

Pembuatan Game Puzzle Gambar Dengan Gerakan Tangan dan Perintah Suara

Yusuf¹, Nana Ramadijanti, S.Kom, M.Kom², Setiawardhana, ST².

Mahasiswa D4LJ Jurusan Teknik Informatika, Dosen Politeknik Elektronika Negeri Surabaya²
Politeknik Elektronika Negeri Surabaya Institut Teknologi Sepuluh Nopember

Kampus ITS, Keputih Sukolilo, Surabaya 60111
Telp. (+62)-31-5947280 Fax. (+62)-31-5946114
E-mail: exyl_06@yahoo.com

Abstract : Perkembangan game terus berjalan dengan cepat, dengan meningkatnya teknologi grafik dan meningkatnya hardware untuk mempercepat proses kecepatan visual yang cukup tinggi, maka dengan itu game lebih interaktif dan lebih nyata. Game teka-teki gambar (Puzzle Gambar) sudah tidak asing lagi dan game tersebut masih dibuat dan terus diperbarui dengan menambah fitur-fitur agar lebih menarik. Sebagian besar permainan yang telah ada masih menggunakan joystick, keyboard dan alat bantu lainnya. Pada proyek akhir ini dikembangkan permainan yang memanfaatkan gerakan tangan dan ucapan pemain sebagai kontrol. Kontrol gerakan tangan dilakukan dengan mendeteksi gerakan tangan menggenggam menggunakan training haar cascade pada library opencv, sedangkan kontrol ucapan dilakukan dengan pengenalan ucapan menggunakan fungsi speech to text pada speech api sdk 5.1.

Kata Kunci : Game puzzle image, Speech API, OpenCV

Abstract : The development of the game running fast with the increasing of technology and graphics hardware to accelerate the process of visual speed, so the game become more interactive and more real. Puzzle game (Puzzle Image) is familiar games and still developed by adding more features to make it more attractive. Most of the existing games are still using the joystick, keyboard and other tools. This final project is to develop a game that uses hand movement and speech of the player as the control. Hand movement control performed by detecting holding form of hand using haar cascade training on the openCV library, while speech recognition control performed by using of speech to text function on the Speech API 5.1 SDK.

Keyword : Puzzle Image Game, Speech API, OpenCV

1. Pendahuluan

Perkembangan game terus berjalan dengan cepat dengan meningkatnya teknologi grafik dan meningkatnya hardware untuk mempercepat proses kecepatan visual yang cukup tinggi maka dengan itu game lebih interaktif dan lebih nyata, tetapi cukup disayangkan game tersebut bukan untuk anak-anak dan memiliki nilai pendidikan yang kecil. Dengan itu proyek akhir ini akan mengembangkan game yang interaktif dan bisa digunakan pada anak-anak dan diberi banyak pendidikan didalam game tersebut. Seperti game teka-teki gambar (Puzzle Gambar) yang bergambar pengetahuan yang baik.

Game teka-teki gambar (Puzzle Gambar) sudah tidak asing lagi dan game tersebut masih dibuat dan terus diperbarui dengan menambah fitur-fitur agar lebih menarik. Yang menarik dari game tersebut masih tetap digemari oleh pengguna dan dimainkan di beberapa game platform yang ada seperti playstation, xbox, game java di hp, dan Wii.

Game ini akan menampilkan gambar-gambar hewan kemudian dijadikan beberapa potongan kemudian acak, dan tugas pemain mengembalikan potongan gambar menjadi gambar utuh. Setelah gambar kembali utuh maka akan ada "pertanyaan soal gambar hewan apakah ini?". Pemain akan menjawab dengan keras nama hewan tersebut dengan menggunakan bahasa inggris.

Dengan begitu pemain khususnya anak-anak akan mendapat pengetahuan tentang nama hewan beserta

bahasa inggris nama hewan tersebut. Dengan metode ini game akan menjadi interaktif.

