

The velocity of the electrical slider was set to ± 1.0 m/s and motion data for the transmitters was obtained over a period of 51.8 s. The success rate for the motion capture at each receiver

TABLE I
SUCCESS RATE FOR MOTION CAPTURE [%]

	Tx1	Tx2	Tx3	Tx4	Tx5
Rx1	100	0	0	0	0
Rx2	0	78.0	0	62.2	0
Rx3	0	0	64.1	0	69.3
Rx4	0	99.6	0	100	0
Rx5	0	0	85.1	0	100

TABLE II
ERRORS IN THE ESTIMATED TRANSMITTERS' POSITIONS WHEN ONE OR TWO RECEIVER UNITS ARE AVAILABLE [MM]

Num. of Rx	Tx1	Tx2	Tx3	Tx4	Tx5
1	78.26 (Rx1)	349.8(Rx2) 233.6(Rx4)	304.1(Rx3) 340.3(Rx5)	351.4(Rx2) 79.64(Rx4)	315.0(Rx3) 96.93(Rx5)
2		67.67	91.17	54.90	78.89

is shown in Table I. The table indicates that the success rate for localizing each transmitter by a single receiver unit was more than 60% and that transmitters other than Tx1 were tracked by two receiver units simultaneously. Because of the directivity of the ultrasonic transmitter and receiver sensors, it was found that the transmitters were successfully tracked only within limited areas. However, no transmitter that was not tracked by any receiver unit was found. When multiple receiver units received signals from a transmitter, the system could estimate its position by using the longer baselines between the units. The errors and their standard deviations for the position estimates are shown in Table II and III, respectively. From these results, when only one receiver unit was available, the accuracy level of the position estimations was not always sufficient for tracking the transmitters. To conduct a deeper analysis on the performance of the system, we tested motion tracking with the transmitters moving at speeds of ± 0.1 and ± 1.0 m/s. The trajectory of Tx2 shown in Fig. 10 indicates that as the incident angles of ultrasonic signals arriving at the receiver units increased, the tracking accuracy deteriorated, with its accuracy being less affected by the velocity. Because the EPAM estimates distance by using the phases of received signals, we think that one of the causes of the deterioration was that the phase response of the ultrasonic sensors depends on its directivity.

Despite relatively large errors caused by the phase responses being found, the positions of transmitters were identified with

TABLE III
STANDARD DEVIATIONS FOR THE ESTIMATED TRANSMITTERS' POSITIONS WHEN ONE OR TWO RECEIVER UNITS ARE AVAILABLE [MM]

Num. of Rx	Tx1	Tx2	Tx3	Tx4	Tx5
1	28.05 (Rx1)	191.8(Rx2) 128.2(Rx4)	171.6(Rx3) 197.8(Rx5)	195.0(Rx2) 58.50(Rx4)	196.0(Rx3) 54.49(Rx5)
2		35.80	35.87	21.29	30.71

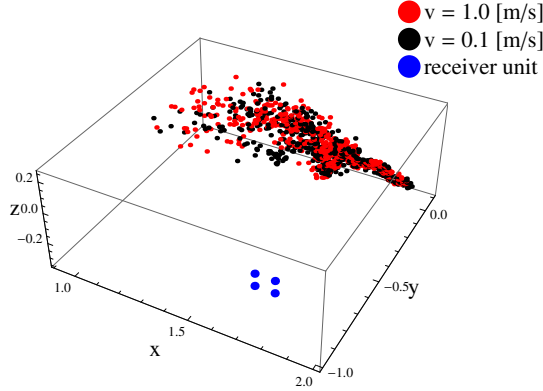


Fig. 10. Tx2's trajectory detected at Rx4

TABLE IV
POSITION ERRORS AND STANDARD DEVIATIONS FOR USER MOTION IN
THE PROPOSED SYSTEM [MM]

	Right hand	Left hand	Chest	Right leg	Left leg
Error	47.05	45.50	39.40	53.96	54.83
S.D.	24.41	41.74	17.50	12.33	9.93

a standard deviation of less than 35.9 mm by two receiver units.

B. Tracking a conductor's motion

In the second experiment, a user was asked to place five markers, on both arms, chest and both legs (Fig. 11). His motion was captured by ultrasonic receiver units and a stereo camera as shown in Fig. 12. The user behaved like an orchestral conductor and his motion data was captured over a period of 10 s. Table IV shows the position errors and standard deviations of the proposed system calculated by using the reference data. From the table, it is found that his motion was captured with an error of less than 55 mm and a standard deviation of 42 mm. The CDF (cumulative distribution function) in Fig. 13 shows that the 90th-percentile errors are 81.8 mm (right hand), 89.9 mm (left hand), 55.6 mm (chest), 76.3 mm (right leg) and 56.7 mm (left leg), respectively. From this experiment, we found that each marker was tracked to a moderate level of accuracy, even when its position was estimated by a single receiver unit. Overall, in the captured data, which included data for 100 positions for each marker, only three items of data (for the left hand) were missing, with the other position data being always obtained through one or more receiver units. This small failure rate ($3/500 = 0.6\%$) was caused by the compact configuration of the receiver unit whose baseline was much smaller than that for conventional 3D localization systems.

