

STUDI PERBANDINGAN WARNA MARKER PADA TRACKING GERAK TANGAN BERBASIS VIDEO

Affan Mahtarami

Jurusan Teknik Informatika,
Fakultas Teknologi Industri, Universitas Islam Indonesia
Jl. Kaliurang Km. 14,5, Yogyakarta
mahtarami@hotmail.com

ABSTRACT

More natural interaction and communication methods is a demand of advancing technology. Interaction methods for application like virtual reality, wearable computer, and augmented reality is not proper if using traditional input peripherals, e.g. keyboard, mouse, and joystick.

Vision-based interaction technique is a candidate for natural interaction. Using this technique, user is not do a direct-contact with input peripherals, but use one or more cameras to capture user motion and then system interpret it. Marker tracking in video sequences is base-technique of vision-based interaction technique. One of human motion that used to augment human communication is hand motion.

In this research, marker tracking was done. Marker was positioned in six location of hand, i.e. five in finger tips and one in the center base of palm. Using three different color marker, i.e. red, green, and blue, the research showed that there is no significant different tracking result for these three different marker color.

Keywords: *Hand tracking, Interaction technique, Human-Computer Interaction*

INTISARI

Kebutuhan akan metode interaksi dan komunikasi yang lebih alami antara user dan komputer adalah salah satu tuntutan dari perkembangan teknologi. Metode interaksi pada aplikasi *Virtual Reality*, *Wearable Computer*, dan *Augmented Reality* menyarankan bahwa perangkat interaksi tradisional seperti keyboard, mouse, dan joystick, tidak lagi nyaman digunakan.

Teknik interaksi berbasis visi komputer menjadi kandidat teknik interaksi yang bersifat alami. Teknik ini tidak membutuhkan kontak langsung pengguna dengan peralatan input, melainkan komputer menangkap gerakan pengguna melalui kamera video dan menginterpretasikannya. Tracking terhadap suatu marker pada sekuen video menjadi dasar teknik interaksi berbasis visi computer. Salah satu gerakan dari tubuh manusia yang sering digunakan sebagai alat komunikasi adalah gerakan tangan.

Pada penelitian ini, *tracking* dilakukan terhadap *marker* yang terpasang pada tiap posisi ujung jari dan telapak tangan. Dengan menggunakan tiga buah warna *marker*, yaitu merah, hijau, dan biru, hasil penelitian menunjukkan tidak terdapat perbedaan signifikan pada hasil tracking terhadap ketiga warna *marker* tersebut.

Kata kunci : *tracking* gerak tangan, teknik interaksi, interaksi manusia dan komputer

PENDAHULUAN

Seiring dengan terus meningkatnya kemajuan teknologi komputer, peranan teknologi komputer dalam mendukung kepentingan manusia juga semakin meningkat. Termasuk salah satunya adalah bidang interaksi manusia dan komputer (IMK), atau sering disebut *Human Computer Interaction* (HCI). Perangkat keras interaksi antara manusia dan komputer yang digunakan sekarang ini lebih bersifat mekanis, seperti: keyboard, mouse, dan joystick. Meskipun peralatan-peralatan tersebut akurat dan handal

(*reliable*), tetapi model interaksi yang dilakukan tidak bersifat alami sebagaimana manusia berinteraksi dengan sesamanya. Penggunaan peralatan-peralatan tersebut untuk mengoperasikan komputer membutuhkan kontak langsung antara *user* dengan komputer. Kebutuhan akan komunikasi yang lebih alami antara user dan komputer adalah tuntutan dari perkembangan teknologi komputer itu sendiri. Ketika user berinteraksi dengan komputer tidak hanya di depan meja, melainkan dalam lingkungan virtual (*Virtual Reality*), menggunakan komputer yang terpasang di

badan (*Wearable Computer*), ataupun dengan perangkat *Augmented Reality*, maka perangkat-perangkat interaksi tradisional tidak lagi handal untuk digunakan. Begitu pula ketika seseorang melakukan presentasi, maka dibutuhkan pengoperasian komputer dari suatu jarak tertentu dan dilakukan secara alami, tanpa harus memegang *remote* atau *pointer* secara terus menerus selama melakukan presentasi.

