

SISTEM INTERAKSI MANUSIA DAN ROBOT MENGGUNAKAN DETEKSI WAJAH *REAL TIME* DENGAN METODE *HAAR CASCADE CLASSIFIER* UNTUK *SALES PROMOTION ROBOT*

HUMAN-ROBOT INTERACTION USING REAL TIME FACE DETECTION WITH HAAR CASCADE CLASSIFIER METHOD FOR SALES PROMOTION ROBOT

Budi Permana¹, Angga Rusdinar, Ph.D², Dr.-Ing. Fiky Yosef Suratman³

¹Prodi S1 Teknik Elektro, ³Prodi S2 Teknik Elektro, Fakultas Teknik Elektro, Universitas Telkom
Jl. Telekomunikasi, Dayeuh Kolot Bandung 40257 Indonesia

¹budipermana@students.telkomuniversity.ac.id, ²anggarusdinar@telkomuniversity.ac.id,

³fvsuratman@telkomuniversity.ac.id

Abstrak

Promosi adalah suatu kegiatan pemasaran yang bertujuan untuk mempercepat proses penjualan dan memaksimalkan penjualan produk yang ditawarkan. Didalam sebuah *event* seperti pameran, biasanya terdapat seorang *sales promotion girl/boy*. Tugas dari *sales promotion* adalah membangun interaksi langsung dengan konsumen. Dengan berkembangnya teknologi robotika, praktik promosi juga bisa dilakukan oleh sebuah robot. Untuk membuat robot dapat berinteraksi kepada manusia layaknya seorang *sales promotion*, dibutuhkan sebuah sistem interaksi antara robot dan manusia. Pada tugas akhir ini dibuat sebuah sistem interaksi antara robot dan manusia dengan menggunakan deteksi wajah (*face detection*). Namun pencapaian pada tugas akhir ini masih sebatas interaksi non-verbal, yaitu berupa pergerakan robot menghadap kanan atau menghadap kiri. Sistem ini dibuat dengan teknik pengolahan citra menggunakan *webcam* sebagai *visual sensor*, yang kemudian diimplementasikan pada sebuah robot *line follower* yang telah diprogram dengan algoritma *fuzzy logic*. Hasil dari penelitian ini adalah sebuah sistem deteksi wajah dengan tingkat akurasi sebesar 100% pada jarak $\pm 30\text{cm}$, 83,3% pada jarak $\pm 1\text{m}$, 80% pada jarak $\pm 1.5\text{m}$, 73% pada jarak $\pm 2\text{m}$ dan $\pm 2.5\text{m}$ serta 60% pada jarak $\pm 3\text{m}$. Deteksi wajah pada tugas akhir ini diimplementasikan untuk membuat robot dapat melakukan sebuah interaksi berupa promosi kepada manusia berdasarkan wajah yang terdeteksi.

Kata kunci : *Sales promotion, sales promotion robot, robot vision, face detection, OpenCV*

The promotion is a marketing activities that have the purpose to accelerate the sales process and maximizing sales of products offered. In an event like the exhibit, there is usually a sales promotion girl/boy. The duty of the sales promotion is build direct interaction with consumers. With the development of Robotics technology, promotional practices can also be done by a robot. To make the robot can interact to human as a sales promotion system, it takes a Human Robot Interaction. In this final project, a non-verbal interaction system between human and robot using face detection has been created. The system was created with the technique of image processing using the webcam as visual sensors, which are then implemented on a line follower robot that has been programmed with fuzzy logic algorithm. The result of this final project is a face detection system with 100% accuracy at a distance of approximately 30cm, 83,3% at a distance of approximately 1m, 80% at a distance of approximately 1.5m, 73% at a distance of approximately 2m and 2.5m and also 60% at a distance of approximately 3m. Face detection on this final project is implemented to make the robot can do an interaction in the form of promotion to the human based on detected face.

