

Pendeteksian Gerak Menggunakan Sensor Kinect for Windows

Hartono¹, Liliana², Rolly Intan³

Program Studi Teknik Informatika Fakultas Teknologi Industri Universitas Kristen Petra

Jl. Siwalankerto 121 – 131 Surabaya 60236

Telp. (031) – 2983455, Fax. (031) – 8417658

E-mail: m26411121@john.petra.ac.id¹, lilian@petra.ac.id², rintan@petra.ac.id³

ABSTRAK

Seiring dengan perkembangan teknologi yang sangat pesat, penggunaan alat input komputer juga ikut berkembang. Bermula dari menggunakan *keyboard* yang memiliki pesan tertentu dari tiap tombolnya, berkembang ke layar sentuh yang membuat manusia dapat berinteraksi langsung dengan layarnya; sekarang sedang dikembangkan teknologi dimana manusia dapat memberikan input ke komputer dengan menggunakan *gesture*. Pendeteksian *gesture* dapat menggunakan beberapa alat, salah satunya adalah Kinect. Kinect dilengkapi dengan *depth sensor* yang digunakan untuk mengenali dan melacak sendi-sendi tubuh manusia.

Sistem yang dibuat meliputi proses *skeletal tracking* pada *user*, pendeteksian posisi tangan, dan pendeteksian gerak. Proses *skeletal tracking* sudah disediakan oleh Kinect SDK. *Skeletal tracking* tersebut berguna untuk mendeteksi posisi tangan *user*. Pendeteksian tangan *user* digunakan untuk mendeteksi *gesture* yang telah dilakukan oleh *user*. Pembuatan aplikasi ini menggunakan Microsoft Visual Studio C# WPF 2010 serta Kinect Software Development Kit.

Oleh karena itu, pada skripsi ini dilakukan pembuatan aplikasi untuk melakukan pendeteksian gerak serta pendeteksian posisi tangan dalam area-area tertentu. Aplikasi ini menggunakan sensor Kinect sebagai media untuk memberi inputan. Proses dimulai dari pengecekan pergeseran posisi yang nantinya akan diidentifikasi sebagai *swipe gesture*. Selain itu, pendeteksian posisi tangan dalam area-area tertentu juga digunakan untuk mengetahui seberapa banyak area-area tersebut bisa disentuh, ketika pengguna berada ditengah sensor Kinect.

Kata Kunci

Kinect, *Virtual Reality*, *Skeletal Tracking*, Pendeteksian posisi, Pendeteksian gerak

ABSTRACT

Along with the rapid technology development, the usage of computer input tools is also on the rise. Starting from a keyboard which had specific message for each button, developed into touch screen technology which enables human to interact directly with the screen; now technology in which a mere gesture is enough to input a command to the computer is being developed. Gesture detection can be done by the utilization of several tools. For instance, Kinect is one of the tools which are equipped with depth sensor that is capable to recognize and tracking human joints.

The system being made is consisted of skeletal tracking process toward user and motion detection. Skeletal tracking process had been provided by kinect sdk. That skeletal tracking is utilized to

detect user's hand position. User's hand detection is used to detect the gesture that is done by the user. The making of this application is done by Microsoft Visual Studio C# WPF 2010 and Kinect Software Development Kit.

Therefore, this thesis is aimed to make an application which is capable in gesture and hand position detection on certain areas. This application utilized Kinect as the media for command input. The process starts from the checking of position movement, which will be identified as swipe gesture later on. In addition, hand position detection in specific areas is also being utilized, in order to determine how wide those areas is able to be touched by hand, when the user's position is exactly in the centre of Kinect's sensor

Keywords

Kinect, Virtual Reality, Skeletal Tracking, Position Detection, Motion Detection.

