Proposal for A KINECT-Based Auscultation Practice System

Yoshitoshi Murata, Kazuhiro Yoshida
Faculty of Software and Information Science
Iwate Prefectural University
Takizawa, Japan
e-mail: y-murata@iwate-pu.ac.jp, kyoshida@ipu-office.iwate-pu.ac.jp

Natsuko Miura, Yoshihito Endo
Faculty of Nursing
Iwate Prefectural University
Takizawa, Japan
e-mail: natsuko@iwate-pu.ac.jp, y-endo@iwate-pu.ac.jp

***Abstract*—Students in medical and nursing schools have to practice auscultation. A humanoid simulator is effective for learning disease sounds and correct stethoscope location. Such humanoid simulators, however, are too expensive for most nursing schools to buy. In this paper, we propose a low-cost system for the practice of auscultation. In this system, students themselves play the role of a patient, instead of a humanoid, and stethoscope locations on the body are measured with KINECT. Practicing students hear disease sounds, synchronized with the movement of breathing, through earphones. Movements of the upper body from breathing are also detected by KINECT. Experimental results with a prototype showed that our system could perfectly detect stethoscope locations on a body, except for a few lower points, and it could also detect respiratory changes. There are a few challenges, however, to be solved in future work.**

Keywords-simulator; auscultation; nursing; KINECT.

I. INTRODUCTION

Generally, practicing auscultation is a required subject for students in medical and nursing schools. Humanoid-type simulators have been developed [1][2][3][4][5], and there are several reports that such simulators improve auscultation skills [6]. Unfortunately, these simulators are generally too expensive for most nursing schools to buy.

We are developing a new, low-cost auscultation simulator whose concept is different from that of the existing humanoid simulators. In this simulator, students themselves are the practice subjects instead of a humanoid model, and it is possible to detect the location of a stethoscope with KINECT that is a line of motion sensing input devices by Microsoft [7].

In the case of existing humanoid simulators, it is possible to know correct stethoscope locations by marking these points on a mannequin, but it is impossible to detect whether a stethoscope is actually placed correctly on a mannequin. Moreover, correct locations vary among patients according to body size. Our proposed simulator can both show correct locations on a body and detect whether a stethoscope is placed on correct points. The correct locations are normalized with respect to the positions of both shoulder joints and both hip joints.

In addition, most humanoid simulators cannot simulate the timing of breathing or the synchronized forward and backward movements of the upper body. Our simulator, however, can detect these forward and backward movements and provide exhalation and inhalation sounds synchronized with those movements.

We have developed a prototype system and evaluated it experimentally. The results showed that our system could perfectly detect stethoscope placement on a body at eight of ten points. Because two lower points were easily shadowed from KINECT by the T-shirt worn by a student acting as patient, our system could not always detect a stethoscope. Moreover, our system could detect changes in breathing, but it sometimes made a mistake in counting the number of them and the detection delay for respiratory changes was slightly larger than expected.

After introducing related works in Section II, we describe the concepts and features of our system in Section III. The key technologies of our simulator and the evaluation results are described in Section IV. The key points are summarized in Section V.

II. RELATED WORKS

Many kinds of patient simulators have been developed and provided as medical and nursing training tools [1][2][3][4][5]. Since we propose a new type of simulator for practicing auscultation, we first discuss existing auscultation simulators, which are divided into two groups: the humanoid model type, and the virtual reality type.

A. Humanoid model type

Kyoto Kagaku Co., Ltd. provides the Lung Sound Auscultation Trainer (LSAT) [2], shown in Figure 1, for respiratory auscultation. There are several small speakers inside a mannequin. Disease sounds are recorded from real patients. This simulator also works for cardiac auscultation by changing from respiratory sounds to cardiac sounds. Sakamoto Model Corporation provides the Sakamoto auscultation simulator [3]. This simulator also works for both respiratory and cardiac auscultation. Sakamoto provides a transparent cover for this simulator, as shown in Figure 2, to illustrate correct stethoscope locations.