Teknik interaksi berbasis visi komputer menjadi kandidat teknik interaksi yang bersifat alami. Teknik ini tidak membutuhkan kontak langsung pengguna dengan peralatan input, melainkan komputer menangkap gerakan pengguna melalui kamera video dan menginterpretasikannya. Input secara visual dapat memberikan kemampuan penginderaan pada komputer sebagaimana manusia melakukan penginderaan menggunakan mata. Interaksi manusia secara alami tidak membutuhkan mouse atau keyboard, melainkan diantaranya menggunakan mata dan telinga untuk mendapatkan informasi dari lingkungan. Komputer, pada prinsipnya, dapat menginterpretasikan gerakan manusia melalui citra yang ditangkap menggunakan kamera video. Dalam mengadopsi kemampuan visual manusia pada komputer, beberapa faktor harus diperhatikan:

- *Robustness* dan *akurasi*. Input dari kamera video seringkali memiliki *noise* yang disebabkan oleh berbagai hal, seperti: sensor, pencahayaan, komposisi latar, dan *occlusion*. Algoritma yang digunakan untuk menyusun teknik interaksi berbasis kamera video harus mampu mengambil informasi yang relevan secara *robust* dan akurat.
- *Kecepatan*. Agar dapat digunakan sebagai teknik interaksi, maka komputasi harus bersifat *real-time*, atau mendekati video-frame rate.
- *Implementasi*. Teknik interaksi harus bersifat umum, dan dapat diadaptasikan pada berbagai skenario.

Sedangkan beberapa keuntungan penggunaan teknik interaksi berbasis visi komputer adalah:

- *Nonintrusiveness*: Keuntungan penggunaan teknik interaksi berbasis visi komputer adalah sifat non-intrusif-nya, yaitu user tidak perlu melakukan kontak secara fisik dengan perangkat interaksi. Hal ini menjadikan interaksi yang dilakukan dengan komputer dapat bersifat alami dan memberikan pengalaman yang lebih imersif pada lingkungan virtual.

- *Kaya informasi*: Dibandingkan dengan data input yang didapatkan dari perangkat input tradisional seperti mouse, keyboard, dan joystick, kamera dapat menangkap informasi yang jauh lebih banyak dari sinyal warna 2 dimensi. Lebih jauh lagi, interpretasi juga dapat dilakukan terhadap sekuen video yang didapatkan dari kamera. Hal tersebut menjadikan kamera sebagai perangkat input yang serba guna. Teknik berbasis visi komputer juga memiliki beragam aplikasi lain, seperti *teleconference* dan *biometric*.
- *Murah*: Saat ini, kamera digital dipasarkan kepada konsumen dengan harga yang semakin terjangkau dengan teknologi dan kualitas yang terus meningkat. Banyak vendor-vendor dan pabrik-pabrik komputer dan laptop telah menyertakan kamera sebagai perangkat standar yang terintegrasi didalam komputer produksinya. Begitu pula pada produk-produk telepon seluler atau *handphone*.

Pengambilan data gerakan dengan menggunakan kamera memiliki beragam aplikasi. Salah satunya adalah dibidang interaksi manusia dan komputer. Di bidang ini, pengambilan data dengan menggunakan kamera dapat digunakan sebagai proses input data. Beragam gerakan manusia, baik anggota badan yang melakukan gerakan ataupun pose yang dihasilkan, dapat diinterpretasikan terhadap beragam makna. Salah satu alternatif anggota badan manusia yang dapat digerakan untuk menghasilkan beragam pose adalah tangan. Manusia secara alami terbiasa menggunakan tangan untuk melakukan komunikasi. Manusia menggerakkan tangan untuk mendapatkan beragam pose. Tiap pose memiliki maksud dan makna tersendiri sesuai dengan kesepakatan umum ataupun kesepakatan diantara personal yang melakukan komunikasi. Untuk dapat menangkap informasi yang diberikan, seseorang harus mampu melihat pose tangan dari lawan komunikasinya secara benar, sebelum kemudian menginterpretasikan maknanya. Adopsi teknik interaksi menggunakan isyarat tangan pada komputer menuntut hal serupa. Sebelum proses interpretasi, komputer pertama kali harus mampu menangkap pose tangan pengguna dengan benar. Metode *tracking* yang digunakan menjadi salah satu faktor penting untuk memastikan komputer dapat menangkap pose tangan pengguna dengan benar.

