

RANCANG BANGUN SISTEM PENGENALAN SIMBOL TOILET DI DALAM RUANGAN DENGAN METODE SURF UNTUK TUNA NETRA BERBASIS KAMERA

Tonny Feriandi, Ronny Mardiyanto, Djoko Purwanto
Teknik Elektro, Fakultas Teknologi Industri, Institut Teknologi Sepuluh Nopember (ITS)
Jl. Arief Rahman Hakim, Surabaya 60111
E-mail: ronny@ee.its.ac.id

Abstrak— Penginderaan visual merupakan suatu proses pengolahan data visual menjadi data referensi yang dijadikan dasar untuk menggambarkan objek visual. Proses tersebut dapat digunakan untuk membantu penyandang tuna netra mengenali ruang toilet. Pada dasarnya penyandang tuna netra tidak mampu mengenali lingkungan sekitarnya, sehingga dengan bantuan pengenalan simbol toilet maka penyandang tuna netra dapat mengetahui jika di sekitarnya terdapat ruang toilet.

Dalam tugas akhir ini, dirancang sistem pengenalan simbol toilet untuk tuna netra di dalam ruangan berbasis kamera untuk mempermudah penyandang tuna netra mengenali ruang toilet. Kamera diletakkan pada tas bagian depan pengguna dengan sudut pengambilan gambar tegak lurus kedepan, ini bertujuan untuk mempermudah dalam proses pengambilan data visual. Kemudian data yang didapat diolah dengan beberapa metode pengolahan citra seperti grayscale dan SURF. Hasil pemrosesan data citra akan dikomunikasikan dalam bentuk suara melalui earphone pengguna.

Pengujian dilakukan dengan memasang alat pada pengguna dan pada waktu yang berbeda yakni pada jam 08.00 pagi dan jam 08.00 malam dan dilakukan dengan 5 cara yakni yang pertama pengujian pengenalan simbol dengan membandingkan keypoint citra yang ditangkap kamera dengan keypoint citra referensinya sendiri, pengujian di tempat umum, pengujian dengan menggabungkan seluruh simbol sebagai citra referensi, pengujian kecepatan pengguna menuju toilet ketika menggunakan sistem, dan pengujian jika terdapat gangguan pada simbol. Sistem ini memiliki error rata-rata sebesar 20% dan akurasi rata-rata sebesar 80%. Error pengenalan simbol disebabkan karena jarak tangkap efektif citra kamera yang digunakan pada sistem ini hanya sebesar 1 meter hingga 3 meter.

Kata kunci: simbol toilet, *grayscale*, *SURF*, kamera, *keypoint*

I. PENDAHULUAN

Penyandang tuna netra memiliki banyak keterbatasan dalam menjalankan kehidupan sehari-hari. Keterbatasan itu didasari dari ketidakmampuan mereka dalam mengenali lingkungan secara baik. Misalnya mengenali ruang toilet. Toilet pada ruangan umum memiliki karakteristik tersendiri, salah satunya simbol toilet.

Untuk itu, dalam tugas akhir ini diimplementasikan rancang bangun sistem pengenalan simbol toilet untuk tuna netra, dimana alat ini akan diletakkan pada tubuh pengguna. Kemudian dengan menggunakan teknik pengolahan citra (*Image Processing*) yang dilakukan oleh *Processing Unit* akan mempermudah pengenalan simbol terjadi, hasil pengenalan simbol akan dikomunikasikan dalam bentuk suara yang sesuai dengan simbol yang telah dikenali. Sehingga penyandang tuna

netra dapat secara cepat memenuhi kebutuhan dasarnya dan mengurangi ketergantungan orang lain.

II. TEORI PENUNJANG

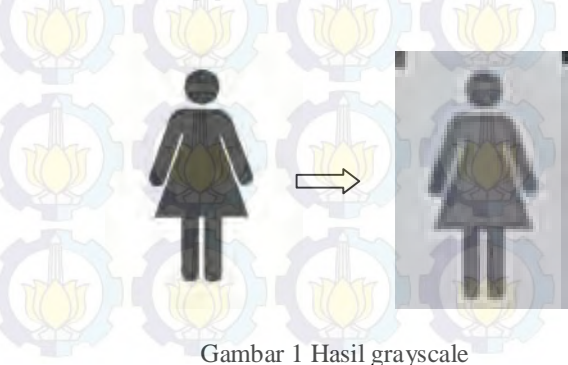
A. Pengolahan Citra (Image Processing)

Pengolahan citra adalah proses olah citra untuk mendapatkan informasi-informasi penting dari suatu objek yang nantinya akan diproses lebih lanjut sesuai dengan tujuan yang diinginkan. Pada bidang industri pengolahan citra dapat digunakan sebagai pengganti sensor, misalnya untuk menghitung jumlah botol dan memilah botol. Dengan menggunakan pengolahan citra dapat menggantikan beberapa sensor, sehingga lebih efisien, efektif dan hemat biaya.