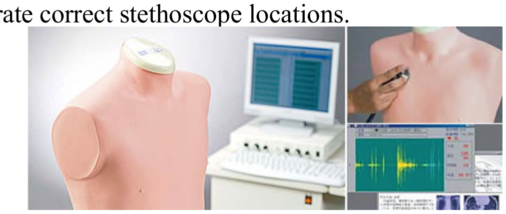


Figure 1. Lung Sound Auscultation Trainer (LSAT) by Kyoko Kagaku

 ISBN: 978-1-61208-470-1

Figure 2. Transparent chest cover, by Sakamoto Model, to illustrate correct stethoscope locations

Although the above two simulators are focused on the upper body, they simulate disease sounds, not the motion of the upper body. On the other hand, the SimMan® 3G [4] by Laerdal is an advanced patient simulator that can simulate the characteristics of a real patient, including the blood pressure, heart beat, chest motion, and so on. It is too expensive, however, for a general nursing school to buy.

B. Virtual reality type

Zadow, et al. developed the SimMed system for medical education [5]. By using an interactive multi-touch tabletop to display a simulated patient, as shown in Figure 3, they have created an immersive environment that supports a large variety of learning scenarios. The simulated patient can show skin changes and be animated to show realistic bodily and facial movements. By its nature, the setup allows scenarios to be repeated easily and to be changed and configured dynamically.

Figure 3. The SimMed system

SimMed is substantially lower in cost than a full-scale humanoid simulator. It has many functions, however, and is still too expensive for most nursing schools. Moreover, while students can touch the virtual patient on a display, they cannot physically feel the motion of the virtual patient.

III. CONCEPT OF KINECT-BASED SIMULATOR

Among the nursing skills that students have to learn, are: the recognition of different sounds between different kinds of disease and the knowledge about placing correct points for locating a stethoscope on a body. Moreover, in the case of respiratory auscultation, students have to listen to respiratory sounds for more than one cycle. Therefore, an auscultation practice system requires the following functions:

- Simulating real disease sounds at different points on the body.
- Showing correct points for locating a stethoscope on an operation display.
- Judging whether a stethoscope is located on showed points.
- Judging whether a stethoscope is fixed on a body for more than one cycle in respiratory.

As introduced in the above section, existing auscultation simulators represent patients with mannequins or virtual reality technology, and disease sounds played through speakers mounted on a humanoid model or an external speaker. Most disease sounds are recorded from real patients. Students learn differences in disease sounds and correct stethoscope locations on the body by marking them on a humanoid model or hearing lectures by a teacher. Since most such models, however, do not have functions for detecting stethoscope locations, they cannot show whether these locations are correct. Furthermore, since the humanoid model has only a single size, it is difficult to learn differences depending on body size.

With our practice system, on the other hand, students themselves act as patients, instead of mannequins. The stethoscope locations and forward and backward movement of a body from breathing are measured with KINECT, as shown in Figure 4. Students hear disease sounds, generated by a PC, through earphones. The sound volume for each point is different for locating a stethoscope as with a real patient.

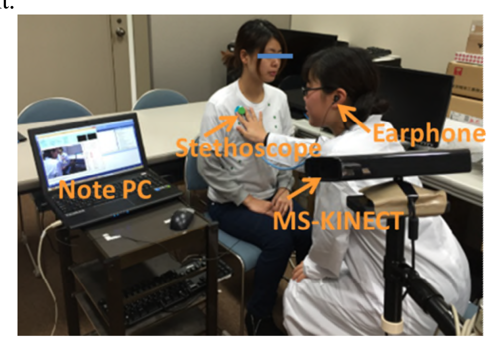


Figure 4. Auscultation practice with our proposed system

As introduced in the next section, color tracing technology is used to trace a stethoscope, and select “yellow green” for tracing. Therefore, we decided a white T-shirt for a student acting as a patient and a white coat for a student acting as a nurse wearing to remove “yellow green” and likeness in experiments.

IV. DETECTION TECHNOLOGIES

The following capabilities are required to implement our proposed auscultation practice tool:

- Tracing a stethoscope.
- Detecting a stethoscope’s location.
- Detecting a stethoscope on a body.

 ISBN: 978-1-61208-470-1

- Detecting forward and backward movements of the upper body from breathing.
- Automatically adjusting correct points for locating a stethoscope on a body according to body size.

Since KINECT automatically generates position data for a stethoscope while it is traced, we examine each of the above four issues except the second one.