Banyak pendekatan yang dilakukan untuk melakukan *tracking* terhadap gerak tangan manusia. Karena kompleksitas gerakan tangan yang memiliki derajat kebebasan, atau

Degree of freedom (DOF), dari keseluruhan sendinya kurang lebih berjumlah 27, maka penggunaan sarung tangan khusus yang dilengkapi sensor-sensor elektromagnetis pada tiap sendi ruas jari, atau pada posisi-posisi tertentu pada tangan, menjadi pendekatan yang paling mudah dengan hasil yang relatif akurat. Namun pendekatan tersebut membutuhkan perangkat khusus dengan kerumitan pada proses instalasi alat, dan memiliki kelemahan dalam hal ketidak-leluasaan tangan saat bergerak. Metode berbasis visual dengan menggunakan satu atau beberapa kamera merupakan metode *tracking* yang lebih mudah, khususnya dalam proses instalasi perangkat, dan juga memiliki kelebihan pada keleluasaan tangan ketika bergerak, jika dibandingkan dengan proses *tracking* menggunakan sarung tangan khusus. Meskipun demikian tantangan yang dihadapi metode *tracking* berbasis visual adalah terkait dengan masalah akurasi. Hingga saat ini, interpretasi pose tangan berbasis visual menggunakan kamera dilakukan dengan melakukan estimasi berdasar data-data geometri dari fitur-fitur pada tangan yang teridentifikasi.

Salah satu metode estimasi pose tangan berbasis visual adalah dengan menggunakan ujung jari sebagai fitur yang menjadi titik acuan proses *tracking*. Metode ini melakukan generalisasi pose tangan 3D menjadi 2D berdasarkan posisi ujung jari relatif terhadap posisi suatu acuan utama pada telapak tangan. Dari informasi tersebut maka sudut rotasi dari tiap-tiap sendi ruas tangan dapat diketahui dengan mengimplementasikan rumusan umum gerak ruas-ruas tangan (Lee, 1995). Terkait dengan metode *tracking* objek, terdapat beberapa algoritma yang dapat digunakan. Salah satunya adalah metode *tracking* berbasis *optical flow* dengan menggunakan algoritma Lucas-Kanade.

Secara umum, permasalahan yang ingin dipecahkan pada penelitian ini adalah bagaimana melakukan *tracking* terhadap gerakan tangan pada sebuah sekuen video. Dengan mengimplementasikan algoritma *optical flow* Pyramidal Lucas-Kanade, dan membatasi bagian tangan yang ingin diketahui adalah 6 posisi acuan, yaitu 5 acuan pada ujung jari, dan 1 acuan pada telapak tangan, maka rumusan masalahnya secara khusus adalah bagaimana melakukan *tracking* 5 ujung jari dan 1 acuan pada telapak tangan dengan menggunakan algoritma *Optical Flow* Pyramidal Lucas-kanade pada sebuah sekuen video. Studi secara spesifik difokuskan pada perbandingan penggunaan tiga warna marker

yang berbeda, yaitu merah, hijau, dan biru. Pemilihan warna yang digunakan didasarkan pada tiga *channel* warna primer, yaitu RGB (*Red, Green, Blue*)

Beberapa penelitian telah dilakukan berkaitan dengan deteksi ujung jari dengan implementasi pada beragam aplikasi. Deteksi ujung jari dengan pendekatan teknik korelasi dengan menggunakan mask berbentuk lingkaran dilakukan oleh Koike (Koike, 2001), Oka (Oka, 2001), dan Letessier (Letessier, 2004). Sedangkan deteksi ujung jari menggunakan teknik *template matching* dilakukan oleh Crowley (Crowley, 1995), dan O'Hagan (O'Hagan, 1997). Pendekatan yang dilakukan oleh Malik (Malik, 2004), Segen (Segen, 1998), dan O'Hagan (O'Hagan, 2002) menggunakan *local-maxima* bentuk kurva dari batas-luar (*boundary*) silhuet tangan.