Keyword: *Sales promotion, sales promotion robot, robot vision, face detection, OpenCV*

1. Pendahuluan

Promosi merupakan salah satu kegiatan pemasaran yang sangat penting manfaatnya bagi sebuah perusahaan dalam penawaran sebuah produk atau jasa yang ditawarkan. Dalam praktik promosi, dibutuhkan adanya suatu bagian yang memiliki interaksi langsung dengan konsumen dengan maksud mempercepat dan memaksimalkan proses penjualan, bagian tersebut adalah *sales promotion*. Seiring dengan berkembangnya teknologi dalam bidang

robotika, praktik promosi tidak hanya bisa dilakukan oleh seorang *sales promotion boy/girl*, namun juga bisa dilakukan sebuah robot. Dengan adanya robot promosi ini, terdapat beberapa keuntungan yang bisa didapatkan, yaitu tidak memerlukan banyak tenaga *sales promotion boy/girl* sehingga *budget* biaya untuk upah para *sales promotion* bisa diperkecil karena ada bantuan dari robot. Salah satu keuntungan lainnya yaitu robot cenderung menarik perhatian. Pada tahun 2011 di Jepang telah dilakukan verifikasi bahwa praktik promosi dengan robot memiliki keuntungan dan keunikan tersendiri serta dapat menarik perhatian pelanggan atau pengunjung sebesar 21,8%[1].

Untuk membuat robot dapat melakukan interaksi kepada manusia layaknya seorang *sales promotion*, dibutuhkan sebuah sistem interaksi berdasarkan pengelihatannya. Robot diharuskan untuk mengetahui terlebih dahulu apakah ada manusia disekitarnya atau tidak. Dari permasalahan yang muncul ini, dibuatlah sistem interaksi manusia dan robot dengan deteksi wajah yang menggunakan metode *haar cascade classifier*. Tujuan dibuatnya sistem interaksi ini adalah, untuk membuat robot dapat melakukan interaksi kepada manusia berdasarkan ada atau tidaknya wajah yang terdeteksi. Dengan begitu, robot dapat melakukan interaksi layaknya seorang *sales promotion girl/boy*. Sistem ini kemudian diimplementasikan pada sebuah robot *line follower* yang telah diprogram dengan algoritma *fuzzy logic*. Pencapaian pada penelitian masih sebatas interaksi non-verbal, yaitu berupa pergerakan robot menghadap kanan atau menghadap kiri.

2. Dasar Teori

2.1 Pengertian Citra

Kata citra atau yang lebih dikenal secara umum dengan kata “gambar” dapat didefinisikan sebagai fungsi dua dimensi (2D) (x, y) dimana x dan y adalah koordinat spasial dan amplitudanya pada setiap koordinasi atau dimensi merupakan tingkat kecerahan dari sebuah citra pada titik tersebut [2]. Citra yang dihasilkan oleh alat-alat optik seperti kamera merupakan citra analog. Sedangkan citra digital merupakan sebuah citra analog (x, y) yang telah didiskritkan, yaitu citra yang mengalami proses sampling dan juga kuantisasi. Masing-masing element dari array disebut *pixel* atau *pel*, berasal dari istilah “*picture element*”. Selain citra analog dan digital, citra lainnya yaitu citra bergerak, merupakan kumpulan dari banyak *frame* citra yang diakuisisi dengan kecepatan yang tinggi secara terus menerus dalam rentang waktu yang diinginkan. Kecepatan akuisisi citra secara terus menerus akan menghasilkan citra sekuensial yang kemudian akan membuat citra tampak seperti bergerak. Citra bergerak (video) memiliki sebuah satuan yaitu *Frame per Second (FPS)*.

2.2 Ruang Warna

Ruang warna RGB merupakan tiga warna pokok yang memiliki panjang spektrum paling dominan, yaitu merah (*red*), hijau (*green*), dan biru (*blue*). Masing-masing warna pokok ini memiliki intensitas dengan jumlah 8-bit (0-255). Sehingga ruang warna RGB sendiri memiliki intensitas dengan jumlah 24-bit. Berbeda dengan citra RGB, *grayscale* merupakan jenis citra yang setiap piksel dalam arraynya hanya terdiri dari satu komponen warna saja, oleh karena itu citra *grayscale* memiliki intensitas sebesar 8-bit (0-255), interval tersebut menunjukkan perbedaan warna keabuan pada setiap interval, semakin mendekati 0 maka citra akan cenderung berwarna hitam, ketika mendekati nilai 255 maka citra akan cenderung berwarna putih. Citra RGB dikonversikan menjadi citra *grayscale* dengan cara merata-ratakan ketiga nilai R, G dan B yang dimiliki, seperti yang ditunjukkan pada persamaan 2.1[3]:

$$I = (0.299)R + (0.587)G + (0.114)B \quad (1)$$

2.3 Wajah Manusia

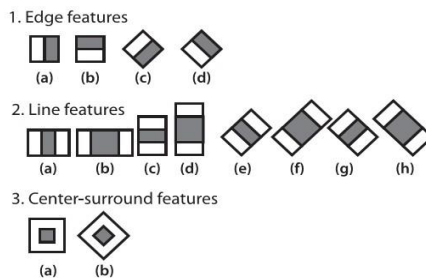
Wajah pada manusia memiliki peranan yang sangat penting, karena wajah merupakan ciri khusus yang dimiliki setiap manusia. Dengan adanya wajah, manusia juga bisa saling mengenal dan mengingat satu sama lain. Di dalam hubungan sosial dan komunikasi langsung, wajah mempunyai peranan yang sangat penting, karena wajahlah yang membuat manusia bisa dibedakan dari makhluk hidup lainnya. Wajah juga memiliki peranan yang sangat penting bagi hubungan interaksi antara manusia dan robot, karena wajah dapat digunakan oleh robot untuk membedakan manusia dengan objek lain[4].

2.4 Face detection

Adalah salah satu bidang kajian dalam ilmu *computer vision* yang bertujuan untuk mendeteksi ada atau tidaknya wajah didalam sebuah cakupan *field of view* dari webcam. *Face detection* merupakan langkah awal dari pengenalan sebuah wajah yaitu *face recognition*. Pengolahan citra wajah belakangan telah banyak dijadikan studi, salah satu tujuannya adalah dengan maksud menawarkan service kepada manusia[5]. Pengolahan citra wajah manusia kemudian diterapkan pada robot untuk membangun sebuah sistem interaksi manusia dan robot. Metode *face detection* yang digunakan pada tugas akhir ini adalah *haar cascade classifier*. Metode tersebut terdiri dari fitur *haar*, *integral image*, algoritma *adaboost*, dan *cascade classifier*[3].

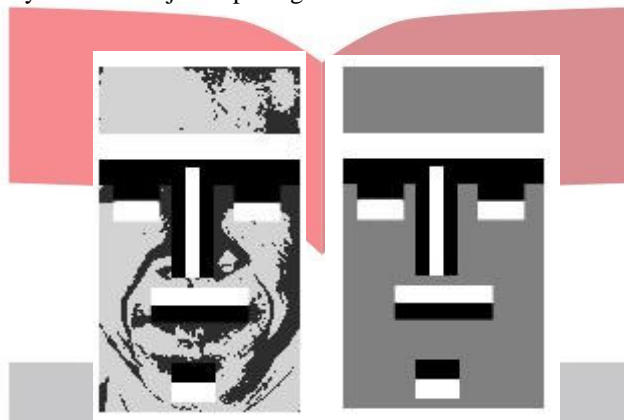
2.4.1 Fitur Haar

Fitur ini merujuk kepada sebuah fungsi matematika yang bernama *haar wavelet*



Gambar 1. Bentuk Fitur Haar[3]

Jika fitur-fitur tersebut digabungkan, sebuah fitur yang menyerupai wajah manusia akan terbentuk. Ilustrasinya akan ditunjukkan pada gambar dibawah ini



Gambar 2. Pencarian Wajah Dengan *Haar Features*[6]