A. *Tracing a stethoscope*

Two candidate tracing methods are shape tracing and color tracing. Since a stethoscope is held by hand, the shape of a stethoscope as viewed through a video camera changes over time. Therefore, we chose color tracing, rather than shape tracing. The process of color tracing is shown schematically in Figure 5. Video data from the BGR (Blue, Green, and Red) 32 output of KINECT is converted to HSV (Hue, Saturation, and Value) color data. First, the traced object, i.e., a stethoscope, is pointed, and its hue histogram is generated and stored. Then, masking data are generated for each frame from the HSV data, the minimum saturation, and the maximum and minimum brightness.

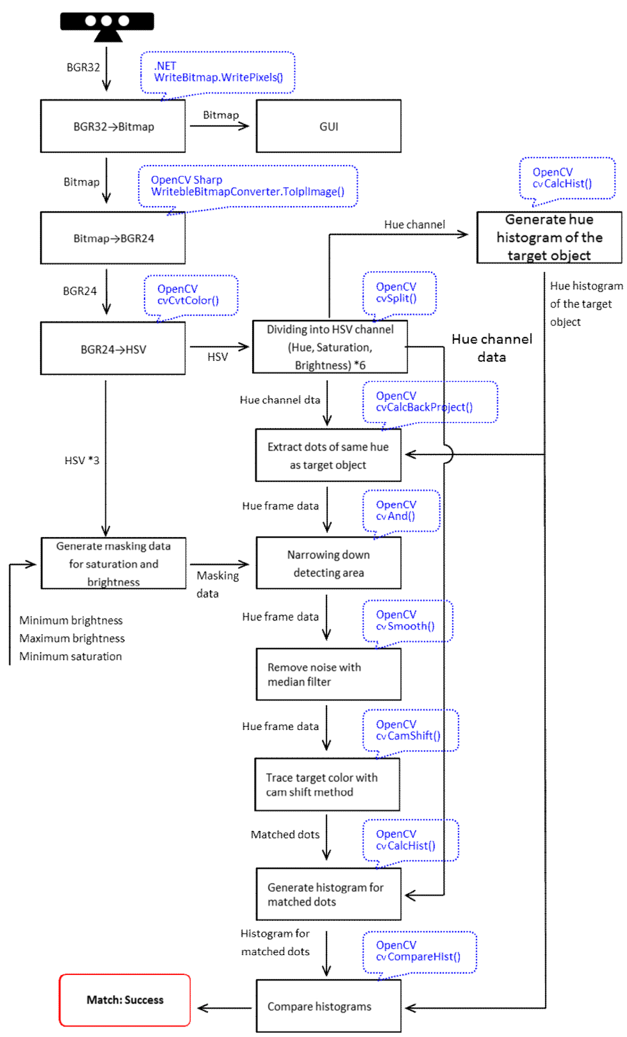


Figure 5. Process of color tracing

The video data for practice is also converted to HSV data, and target areas are separated with the masking data. Noise, including the hue data of the target area, is reduced by a median filter. The output data from the median filter is then traced by the cvCamShift function of Open CV [8]. Since cvCamShift sometimes outputs incorrect data, the hue histogram for the output data is repeatedly compared with the pre-stored hue histogram. When these histograms are equivalent, the color tracing is successful.

We used an experiment to select a target color from among seven choices: "red", "green", "light blue", "yellow", "yellow-green", "pink", and "orange". We used KINECT v1 and examined whether it could detect only a target color. We show the resulting data for the top three colors in Figure 6. "Yellow-green" had the best performance, with no portions having the target color except the target. "Light blue" also performed well, but since the color of a stethoscope tube is "light blue", that was detected. In the case of "yellow", dots of the same color appeared in the bottom-left region of the image. The other colors exhibited more dots of the same color or had a smaller target size. Hence, we chose "yellow-green" as the target color for our experiments.

