An Effective Scheme of a Depth Sensor Set Up for a Real-Time Ergonomics Assessment by the Gesture Confidence Level

Jauharah Hasna Dzakiyyah Adani[1, a], Teddy Sjafrizal[1, b], Rino Andias Anugraha[1, c], Muhammad Iqbal[1, d] and Ilma Mufidah[1, e]

[1]Industrial Engineering Department, School of Industrial Engineering, Telkom University, Bandung 40257, Indonesia.

[a]joo.gtxlab@gmail.com, [b]teddysjafrizal@telkomuniversity.ac.id , [c]rinoandias@telkomuniversity.ac.id , [d]muhiqbal@telkomuniversity.ac.id , [e]Ilmamufidah@telkomuniversity.ac.id

Keywords: Ergonomic Assessment, Depth Sensor, Sensor Placement, Visual Gesture Builder, Confidence Level

Abstract. Ergonomics assessment of human body movement demands a comprehensive and systematic data collection. The easy-to-use self-report (e.g. questionnaires, checklist, interview) and the observational technique (e.g. pose rating) are the commonly practiced techniques. However, these methods suffer from a high bias across different respondents and observers. Recently, the direct measurement technique by utilizing a depth sensor equipped with a modelling software is the alternative tool to facilitate a real-time digital human modelling. It gathers the 3-D human motion data with real-time ergonomics analysis and intervention features. This study aims to obtain the effective sensor setup by examining its parameters (object-to-sensor distance, horizontal field of view (FOV), and light intensity) to reach the acceptable gesture confidence level using a Kinect SDK V2.0. The standing position with a hand overhead was selected as the investigated gestures. The result showed that distance and horizontal FOV were statistically significant parameters. Thus, it proposes to place the sensor within 2 or 3 m away from the investigated object and to limit the horizontal FOV to 0 or 10°. Eventually, this proposal could be set as the reference in setting up a direct measurement studio for acquiring the human body movement data.

Introduction

Developing ergonomic equipment is intended to optimize the interaction between human, as an operator, and the equipment. It ultimately leads to an improvement in works performance. If products, equipment, work stations, and work methods are designed accordingly to human abilities and limitations, the performance and the results given will be better [1]. Conversely, if ergonomics assessment is ignored in designing equipment, work stations, and work methods, it results oppositely. Eventually, a designed work condition impact the operator in several circumstances, among them, (1) decreasing production output, (2) increasing costs and material for health, (3) increasing operator absence rates, (4) decreasing the quality of work, (5) increasing the probability of work accidents to the operator [2]. An instance of potential health injury to the operator is a condition cited as Musculoskeletal Disorders (MSDs). They are injuries that are often experienced by a human in carrying out activities, typically injuries to muscles, nerves, tendons, bones, joints of bones, cartilage caused by work activities [3]. When someone works in a standing or sitting position, the movement of the spine, especially the waist which is vulnerable to extreme movements, can cause injury [4]. Those impacts occur not only in the static posture but also in dynamic posture, as people need to move around at work or while using some particular products. Hence, ergonomics design indicted to support body movement, not only focused on body positions [5].

A comprehensive and systematic data collection of human body movement is necessary for an excellent ergonomics assessment. The data validity is the primary requirement. Failure to obtain valid data affects the whole proceeding engineering process [6]. The common data collection methods are compiled in Table 1.

Table 1. Data Collection Techniques for Ergonomics Evaluation [7]

Techniques	Input Activity(s)	Advantage(s)	Disadvantage(s)
Self-Report	rating scales questionnaires checklists interviews	• easy-to-use	• not reliable enough • misleading interpretations • difficult to objectively measure • need a certain time
Observational Technique	Pose rating	• easy-to-use • evaluate a wide range of work tasks	• do not require complex setups • inaccuracy or bias across different observers • need a certain time
Direct Measurement	collect data directly from sensors attached to the worker's body	• provide high accuracy for epidemiologic studies • markerless • need a shorter time	• difficult to implement in real work situations • limited to planar movements • problematic for complex joints

Table 1 describes the available techniques in gathering the human body movement for ergonomics assessments. Despite the advantages and disadvantages of those methods, it is worth to note that the self-report and observational technique suffers from a high bias across different respondents and observers. In that sense, a direct measurement technique offers high accuracy data by utilizing the sensors attached to the investigated human body. The use of such technologies has contributed significantly to the reduction of ineffective output in terms of cost, time, and quality [8]. Thus, direct measurement or real-time evolution is a reasonable method to be further utilized for collecting the body movement data.

