

PEMANFAATAN REAL-TIME FACE TRACKING DALAM APLIKASI AUGMENTED REALITY FRAME KACAMATA

¹Endang Setyati, ²David Alexandre

Teknik Informatika

Sekolah Tinggi Teknik Surabaya

¹endang@stts.edu, ²v.davidalexandre@gmail.com

ABSTRAK

Augmented Reality adalah penggabungan antara objek buatan hasil penghitungan komputer dengan dunia nyata. Objek buatan tersebut berasal dari grafik komputer, misalnya grafik dalam bentuk dua dimensi maupun tiga dimensi. Aplikasi yang dibuat dengan augmented reality dituntut untuk dapat bekerja secara real-time dan dapat berinteraksi dengan pengguna. Dengan demikian aplikasi augmented reality dapat menjadi daya tarik tersendiri bagi pengguna.

Penerapan augmented reality ini akan dilakukan untuk pembuatan aplikasi simulasi model frame kacamata. Pengguna dapat mencoba berbagai jenis model frame kacamata dengan konsep augmented reality sehingga pengguna seakan-akan sedang menggunakan model frame kacamata yang sedang dipilih dalam aplikasi. Pengguna hanya perlu berada dalam jangkauan webcam saat mengaktifkan fitur pemasangan frame kacamata pada aplikasi dan model frame kacamata buatan akan terpasang dengan sendirinya pada wajah pengguna. Aplikasi simulasi ini didukung oleh berbagai macam pilihan model frame kacamata yang dapat dicoba oleh pengguna.

Dengan menggunakan pendeteksian gerakan pengguna secara otomatis dan pembuatan frame kacamata dengan OpenGL maka aplikasi simulasi dengan augmented reality ini menjadi lebih realistis dan memiliki daya tarik tersendiri untuk pengguna. Melalui aplikasi simulasi model frame kacamata ini pengguna dapat mencoba berbagai macam model frame kacamata dengan praktis daripada harus mencoba model frame kacamata dalam bentuk fisik satu-persatu.

Kata kunci: *Augmented Reality, Face Tracking, Pose Detection*

ABSTRACT

Augmented reality is a term for a merging of real-world and augmented object from computer such as two-dimensional and three-dimensional graphics, video and sound. Augmented reality application is expected to work in real-time environment and have interaction with user. Thus the augmented reality application can show the reality to the user just like the concept itself.

The implementation of augmented reality in this final project is used to make eyeglasses frame simulation. User can try different types of eyeglasses frame models in the application with the concept of augmented reality as if the user really use the eyeglasses frame. The user must be in the input image from webcam when activate the augmented frame and the eyeglasses frame will be fitted on the user's face. This

application is also supported with many features like capture image, add frame and so on.

With automatic detection of user movements and eyeglasses frame generated by OpenGL, this application become more realistic. User can simulate various models of eyeglasses frame on his face.

Keywords: Augmented Reality, Face Tracking, Pose Detection

1. PENDAHULUAN

Pendeteksian wajah atau sering disebut *face detection* atau *face tracking* adalah salah satu hal yang mengalami perkembangan pesat dalam dunia *computer vision*. Berbagai macam algoritma dan metode telah diterapkan untuk melakukan pendeteksian wajah. Hal ini didukung dengan semakin luasnya manfaat yang dapat diperoleh dari pengembangan *face detection*. Aplikasi dari *face detection* dapat menggunakan media video-camera dan wajah pengguna aplikasi sebagai *input*.

Salah satu perkembangan yang cukup berpengaruh adalah pada pengolahan data wajah secara dua dimensi. Dari hasil penelitian diperoleh fakta bahwa pengubahan data wajah dari dua dimensi dimodelkan pada sebuah ruang tiga dimensi adalah nyata. Melihat dari hasil tersebut berdampak pada semakin besar dan luas hal yang dapat dilakukan oleh *face detection*. Ide dasar dari pemodelan ini adalah untuk memperbaiki kelemahan-kelemahan yang didapatkan apabila hanya dilakukan pendeteksian secara dua dimensi. Contohnya, dengan pendeteksian secara dua dimensi penghitungan tidak dapat memperkirakan bentuk dan arah wajah sebaik pendeteksian tiga dimensi. Melalui beberapa pendekatan yang tepat pendeteksian wajah secara tiga dimensi tersebut dapat semakin melebarkan manfaat dari pendeteksian wajah seperti *face recognition* yang lebih akurat maupun untuk penggunaan *augmented reality*.