1. PENDAHULUAN

Virtual Reality istilah *Virtual Reality* dikemukakan ketika Morton Heilig menulis tentang “*Experience Theatre*” pada tahun 1950. Dalam hal ini, yang dimaksud dengan *Experience Theatre* adalah suatu hal yang meliputi semua indra dengan suatu cara yang efektif. Lalu pada tahun 1968, dengan bantuan dari siswanya yang bernama Bob Sproull, Ivan Sutherland menciptakan *Virtual Reality* dan *Augmented Reality* dengan bantuan alat *head-mounted display system* (HMD). Pada tahun 1975, seorang ilmuwan bernama Myron Krueger menemukan videoplace yang memungkinkan *user* untuk pertama kalinya dapat berinteraksi dengan objek virtual. Pada tahun 1989, Jaron Lanier memperkenalkan *Virtual Reality* dan menciptakan bisnis komersial pertama kali di dunia maya. Di tahun 1990an riset mengenai *Virtual Reality* mulai banyak berkembang, riset-riset ini termotivasi dari buku *Virtual Reality* yang dibuat oleh Howard Rheingold [9].

Kinect memiliki sebuah kamera RGB, sensor kedalaman dan *multi-array microphone* yang dilengkapi *software development kit*. Kinect mampu menyajikan pengenalan 3D pada seluruh tubuh dan juga kemampuan pengenalan suara. Kamera RGB ini membantu dalam pengenalan wajah dan fitur deteksi lainnya dengan mendeteksi tiga komponen warna yaitu *Red*, *Green*, dan *Blue*. Microsoft menamakannya “*RGB Camera*” dengan mengacu pada tiga komponen warna tersebut. *Depth sensor* atau sensor kedalaman merupakan sebuah proyektor infrared dan sebuah sensor monochrome CMOS yang bekerja secara bersama-sama untuk “melihat” ruangan atau area dalam bentuk 3D dengan tanpa memperdulikan kondisi cahaya. Inputan Kinect biasanya berupa gerakan pengguna yang akhirnya

pengguna tidak perlu menyentuh secara langsung *device controller*, cukup dengan melakukan gerakan-gerakan atau *gesture*. *Gesture* (sikap yang ditunjukkan oleh gerakan anggota tubuh) merupakan salah satu unsur dari komunikasi non-verbal[7].

Seiring dengan perkembangan teknologi yang sangat pesat, penggunaan alat input komputer juga ikut berkembang. Bermula dari penggunaan *keyboard* yang memiliki pesan tertentu dari tiap tombolnya, berkembang ke layar sentuh yang membuat manusia dapat berinteraksi langsung dengan layar, dan sekarang sedang dikembangkan teknologi dimana manusia dapat memasukan input ke komputer dengan menggunakan *gesture*.

Oleh karena itu terpikirkan ide untuk pembuatan aplikasi pendeteksian gerak menggunakan sensor Kinect for Windows sebagai *device* untuk memberi inputan yang digunakan untuk pengoperasian pada beberapa program (Contoh: Microsoft Power Point, *Image Viewer*).

2. TEORI PENUNJANG

2.1 Virtual Reality

Virtual Reality adalah teknologi yang membuat *user* dapat merasakan keberadaannya dalam dunia maya. Dalam *virtual reality user* juga dapat berinteraksi dengan dunia tersebut berdasarkan objek nyata maupun animasi [2]. Dunia *Virtual Reality* ditampilkan pada sebuah monitor atau melalui sebuah *stereoscopic display*. Banyak anggapan bahwa *Virtual Reality* sama dengan *Augmented Reality*, tetapi nyatanya berbeda. *Virtual Reality* menggantikan kenyataan dengan dunia *virtual* secara keseluruhan, sedangkan *Augmented Reality* menambahkan atau melengkapi kenyataan dengan benda-benda yang tidak nyata[8].

2.2 Gesture

Gesture adalah sikap atau pola suatu gerakan yang terlihat memperlihatkan suatu pesan. *Gesture* sendiri termasuk gerakan tangan, wajah atau bagian tubuh lainnya. *Gesture* memungkinkan individu untuk berkomunikasi berbagai perasaan dan pikiran, dari penjelasan, persetujuan dan kasih sayang dengan bahasa tubuh ketika berbicara [6].