Pada bidang militer pengolahan citra dapat digunakan untuk mendeteksi keberadaan musuh, sehingga dapat mengurangi kontak langsung dan korban jiwa.

Salah satu metode pengolahan citra yang digunakan pada tugas akhir ini antara lain adalah *grayscale*. *Grayscale* adalah citra abu, bervariasi pada warna hitam pada bagian yang intensitas terlemah dan warna putih pada intensitas terkuat. Citra *grayscale* berbeda dengan citra "hitam-putih", dimana pada konteks komputer, citra hitam putih hanya terdiri atas 2 warna saja yaitu "hitam" dan "putih" saja. Pada citra *grayscale* warna bervariasi antara hitam dan putih, tetapi variasi warna diantaranya sangat banyak.

Citra *grayscale* disimpan dalam format 8 bit untuk setiap sample piksel, yang memungkinkan sebanyak 256 intensitas. Format ini sangat membantu dalam pemrograman karena manipulasi bit yang tidak terlalu banyak. Untuk mengubah citra berwarna yang mempunyai nilai matrik masing-masing R, G dan B menjadi citra *grayscale* dengan nilai X, maka konversi dapat dilakukan dengan mengambil rata-rata dari nilai R, G dan B. [1]



Gambar 1 Hasil grayscale

B. Pengenalan Simbol

Pengenalan simbol merupakan proses utama pada tugas akhir ini. Banyak metode yang dapat digunakan untuk mengenali simbol, misalnya LBP (Local Binary Pattern), PCA (Principal Component Analysis), Chi Square dan SURF(Speeded-Up Robust Feature). Semua metode tersebut memiliki kelebihan dan kekurangan masing-masing, pada tugas akhir ini digunakan metode SURF dengan alasan metode ini sangat akurat walaupun objek yang dideteksi mengalami perubahan skala dan kemiringan namun sedikit kekurangannya adalah metode ini agak terasa lambat saat memproses. Proses pengenalan SURF menggunakan titik-titik karakteristik suatu citra yang biasa disebut *keypoint*, tiap *keypoint* memiliki *scale space* dan nilai deskriptor. Nilai vektor deskriptor inilah yang digunakan untuk mengenali citra dengan membandingkan nilai vektor citra referensi dengan nilai vektor citra yang didapat dari kamera. Untuk lebih jelasnya perhatikan gambar 2 berikut ini.



Gambar 2 Hasil pemrosesan SURF

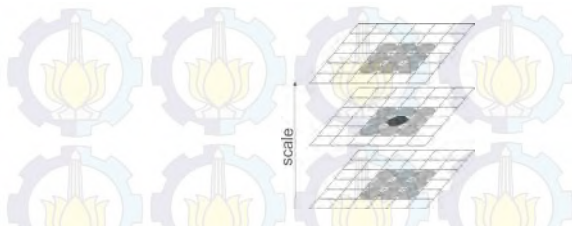
❖ SURF

Sebelum dapat mengenali suatu citra, metode SURF memiliki beberapa tahapan yang pertama adalah lokalisasi *keypoint*, deskriptor *keypoint*, dan proses *matching*. Lebih jelasnya berikut penjelasannya.

1. Lokalisasi *Keypoint*

Keypoint adalah titik-titik yang dihasilkan dari metode *surf*. Pada metode ini *keypoint* digunakan sebagai karakteristik suatu citra atau gambar. Sehingga setiap gambar atau citra memiliki jumlah *keypoint* yang berbeda. Lokalisasi *keypoint* dilakukan dengan beberapa proses. Proses pertama menentukan *threshold* untuk *keypoint*.