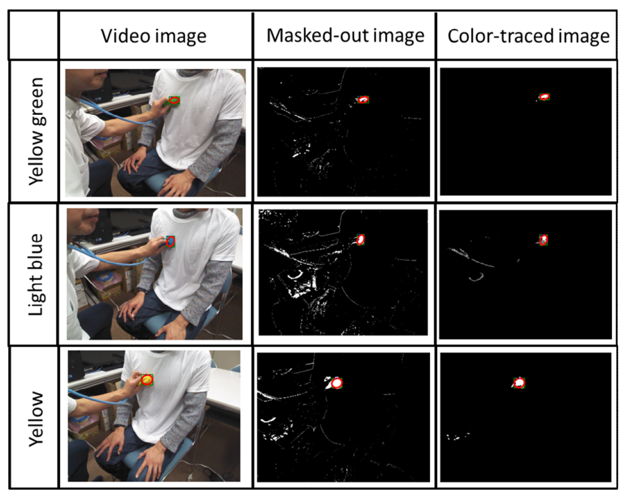


Figure 6. Experiment for deciding a target color

B. *Detecting a stethoscope on a body*

At this time, we think it is unnecessary to determine whether a stethoscope is exactly placed on a body, so we do not use any sensors on the stethoscope. Instead, we estimate a stethoscope is located on a body, where a stethoscope is placed and fixed within the distance S from a body surface for T seconds. For now, we use S = 10 cm, and T = 0.3 second to achieve balance between certainty and fast recognition. If stricter detection is required, we can change these parameters. Before measuring length L_{st} between a stethoscope and a KINECT, we determine whether the stethoscope is within the outline of a body, by using a pre-installed program on a KINECT to get an outline.

We experimentally examined whether this method was useful. A student acting as patient wore a white T-shirt with dots marking correct stethoscope locations; and a student acting as nurse placed a stethoscope on these marked dots.

 ISBN: 978-1-61208-470-1

We marked 10 dots on the T-shirt, as shown in Figure 7. The participants were five male students and five female students. The experimental results are listed in Table I.

Figure 7. T-shirt with dots marking stethoscope locations for a female participant

TABLE I. COUNT OF DETECTING A STETHOSCOPE PLACED ON A BODY

Point #	1	2	3	4	5
Male	5	5	5	5	5
Female	5	5	5	5	5
Point #	6	7	8	9	10
Male	5	5	5	5	3
Female	5	5	4	3	2

The proposed system sometimes missed when the stethoscope was placed at certain points (points 8, 9, 10). As seen from Figure 4, a stethoscope placed on one of these points would sometimes be shadowed from KINECT, especially for women, since these points are below the breast. Possible solutions for this problem are the follows:

- Switch from a T-shirt to clothing more fitted to the body.
- Attach some dimensional marker to the stethoscope.

C. *Detecting upper body motion*

Burba et al. used chest motion to detect breathing [9], since there are two main types of breathing: chest respiration and abdominal respiration. Therefore, we select six points for measuring movement of the chest and abdomen in this experiment, as shown in Figure 8. Since the stethoscope location is not always fixed, we do not consider it in this experiment. The upper three measuring points are the inner junctions of five vertical lines equally dividing the space between both shoulders into four regions and a horizontal line halfway between the height of the center of the spine and the average height of both shoulders. The lower three measuring points are the inner junctions of the above-mentioned vertical lines and a horizontal line through the hip center. We adopt the same direction of more than 4 points as the resultant direction.

We designed our system to detect the changes from expiration to inhalation and vice versa as quickly as possible. Since there were small but rapid changes in the output data, we used a moving average of 30 samples, with a sampling period of 10 ms. We then have specified that when the sampling data continued to increase 3 times, the breathing mode was expiration; and when the sampling data continued to decrease 3 times, the breathing mode was inhalation.

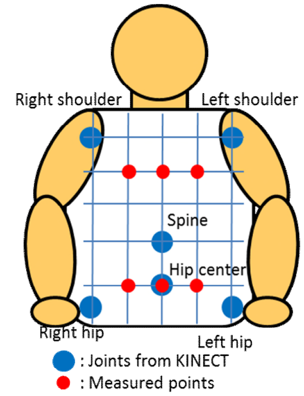


Figure 8. Measuring points for breathing motion

We measured the number of breathings and their periods with our proposed system and compared the results with other data obtained by participants keying the up and down arrow keys on a keyboard. The experimental results are shown in Table II. We measured them for two KINECTs that are K-1 and K-2. Both of them were the same model. At first, we measured for many participants with K-1. Our system counted more and more breaths than did keying for most participants. Data shown in Table II are some of them. In addition, the output data for some measuring points had extraordinary values, as shown in Figure 9 (2). We changed K-1 to K-2 to clear the reason of this problem. Our system with K-2 could count more accurately than that with K-1. However, our system counted fewer breaths than did keying for participant D. And, the extraordinary values were sometimes measured with K-2, too.