Sebagian besar permainan yang telah ada masih menggunakan joystick, keyboard dan alat bantu lainnya. Maka dengan ini kami dituntut untuk mengembangkan permainan yang memanfaatkan gerakan tangan dan suara dari pemainnya sendiri sebagai kontrol karakter. Hal ini bisa dilakukan dengan mendeteksi gerakan tangan dengan WebCam dan mendeteksi suara pada computer.

Karena pengolahan citra dan pengolahan suara menggunakan api library windows merupakan salah satu proses intelligent dengan fleksibilitas yang sangat tinggi dibidang aplikasi apapun, maka berdasarkan persoalan diatas kami mencoba untuk mewujudkan pengenalan deteksi tangan dan deteksi suara hingga dapat membuat game yang mampu membuat pemain bergerak aktif dan melantunkan kata selama permainan berlangsung.

1.1 TUJUAN

Tujuan dari proyek akhir ini adalah membuat aplikasi, yang dapat mendeteksi gerakan tangan yang tertangkap pada web kamera dan pendeteksian suara yang menggunakan microphone. dan kemudian gerakan dan suara tersebut digunakan sebagai pengganti fungsi keyboard ataupun controlling device

lainnya untuk mengenali gerakan dan suara pemain sesuai perintah pada game.

1.2 Batasan Masalah :

Adapun beberapa hal yang membatasi pemakaiannya dalam proyek akhir ini, antara lain :

- Satu pemain.
- Kata yang diucapkan hanya terbatas dan sudah tersedia di game.
- Bentuk tangan harus seperti yang telah ditentukan.
- Intensitas cahaya harus cukup
- Jarak kamera dengan objek terbatas.
- Tangan Kanan
- Tidak ada dominan suara yang lain.

2. Teori penunjang

Pada bab ini akan membahas mengenai teori-teori yang akan dijadikan materi penunjang dalam pembuatan proyek akhir ini. Adapun teori-teori yang akan dibahas di dalam bab ini adalah :

Dari teori diatas akan dijadikan acuan dalam perencanaan dan pembuatan perangkat lunak.

2.1 Pengolahan Citra

Pengolahan citra adalah suatu metode yang digunakan untuk mengolah gambar sehingga menghasilkan gambar yang sesuai dengan keinginan kita. Pengambilan gambar bisa dilakukan dengan menggunakan kamera atau alat lain yang bisa digunakan untuk mentransfer gambar misalnya scanner atau kamera digital. Bahasan kali ini berfokus pada pengambilan gambar menggunakan web kamera. Sehingga citra yang dihasilkan sudah berbentuk sinyal digital dan mudah dikenali atau dibaca oleh komputer. Citra digital adalah citra kontinyu yang sudah didiskritkan baik koordinat spasial maupun kecerahannya. Citra digital dianggap matrik dengan ukuran $M \times N$ dimana baris dan kolom menunjukkan titik-titiknya. Citra berwarna menggunakan metode RGB. Adapun masing-masing warna dalam tabel memiliki 3 buah kombinasi angka yaitu R, G, dan B yang menentukan proporsi warna merah, warna hijau dan warna biru dari warna tersebut. RGB masing-masing memiliki range antara 0 hingga 63 sehingga jumlah warna yang dapat kita pilih untuk mengisi warna pada sebuah cell ditabel ialah $63 \times 63 \times 63 = 16$ juta warna. Tetapi seluruh tabel hanya dapat diisi dengan 256 pilihan warna.

2.2 Pengertian Gesture

Dalam studi pustaka, Andy Wilson dan Aaron Bobick dalam makalahnya menjelaskan tentang cara kerja gerak tangan manusia dimana gerak tangan manusia setiap gerakannya mengalami moment-moment tertentu misalnya pada saat seseorang tersebut

istirahat maupun menggerakkan tangannya. Isyarat dari setiap interaksi nonverbal, menggunakan bahasa gerak baik wajah, tangan, kaki maupun otot-otot yang lainnya sebagai alat untuk berinteraksi dengan komputer.