Proses berikutnya *non-maxima suppression*, proses ini merupakan proses mencari calon *keypoint* dengan membandingkan tiap-tiap piksel gambar pada *scale space* dengan 26 tetangga yang terdiri dari 8 titik di *scale* asli dan 9 titik di *scale* atas bawahnya. Proses inilah yang menghasilkan *keypoint*, gambar 3 menunjukkan *non-maxima suppression*



Gambar 3 *non-maxima suppression*

Proses terakhir adalah mencari lokasi *keypoint* menggunakan interpolasi data yang dekat dengan *keypoint* hasil proses sebelumnya. Untuk mencari nilainya dapat dilakukan dengan determinan *Hessian* $H(x,y,\sigma)$ seperti berikut ini[2]:

$$H(x) = H + \frac{\partial H^T}{\partial x} x + \frac{1}{2} x^T \frac{\partial^2 H}{\partial x^2} x$$

Lokasi ekstrim yang diinterpolasi ditemukan dari turunan fungsi diatas, sehingga menjadi :

$$x = - \frac{\partial^2 H}{\partial x^2} \frac{\partial H}{\partial x}$$

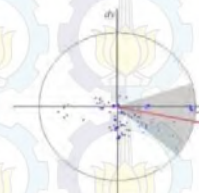
2. Deskriptor *Keypoint*

Deskriptor merupakan daerah piksel disekitar *keypoint* yang dihasilkan. Deskriptor menggambarkan sebaran intensitas piksel tetangga di sekitar *keypoint*. Sehingga Deskriptor sendiri dapat dikatakan sebagai ciri dalam sebuah citra karena tiap citra memiliki sebaran intensitas piksel yang berbeda.

Deskriptor *keypoint* terdiri dari dua proses perhitungan, yang pertama adalah Pembuatan Orientasi. Pembuatan orientasi dilakukan dengan menghitung respon *Haar Wavelet* arah x dan y pada daerah piksel tetangga *keypoint* yang berjarak 6 *scale*. Orientasi dominan dihitung menggunakan jumlah semua respon pada pergeseran orientasi dengan ukuran $\pi/3$ seperti gambar 4, kemudian respon horisontal dan vertikal dijumlahkan sehingga menghasilkan sebuah vektor orientasi lokal. Vektor terpanjang inilah merupakan orientasi *keypoint*.

Proses kedua adalah berdasarkan jumlah respon *Haar Wavelet*, pada proses ini dilakukan dengan membuat daerah kotak disekitar *keypoint*, dimana titik pusatnya *keypoint* itu sendiri. Kemudian, respon *wavelet* dx dan dy dijumlahkan pada setiap *sub-region*. Maka akan didapatkan nilai absolut dari respon *wavelet*. Masing-masing *sub-region* memiliki 4 dimensi deskriptor sehingga vektornya berjumlah 64. Maka didapat[2] :

$$v = \left(\sum d_x, \sum d_y, \sum |d_x|, \sum |d_y| \right)$$



Gambar 4 Deskriptor *Keypoint*

3. Proses *Matching*

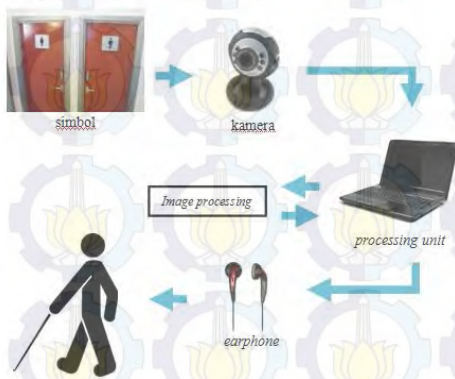
Pada sistem ini proses *matching* dilakukan menggunakan teknik *FLANN* (*Fast Library Approximated Nearest Neighbor*). Fitur *SURF* ini terdiri dari *keypoint* dan deskriptor yang berupa vektor. Pada intinya sistem *matching FLANN* adalah dengan mencari data dari vektor *keypoint* dan deskriptor pada citra yang didapat dari kamera lalu membandingkannya dengan data vektor *keypoint* dan deskriptor pada citra referensi, sehingga nilai vektor pada citra yang didapat dari kamera yang mendekati nilai vektor pada citra referensi akan dianggap cocok atau sama. Pada sistem tugas akhir ini jika nilai vektor dari citra kamera kurang dari 0,6 dari citra referensi maka dianggap sama.

C. Perintah Suara

Proses ini merupakan tahap akhir dalam sistem tugas akhir ini. Perintah suara pada sistem ini digunakan untuk mempermudah mengkomunikasikan hasil proses pengenalan simbol kepada pengguna. Keluaran suara dilakukan dengan menggunakan earphone satu sisi, sehingga suara yang dikeluarkan tidak mengganggu sekitar dan pengguna masih bisa mendengar lingkungan sekitarnya dengan telinga lainnya.