2.3 Kinect

Kinect adalah suatu teknologi yang membuat *user* dapat berinteraksi secara natural dengan komputer tanpa menggunakan *controller*. Dengan kata lain *user*, dapat melakukan suatu pengoperasian komputer hanya dengan menggunakan gerakan tangan atau gerakan tubuh lainnya[8].

Kinect memiliki fitur-fitur yang meliputi kamera RGB, *depth sensor* atau sensor kedalaman, dan *multi-array microphone*. Dari fitur-fitur tersebut, *depth sensor* merupakan fitur yang berperan dalam membedakan Kinect dengan kamera. *Depth sensor* digunakan untuk mendapatkan data sebuah area dalam bentuk 3D tanpa memperdulikan kondisi cahaya pada area tersebut.

Sensor Kinect berdasarkan pada *optical lenses* dan juga sensor Kinect memiliki beberapa kekurangan, sensor Kinect dapat bekerja dengan baik pada kondisi sebagai berikut [4]:

- *Horizontal viewing angle*: 57°
- *Vertical viewing angle*: 43°
- Jarak terbaik untuk *user* menggunakan *Kinect* adalah 1.2 meter sampai 4 meter.
- *Depth range*: 400mm (*near mode*) sampai 8000mm (*standard mode*)

- Suhu 5° sampai 35° Celcius (41° sampai 95° Fahrenheit)

2.3.1 RGB Camera

Kamera RGB berfungsi untuk pengenalan wajah dan fitur deteksi lainnya dengan mendeteksi tiga komponen warna yaitu *Red*, *Green*, dan *Blue*. Microsoft menamakannya “RGB Camera” dengan karena ditunjukan pada tiga komponen warna tersebut. Kamera RGB ini mirip dengan webcam, yang dapat menangkap video pada resolusi 640 x 480 dengan warna 32bit pada 30 *frames* per detik. Beberapa resolusi dan *frame* standard yang dapat digunakan Kinect :

- Resolusi 640 x 480Fps30 dengan format RGB
- Resolusi 1280x960Fps12 dengan format RGB
- Resolusi 640x480Fps12 dengan format RAW YUV
- Resolusi 640x480Fps15 dengan format YUV

2.3.2 Depth Sensor

Depth sensor terdiri atas kombinasi dari *infrared laser projector* dan *monochrome CMOS sensor* yang mengambil data video dalam 3D tanpa memperdulikan kondisi cahaya [5].

Infrared pada *Kinect* tidak dapat dilihat secara kasat mata serta tidak berbahaya bagi tubuh manusia. *Infrared* mengirimkan ribuan sinar yang memantul untuk mengetahui objek yang ada didepannya [6].

Sinar yang dihasilkan oleh *infrared* inilah yang ditangkap oleh kamera *monochrome CMOS sensor* dan mengukur waktu sinar yang ada setelah terpantul oleh objek yang ada di depannya. Hal ini memungkinkan Kinect untuk memetakan gambar 3D yang ada didepannya sampai kedalaman satu centimeter dan tiga milimeter untuk lebar dan tinggi [3].