Augmented reality adalah sebuah konsep penggabungan antara objek hasil kalkulasi komputer dengan dunia nyata. Penggunaan *augmented reality* sangat erat kaitannya dengan dunia *computer vision*. Pada umumnya media yang digunakan adalah video-camera dan sebuah *marker*. Dengan posisi dari *marker* yang sudah dideteksi tersebut maka *augmented reality* dapat dilakukan.

Dalam implementasinya aplikasi *augmented reality* akan menggunakan *face tracking* untuk melakukan simulasi *frame* kacamata. Cara yang digunakan adalah dengan memanfaatkan hasil *face tracking* sebagai *marker* dan kemudian dilakukan proses untuk pengenalan model wajah dalam tiga dimensi untuk memunculkan *augmented reality* berupa *frame* kacamata. Dengan demikian program yang dihasilkan akan menjadi simulator bagi pengguna untuk melakukan simulasi pemakaian *frame* kacamata.

Untuk tujuan dari implementasi ini adalah untuk melakukan penelitian mengenai aplikasi *augmented reality* dan tahap pembuatannya. Aplikasi tersebut diharapkan dapat menjadi media hiburan serta menambah pengetahuan mengenai pendeteksian wajah dan *augmented reality*.

Masalah yang dihadapi pada implementasi adalah bagaimana cara untuk melakukan proses pendeteksian wajah, perkiraan pose wajah dan penggambaran *frame* kacamata secara cepat dan akurat. Dikarenakan aplikasi *augmented reality* ini dituntut untuk dapat bekerja secara real-time sehingga pemilihan metode akan sangat berpengaruh pada hasil akhir dari implementasi.

2. LANDASAN TEORI

Pada penerapan real-time face tracking dalam aplikasi terdapat beberapa teori pendukung. Adapun teori-teori yang digunakan saat proses pembuatan adalah:

2.1. Haar-Like Features

Haar-like features adalah salah satu cara untuk melakukan pendeteksian wajah. Cara kerja dari metode ini adalah dengan melakukan perubahan suatu bagian image (Region of Interest) menjadi sebuah nilai. Perubahan suatu daerah image yang memiliki wajah manusia menjadi nilai tertentu dapat menjadikan metode ini salah satu metode pendeteksian wajah yang cepat dan akurat. Nama dari haar-like features didapatkan dari Viola dan Jones yang melakukan penghitungan yang memiliki kemiripan dengan metode haar wavelet transform.

Secara lebih detail, Haar-like features bekerja dengan menandai daerah yang berdekatan pada sebuah image menjadi area pixel berdasarkan classifier-nya. Kemudian dilakukan penghitungan intensitas area pixel pada daerah tersebut dan menghitung perbedaan masing-masing daerah tersebut. Perbedaan yang didapatkan digunakan untuk mengkategorisasi tiap daerah sehingga dapat digunakan untuk melakukan pendeteksian objek.

2.2. Statistical Shape Model

Pada dasarnya bentuk dari sebuah objek dapat direpresentasikan dengan himpunan dari n titik yang dapat berada dalam dimensi berapapun (biasanya dalam dua atau tiga dimensi). Bentuk dari sebuah objek tidak akan berubah walaupun mengalami translasi, rotasi, maupun skalasi.

Dalam Statistical Shape Model didapatkan sebuah model yang memungkinkan untuk menganalisa bentuk yang baru dan menghasilkan bentuk yang sesuai dengan yang ada pada training set. Training set biasanya didapatkan dari sejumlah training images yang ditandai secara manual. Dengan menganalisa variasi bentuk dari training set, dibentuk sebuah model yang dapat menyamai variasi tersebut.