As shown in Figure 9 (1), our system detected changes in breathing with a delay of about 1 s relative to keying when breathings were correctly detected. The measured delay is bigger than we expected.

We think these problems can be derived from the above extraordinary output data, sagging T-shirt, and the above algorithm. The detecting algorithm would detect fluctuation of sagging T-shirt as movement from breathing. A sagging T-shirt would sometimes hide upper body motion from the detecting algorithm. We are re-programming the proposed system to KINECT v2 [7] and we are adding an algorithm to remove extraordinary data. We plan to switch from a T-shirt to clothing more fitted to the body; and optimize parameters of the detecting algorithm.

TABLE II. NUMBER OF BREATHS

	Participant	Keying		Proposed system	
		Inhalation	Expiration	Inhalation	Expiration
K-1	A	12	12	16	16
	B	10	10	39	39
	C	14	14	15	15
K-2	C	11	11	11	11
	D	10	10	7	7
	E	15	15	15	15

 ISBN: 978-1-61208-470-1

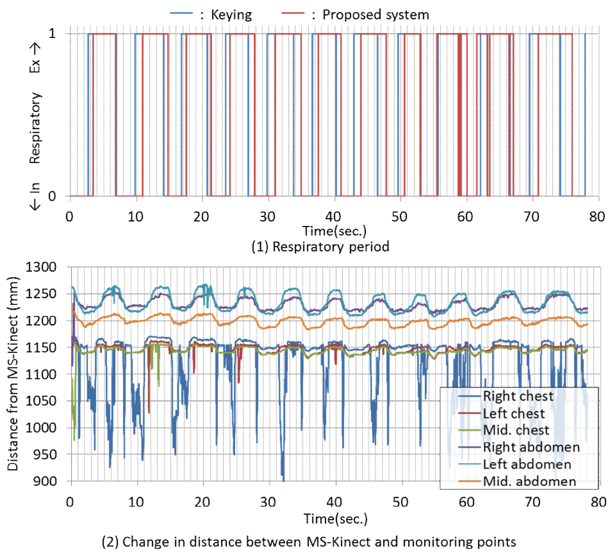


Figure 9. (1) Respiratory period, and (2) change in distance between KINECT and measuring points

D. *Automatically adjusting according to body size*

The correct points for placing a stethoscope depend on the size of the body; they differ a little between men and women. Also, their X and Y values vary according to the distance between a student playing a patient and a KINECT. Therefore, we estimate correct positions for placing a stethoscope with respect to the positions of both shoulder joints and both hip joints.

Two sets of correct position data and the positions of both shoulders and both hips are measured and stored as standard man and woman data. Since the origin of the output of KINECT's depth camera is at the upper left, it is difficult to compare body sizes among people. Therefore, for each person we reset the origin to the junction of a horizontal line connecting the average right and left hip heights and a vertical line passing through the midpoint between the right and left hip heights. Here, we assume that a person is sitting upright, and that the human body is a little unsymmetrical. The result of resetting the origin for each person is illustrated in Figure 10. The estimated left-side (X_{LE}, Y_{LE}) and right-side (X_{RE}, Y_{RE}) locations are calculated with the following equations by using the above stored standard locations and the measured positions of both shoulders and both hips:

$$X_{LE} = X_{LS} * X_{mls}/X_{sls} \quad , \qquad (1)$$
$$Y_{LE} = Y_{LS} * Y_{mls}/Y_{sls} \quad , \qquad (2)$$
$$X_{RE} = - X_{RS} * X_{mrs}/X_{srs} \quad , \qquad (3)$$
$$Y_{RE} = Y_{RS} * Y_{mrs}/Y_{srs} \quad , \qquad (4)$$

where

(X_{LS}, Y_{LS}) is the left-side standard location,
(X_{sls}, Y_{sls}) is the standard left shoulder position,
(X_{srs}, Y_{srs}) is the standard right shoulder positon,
(X_{pls}, Y_{pls}) is the measured left shoulder position of the patient, and
(X_{prs}, Y_{prs}) is the measured right shoulder positon of the patient.