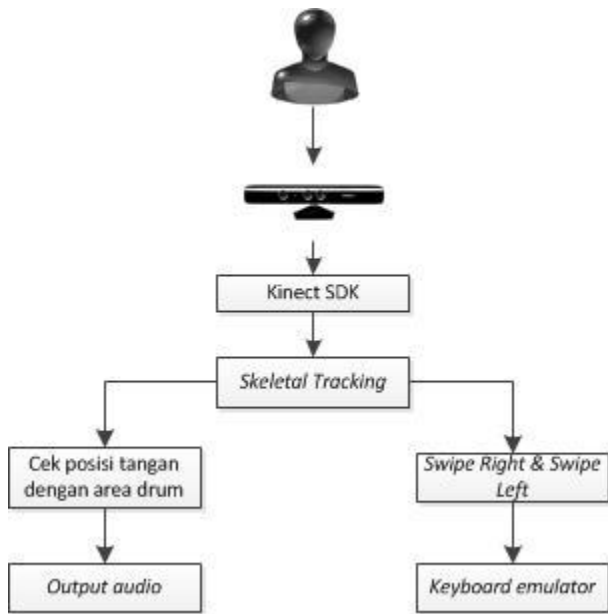
2.3.3 Skeletal Tracking

Skeletal tracking adalah fitur yang disediakan oleh Kinect SDK, fitur ini dapat melacak titik sendi utama tubuh manusia. *Skeletal tracking* ini tidak lepas dari penggunaan *depth sensor*. *Depth sensor* yang memetakan objek-objek berdasarkan jarak yang akan dibandingkan dengan data hasil *training* sebelumnya. Data *training* tersebut dibuat menggunakan 100.000 *frame* gambar manusia yang diambil dari posisi yang berbeda-beda pada saat berdiri, oleh karena itu ketika pada saat duduk Kinect tidak dapat mendeteksi kaki manusia tersebut [1].

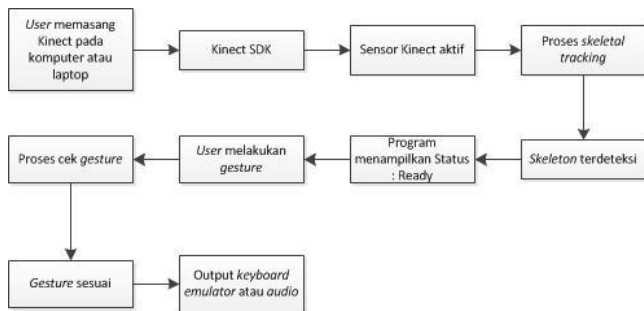
3. DESAIN SISTEM

Blok diagram sistem dan proses kerja sistem digambarkan dalam pada Gambar 1 dan Gambar 2. Pertama, *user* diminta untuk memasang Kinect pada komputer atau laptop. Setelah itu program dijalankan, dan program akan mendeteksi apakah status Kinect tersebut *connected* atau belum. Bila status Kinect sudah *connected*, program akan menjalankan sensor Kinect. Program akan menampilkan *message box* yang memberitahukan *user* bahwa Kinect belum terdeteksi jika Kinect belum *connected*; sedangkan jika Kinect sudah terdeteksi, tampilan pada program akan muncul gambar yang diambil oleh Kinect. Setelah itu, SDK Kinect akan melakukan proses *skeletal tracking* yang nantinya digunakan untuk mengambil data kordinat *skeleton*. Data kordinat *skeleton* inilah yang akan digunakan untuk mengatur proses pengambilan gerak. Saat *skeleton* sudah dideteksi, maka akan muncul tulisan *ready* pada status yang berada di bagian atas. Bila *skeletal tracking* tidak mendeteksi adanya *skeleton*, maka proses tersebut akan terus berjalan sampai menemukan *user skeleton*. Setelah menemukan *user skeleton*, barulah *user* melakukan

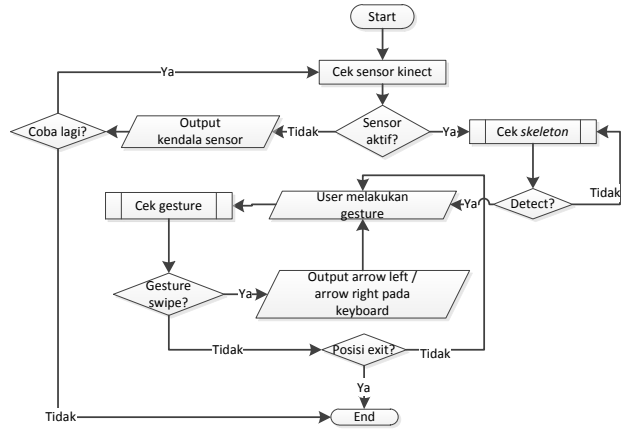
gesture. *Gesture* ini yang nantinya ditangkap dan diproses pada cek *gesture*. Proses pengecekan *gesture* sendiri meliputi pengambilan setiap gerakan *user*, di mana pada gerakan tersebut akan dicek apakah *gesture* tersebut merupakan *gesture swipe* atau bukan. Pengecekan untuk *gesture* tersebut dilakukan dengan cara memberi batasan pada durasi awal pengambilan dan durasi maksimal pengambilan gerakan *swipe*. Selain itu, terdapat pengecekan lain di mana tangan *user* dicek apakah bergeser sesuai arah dan tidak melebihi batas tinggi *swipe* tersebut. Bila proses tersebut benar, maka program mengambil data tersebut sebagai *swipe right* atau *swipe left*. Di mana nantinya data tersebut yang akan digunakan untuk melakukan proses *keyboard emulator*. Selain itu, untuk aplikasi pendeteksian posisi dilakukan pengecekan pada posisi tangan, apakah berada di area yang sudah ditentukan. Apabila benar maka, respon program akan mengeluarkan *audio*. Pada Gambar 3 dan Gambar 4 dapat dilihat cara kerja program *swipe* dan cara kerja program pendeteksian posisi.