2.3. Active Shape Model

Active Shape Model (ASM) adalah sebuah metode dimana model dari sebuah objek secara iteratif berubah bentuk untuk mencocokkan model tersebut pada sebuah objek pada gambar. Metode tersebut menggunakan model yang fleksibel yang didapatkan dari sejumlah training example.

Diberikan sebuah posisi perkiraan pada gambar, active shape model dapat secara iteratif mencocokkan pada gambar. Dengan memilih satu set shape parameter untuk Point Distribution Model, bentuk dari model dapat didefinisikan dalam frame koordinat yang berpusat pada objek. Instance X pada model dalam gambar dapat dibuat dengan mendefinisikan posisi, orientasi, dan skala.

Secara garis besar, active shape model bekerja dengan mengulangi langkah-langkah berikut:

1. Cari pada gambar di sekeliling setiap titik untuk lokasi yang lebih baik bagi titik tersebut
2. Meng-update Pose dan shape parameter (X_b, Y_b, s, θ, b) pada model sehingga sesuai dengan posisi baru yang ditemukan
3. Memberikan konstrain pada parameter b untuk memastikan bentuk yang sesuai (Misal $|b_i| < 3 \sqrt{\lambda_i}$)

4. Ulangi sampai konvergen (konvergensi dicapai apabila tidak terdapat perbedaan berarti antara sebuah iterasi dengan iterasi sebelumnya).

2.4. Pose From Orthography And Scaling With Iteration

Pose from Orthography and Scaling with Iteration atau disebut POSIT adalah algoritma yang dapat digunakan untuk mengestimasi posisi dari objek tertentu dalam bidang tiga dimensi. Untuk mengukur pose dari sebuah objek, diperlukan paling sedikit empat titik non-coplanar objek dari image. Karena itu akan digunakan proyeksi bidang 3D dari titik-titik 2D yang ditemukan sebagai input dari algoritma ini. Algoritma ini memperkirakan proyeksi perspektif dari proyeksi ortografi objek pada image yang sudah disesuaikan skalanya. Dari hasil proyeksi tersebut ditemukan *rotation matrix* dan *translation vector* dari objek.

Dengan demikian tujuan dari algoritma ini adalah untuk menghitung rotation matrix dan translation vector dari sebuah objek. Rotation matrix \mathbf{R} dari sebuah objek adalah sebuah matriks yang terdiri dari koordinat tiga buah vektor \mathbf{i} , \mathbf{j} , dan \mathbf{k} dari sistem koordinat kamera yang diubah menjadi sistem koordinat objek. Fungsi dari rotation matrix tersebut adalah untuk melakukan penghitungan dari sistem koordinat objek dari vektor seperti M_0M_i menjadi sistem koordinat kamera. Hasil perkalian dari $M_0M_i \cdot \mathbf{i}$ antara baris pertama matriks dan vektor M_0M_i akan memberikan proyeksi vektor tersebut pada vektor \mathbf{i} yang berada pada koordinat kamera. Sebagai contoh koordinat $X_i - X_0$ dari M_0M_i , selama koordinat M_0M_i dan barisan vektor \mathbf{i} berada pada sistem koordinat yang sama maka akan didapatkan hasil sistem koordinat objek sebagai berikut.

$$\mathbf{R} = \begin{bmatrix} i_u & i_v & i_w \\ j_u & j_v & j_w \\ k_u & k_v & k_w \end{bmatrix}$$

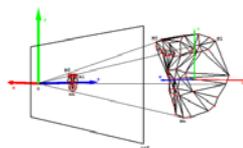
dimana i_u , i_v , i_w adalah koordinat \mathbf{i} dalam sistem koordinat objek (M_0u , M_0v , M_0w).

Untuk mendapatkan rotation matrix, kita hanya memerlukan penghitungan \mathbf{i} dan \mathbf{j} pada sistem koordinat objek. Vektor \mathbf{k} akan didapatkan melalui cross-product dari $\mathbf{i} \times \mathbf{j}$. Sedangkan untuk translation vector, \mathbf{T} adalah vektor OM_0 diantara titik proyeksi O dan titik referensi objek yaitu M_0 . Koordinat translation vector yang didapatkan berupa X_0 , Y_0 , dan Z_0 .