Penggunaan marker dalam deteksi ujung jari juga dilakukan pada beberapa penelitian guna meningkatkan kecepatan pemrosesan. Maggioni et al. (Maggioni, 1998) menggunakan marker berbentuk ellips untuk melakukan estimasi terhadap frame tangan 3D. Metode serupa juga dilakukan oleh Usabiaga (Usabiaga, 2005), namun dengan menggunakan pendekatan *multipleview*. Sedangkan Kim dan Felner (Kim, 2004) menggunakan marker berwarna putih yang ditempelkan pada ujung jari dengan latar belakang hitam.

Kajian terhadap penelitian-penelitian terkait dengan estimasi terhadap pose tangan berbasis visi komputer telah dilakukan oleh Erol (Erol, 2007). Dalam kajian tersebut diuraikan bahwa kesulitan-kesulitan dalam mendesain sistem untuk melakukan estimasi pose tangan adalah sebagai berikut:

- *Masalah dimensi yang tinggi*. Tangan manusia adalah sebuah obyek artikulasi yang memiliki lebih dari 20 DOF. Karena ada interdependensi diantara ruas-ruas jarinya ketika bergerak, maka derajat kebebasan tangan pada kenyataannya tidak sampai 20 DOF. Meskipun demikian, berdasarkan penelitian, tidak mungkin menggunakan derajat kebebasan kurang dari 6 DOF. Terkait dengan lokasi dan orientasinya, maka terdapat banyak parameter yang harus dicari.
- *Self-occlusion*. Karena merupakan obyek artikulasi maka proyeksi terhadap pose tangan menghasilkan beragam bentuk dengan banyak oklusi oleh bagian-bagian tangan itu sendiri. Hal ini membuat proses segmentasi diantara bagian-bagiannya

dan ekstraksi fitur-fitur level tingginya menjadi sulit dilakukan.

- *Kecepatan pemrosesan.* Meskipun pada sekuen tunggal, sebuah sistem visi komputer real-time harus memproses data yang besar. Sehingga, aplikasinya membutuhkan tenaga komputasi yang besar pula. Beberapa diantaranya menggunakan *hardware* khusus, dan juga pemrosesan paralel.
- *Lingkungan yang tidak terkendali.* Untuk penggunaan secara luas, banyak sistem IMK digunakan pada lingkungan dengan beragam keadaan latar belakang (*background*) dan kondisi pencahayaan. Hal ini hingga sekarang masih menjadi tantangan dibidang visi komputer.
- *Kecepatan gerakan tangan.* Tangan sanggup melakukan gerakan dengan sangat cepat. Kecepatan gerakan tangan dapat mencapai 5 meter/detik untuk gerakan tipe translasi, dan 300°/detik untuk gerakan tipe orientasi. Saat ini, kamera hanya sanggup mendukung frekuensi frame 30-60 Hz. Disamping itu, banyak algoritma *tracking* visi komputer kesulitan memproses kecepatan *tracking* 30 Hz. Pada kenyataannya, kombinasi antara kecepatan gerakan tangan yang tinggi dan rendahnya *sampling rate* menghasilkan kesulitan yang lebih bagi algoritma *tracking*.

Kesulitan-kesulitan tersebut diatas berkaitan erat dengan proses *tracking* gerak tangan karena pada dasarnya estimasi pose tangan berlandaskan pada hasil proses *tracking*. Karena adanya kesulitan-kesulitan itu pula, maka beberapa penelitian menerapkan batasan-batasan terhadap variabel *user* dan lingkungan. Sebagai contoh, beberapa penelitian menggunakan latar seragam atau statis, dan tangan dianggap sebagai objek dengan warna kulit (*skin-colored*). Namun demikian, untuk aplikasi-aplikasi nyata, batasan-batasan tersebut tidak dapat digunakan. Beberapa aplikasi tidak dapat bertoleransi terhadap kecepatan *tracking* yang lambat, meskipun dapat menyuruh *user* untuk menggerakkan tangannya dengan lambat.