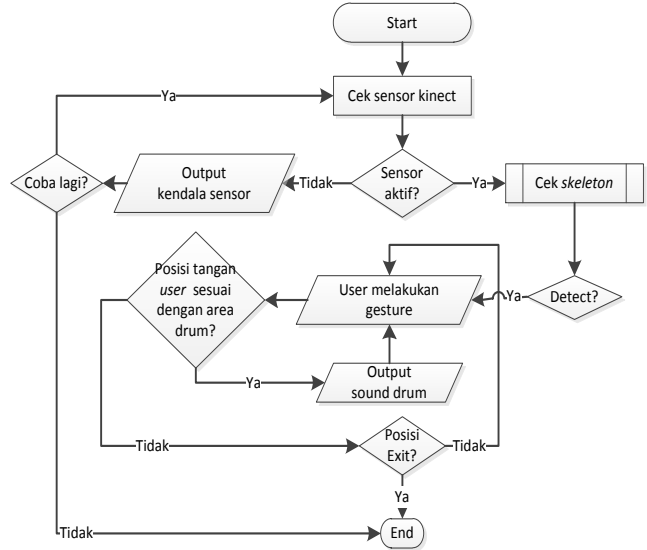
Gambar 1. Blok diagram sistem.



Gambar 2. Proses kerja sistem secara garis besar



Gambar 3. Cara kerja program *swipe*



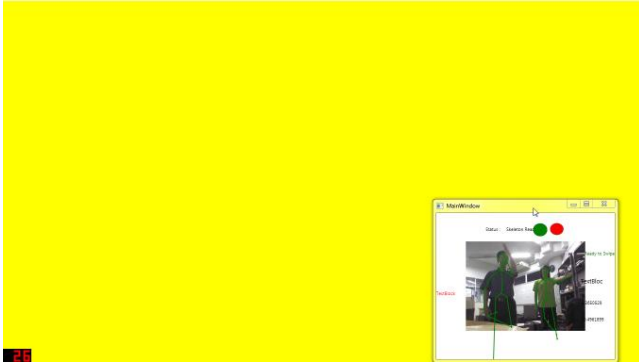
Gambar 4. Cara kerja program pendeteksian posisi

4. PENGUJIAN SISTEM

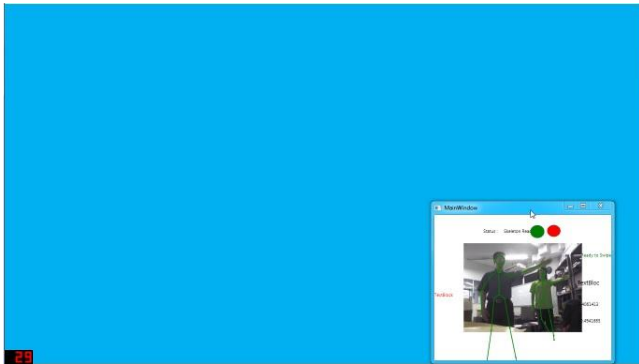
Pengujian terhadap aplikasi dibagi menjadi beberapa bagian, di antaranya adalah sebagai berikut:

4.1 Pengujian deteksi *swipe*

Pada aplikasi yang dibuat, kordinat *user* yang didapat dari hasil deteksi *skeleton* digunakan untuk melakukan pengecekan terhadap *swipe gesture*. Sebelum melakukan *swipe gesture*, *user* diharuskan untuk mengangkat telapak tangan. Setelah pengecekan tersebut selesai, barulah *user* bisa melakukan *swipe gesture* dan akan direspon oleh program. Bila ada dua *user* dengan jarak yang berbeda, program hanya akan mendeteksi *gesture* dari *user* terdekat saja. Pada Gambar 5 dan Gambar 6, dapat dilihat penerapan aplikasi pada Microsoft Power Point dan respon dari program Microsoft Power Point setelah *user* melakukan *swipe*.