III. PERANCANGAN SISTEM

Pada tugas akhir ini, sistem dilakukan di sebuah *processing unit* berupa *notebook*. Input data citra didapat dari sensor kamera berupa *webcam*. Sensor kamera diletakkan pada tali tas pengguna bagian depan. Hal ini dilakukan untuk mempermudah pengambilan data citra yang ada didepan pengguna. Kemudian dilakukan pengolahan citra gambar untuk mengenali ada tidaknya simbol di depan pengguna. Ketika simbol ditemukan maka akan dikomunikasikan dalam bentuk suara. Ilustrasi cara kerja sistem ditunjukkan pada gambar 5.



Gambar 5 Ilustrasi Cara Kerja Sistem

Berikut ini merupakan cara kerja keseluruhan sistem :

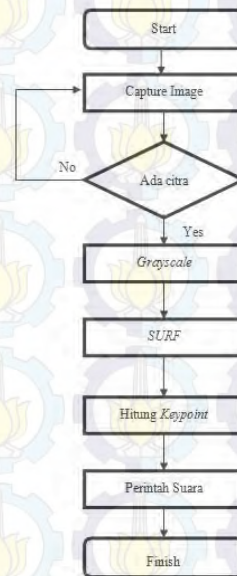
- Memastikan *processing unit* telah aktif memproses.
- Kamera menangkap citra yang berada di depan pengguna, kemudian data citra diolah oleh *processing unit* dengan metode-metode pengolahan citra.
- Metode pengolahan citra antara lain dilakukan proses *SURF* (*Speeded-Up Robust Feature*) untuk mengenali

simbol ada tidaknya simbol. Apabila simbol dikenali maka akan muncul banyak *keypoint* pada citra input.

- Memberikan nilai pada data *SURF* untuk memberikan batas kesamaan antara data dari input citra dengan data dari citra simbol.
- Pengenalan simbol dilakukan dengan cara membandingkan jumlah *keypoint* pada input citra dengan data dari citra simbol.
- Memberikan batasan jumlah *keypoint* yang terdapat pada input citra untuk referensi kapan perintah suara bekerja.
- Apabila jumlah *keypoint* memenuhi keberadaan simbol, maka dikomunikasikan melalui perintah suara.

A. Perancangan Software

Tahapan ini terdiri dari beberapa tahapan yakni, pengambilan citra, *grayscale*, *surf*, membandingkan jumlah *keypoint*, perintah suara. Setiap tahapan memiliki fungsi masing-masing yang saling mendukung serta memiliki tahapan yang berurutan. Lebih jelasnya perhatikan *flowchart* berikut.



Gambar 6 *Flowchart Image processing* pada Sistem

B. Perancangan Hardware

Perancangan *hardware* pada sistem ini terdiri dari *web camera*, *processing unit* dan *earphone*. Lebih jelasnya perhatikan gambar berikut.



Gambar 7 Ilustrasi Penggunaan Alat

Kamera dan *Earphone* dihubungkan pada *notebook* dan dipasangkan kepada pengguna. Kamera menangkap citra, citra di proses oleh *notebook* dengan bantuan *image processing* setelah itu hasil yang didapat dikomunikasikan kepada pengguna melalui *earphone*.

IV. PENGUJIAN (TIDAK DIPUBLIKASIKAN)

V. KESIMPULAN

Dari hasil pengujian dan analisis tugas akhir ini dapat diambil beberapa kesimpulan yakni :

1. Proses ekstraksi fitur SURF pada suatu citra tidak terpengaruh cahaya.
2. Metode SURF membutuhkan waktu ekstraksi sebesar 2 hingga 3 detik.
3. Kemampuan efektif kamera menangkap citra pada sistem ini berjarak 1 hingga 3 meter.
4. Dari hasil pengujian pada sistem tugas akhir ini memiliki error sebesar 25.20% pada pengujian simbol dengan simbol referensinya sendiri dan error ditempat umum sebesar 22.22% sedangkan untuk pengujian citra referensi gabungan memiliki error yang paling rendah yakni 12.84%. Sehingga sistem ini memiliki error rata-rata 20% dan akurasi sebesar 80%

DAFTAR PUSTAKA

- [1] Achmad Rizal, "*Pengolahan Citra*", <URL: <http://achmadrizal.staff.telkomuniversity.ac.id/2014/06/19/pengolahan-citra/>>, Juni 2014.
- [2] Marius Muja and David G. Lowe. "*Fast approximate nearest neighbors with automatic algorithm configuration*". In International Conference on Computer Vision Theory and Applications (VISSAPP'09), pages 331–340, 2009