Among the developed tools for direct measurement, the introduction of Microsoft Kinect V1.0 as a low-cost sensor depth cameras provides a reliable device to collect data at high frequencies and suggests semiautomatic functions [9]. The accuracy of the skeleton tracking is promising than the marker-based optical motion capture system in capturing human posture during working tasks for ergonomic assessment [10]. Furthermore, the digital real-time data captured by Microsoft Kinect V1.0 facilitates human digital modelling in the virtual reality area. Besides the great potential of digital data collection for the real-time ergonomic assessment, some studies suggest that the accuracy of the captured data relates to its placement [11], [12]. Evaluation by [11] proposes that placing Kinect V1.0 in between 1 m to 3 m from the object and at 54° and 39.1° horizontal and vertical field produces accurate skeletal data for ergonomic assessment. In the view of the sensor placement away from the object, the investigation by [13] recommends installing the sensor between 1 m to 3 m from the object. A similar study by [14] with a different set of sample selected 2.5 m as the optimal position to place the sensor from the scanned object. It was also discussed that Kinect V1.0 was sensitive to the distances of various objects that state if an object is too close or too far from the Kinect, the depth values are distorted [15]. Aside from the distance and the angle field of view factor, the skeleton data tracking may be deformed because of the accent lighting or environment lighting in general [16]. Contrasting the color between clothes and background is preferred while conducting the work under full outdoor lighting is avoidable [16].

From the investigations as mentioned earlier, the identified vital Kinect setup parameters are the object-sensor distance, the field of view (FOV), and light intensity. To the author's best understanding, the study of those parameters using Kinect V2.0 has not been conducted. Therefore, this study aims to determine the optimum specification parameter of the direct measurement sensor according to its object-sensor distance, horizontal field of view (FOV), and light intensity, using the newest Kinect V2.0. In this study, the new version of Kinect sensor, Microsoft Kinect V2.0, as machine learning technologies by approaching to the confidence level of gesture recognition which provided by Microsoft software development Kit (SDK V2.0) – Visual Gesture Builder (VGB). Finally, this research aims to determine the important parameter that influences the confidence level of gesture recognition of Kinect sensor V2.0 according to the distance, horizontal field of view (FOV), and light intensity. Thus, an effective parameter setup for Microsoft Kinect V2.0 in data gesture collecting process is proposed.

Procedures

A mock-up mini studio (4m x 4.5m) was prepared for this experiment (Fig.1). The Kinect for Windows V2.0 sensor was utilized, and it was placed on a table with 73 cm height. To minimize the noise from the environment, a white background was placed behind the observed object. The clip motion data from Kinect V2.0 was captured using Kinect Studio. To obtain the gesture data recognition, the device was supported with official Microsoft software development kit (SDK) [17].