PEMBAHASAN

Langkah-langkah yang dilakukan dalam penelitian ini secara umum dapat dibagi menjadi dua kegiatan utama, yaitu pengambilan data, dan pengolahan data. Pengambilan data adalah pekerjaan yang dilakukan pertama kali. Hasilnya adalah sekuen video gerak tangan.

Sekuen video tersebut kemudian menjadi input bagi proses pengolahan data. Eksperimen dilakukan terkait dengan penggunaan beberapa warna marker, yaitu merah, hijau, dan biru. Alur pekerjaan diperlihatkan pada gambar 1.

Tahap pengambilan data dimulai dengan beberapa persiapan. Persiapan-persiapan tersebut meliputi persiapan alat dan tempat. Penjelasannya adalah sebagai berikut:

- *Setting alat.* Peralatan yang disiapkan antara lain: sarung tangan, marker, *software*, dan kamera *web-cam*. Sarung tangan yang digunakan adalah sarung tangan kain berwarna putih. Warna sarung tangan adalah putih. Sedangkan marker berwarna merah, hijau, dan biru. Marker dipasang pada sarung tangan pada posisi tiap ujung jari dan tengah-pangkal tangan pada bagian telapak (Gambar 3.2). *Software* yang digunakan adalah **Video Capture Factory**. Sedangkan kamera yang digunakan adalah kamera *web-cam* merk **Aladdin** buatan **Lexcron**.
- *Setting tempat.* Setting tempat dilakukan berdasarkan kondisi pencahayaan dan latar. Kondisi pencahayaan yang digunakan adalah kondisi pencahayaan ruang (*indoor*). Sedangkan latar yang digunakan adalah latar dengan warna putih.

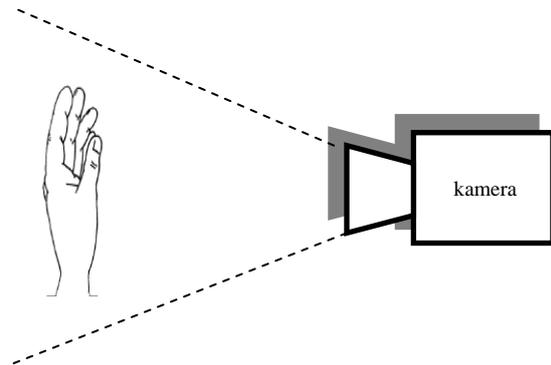
Pemilihan warna marker ditentukan berdasarkan channel-channel warna RGB (*Red, Green, Blue*). Sedangkan posisi-posisi marker ditentukan pada ujung-ujung jari dan satu di telapak tangan karena dengan informasi lokasi posisi-posisi tersebut gerakan tangan secara keseluruhan dapat diestimasi. Sedangkan latar putih digunakan agar marker dapat teridentifikasi tanpa harus menggunakan proses segmentasi.

Proses pengambilan data dilakukan beberapa kali sehingga didapatkan hasil yang paling memuaskan terkait dengan proses segmentasi dan *tracking*. Pada saat pengambilan data, tangan dengan memakai sarung tangan yang telah ditemeli marker digerakkan didepan kamera dengan telapak tangan menghadap kamera (Gambar 2). Gerakan yang dibuat meliputi gerakan bergeser, berotasi, jari-jari membuka dan menutup, serta mencengkram.