Perlu diketahui bahwa penentuan titik referensi objek harus berada pada titik awal koordinat axis.

2.5. Head Pose Estimation

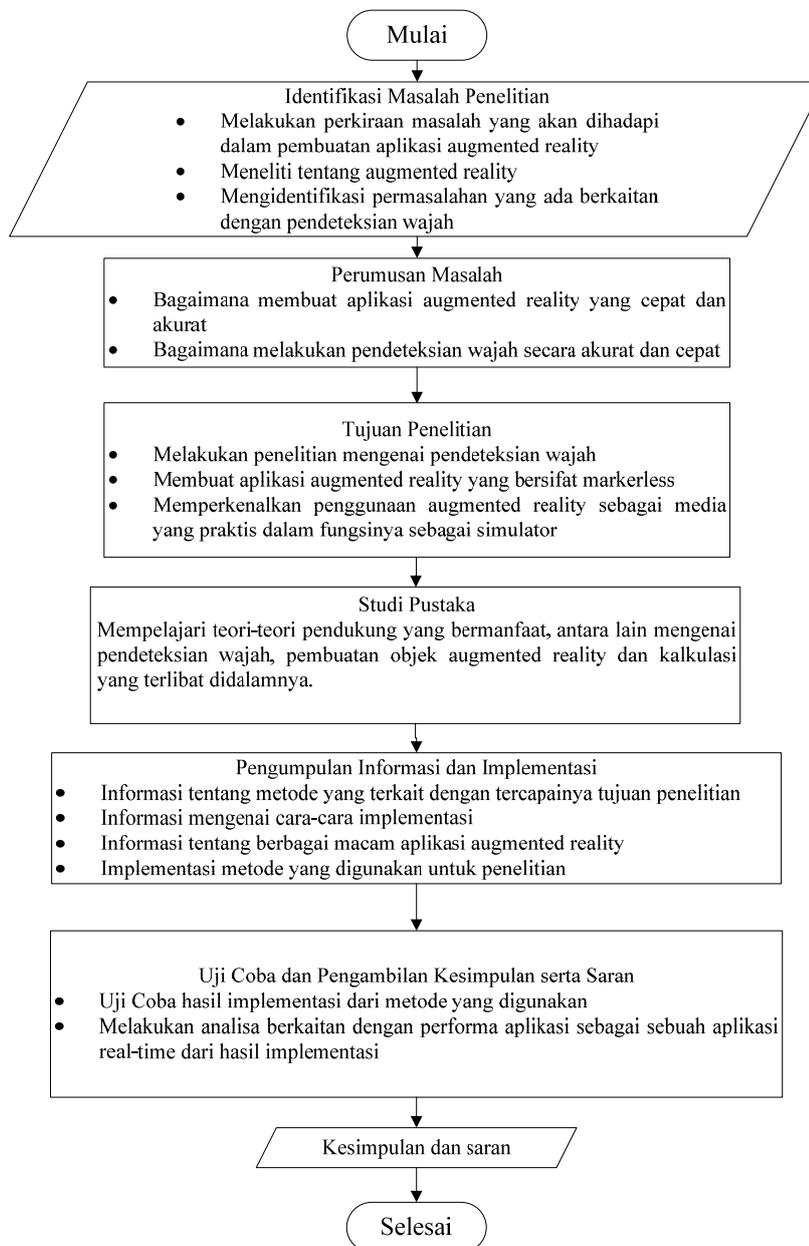
Estimasi dari pose kepala pengguna saat menggunakan aplikasi akan memanfaatkan ketiga algoritma yang sebelumnya sudah disebutkan. Langkah yang digunakan adalah pendeteksian wajah menggunakan Haar-like features, melakukan pendeteksian titik fitur wajah pengguna dengan active shape model dan mengambil pose dari wajah pengguna dengan memanfaatkan algoritma POSIT. Pengambilan pose akan menggunakan titik fitur pada wajah pengguna yang sebelumnya sudah didapatkan dengan active shape model.