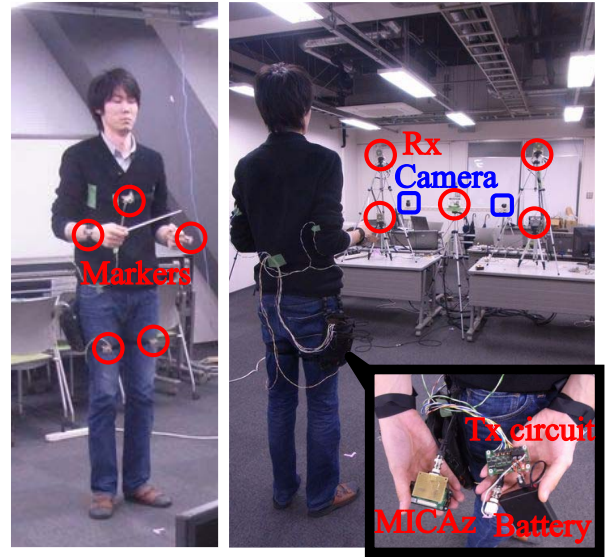


Fig. 11. Experimental setup for motion capture

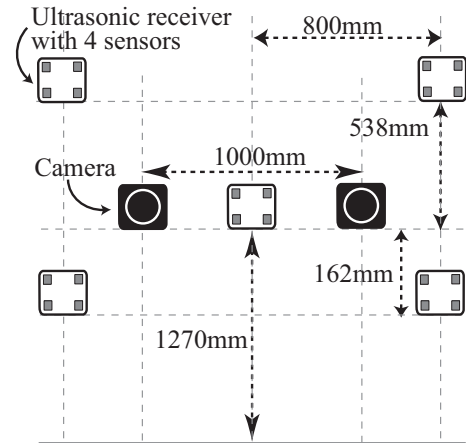


Fig. 12. Arrangement of receiver units and stereo camera

VI. DISCUSSIONS

A. Phase Characteristics

As shown in Fig. 3, 3D position estimates using the EPAM method with a short baseline (76.2mm) was expected to achieve around a few centimeter standard deviation. However, the proposed motion-capture system did not achieve the expected level of the performance in its standard deviations and errors by using a single receiver unit. One of the main causes of this deterioration is, as mentioned in the previous section, related to the phase characteristics of transmitter and receiver sensors. Fig 14 shows 3D position estimate results of a transmitter by using one receiver unit with four receiver sensors. The transmitter was mounted on the electrical slider (as shown in the previous section) moving close to and away from the receiver unit on a straight line. From the figure, it was found that the estimated positions of the transmitter had different errors at different positions. These errors were confirmed to

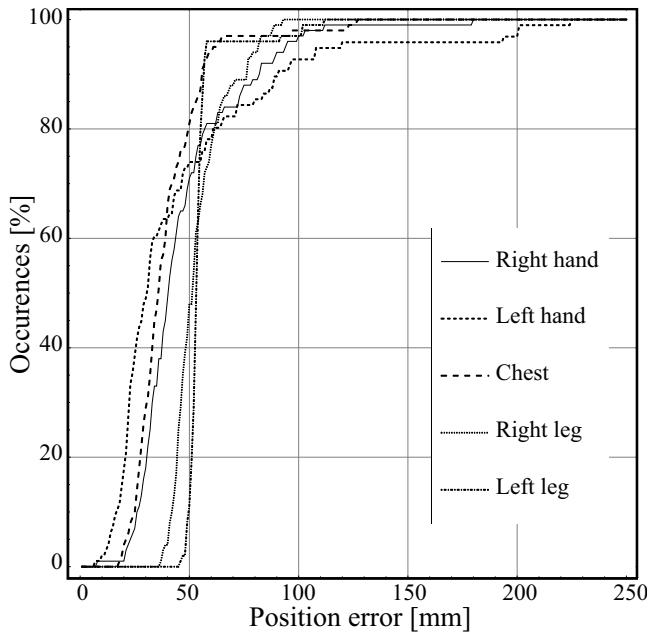


Fig. 13. Error CDF of measured positions of markers

be reproducible through repetitive measurements. When the slider changed its position, orientations between the transmitter sensor and each of the receiver sensors changed accordingly. Therefore, it was thought that the phase characteristics due to the directivity of the sensors affected the performance of the distance measurements between them, which deteriorated the 3D position estimates. Thus, we measured the phase characteristics of each sensor and are now investigating a compensation method for improving the accuracy of the proposed system.