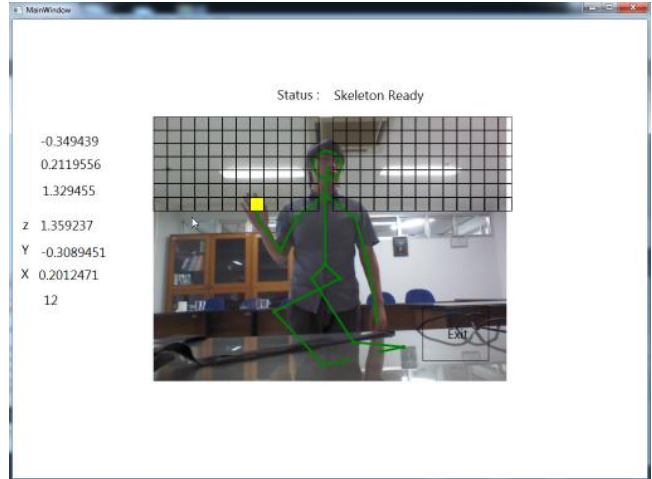
Gambar 5. *User* terdekak bersiap untuk *swipe*



Gambar 6. Respon dari program Microsoft Power Point setelah *user* melakukan *swipe*

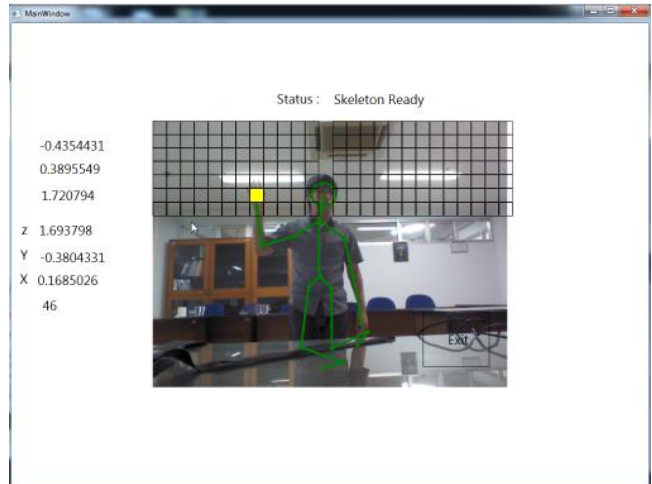
4.2 Pengujian posisi tangan untuk simulasi drum

Pada aplikasi yang dibuat, kordinat posisi tangan yang didapat dari *skeletal tracking* digunakan untuk melakukan pengecekan posisi apakah posisi tangan tersebut mengenai area yang sudah ditentukan, bila benar maka program akan merespon dengan memberi warna pada area tersebut. Pada pengujian ini aplikasi memiliki area yang berbeda dalam beberapa ukuran. Area tersebut yang akan digunakan untuk menentukan apakah pengecekan pergeseran per pixel nya masih presisi sampai jarak yang sudah ditentukan atau tidak. Pada Gambar 7 dapat dilihat program pada saat tangan mengenai area tersebut pada jarak 1.4meter