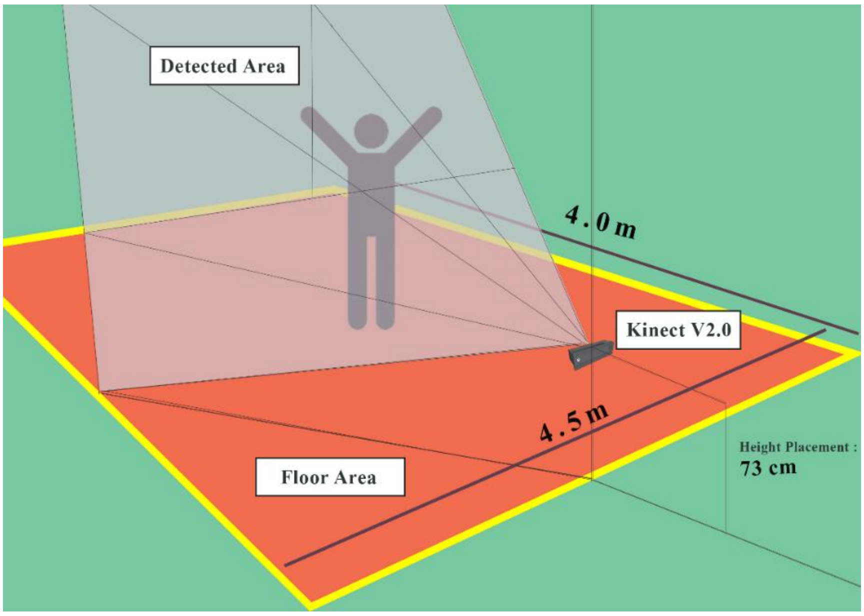

Figure. 1. The experimental setups schematic, showing the studio specifications and the placement of the Kinect V2.0 sensor by reference to the object.

As many as 60 experimental combinations were analyzed to obtain the optimum set up of the new Kinect V.20 for human body movement data collection. Three main parameters were selected: they are distance, horizontal field of view, and light intensity. The level of each factor was derived from the setup specifications of Kinect V.20 (horizontal viewing angle= 71 °, vertical viewing angle= 60°; and distance = 1 m - 4.5 m) [18]. The experiment runs were determined from the full factorial design method, which produced 60 combinations (Table 2).

Table 2. Factor And Level of The Current Study

Factor	Descriptions	Level				
		1	2	3	4	5
A	Distance [m]	1	2	3	4	
B	Horizontal Field of View [°]	0	10	20	30	40
C	Light Intensity [Lux]	132	156	185	---	---

On each experiment run, the object was requested to perform a hand-over-head gesture. After scanning the object by Kinect V.20 and recorded using Kinect Studio (Fig.2A), the captured gestures were tagged by the Visual Gesture Body software (VGB). The first 20 frames were selected and tagged. The adaptive boosting algorithm of the VGB was executed to produce the gesture's confidence level (range = 0 - 1) on each frame. Thus, the mean confidence level (E.q.1) was selected to represent the confidence level (CL) of the gesture on every experiment run

$$\sum_{k=1}^{n} C_k = \frac{C_1 + C_2 + C_3 + \dots + C_n}{\sum n} \quad (1)$$

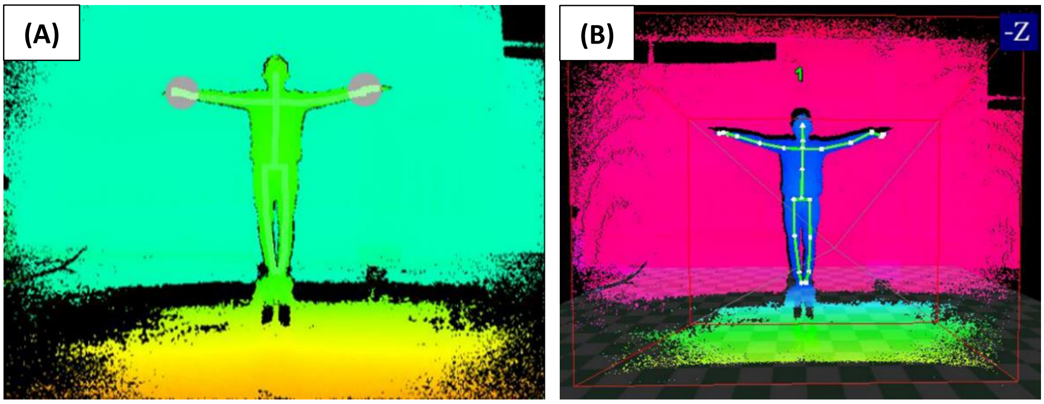

Figure. 2. (A) Clip Body Motion Using Kinect Studio, (B) Tagged Gesture of Body Motion Clip

Finally, there were 60 confidence levels were gathered for further analyses by the statistical tools. The significant factors were determined by the ANOVA technique (CI= 95 %, α = 0.05) whereas selection on the optimum level was defined by the least significant differences (LSD) method of Post-Hoc Test.

Results and Discussions

The average confidence levels of the defined gesture are listed in Table 3. The response on each experiment run resulted in different CL value. Maximum CL was 0.929, obtained at run 34, whereas the minimum Cl was 0.851 (run 58). It depicts that the capability of the sensor in capturing the gesture was sufficiently good (Cl ≥ 0.851).