2.4.2 Integral Image

Pada metode *haar cascade classifier*, *integral image* digunakan untuk mempercepat perhitungan fitur masukan *Haar-like*. Untuk mempermudah proses *integral image*, citra yang akan diolah terlebih dahulu diubah kedalam bentuk grayscale agar citra hanya memiliki nilai 0-255(8-bit), dengan 0 menyatakan intensitas citra paling gelap (hitam) dan 255 menyatakan intensitas citra paling terang (putih). Cara menghitung masukan fitur *haar* yaitu dengan mengurangi nilai piksel area hitam dengan nilai piksel area putih. Berikut persamaan dari *integral image*[7]:

$$I(x, y) = \sum_{x' \leq x} \sum_{y' \leq y} I(x', y') \tag{2}$$

Dimana $I(x, y)$ merupakan *integral image* dari citra asli yaitu $I(x, y)$. Dengan menggunakan pasangan persamaan seperti dibawah ini[7]:

$$I(x, y) = I(x, y-1) + I(x, y) \tag{3}$$

$$I(x, y) = I(x-1, y) + I(x, y) \tag{4}$$

Dimana $I(x, y)$ adalah penjumlahan kumulatif baris piksel.

2.4.3 Algoritma Adaboost

Algoritma ini digunakan untuk menentukan ada atau tidaknya fitur *haar* didalam sebuah *frame* citra yang diolah. Bila ternyata didalam sebuah *frame* terdapat sebuah fitur haar, maka fitur haar yang terdeteksi akan segera dihitung nilainya dengan integral image.

2.4.4 Cascade Classifier

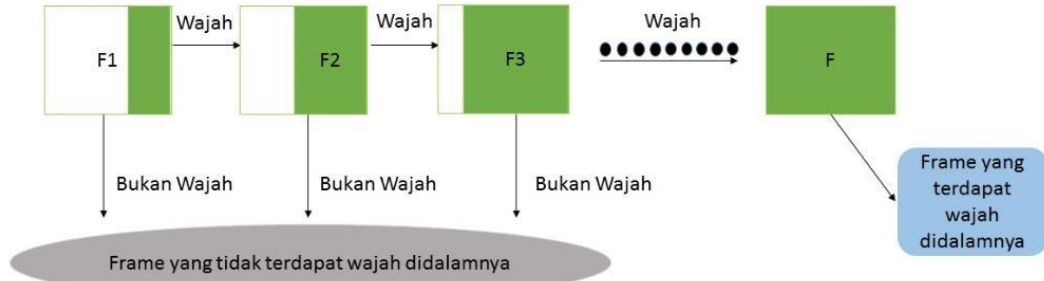
Sifat classifier dan *learning* dari satu fitur *haar* sangatlah lemah, sehingga hasil yang didapatkan juga kurang akurat. *Cascade classifier* adalah sebuah cara untuk mengkombinasikan hasil dari banyak *stage weak classifier* sehingga bisa menjadi sebuah *classifier* yang kuat. Berikut persamaan yang didapatkan[3]:

$$F = F_1 + F_2 + \dots + F_n \tag{5}$$

α (*strong classifier*) merupakan kombinasi dari α (*weak classifier*). Dan α merupakan nilai bobot awal. Dan α memiliki persamaan sebagai berikut:

$$\alpha = \begin{cases} +1, & \alpha \geq \alpha \\ -1, & \alpha < \alpha \end{cases} \quad (6)$$

α merupakan nilai *threshold* pada *stage* ke- n dengan α adalah nilai fitur *haar* pada *stage* ke- n . Nilai *threshold* akan menjadi nilai pembatas untuk mengindikasikan ada atau tidaknya wajah pada *stage* ke- n .



Gambar 3. Bentuk cascade classifier

2.5 Komunikasi Data

Komunikasi data yang digunakan dalam pembuatan sistem interaksi ini adalah komunikasi serial. Komunikasi serial merupakan komunikasi yang bersifat *full duplex* karena memiliki dua jalur berbeda dalam pengiriman dan penerimaan data, satu jalur digunakan untuk media pengiriman (TX) dan satu jalur untuk media penerimaan (RX). Komunikasi serial digunakan untuk pengiriman karakter serial yang memberikan informasi letak wajah.