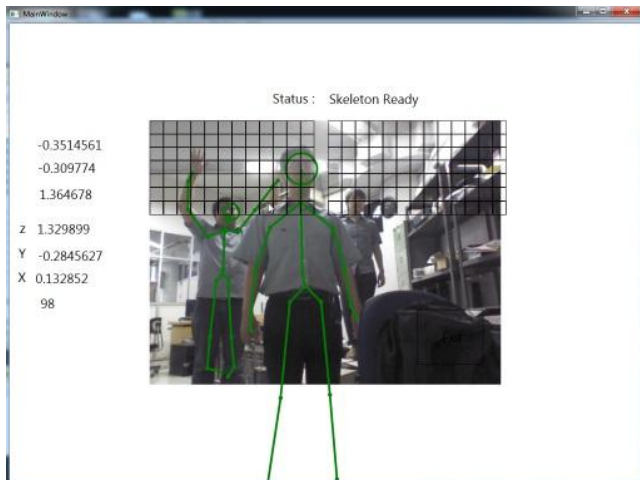
Gambar 7. Respon program pada saat tangan *user* menyentuh area pada jarak 1.1 meter sampai 1.4 meter

Selanjutnya *user* mencoba meletakkan posisi tangan di area lain pada jarak 1.7 meter dengan kondisi x yang berbeda dan tidak sama pada saat berada dalam jarak 1.4meter. Pada Gambar 8, dapat dilihat respon program saat *user* menyentuh area yang berbeda.



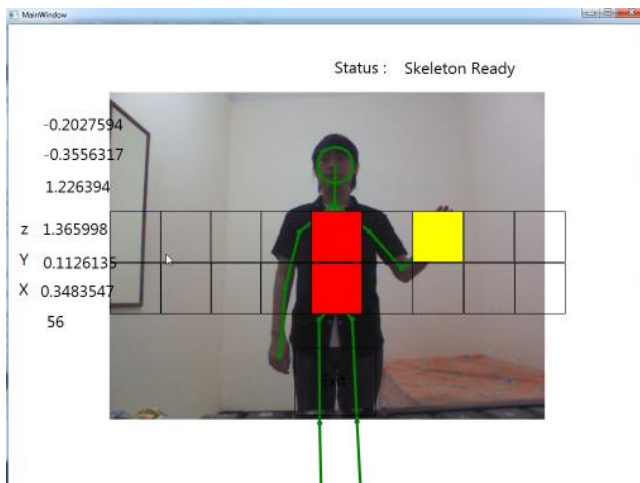
Gambar 8. Respon program pada saat tangan *user* menyentuh area pada jarak 1.7 meter

Apabila bukan *user* terdekak yang menaruh tangan pada area yang sudah disediakan, program tidak akan merespon input dari *user* tersebut. Pada Gambar 9 dapat dilihat percobaan *user* lain meletakkan posisi tangan pada area yang sudah ditentukan.



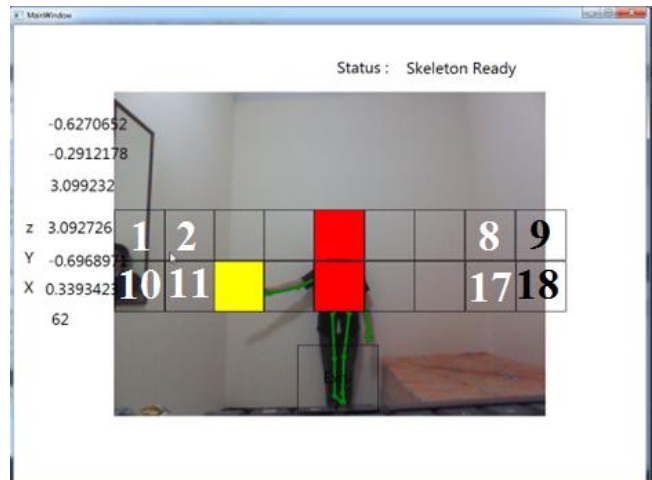
Gambar 9. *User* terdekat kedua meletakkan posisi tangan pada area yang sudah disediakan

Pengujian akurasi penempatan posisi tangan dilakukan berdasarkan jarak *user* dengan sensor Kinect. Pada jarak 1.1 meter sampai dengan 1.4 meter, posisi tangan masih terbaca dengan baik. Pada Gambar 10 dapat dilihat penempatan posisi pada kisaran 1.1 meter sampai 1.4 meter.