Gambar 1. Titik Fitur pada Wajah digunakan untuk Proyeksi pada Algoritma POSIT

3. METODE PENELITIAN

Tahapan dari penelitian yang dilakukan akan dijelaskan sebagai berikut:



Gambar 2. Flowchart Penelitian

1. Tahap perumusan masalah

Pada tahap perumusan masalah dilakukan perkiraan masalah-masalah yang akan ditemukan pada saat dilakukan implementasi. Perumusan masalah ini dilakukan dengan cara mencari informasi literatur yang berkaitan dengan pembuatan augmented reality serta mengumpulkan bahan yang bersumber pada buku, paper, dan lainnya yang berkaitan dengan augmented reality, face tracking, dan pose estimation. Dilakukan pula

pembelajaran cara penghitungan dan cara terkait yang akan digunakan dalam algoritma yang ada.

2. Tahap pengumpulan informasi dan implementasi

Setelah permasalahan dan tujuan ditetapkan maka akan dilakukan studi pustaka. Pembelajaran terhadap studi pustaka akan meliputi teori-teori yang berkaitan dengan permasalahan yang dihadapi dengan maksud untuk mencari solusi yang paling tepat. Teori yang diperoleh berdasar dari buku, jurnal serta penelitian terkait dengan augmented reality, face tracking dan pose estimation. Pada penelitian ini teori yang digunakan antara lain adalah haar-like features, active shape model, shape model statistic dan pose estimation. Selanjutnya akan dilakukan implementasi menggunakan teori yang ada untuk mencapai tujuan penelitian. Implementasi dilakukan secara bertahap mengingat proses dari penelitian ini bersifat berkesinambungan dan satu proses akan mempengaruhi proses yang lainnya.

3. Tahap uji coba dan pengambilan kesimpulan serta saran

Proses setelah dilakukan implementasi adalah dengan melakukan uji coba. Penerapan uji coba akan memfokuskan pada hasil akurasi pendeteksian yang digunakan sebagai dasar bagi penempatan frame kaca mata secara augmented reality. Setelah dilakukan uji coba maka akan diambil beberapa kesimpulan dari penelitian ini. Kesimpulan tersebut berkaitan dengan tujuan dari penelitian. Dari kesimpulan ini kemudian akan diambil saran-saran yang dapat membangun penelitian ini ke arah yang lebih baik lagi.

4. HASIL DAN UJI COBA

Pada ujicoba berikut telah terlihat bahwa pendeteksian wajah atau face tracking yang digunakan dalam program diketahui memiliki perubahan kecepatan kinerja yang signifikan sesuai pada resolusi dari image sebagai input. Dengan melakukan percobaan pada penggunaan image dari berbagai resolusi dengan perbandingan resolusi yang sama yaitu 4 : 3, diketahui bahwa image dengan resolusi yang rendah dapat mempercepat kinerja program dan menambah jumlah frame per second dari hasil output program. Semakin rendah resolusi yang digunakan, maka frame per second yang didapatkan semakin tinggi. Dapat disimpulkan bahwa untuk pendeteksian wajah pada program, resolusi memiliki perbandingan terbalik dengan frame per second. Untuk perbandingan resolusi dapat dilihat pada tabel 1.

Tabel 1. Perbandingan Resolusi dan Frame Per Second

Resolusi	Frame Per Second (FPS)
320 x 240	15-18
480 x 360	12-15
600 x 400	10-12

Ujicoba pada pose kepala pengguna dalam program menggunakan algoritma POSIT dimana didapatkan rotation matrix dan translation vector sebagai hasilnya. Hasil rotation matrix dan translation vector yang didapatkan dari algoritma POSIT dapat dikatakan benar walaupun tidak akurat. Hal ini dikarenakan hasil dari pendeteksian feature point pada wajah dengan active shape model yang tidak bersifat konstan

sehingga menyebabkan hasil yang tidak akurat pada perkiraan pose kepala pengguna. Hasil error pada perkiraan pose kepala pengguna memiliki kesalahan sampai dengan ± 10 derajat. Hal ini menyebabkan pergeseran titik pada hasil yang sebenarnya sehingga mengakibatkan pergeseran letak frame kacamata pada output program. Namun pergeseran tersebut tidak sampai pada tahap yang mengganggu fungsionalitas program.