2.3 Teknik Pengambilan Citra

Permasalahan awal dari pengolahan citra ialah bagaimana cara kita menangkap gambar dari keadaan real untuk kemudian di sampling sehingga terbentuk suatu gambar atau citra digital. Permasalahan ini biasa di sebut dengan *capture*. Proses ini memerlukan alat-alat *capture* yang baik seperti kamera, scanner, light-pen dan lain sebagainya, sehingga kita bisa mendapatkan hasil yang baik pula.

2.4 Contour Finding

Contour sebenarnya adalah sebuah nilai sequences, maka segala sesuatunya harus berada pada koor sequences. Contour ditemukan jika terdapat perbedaan titik-titik yang tinggi dengan tetangganya. Hal ini terjadi karena didasarkan pada hasil deteksi tepi. Maka metode ini kemungkinan dapat mempunyai error yang lumayan tinggi, apabila dihadapkan dengan berbagai macam obyek yang ada di dalam suatu image/gambar. Oleh karena itu, harus diberi batasan bahwa dalam suatu image/gambar hanya boleh terdapat satu obyek, terutama dalam gambar/image yang akan menjadi templatnya.

2.5 Object Detection

Metode yang digunakan untuk melakukan *object detection* pada proyek akhir ini adalah *Haar Classifier*, yaitu metode *object detection* yang membangun sebuah boosted rejection cascade, yang akan membuang data training negative sehingga didapat suatu keputusan untuk menentukan data positif. *Haar Classifier* merupakan metode supervised learning, yaitu membutuhkan data training untuk dapat mendeteksi obyek-obyek tertentu. Untuk itu, *Haar Classifier* membutuhkan data positif (obyek yang akan dideteksi) dan data negatif (bukan obyek yang akan dideteksi).

2.6 OpenCV Library

OpenVC adalah suatu library gratis yang dikembangkan oleh developer-developer Intel Corporation. Library ini terdiri dari fungsi-fungsi computer vision dan image processing tingkat tinggi. OpenCV sangat disarankan untuk programmer yang akan berkuat pada bidang computer vision, karena library ini mampu menciptakan aplikasi yang handal, kuat dibidang digital vision. Karena library ini bersifat cuma-cuma dan sifatnya yang open source, maka dari itu OpenCV tidak dipesan khusus untuk pengguna arsitektur Intel, tetapi dapat dibangun pada hampir semua arsitektur.

Saat ini para developer dari Intel Corporation telah membuat berbagai macam versi, yaitu:

- openCV untuk bahasa pemrograman C/C++,
- openCV untuk bahasa pemrograman C# (masih dalam tahap pengembangan), dan
- openCV untuk bahasa pemrograman Java.

Untuk bahasa pemrograman C# dan Java, karena masih dalam tahap pengembangan, maka kita membutuhkan library lain sebagai pelengkap kekurangan yang ada. Namun untuk bahasa pemrograman C/C++ tidak memerlukan library lainnya untuk pemrosesan pada computer vision.