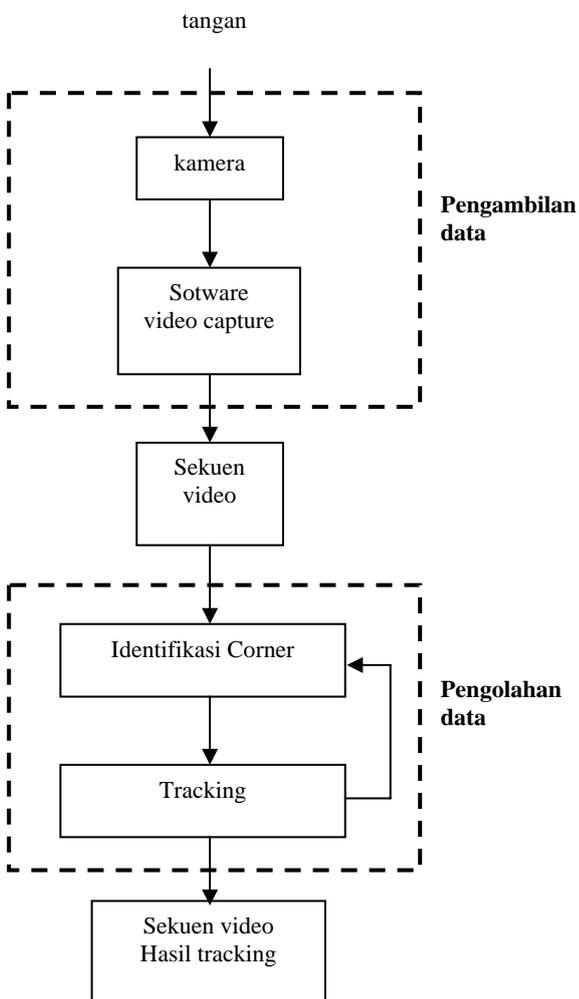
Setelah proses pengambilan data, maka langkah selanjutnya adalah pengolahan data. Dalam tahap ini dibuat sebuah program untuk melakukan *tracking* terhadap marker pada tiap ujung jari tangan dan satu pada pangkal telapak tangan. Metode *tracking* yang digunakan berbasis *optical-flow*, yaitu algoritma Pyramidal Lucas-Kanade. Proses *tracking*

diawali dengan identifikasi marker sebagai *corner* dengan menggunakan algoritma “Good Feature to Track” Shi-Tomasi. *Corner-corner* yang teridentifikasi tersebut digunakan sebagai input bagi algoritma Pyramidal Lucas-Kanade untuk dicari area yang sesuai pada frame berikutnya. Sedangkan perangkat lunak yang digunakan dalam tahap pengolahan data adalah:

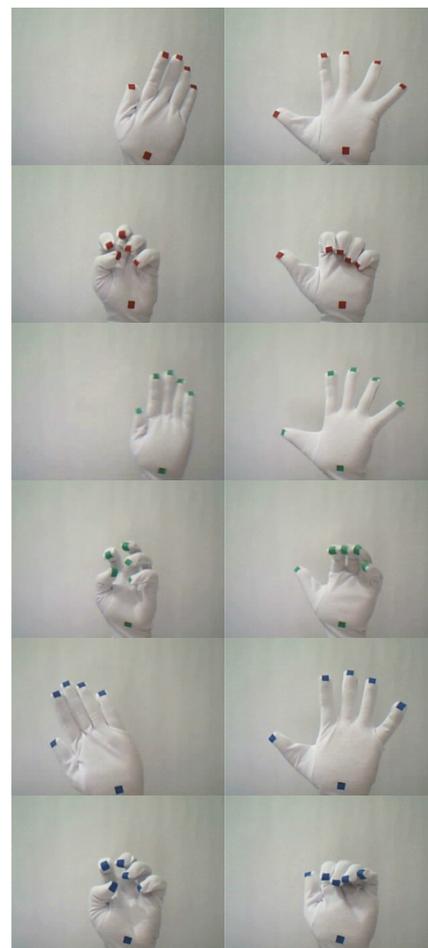
- Bahasa pemrograman **C++**
- Librari visi komputer **OpenCV**
- Lingkungan pengembangan **Dev-C++**
- Perangkat lunak pengurai video **SC Video Decompiler**