Gambar 10. *user* menempatkan posisi tangan pada jarak 1.2 meter sampai 1.4 meter

Selanjutnya pada jarak 2.9 meter sampai 3.2 meter, posisi tangan hanya dapat terbaca 8 kotak. Pada kotak 1,2,8,9,10,11,17 dan 18, posisi tangan tidak terbaca, karena tangan *user* tidak dapat menjangkau area tersebut. Pada jarak ini, kotak pada nomor 3,7,12 dan 14 sudah sangat minim untuk terdeteksi, karena pada jarak tersebut tangan *user* tidak dapat menjangkau area tersebut. Pada Gambar 11 dapat dilihat penempatan posisi pada kisaran 2.9 meter sampai 3.2 meter



Gambar 11. *user* menempatkan posisi tangan pada jarak 2.9 meter sampai 3.2 meter

5. KESIMPULAN

Selama perencanaan, analisa dan desain, pembuatan program sampai dengan penulisan buku ini, terdapat beberapa kesimpulan yang dapat diperoleh. Kesimpulan yang telah diperoleh tersebut antara lain:

- Dengan adanya aplikasi pendeteksian gerak ini, *user* akan dapat mengontrol pengoperasian komputer dengan menggunakan *gesture* yang sudah disediakan. Pengoperasian komputer yang dimaksud tersebut di antaranya adalah pada saat presentasi menggunakan Microsoft Power Point, ataupun pada saat melihat gambar.
- Pada pengujian simulasi ini dapat diketahui bahwa semakin jauh jarak *user* maka semakin berbeda jauh juga kordinat-kordinat pergeseran yang diambil per pixel. Jarak optimal untuk pergeseran yang masih bisa presisi untuk area 75x75 antara 1.1 meter sampai 1.7 meter.
- Aplikasi yang dibuat cukup kompleks untuk pengecekan posisi tangan pada jarak-jarak yang sudah disediakan dengan menggunakan Kinect For Windows.
- Aplikasi simulasi drum dapat dikembangkan menjadi seperti menjadikan tangan sebagai pengganti pointer *mouse* atau membuka program pada saat tangan di area tertentu.

6. REFERENSI

- [1] Aron, J. (2011, March 25). *Microsoft explains the tech behind Kinect*. Retrieved January 18, 2015 from <http://www.newscientist.com>: <http://www.newscientist.com/blogs/onepercent/2011/03/microsoft-explains-the-tech-be.html>
- [2] Brooks, Jr., F.P. (1999). *What's Real About Virtual Reality*. Retrieved October 26, 2014 from <http://www.cs.unc.edu/~brooks/WhatsReal.pdf>
- [3] Carmody, T. (2010, November 3). *How Motion Detection Works in Xbox Kinect*. Retrieved January 18, 2015 from <http://www.wired.com/gadgetlab/2010/11/tonights-release-xbox-kinect-how-does-itwork>

- [4] Catuhe, David. 2012. Programming with the Kinect for Windows Software Development Kit. United States of America
- [5] Klug, B. (2010, December 9). Microsoft Kinect review. Retrieved January 18, 2015 from <http://www.anandtech.com/show/4057/microsoft-kinect-the-anandtechreview/2>
- [6] Mathe, Z. (2011, May 27). Inside Kinect: Skeletal Tracking Deep Dive. Retrieved January 18, 2015 from <http://www.microsoft.com/download/en/confirmation.aspx?id=26098>
- [7] McNeill, D. (2005,September,19). Gesture and Thought. Retrieved January 20, 2015 from <http://mcneilllab.uchicago.edu/pdfs/Gesture&Thought.pdf>
- [8] Metcalf, J. (2009, June 1). E3 2009 : Microsoft at E3 Several Metric Tons of Press Releaseapalloza. Retrieved January 18, 2015 from <http://blog.seattlepi.com/digitaljoystick/2009/06/01/e3-2009-microsoft-at-e3-several-metric-tons-of-press-releaseapalloza/>
- [9] Sherman, W.R dan Craig, A.B (2003). Understanding Virtual Reality: Interface, Application, and Design. San Francisco