2.6 Arduino UNO

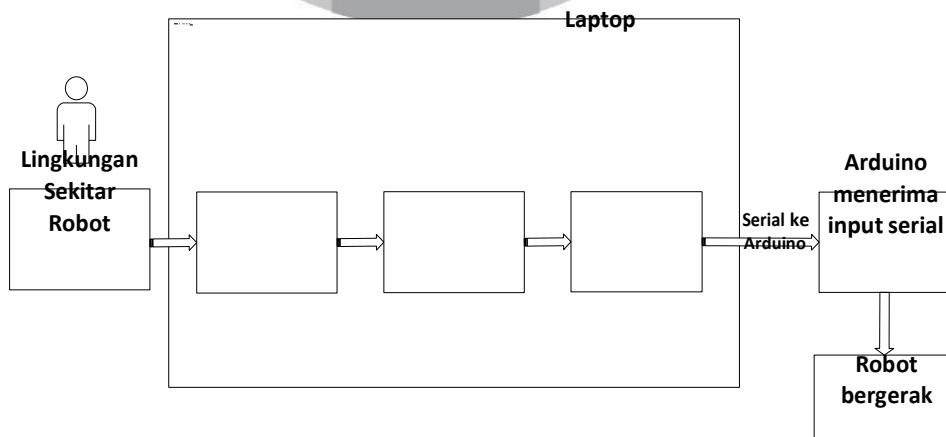
Arduino Uno merupakan sebuah *board* yang berbasis mikrokontroler ATmega328. Arduino memiliki 14 digital I/O, 6 diantaranya dapat digunakan sebagai pin output PWM (*Pulse Width Modulation*). Selain itu juga terdapat 6 pin analog, osilator kristal 16 MHz, sebuah koneksi ke port usb, sebuah soket power, ICSP header, dan sebuah tombol reset. Arduino sama seperti sistem minimum mikrokontroler biasa yang didalamnya terdapat komponen-komponen pendukung yang diperlukan agar mikrokontroler dapat bekerja dengan baik.

2.7 Motor DC

Jenis aktuator yang digunakan robot *line follower* pada pembuatan sistem interaksi ini adalah motor dc. Pada sistem interaksi ini, motor dc berfungsi untuk menggerakkan robot agar dapat bergerak menghadap kanan atau kiri tergantung posisi wajah.

3. Perancangan

Sistem yang akan dibangun pada tugas akhir ini adalah sistem interaksi antara robot dan manusia. Interaksi yang dimaksud interaksi berupa gerakan robot yang akan menghadap ke kanan atau ke kiri tergantung kepada posisi wajah yang terdeteksi. Interaksi dibuat menggunakan deteksi wajah dengan metode *haar cascade classifier*. Karakter serial akan dikirimkan berdasarkan posisi wajah yang berbeda-beda tersebut dan dikirimkan dari laptop ke Arduino via komunikasi serial.



Gambar 4. Diagram Blok

3.1 Mekanik Robot

Robot yang digunakan pada penelitian ini dibuat dari bahan alumunium dengan tebal 2 milimeter. Dengan lebar rangka tanpa roda yaitu 30 cm, tinggi 127 cm, dan lebar rangka dengan roda terpasang yaitu 169 cm.