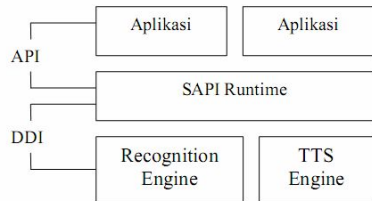
2.7 Sistem Pengenalan Pembicaraan

1. **Mode diktasi.** Pada mode ini pengguna komputer dapat mengucapkan kata/kalimat yang selanjutnya akan dikenali oleh komputer dan diubah menjadi data teks. Kemungkinan jumlah kata yang dapat dikenali dibatasi oleh jumlah kata yang telah terdapat pada database. Pengenalan mode diktasi merupakan speaker dependent. Keakuratan pengenalan mode ini bergantung pada pola suara dan aksentuasi pembicara serta pelatihan yang telah dilakukan.
2. **Mode command and control.** Pada mode ini pengguna komputer mengucapkan kata/kalimat yang sudah terdefinisi terlebih dahulu pada database dan selanjutnya akan digunakan untuk menjalankan perintah tertentu pada aplikasi komputer. Jumlah perintah yang dapat dikenali tergantung dari aplikasi yang telah mendefinisikan terlebih dahulu pada database jenis-jenis perintah yang dapat dieksekusikan. Mode ini merupakan speaker independent karena jumlah kata yang dikenali biasanya terbatas sekali dan ada kemungkinan pembicara tidak perlu melakukan pelatihan.

SAPI 5.1 terdiri dari 2 antar muka yaitu application programming interface (API) dan device driver interface (DDI).

Application Programming Interface(API).

Pada sistem pengenalan pembicaraan, aplikasi akan menerima even pada saat suara yang diterima telah dikenali oleh engine.



Gambar 2.1 Blok Diagram Arsitektur SAPI

3. Perancangan Sistem

3.1 Instalasi dan Konfigurasi OpenCV

OpenCV adalah suatu library computer vision yang dibuat oleh para developer Intel Corporation yang bersifat Open Source, library tersebut merupakan kumpulan fungsi-fungsi C dan beberapa kelas C++ yang mengimplementasikan banyak algoritma-algoritma Image Processing dan Computer Vision. Modul-modul OpenCV, antara lain:

- cv* – fungsi-fungsi utama OpenCV,
- cvaux* – fungsi tambahan (eksperimental) OpenCV,
- cxcore* – dukungan struktur data dan aljabar linear,
- highgui* – fungsi-fungsi yang menangani masalah GUI.

3.2 Instalasi Speech API SDK

Speech API adalah suatu library computer vision yang dibuat oleh para developer Windows, library tersebut merupakan kumpulan fungsi-fungsi C dan beberapa kelas C++ yang mengimplementasikan banyak algoritma-algoritma speech processing .

Proyek akhir ini memakai Speech SDK versi 5.1 .Untuk instalasi hanya seperti install program seperti biasa.Kemudian Speech digunakan di aplikasi speech yang menggunakan VB.NET kemudian digabungkan ke aplikasi game yang menggunakan VC++.NET .

3.3 Perancangan Sistem Game

3.3.1 Perancangan Sistem image Processing

Pra pengolahan citra dilakukan training image pada algoritma haar training.

Mempunyai 1800 data gambar positif masing-masing objek yang dimaksud gambar positive adalah gambar yang didalamnya terdapat gambar suatu objek yang akan menjadi deteksi gerakan, seperti gambar manusia yang mengenggam telepek tangannya. Kemudian mempunyai 2000 gambar negatif, yang dimaksud dengan gambar negative adalah suatu gambar yang didalamnya tidak terdapat suatu objek yang akan dideteksi seperti gambar telepek tangan mengenggam. Kemudian melakukan crop gambar objek yang terdapat di gambar positif.

Setelah proses cropping maka akan mendapat nilai suatu variable cascade dikonvert menjadi xml database.dengan database tersebut dapat melakukan load objek yang akan dideteksi oleh webcam.



Gambar 3.1 Gambar Negatif



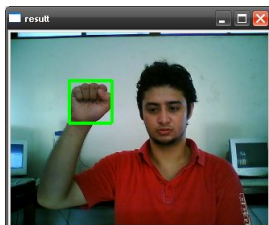
Gambar 3.2 Gambar positif

3.3.2 Deteksi Objek

Setelah training selesai maka xml akan menjadi suatu database objek deteksi sudah.Kamera akan menangkap gambar kemudian aplikasi melakukan load database pada xml yang sudah terbuat maka aplikasi akan mencari objek yang akan dideteksi .

Aplikasi akan mengcapture image dari webcam hasil capture tersebut akan melakukan konversi warna RGB ke Gray.

Setelah itu gambar/image grayscale tersebut dirubah skalanya.