Table 3. Average Confidence Level (CV) Of Every Experiment Runs

Runs	Avg. CL	Runs	Avg. CL	Runs	Avg. CL	Runs	Avg. CL
1	0.908	16	0.920	31	0.906	46	0.911
2	0.897	17	0.913	32	0.906	47	0.896
3	0.897	18	0.928	33	0.921	48	0.917
4	0.906	19	0.911	34	0.929	49	0.903
5	0.898	20	0.916	35	0.913	50	0.904
6	0.903	21	0.923	36	0.920	51	0.921
7	0.883	22	0.911	37	0.902	52	0.911
8	0.897	23	0.904	38	0.925	53	0.898
9	0.888	24	0.915	39	0.909	54	0.916
10	0.888	25	0.898	40	0.892	55	0.880
11	0.890	26	0.886	41	0.902	56	0.885
12	0.894	27	0.898	42	0.927	57	0.892
13	0.898	28	0.897	43	0.880	58	0.851
14	0.861	29	0.886	44	0.900	59	0.881
15	0.886	30	0.890	45	0.875	60	0.870

The response data in Table 3 were then analyzed using ANOVA. Table 4 describes the ANOVA analysis result. The model's factors that were observed and analyzed had R^2 = 84.06 %, thus the model was qualified as valid (R^2 = min 80%). It is found (F (3, 56) = 7.6, P = 0.001) of distance factor and (F (4, 55)) = 18.71, P = 0.000) of Horizontal FOV factor were statistically significant with P-value were 0.001 and 0.000 respectively. In other words, the light intensity factor condition and the two-way factor interaction terms of the studio were not significantly vital to control the CL of the image captured. Being less sensitive to lighting is probably the improvement of KINECT 2.0 than its predecessor.

Table 4. Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	*P*-Value
Model	35	0.013151	0.000376	3.62	0.001
Linear	9	0.010627	0.001181	11.37	0.000
Distance	3	0.002368	0.000789	7.60	0.001
Horizontal FOV	4	0.007773	0.001943	18.71	0.000
Light Intensity	2	0.000486	0.000243	2.34	0.118
2-Way Interactions	26	0.002524	0.000097	0.93	0.569
Distance*Horizontal FOV	12	0.001467	0.000122	1.18	0.352
Distance*Light Intensity	6	0.000559	0.000093	0.90	0.514
Horizontal FOV*Light Intensity	8	0.000498	0.000062	0.60	0.769
Error	24	0.002493	0.000104		
Total	59	0.015644			

Note:
(1) P value < 0.05 is significantly otherwise for P value > 0.05 is not significant
(2) $F(df,df_{within})$ = F Ratio, P- value*
** df = degree of freedom*

The least significant differences (LSD) method of Post-Hoc was conducted on significant factors. The test was conducted to determine the comparison between levels on factors that had a significant difference. There were 3 levels that had significant differences, namely level 1-2 (*P*-value = 0.02), level 1-3 (*P*-value 0.015), and level 3-4 at (*P*-value =0.048. The three comparison levels were said to have a significant difference because the *P*-value produced had a value of less than 0.05. While the comparison between level 1-4, 2-3, and 2-4 did not show a significant difference because the *P*-value obtained was greater than 0.05.

Table 5. LSD Method Results of Distance Factor

(I) Distance	(J) Distance	Mean Difference (I-J)	Std. Error	Sig.
1	2	-0.013467*	0.005645	0.02
	3	-0.014200*	0.005645	0.02
	4	-0.002800	0.005645	0.62
2	3	-0.000733	0.005645	0.90
	4	0.010667	0.005645	0.06
3	4	.0114000*	0.005645	0.05

The result of the Post Hoc LSD Test of the horizontal FOV (Table 6) shows that there were 7 levels that had significant differences, namely level 0-30, level 0-40, level 10-30, level 10-40, level 20-30, level 20-40 and the last one 30-40. The seven comparison levels were said to have a significant difference because the *P-v*alue produced had a value of less than 0.05. While the rest of the three levels, which were 0-10, 0-20, and 10-20 did not have a significant difference because the *P*-value obtained was greater than 0.05.