Gambar 5. Rangka Robot yang Digunakan

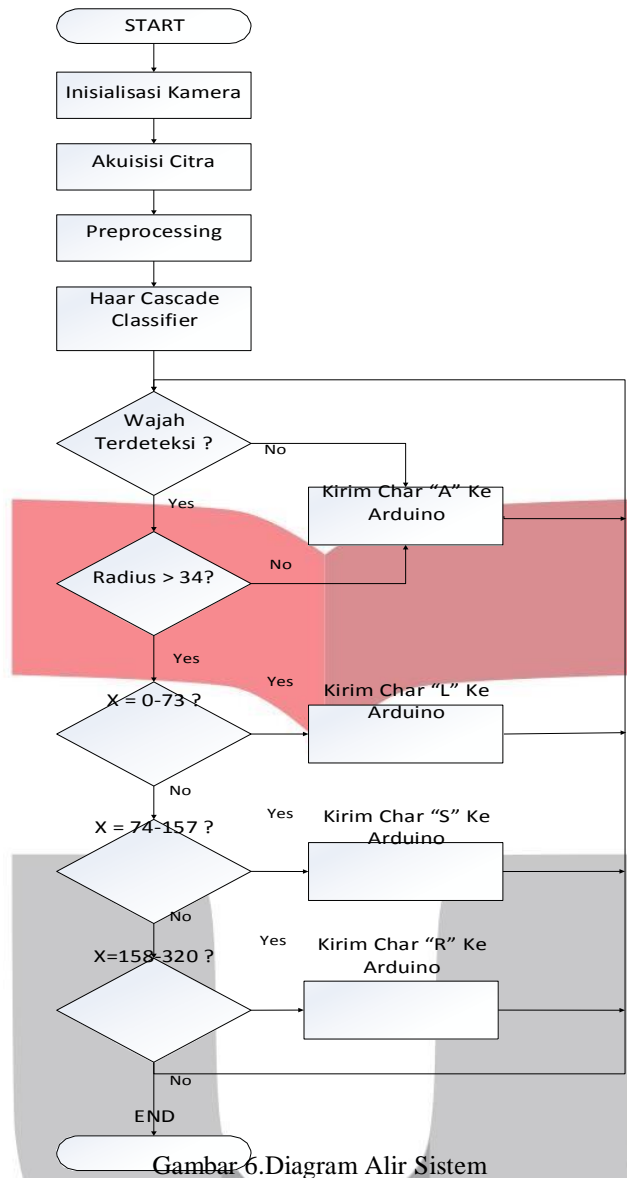
3.2 Pembagian Koordinat Sistem

Interaksi non-verbal yang mampu dilakukan robot pada tugas akhir ini hanyalah sebuah gerakan ke kanan atau ke kiri tergantung dari letak koordinat wajah yang terdeteksi didalam *frame*. Resolusi *frame* yang digunakan yaitu 320×240 . 320 mewakili nilai untuk X, dan 240 mewakili Y. Nilai yang dijadikan dasar untuk mengetahui lokasi wajah yang terdeteksi adalah nilai X.

Untuk menentukan dimana posisi wajah berada maka pada algoritma dibuatlah pembagian tiga wilayah dari nilai X tersebut. Tiga wilayah yang dimaksud adalah kanan, kiri dan tengah. X bernilai 0-73 untuk menyatakan wajah berada disebelah kiri dari robot. X bernilai 74-157 untuk menyatakan wajah wajah berada di tengah, dan X bernilai 158-320 untuk menyatakan wajah berada disebelah kanan dari robot. Penentuan nilai-nilai X tersebut didapatkan melalui metode *trial and error*.

Sistem interaksi yang dibuat ini tidak hanya bisa melakukan interaksi pada satu wajah saja. Jika terdeteksi lebih dari satu wajah, maka nilai X yang merepresentasikan wajah juga ada lebih dari satu, dan muncul sesuai dengan jumlah wajah yang terdeteksi. Jika wajah yang terdeteksi lebih dari satu, maka akan dilakukan perhitungan rata-rata sebagai berikut:

$$\text{Rata-rata} = \frac{x_1 + x_2 + \dots + x_n}{n} \quad (7)$$



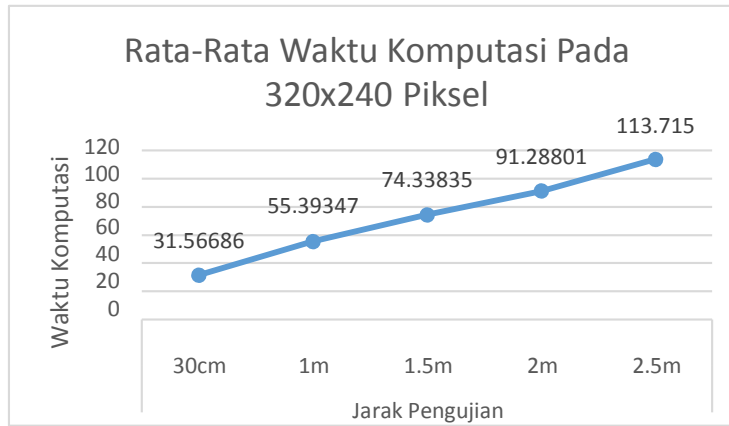
4. Pengujian dan Analisis Sistem

4.1. Waktu Komputasi Deteksi Wajah

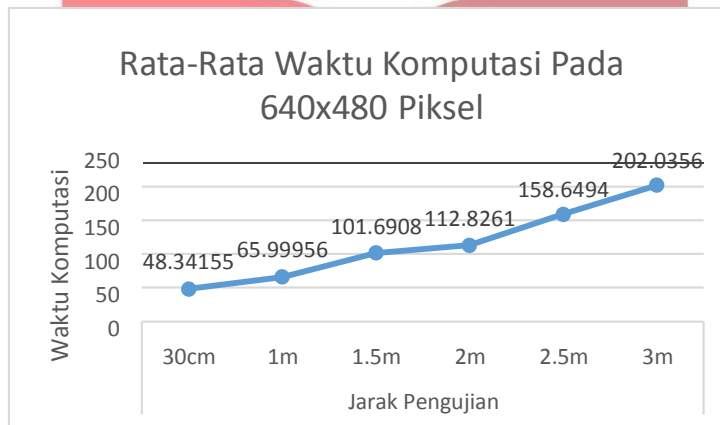
Pengujian ini dilakukan dengan menggunakan dua resolusi *frame* yang berbeda, yaitu 320x240 piksel dan 640x480 piksel. Tujuannya dilakukannya perbandingan dua resolusi ini adalah untuk memilih resolusi *frame* mana yang memiliki waktu komputasi yang lebih cepat. Semakin cepat waktu komputasi, maka semakin *up to date* pula informasi dari pengolahan citra yang citra yang didapatkan.

Pengujian ini dilakukan dengan jarak yang berbeda-beda dengan pencahayaan yang cukup, jarak pengujiannya yaitu ±30cm, ±1m, ±1,5m, ±2m, ±2,5cm dan ±3m. Waktu komputasi rata-rata dapat dihitung dengan rumus seperti berikut:

$$R = \frac{\sum_{i=1}^n t_i}{n} \tag{8}$$



Gambar 7. Grafik Waktu Komputasi pada 320x240 piksel



Gambar 8. Grafik Waktu Komputasi pada 640x480 piksel

Dari dua grafik diatas, dapat disimpulkan bahwa semakin jauh jarak pengujian maka waktu komputasi akan semakin lama. Dan resolusi 320x240 memiliki waktu komputasi yang lebih cepat jika dibandingkan dengan resolusi 640x480.

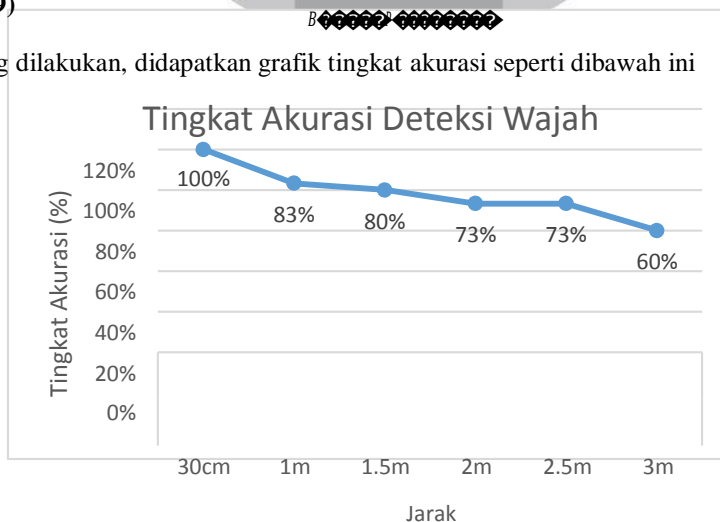
4..2 Akurasi Deteksi Wajah

Pengujian ini dilakukan untuk mengetahui akurasi deteksi wajah saat jarak wajah berbeda-beda. Jarak yang digunakan untuk pengujian akurasi sama dengan pengujian waktu komputasi. Dari pengujian yang dilakukan, akan dicari akurasi dari sistem deteksi wajah dengan persamaan sebagai berikut

$$A = \frac{B}{C} \times 100\%$$

(9)