Pada ujicoba penggunaan program didapatkan hasil augmented reality yang cukup layak untuk digunakan dari segi kecepatan dan akurasi. Frame kacamata yang digunakan dibuat menggunakan OpenGL dan diletakkan pada lokasi sesuai dengan hasil pendeteksian wajah dan pose yang dilakukan sebelumnya. Gambaran hasil aplikasi dapat dilihat pada gambar 3.



Gambar 3. Hasil pada Aplikasi

5. KESIMPULAN DAN SARAN

Dari hasil analisa terhadap implementasi yang sudah dilakukan, maka dapat ditarik kesimpulan sebagai berikut:

1. Pemilihan image input yang tepat dalam penerapan algoritma active shape model dapat mempengaruhi hasil yang diperoleh. Filtering dan pemilihan tempat pengambilan gambar pada input sebelum dilakukan penerapan algoritma active shape model akan sangat berguna untuk mendapatkan hasil yang lebih baik.
2. Algoritma active shape model dalam implementasinya memiliki tingkat error yang tidak dapat dihindari dikarenakan faktor shape model yang digunakan maupun efisiensi algoritma. Sampai sekarang algoritma ini masih terus dikembangkan untuk mendapat hasil yang lebih baik dari sebelumnya.
3. Algoritma POSIT memiliki hasil yang sangat bergantung pada titik pada image yang digunakan sebagai input-nya.
4. Aplikasi augmented reality harus didukung pemilihan algoritma yang tepat karena dituntut untuk dapat berjalan secara real-time.
5. Dalam penggunaan aplikasi akan lebih baik apabila dilakukan pada tempat dengan pencahayaan yang baik serta menggunakan webcam yang memiliki fokus dan ketajaman gambar yang baik.

Pada tahap selanjutnya, berikut adalah saran yang dapat mungkin dipakai untuk pengembangan lebih lanjut.

- Mempelajari lebih lanjut tentang algoritma active shape model dan algoritma efisiensi terkait dengan model statistik yang sekiranya dapat digunakan untuk mengembangkan algoritma active shape model yang digunakan.
- Mempelajari dan melakukan percobaan terkait pada input image dan shape model untuk meningkatkan akurasi dan efisiensi dalam pendeteksian titik feature pada wajah.

6. DAFTAR PUSTAKA

- Baker, K. *Singular Value Decomposition Tutorial*. The Ohio State University. 2005.
- Cootes, T. F. *Active Shape Models – Smart Snakes*. British Machine Vision Conference. 1992.
- Cootes, T. F., C.J. Taylor. *Statistical Models of Appearance for Computer Vision*. University of Manchester. 2004.
- Cootes, T. F., C.J. Taylor, D.H. Cooper, J. Graham. *Active shape models - their training and application*. 1995.
- Dementhon, D., Larry S. D. *Model-Based Object Pose in 25 Lines of Code*. Paper presented at International Journal of Computer Vision. 1995.
- Gonzales, R., Woods, R. *Digital Image Processing Second Edition*. Prentice Hall. 2002.
- Katherine, A., G. Arceneaux, A. John. *Real Time Face And Facial Feature Detection and Tracking*. Paper presented at Midwest Instruction and Computing Symposium. 2010.
- Martins, P., J. Batista. *Monocular Head Pose Estimation*. Paper presented at International Conference on Image Analysis and Recognition. 2008.
- Treiber, M. *An Introduction to Object Recognition*. Springer. 2010.
- Viola, Jones. *Rapid Object Detection using Boosted Cascade of Simple Features*. Paper presented at Computer Vision and Pattern Recognition. 2001.