B. Possible Applications

It was proved through the experiments that the performance of the proposed ultrasonic motion-capture system was lower in its update rate and accuracy than recent commercial systems using optical or magnetic sensing techniques. However, there are possibilities to explore applications suitable for the proposed system. The system will be useful for capturing not-so-fast body movements in our daily life. For example, it is possible to capture user's gestures in order to control video games or home appliances. Using the system for rehabilitation purposes such as learning or recovering body movements by elderly persons or patients is another possible option. By utilizing features of its cost-effective implementation with moderate accuracy, the system is expected to extend possible applications.

VII. CONCLUSION AND FUTURE WORK

In this paper, we have proposed an innovative motion-capture system using ultrasonic communications. The proposed system was found to capture user motion robustly at a moderate level of accuracy when using compact receiver units. The current system achieved around 10 fps with a standard

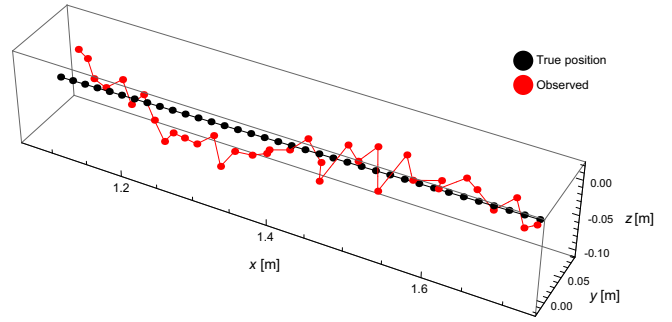


Fig. 14. Positioning errors due to phase characteristics of sensors

deviation of less than 42 mm, and will therefore be useful for applications that do not need to capture rapid motion.

There are several issues to be investigated. We have to compensate for differences in sensor phase response for signals coming from a variety of incident angles. In the present implementation, we do not utilize the velocity data fully to improve the position estimate. By applying an estimation filter, such as the Kalman filter, improved position accuracy becomes possible, which will be described in a later paper [10]. In future, we aim to implement and evaluate both ultrasonic motion-capture systems based on FDM and passive systems.

ACKNOWLEDGMENT

This research is sponsored by a Grant-in-Aid for Scientific Research by the Japanese Ministry of Education, Science, Culture, Sports and Technology, and Tateishi Science and Technology Foundation.

REFERENCES

- [1] VICON, <http://www.vicon.com/>.
- [2] MotionStar: <http://www.ascension-tech.com/>.
- [3] J. Hightower, G. Borriello: Location Systems for Ubiquitous Computing, *Computer*, vol. 34, no. 8, 2001, pp.57-66.
- [4] D. Vlastic, R. Adelsberger, G. Vannucci, J. Barnwell, M. Gross, W. Matusik, J. Popovic: Practical Motion Capture in Everyday Surroundings, *ACM Transactions on Graphics*, vol. 26, No. 3, Article 35, 2007.
- [5] T. Stiefmeier, G. Ogris, H. Junker, P. Lukowicz, G. Tröster: Combining Motion Sensors and Ultrasonic Hands Tracking for Continuous Activity Recognition in a Maintenance Scenario, in *Proceedings of the 10th IEEE International Symposium on Wearable Computers (ISWC 2006)*, Montreux, Switzerland, 2006, pp. 97-104.
- [6] A. Ward, A. Jones, A. Hopper: A New Location Technique for the Active Office, *IEEE Personal Communications*, Vol. 4, No. 5, 1997, pp.42-47.
- [7] N. B. Priyantha, A. K. L. Miu, H. Balakrishnan, The Cricket Location-Support System, in *Proceedings of the 6th Annual International Conference on Mobile Computing and Networking (MobiCom 2000)*, Boston, MA, 2000, pp. 32-43.
- [8] M. Hazas, A. Hopper: Broadband ultrasonic location systems for improved indoor positioning, *IEEE Transactions on Mobile Computing*, Vol.5, No.5, 2006, pp.536-547.
- [9] T. Sato, S. Nakamura, M. Sugimoto, H. Hashizume: Extended Phase Accordance Method : A Realtime and Accurate Technique for Estimating Position and Velocity of Moving Node using Ultrasonic Communication, *Sensors and Transducers Journal*, Vol. 9, Special Issue, Dec. 2010, pp.56-70.
- [10] S. Nakamura, T. Sato, M. Sugimoto, H. Hashizume: An Accurate and Compact 3D Positioning System for a Moving Target by Integrating Extended Phase Accordance Method and Particle Filter, in *Proceedings of IEEE International Conference on Indoor Positioning and Indoor Navigation (IPIN 2011)*, Guimarães, Portugal (to appear).