Dari pengujian yang dilakukan, didapatkan grafik tingkat akurasi seperti dibawah ini



Gambar 9. Grafik Tingkat Akurasi Deteksi Wajah

5. Kesimpulan

Setelah melakukan pengujian dan analisis terhadap interaksi manusia dan robot menggunakan deteksi wajah *real time* dengan metode *haar cascade classifier* untuk *sales promotion robot*, didapatkan kesimpulan seperti berikut :

1. Dari data-data yang didapatkan pada pengujian waktu komputasi, dapat diambil kesimpulan bahwa waktu komputasi resolusi 640×480 lebih cepat jika dibandingkan dengan waktu komputasi pada resolusi 320×240 piksel. Hal inilah yang menyebabkan dipilihnya resolusi 640×480 piksel untuk diimplementasikan ke robot untuk informasi yang *up to date*.
2. Dan dari data-data yang didapatkan pada pengujian waktu komputasi pula, dapat diambil kesimpulan bahwa waktu komputasi resolusi pada kedua *frame* akan semakin lama seiring dengan semakin jauhnya jarak pengujian. Untuk resolusi 320×240 piksel pada jarak ± 30 cm, waktu komputasinya sebesar $31.56686ms$, $55.39347ms$ pada jarak $\pm 1m$, $74.33835ms$ pada jarak $\pm 1.5m$, dan akan terus meningkat seiring dengan semakin jauhnya jarak pengujian. Dan hal yang sama juga berlaku pada resolusi 640×480 piksel.
3. Untuk menganalisis akurasi sistem deteksi wajah, dilakukan 30 kali percobaan pada masing-masing jarak pengujian. Resolusi *frame* yang digunakan pada pengujian akurasi adalah 640×480 piksel. Pada jarak ± 30 cm, akurasi sistem deteksi wajah mencapai 100%, 83% pada jarak $\pm 1m$, 80% pada jarak $\pm 1.5m$, 73% pada jarak $\pm 2m$ dan $2.5m$, serta 60% pada jarak $\pm 3m$. Nilai-nilai tersebut menunjukkan penurunan akurasi seiring dengan menjauhnya jarak pengujian karena semakin banyaknya *false positive* seiring jauhnya jarak pengujian.
4. Dalam pembuatan sistem interaksi manusia dan robot, robot hanya akan melakukan interaksi dengan manusia yang wajahnya terdeteksi pada jarak ± 55 cm. Jarak tersebut masih terdapat didalam rentang ± 30 cm dan $\pm 1m$, dimana pada jarak tersebut menunjukkan akurasi yang baik yaitu 100% dan 83%.

Daftar Pustaka :

- [1] Murakawa, Y. dkk. (2011). "Verification of the Effectiveness of Robots for Sales Promotion in Commercial Facilities". *IEEE/SICE International Symposium on System Integration*.
- [2] Gonzalez, C. Rafael. dan Woods, E. Richard. (2008). *Digital Image Processing*, 3rd Ed. USA: Prentice Hall, New Jersey.
- [3] Bradski, Garry dan Adrian Kaehler. (2008). *Learning OpenCV*. O'reilly.
- [4] Pan, Yaozhang. Dkk . (2009). "Real-time Face Detection for Human Robot Interaction". *The 18th IEEE International Symposium on Robot and Human Interactive Communication*.
- [5].Fahn, Chin-Shryung., dan Lin, Yu-Ta. (2010). "Real-time face tracking techniques used for the interaction between humans and robots". *The 2th IEEE Conference on Industrial Electronics and Applications*.
- [6]. <http://eyalarubas.com/face-detection-and-recognition.html> (Diakses pada tanggal 30 Mei 2015)
- [7]. Viola, P. & Jones, M.J. (2001). "Robust Real-Time Object Detection".