Table 6. LSD Method Results of Horizontal FOV

(I) Horizontal FOV	(J) Horizontal FOV	Mean Difference (I-J)	Std. Error	Sig.
0	10	-0.00225	0.004889	0.647
	20	0.005083	0.004889	0.303
	30	.015667*	0.004889	0.002
	40	.028750*	0.004889	0.000
10	20	0.007333	0.004889	0.139
	30	.017917*	0.004889	0.001
	40	.031000*	0.004889	0.000
20	30	.010583*	0.004889	0.035
	40	.023667*	0.004889	0.000
30	40	.013083*	0.004889	0.010

Conclusions

An experimental study on optimizing the parameter setup of KINECT 2.0 as a direct measurement tool for ergonomic assessment has been conducted. The significant parameters that influence the confidence level of gesture recognition were distance and horizontal field of view. The effective placement specification for Sensor Kinect V2.0 in tracking the human body movement data process based on the confidence level was 2 m or 3 m away from the investigated object and within 0° or 10° of the horizontal field of view. The proposed parameter setup is expected to be beneficial for implementing a real-time ergonomics risk assessment within the optimum accuracy data captured points.

Acknowledgements

This work was supported by a research grant (Penelitian Dana Internal 2019), funded by Telkom University. We would like to give our gratitude to the Research and Community Service Directorate of Telkom University for providing this research program.

References

[1] L. Susanti, H. Zadry, and B. Yuliandra: *Pengantar Ergonomi Industri*, Andalas University Press, Padang (2015).

[2] Alexander, C. David, Pulat, and M. Babu, *Industrial Ergonomics : A. Practitioner's Guide*. (Industrial Engineering and Management Press, 1985).

[3] A. N. Bintang and S. K. Dewi: J. Tek. Ind. Vol. 18 (2017), p. 43

[4] M. A. Wahyudi, W. A. P. Dania, and R. L. R. Silalahi: Agric. Agric. Sci. Procedia Vol. 3 (2015), pp. 195–199

[5] K. E. K. Elbert, H. B.Kroemer, and A. D. K. Hoffman, in: *Ergonomics : How to Design for Ease and Efficiency*, 3rd ed., Elsevier (2018).

[6] S. Ramesh, "Data Collection," in *Pipeline Integrity Handbook (Management and Risk Evaluation)*, Elsevier (2017).

[7] P. Plantard, E. Auvinet, A. S. Le Pierres, and F. Multon: Sensors Vol. 15 (2015), pp. 1785–1803

[8] R. Roy, I. . C. Kerr, and P. J. Sackett: in *Advances in Designs*, edited by H. Elmaraghy and

W. Emaraghy, UK (2006), p. 269

[9] V. M. Manghisi, A. E. Uva, M. Fiorentino, V. Bevilacqua, G. F. Trotta, and G. Monno: Appl. Ergon., (2017), pp. 1–11

[10] P. Plantard, H. P. H. Shum, A. Le Pierres, and F. Multon: Appl. Ergon. Vol. 65 (2017), pp. 562–569

[11] T. Dutta: Appl. Ergon. Vol. 43 (2012), no. 4, pp. 645–649

[12] T. D. Nguyen and M. Kleinsorge, in: *ErgoAssist : An assistance system to maintain ergonomic guidelines at workplaces* (2014)

[13] K. Khoshelham and S. O. Elberink: Sensors Vol.12 (2012), pp. 1437–1454

[14] B. Bonnechère *et al.*.: Ergonomics. (2014), pp. 37-41

[15] T. Banerjee, J. M. Keller, M. Popescu, and M. Skubic, in: *Computer Vision and Image Understanding* (2015).

[16] P. Horejsi *et al.*: Sci. J., (2013), pp. 389–392

[17] R. A. Clark *et al.*: Gait Posture Vol. 36 (2012), no. 3, pp. 372–377

[18] M. Bloesch, D. Rodriguez, R. Kaestner, M. Hutter, and R. Siegwart, in: *Kinect V2 for Mobile Robot Navigation : Evaluation and Modeling*, (2015).