Gambar 2. Posisi kamera dan tangan pada arah *frontal view*



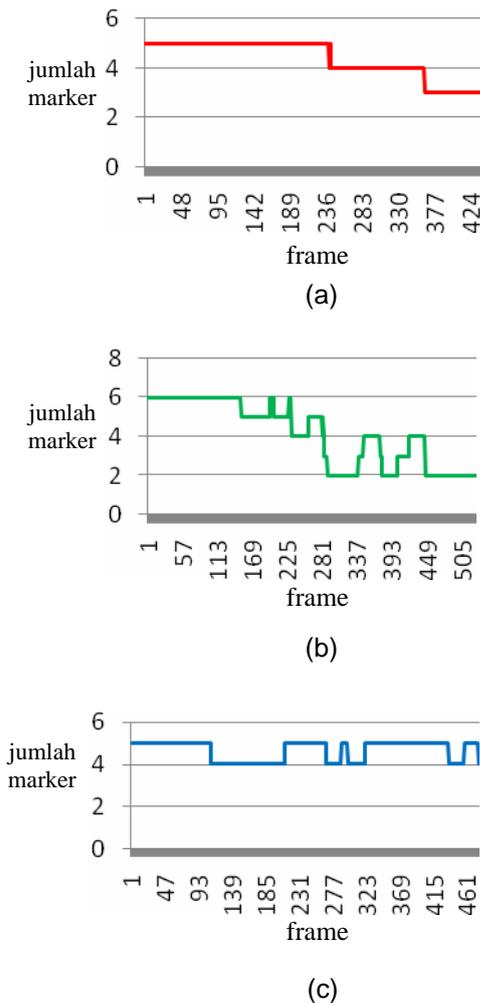
Gambar 1. Alur Kerja Penelitian



Gambar 3. Beberapa Frame Sekuen Video Gerak Tangan

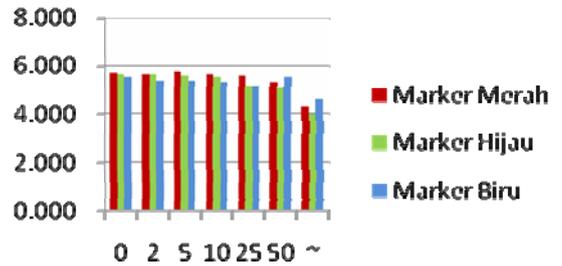
Output dari proses pengambilan data adalah tiga buah sekuen video gerakan tangan, masing-masing untuk marker berwarna merah, hijau, dan biru. Gerakan yang dilakukan tangan meliputi gerakan berputar, bergeser, membuka dan menutup jari, serta mencengkram. Gerakan dilakukan dengan telapak tangan menghadap kamera (*frontal view*). Pengambilan data sekuen video menggunakan bantuan *software Video Capture Factory*. Gambar 3 memperlihatkan beberapa frame dari sekuen video tersebut. Sedangkan variabel-variabel dari sekuen video tersebut adalah sebagai berikut:

- Ukuran frame = 320x240 piksel
- *Frame rate* = 24 fps
- Jumlah frame = 463 (merah)/ 556 (hijau)/ 499 (biru)
- Durasi = 19,3 detik (merah)/ 23,1 detik (hijau)/ 20.8 detik (biru)



Gambar 4. Jumlah marker yang dapat *ditracking* dengan tepat tiap-tiap frame pada Jurnal Teknologi, Volume 2 Nomor 2, Desember 2009, 194-200

ketiga sekuen video dengan warna marker (a) merah, (b) hijau, (c) biru



Gambar 5. Rata-rata jumlah marker yang dapat *ditracking* pada ketiga sekuen video dengan proses *re-identifikasi corner* setelah *n* frame

Jumlah marker yang dapat *ditracking* dengan tepat tiap-tiap frame pada ketiga sekuen video dengan warna marker masing-masing merah, hijau, dan biru ditunjukkan oleh diagram pada Gambar 4. Perhitungan ketepatan *tracking* berdasarkan pada posisi titik hasil *tracking*. Jika posisi titik tersebut berada pada area marker maka *tracking* dianggap tepat pada marker (marker berhasil *ditracking*). Jika tidak maka *tracking* dianggap gagal. Dari ketiga diagram tersebut terlihat bahwa marker yang berhasil *ditracking* jumlahnya tidak tentu dan memiliki kecenderungan semakin sedikit jumlahnya seiring dengan bertambahnya jumlah frame yang *ditracking*. Jika dianalisa hal tersebut dapat disebabkan karena titik-titik hasil *tracking* pada suatu frame digunakan sebagai *corner* bagi proses *tracking* pada frame berikutnya. Sehingga kesalahan *tracking* suatu titik pada suatu frame akan berimbas kesalahan pula pada *tracking* berikutnya. Untuk itu, sebagai eksperimen, dilakukan proses *tracking* dengan menggunakan proses *re-identifikasi corner* setelah proses *tracking* dilakukan pada sejumlah frame. Perbandingan rata-rata jumlah marker yang dapat *ditracking* tiap frame pada proses *tracking* menggunakan proses *re-identifikasi* setelah *n* frame ditampilkan pada Gambar 5.