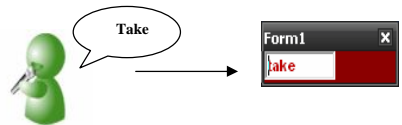


Gambar 3.3 Tangan sudah terdeteksi

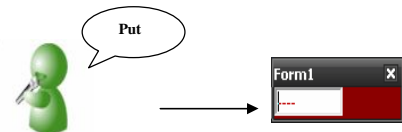
3.4 Speech Recognition

Speech Recognition dilakukan oleh library SpeechAPI (SAPI) SDK 5.1 dan dibuat di aplikasi yang berbeda menggunakan VB.NET .Dalam pendeteksian suara proyek akhir ini akan memanfaatkan SAPI menjadi dalam bentuk perintah suara yang telah ditentukan .

Suara akan diucapkan melalui microphone dan akan diproses untuk mencari mana kata yang paling sesuai diantara kata-kata yang ditentukan .



Gambar 3.4 Pemain Mengucapkan kata “Take” dengan benar



Gambar 3.5 Pemain Mengucapkan kata yang tidak ditentukan

Aplikasi speech dibentuk memakai vb.NET sedangkan aplikasi game menggunakan VC++ ,maka aplikasi speech akan menulis hasil output dari deteksi kata tersebut ke notepad bernama sample.txt.

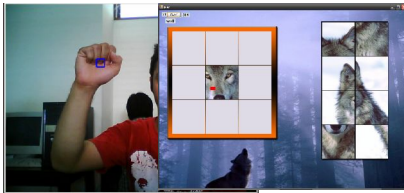
3.5 Penggabungan Speech ke Aplikasi Game

Setelah notepad terbuat maka aplikasi game akan melakukan pemanggilan aplikasi speech yang berupa .exe untuk dapat digabungkan ke aplikasi game.

Setelah aplikasi speech terpanggil kemudian aplikasi game membaca (read).hasil file notepad tersebut .kemudian hasil notepad akan digunakan untuk pengondisian suatu perintah pada game.

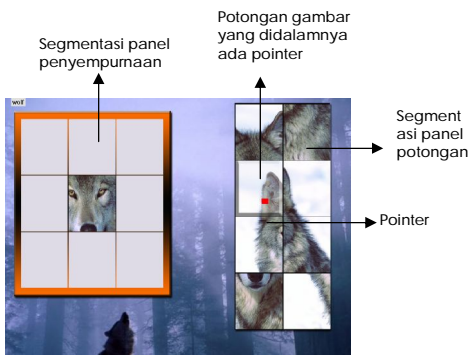
3.6 Implementasi System Pada game

Setelah deteksi tangan menggenggam berhasil dan perintah suara sudah berhasil maka akan digabungkan dengan system game.Aplikasi Game akan membuat sebuah pointer layaknya mouse dengan menggunakan deteksi tangan menggenggam yang berguna untuk memilih dan membawah sebuah potongan gambar.Tangan pemain mengarahkan di area gambar yang akan diambil.



Gambar 3.6 Tangan dapat menjalankan pointer

Untuk mendeteksi pointer game yang digerakan tangan pemain sudah masuk diarea potongan gambar dilakukan dengan cara segmentasi windows



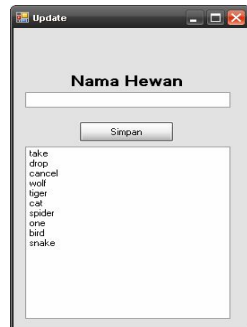
Gambar 3.7 Pointer berada di dalam potongan gambar

Setelah semua potongan gambar terletak di panel penyempurnaan dan gambar menjadi sempurna maka akan tampil pertanyaan seperti gambar berikut :



Gambar 3.8 Pertanyaan nama hewan

Pada game ini dapat menambah nama hewan baru dengan mengkopikan folder gambar yang didalamnya terdapat potongan-potongan gambar yang akan menjadi gambar sempurna. Untuk penambahan gambar pemain harus melakukan penambahan nama hewan untuk perintah kata diwaktu gambar tersebut menjadi sempurna.