We measured the ten points marked on a T-shirt and skeleton data for three men and three women to validate the proposed automatically adjusting algorithm. Each man wore the same T-shirt like that in Figure 8. Since the T-shirt was relatively small, it should have closely fit each person, and we think the marked points should have adjusted to each person correctly. After selecting one man and one woman each as the standard, we estimated correct points a stethoscope placing by using the above equations and the data for the standard person.

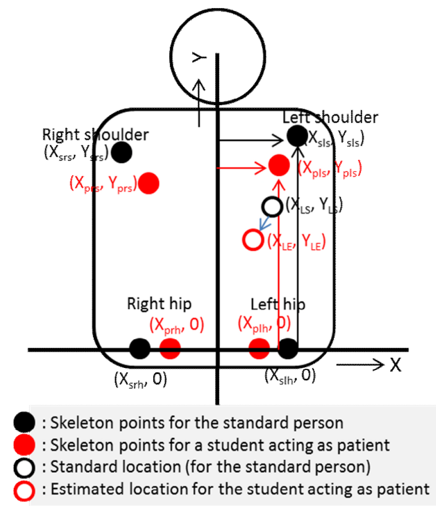


Figure 10. Illustration of adjusting locations for body size

Since there was not a big difference between the left-side data and the right-side data, we showed only the left-side data in Table III. The location relationships between measured and estimated correct points for a female and male participant are shown in Figure 11. In this figure, shoulder points and correct points a stethoscope placing for a standard person, and shoulder points and estimated and measured correct points a stethoscope placing for other participant are presented.

The data unit in Figure 11 is not the millimeter, but the pixel. In these figures, 10 pixels correspond roughly to 4 cm. Differences between estimated and measured positions depend on participants and measuring points. Maximum difference is about 8 cm. In case of Figure 11 (1), shoulder and hip positions of Female-1 are the same as those of a standard participant. Therefore, estimated correct positions of Female-1 are the same as those of a standard participant. However, measured points are different from them. The reason for this difference must be that each person did not wear the T-shirt symmetrically between right and left. In case of Figure 11 (2), differences between estimated and measured Y values were bigger for lower points, because the T-shirt was less elastic in the vertical direction.

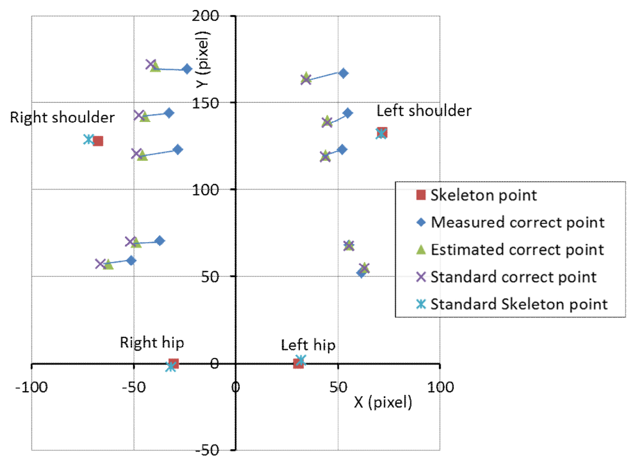
We think the above experiment is not appropriate to evaluate the proposed estimation scheme for adjusting locations to different body sizes after experimenting. If participants wear a T-shirt in bilaterally symmetric and pull it down the hip position, the estimated positions must be