KESIMPULAN

Dari hasil penelitian yang telah dilakukan, tampak bahwa penggunaan proses *re-identifikasi corner* setelah *tracking* dilakukan pada *n* frame meningkatkan jumlah rata-rata marker yang berhasil *ditracking* jika dibandingkan *tracking* tanpa proses *re-identifikasi corner*. Rata-rata peningkatannya adalah 1,291 untuk marker merah, 1,418 untuk marker hijau, dan 0,747 untuk marker biru. Begitu pula ketika *tracking* dilakukan dengan

berbasis pada identifikasi corner saja. Nilai rata-rata jumlah marker yang berhasil diidentifikasi juga lebih besar, yaitu rata-rata 5,610 dibandingkan 4,319 jika marker *tracking* tanpa menggunakan proses *re-identifikasi* marker. Dengan demikian proses *tracking* dapat dilakukan hanya dengan menggunakan proses identifikasi *corner* saja.

Sedangkan jika dibandingkan, penggunaan warna marker berbeda, yaitu merah, hijau, dan biru, pada proses *tracking* ini tidak membuahkan perbedaan hasil yang signifikan. Nilai selisih rata-rata jumlah marker yang berhasil *tracking* terbesar adalah antara marker warna biru dan hijau, yaitu 0,595. Jika dianalisa, perbedaan-perbedaan yang ada lebih disebabkan oleh gerakan-gerakan tangan pada ketiga sekuen video yang tidak persis sama.

DAFTAR PUSTAKA

- Crowley, J., Berard, F., dan Coutaz, J., 1995, "Finger tracking as an input device for augmented reality", IWAGFR '95: International Workshop on Gesture and face recognition, 1995
- Hassanpour, R., Birdal, A., 2008, "Region Based Hand Gesture Recognition", Proceedings 16th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision 2008, Plzen-Bory, Czech Republic
- Kim, H., dan Felner, D.W., 2004, "Interaction with hand gesture for a back projection wall", CGI '04: Computer Graphics International, IEEE Computer Society, Washington, USA, 2004
- Koike, H., Saito, Y., dan Ozawa, S., 2001, "Integrating paper and digital information on enhanced desk: a method for realtime finger tracking on an augmented desk system", ACM Transaction on Computer Human Interface. Vol. 8(4), hal. 307-322
- Letessier, J., dan Berard, F., 2004, "Visual tracking of bare fingers for interactive surfaces", UIST '04: 17th Annual ACM Symposium on User Interface Software and Technology, New York, USA, 2004
- Malik, S., dan Laszlo, J., 2004, "Visual touchpad: a two-handed gestural input device", ICMI '04: 6th International Conference on Multimodal Interfaces, ACM Press, New York, USA, 2004
- Maggioni, C., 2004, "Gesture computer—history, design, and application", Computer Vision and Human-Machine Interaction, Cambridge, hal. 312-325
- O'Hagan, R., dan Zelinsky, A., 1997, "Finger track – a robust and realtime gesture interface", AI '97: 10th Australian Joint Conference on Artificial intelligence, Springer-Verlag, London, UK, 1997
- O'Hagan, R., Zelinsky, A., dan Rougeaux, S., 2002, "Visual gesture interfaces for virtual environments", Interacting with Computers, Vol. 4, hal 231-250
- Oka, K., Sato, Y., dan Koike, H., 2002, "Real-time tracking of multiple fingertips and gesture recognition for augmented desk interface systems", FGR '02: Fifth IEEE International Conference on Automatic Face and Gesture Recognition, IEEE Computer Society, Washington DC, USA, 2002
- Segen, J., dan Kumar, S., 1998, "Gesture VR: vision-based 3D interface for spatial interaction", Sixth ACM International Conference on Multimedia, ACM Press, New York, USA, 1998
- Usabiaga, J.H., 2005, "Global hand pose estimation using multiple camera ellipse tracking", Thesis Master, Department of Computer Science, University Of Nevada, Reno, 1995