Gambar 3.9 Penambahan nama hewan

4. Analisa dan pengujian

Pengujian ini dilakukan oleh 10 pemain dengan berbeda-beda ukuran dan warna kulit pada pemain, berapa nilai keberhasilan untuk melakukan beberapa jenis pengujian. Percobaan dilakukan oleh 10 pemain. Tiap-tiap pemain mempunyai perbedaan masing-masing.

Pengujian dibagi dua macam yaitu :

- Pengujian image processing
- Pengujian speech

4.1 Pengujian Image processing

Pengujian ini terdiri dari :

- Kecepatan terakan tangan
- Posisi objek Tangan
- Jarak Kamera dengan Pemain
- Terhadap Cahaya

Kesimpulan pengujian yaitu cahaya diharapkan terang kulit pemain diharapkan lebih cerah beserta ukuran yang cukup besar, jarak tangan dengan kamera tidak terlalu jauh tidak terlalu dekat.

4.2 Pengujian Speech

Terdiri dari :

- Keras Lemahnya Suara
- Kebisingan Area Pemain
- Kualitas Microphone Pemain

Kesimpulan pengujian yaitu diharapkan suara diucapkan dengan keras, area tidak bising, dan kualitas microphone bagus karena dapat mengurangi noise.

5. Kesimpulan

Setelah melaksanakan pengujian dan analisa pada system "Pembuatan Game Puzzle Gambar Dengan Gerakan Tangan dan Perintah Suara", maka dapat diambil beberapa kesimpulan sebagai berikut:

- Untuk proses deteksi objek cukup baik untuk di area yang terang.
Jarak maksimal yang masih mampu dideteksi adalah 90 cm, selebihnya system tidak dapat mendeteksi adanya tangan.
- Untuk mendapatkan hasil yang optimal dalam pendeteksian tangan, maka posisi tangan harus dalam keadaan tegak dan tidak boleh miring.
- System mampu mengenali obyek tangan dengan baik jika tangan atau obyek berada dalam keadaan diam atau bergerak dalam keadaan lambat.
- Suara pemain lebih baik diucapkan dengan keras.
- Area pemain diusahakan sepi tidak ada noise disekitarnya.
- Microphone berpengaruh melantunkan kata . ini merupakan suatu kejernihan suatu perangkat berpengaruh karena microphone yang bagus sedikitnya noise.
- Kata "Drop" lebih sensitif.

11. Harrington, M, (2006) 'Giving Computers a Voice', <http://msdn.microsoft.com/coding4fun/inthebox/tts-default.aspx>.

DAFTAR PUSTAKA

1. Open computer vision library. [Http://sourceforge.net/project/openclibrary/](http://sourceforge.net/project/openclibrary/).
2. Tom Cassey, Tommy Chheng, Jefery Lien. "Camera Based Hand and Body Driven Interaction with Very Large Display", ucsd lab, June 15, 2007.
3. Neo, Naotoshi. "OpenCV Haar Training". 2007.
4. Sun, Min. "OpenCV and face detection". 2006.
5. Gary Bradski, Adrian Kaehler. "Learning OpenCV". Oreilly team's Intel. 2005.
6. Paul Viola and Michaels J. Jones. "Rapid Object Detection using boosted Cascade of Simple Features". IEEE CVPR, 2001.
7. Robin Hewitt. "seeing with openCV part 2 – Finding Face in image". Jan, 2007.
8. Gary Agam. "introduction to programming with OpenCV". 2006.
9. Ardiansyah Rizkcy. "Kendali Game Tetris Menggunakan Gerakan Tangan", EEPIS-ITS. 2009.
10. Microsoft Speech API 5.1. <http://www.microsoft.com/speech/developers.aspx>.