 ISBN: 978-1-61208-470-1

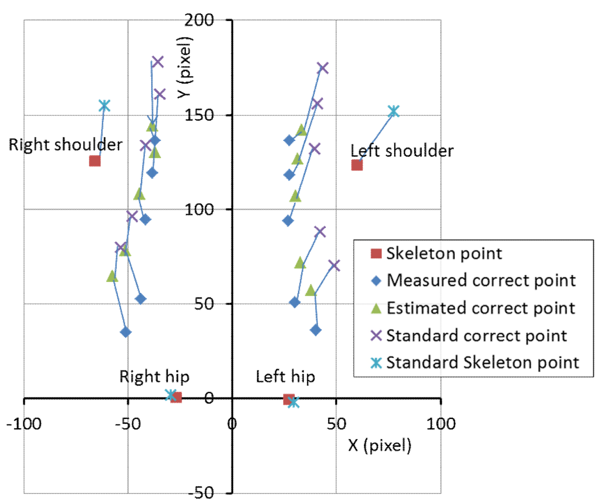
approximately equal to measured positons. However, it is difficult for participants to wear the T-shirt correctly. Therefore, we have to validate the proposed method automatically adjusting algorithm by some other experiment.

TABLE III. COMPARISON BETWEEN MEASURED AND ESTIMATED LOCATIONS

Location		Male-1		Male-2		Female-1		Female-2	
		X	Y	X	Y	X	Y	X	Y
2	measured	28	136	46	157	53	167	52	157
	estimated	33	142	42	154	35	164	36	155
4	measured	28	118	43	140	55	144	53	133
	estimated	32	127	40	137	45	140	46	131
6	measured	27	94	44	118	52	123	50	107
	estimated	31	107	39	116	44	120	45	113
8	measured	30	51	43	78	55	68	50	59
	estimated	33	72	41	77	56	68	57	64
10	measured	40	36	49	59	62	52	65	54
	estimated	38	57	48	62	63	55	65	52



(1) Female-1



(2) Male-1

Figure 11. Example of measured and estimated correct points

V. CONCLUSION

We have proposed a new auscultation practice system for medical and nursing students. In this system, students themselves play the role of a patient instead of a humanoid model, and the locations for stethoscope placement on the body are measured with KINECT. Therefore, this practice system would have low cost.

In addition, the system can judge whether stethoscope locations are correct. Practicing students hear disease sounds, synchronized with the movement of breathing, through earphones or a speaker.

We developed a prototype system and evaluated experimentally. The results showed that our system could perfectly detect stethoscope placement on a body, except for two lower points, and it could detect respiratory changes. However, it sometimes made a mistake to count the number of them and the detection delay for respiratory changes was slightly larger than expected. We have to solve these problems for detecting respiration. After solving them, we plan to develop a real learning system, consisting of a learning unit for teaching correct locations and disease sounds and an evaluation unit for testing.

W used a white-T shirt and white coat for practicing students in experiments to remove "yellow green" and likeness. However, our system works well for usual clothes which have several colors except "yellow green".

REFERENCES

[1] Patient simulators for nurse & nursing care training, Kyoto Kagaku Co., Ltd., https://www.kyotokagaku.com/products/list02.html#cate_head01 [retrieved: March, 2016]

[2] Lung Sound Auscultation Trainer "LSAT", Kyoto Kagaku Co., Ltd., https://www.kyotokagaku.com/products/detail02/m81-s.html [retrieved: March, 2016]

[3] Sakamoto Model Corporation, Sakamoto auscultation simulator, http://www.sakamoto-model.com/product/emergency/m164/ [retrieved: March, 2016]

[4] Laerdal, SimMan 3, http://www.laerdal.com/us/SimMan3G [retrieved: March, 2016].

[5] Ulrich von Zadow, L andet al. , "SimMed: Combining Simulation and Interactive Tabletops for Medical Education," CHI '13, Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 1409-1478, April 2013.

[6] John Butter, William C. McGaghie, Elaine R. Cohen, Marsha E. Kaye, and Daiane B. Wayne, "Simulation-Based Mastery Learning Improves Cardiac Auscultation Skills in Medical Students," Springer, Journal of General Internal Medicine, Volume 25, Issue 8, pp. 780-785, August 2010.

[7] Meet Kinect for Windows, https://developer.microsoft.com/en-us/windows/kinect, [retrieved: March, 2016]

[8] Open CV, http://opencv.org/ [retrieved: March, 2016]

[9] Nathan Burba, Mark Bolas, David M. Krum, and Evan A. Suma, "Unobtrusive Measurement of Subtle Nonverbal Behaviors with the Microsoft Kinect," IEEE, Virtual Reality Short Papers and Posters (VRW), pp. 1-4, 2012.