

Tracking of Pedestrian

Wisnu Bagus P¹⁾, Ir. Dadet Pramadihanto, MEng. Ph.D.²⁾, Setiawardhana, ST.³⁾

¹⁾Jurusan Teknik Informatika, PENS – ITS Surabaya
Kampus ITS Sukolilo, Surabaya 60111

E-mail : wisnubagus@ymail.com

Abstrak

Pada proyek akhir ini, dibuat suatu sistem yang memiliki kemampuan mencari atau mendeteksi adanya pejalan kaki lalu melacaknya (tracking). Sistem ini merupakan sebuah aplikasi dari ilmu *Computer Vision*. Metode-metode yang digunakan adalah metode-metode yang ada dalam bidang ilmu *Image processing*.

Sistem ini memanfaatkan sistem kerja kamera statis yang sering digunakan untuk mengawasi daerah tertentu dengan batasan cahaya normal, stabil dan background yang tidak terlalu kompleks sehingga sistem mampu melakukan *subtraction* dengan tepat.

Metode yang kami gunakan meliputi background subtraction, thresholding, thinning, dan pengenalan objek melalui fitur bentuk kaki.

Seluruh proses sistem berpusat pada sebuah PC yang bertugas mengolah citra dari kamera yang telah diintegrasikan dan mengambil keputusan apakah objek yang terdeteksi merupakan pejalan kaki.

Kata Kunci : *Tracking, detection, pedestrian, image processing, background subtraction*

1. Pendahuluan

Kemampuan untuk mendeteksi pejalan kaki sangatlah menarik dan bermanfaat dalam kehidupan sehari-hari, seperti contoh dalam bidang keamanan, deteksi ini dapat diaplikasikan pada kamera CCTV untuk mengawasi suatu ruang atau daerah tertentu, atau dalam bidang otomotif sebagai driver-otomatis pada sistem bantuan dalam kendaraan.

Pada saat yang sama pula, pejalan kaki merupakan salah satu objek yang paling menantang untuk dilakukan deteksi, banyaknya variabilitas dalam segi tampilan di lokal atau global yang disebabkan oleh berbagai jenis dan gaya pakaian, sehingga hanya beberapa daerah lokal yang benar-benar khas untuk semua kategori. Selain itu secara global, macam dari aksesoris yang digunakan sangat variatif seperti tas, penghangat leher, topi dll yang dapat mengganggu pembentukan siluet bayangan. Perihal di atas menambah kesulitan pada proses deteksi dan akhirnya banyak orang yang diidentifikasi sebagai gambar yang sama pada satu daerah deteksi.

Kemampuan untuk dapat melakukan deteksi pejalan kaki pada tempat yang ramai sangatlah bermanfaat karena hal ini sering terjadi dalam kehidupan nyata. Pada proyek ini diharapkan mampu melakukan tracking pejalan kaki di tempat yang ramai, tujuannya adalah untuk mendeteksi semua

pejalan kaki, pelokalan pada tiap pejalan kaki dan mengetahui arah gerak pejalan kaki tersebut

2. Landasan Teori

Background Subtraction

Proses ini akan melakukan pemisahan antara background dengan foreground dengan menghitung selisih pixel keduanya. Hasil perbedaan image dithreshold untuk mengekstrak moving region. Meskipun pekerjaan ini kelihatan sangat sederhana, pada kenyataannya jarang yang berhasil. Biasanya background jarang yang static dan berubah sesuai dengan waktu dan alasan yang bermacam-macam. Factor yang paling penting adalah perubahan pencahayaan, moving region dan camera noise.

Noise Reduction

Noise yang masih terdapat pada gambar sangat berparuh pada perhitungan selanjutnya, maka dari itu perlu dilakukan noise reduction.

Noise reduction digunakan untuk mengurangi pixel-pixel yang lolos filter background tetapi bukan merupakan foreground utama. Dalam hal ini disebut bukan foreground utama jika pixel tersebut tidak memiliki ketetanggaan 8 pixel disekitarnya

Thresholding

Thresholding bertujuan untuk mendapatkan figure utama atau siluet objek tanpa adanya bayangan objek, dengan meloloskan nilai pixel rendah dan menghilangkan nilai pixel tinggi.

Transformasi Haar wavelet

Haar Wavelet adalah metode wavelet yang pertama kali diajukan oleh Alfred Haar pada tahun 1909. Haar wavelet adalah metode wavelet yang paling sederhana dan mudah untuk diimplementasikan. Untuk mengekstrak ciri-ciri tekstur dengan transformasi Haar Wavelet, dilakukan proses averaging untuk mendapatkan bagian dari gambar yang berfrekuensi rendah dan dilakukan proses differencing untuk mendapatkan bagian dari gambar yang berfrekuensi tinggi. Secara umum, rumus untuk menghitung nilai averaging dan differencing dapat dinyatakan sebagai berikut :

$$H0 : f(n) = (X_n + X_{n+1})/2$$

$$H1 : f(n) = (X_n - X_{n+1})/2$$

dengan :

H0 = nilai averaging

H1 = nilai differencing

Pertama, dilakukan transformasi Haar Wavelet pada setiap baris matriks pada gambar. Kemudian, dilakukan transformasi Haar Wavelet pada setiap kolom. Transformasi ini akan menghasilkan suatu gambar yang seperti terbagi menjadi empat bagian atau empat subband, yaitu subband LL, LH, HL dan HH. Kemudian dilakukan transformasi Haar Wavelet pada subband LL[6]. Subband LL dibagi berdasarkan kolom menghasilkan subband LLL dan LLH. Kemudian subband LL dibagi berdasarkan baris, sehingga menghasilkan empat subband lainnya, yaitu LLLL, LLLH, LLHL, dan LLHH.

Totalnya, ada 7 buah subband yang dihasilkan yaitu LL, LH, HL, LLLL, LLLH, LLHL dan LLHH. Wavelet disini digunakan untuk memperkecil ukuran objek yang akan dideteksi tanpa mengurangi atau merusak fitur yang ada sehingga dapat mempercepat komputasi.

Thinning

Thinning merupakan salah satu image processing yang digunakan untuk mengurangi ukuran dari suatu image (image size) dengan tetap mempertahankan informasi dan karakteristik penting dari image tersebut. Hal ini diimplementasikan dengan mengubah image awal dengan pola binary menjadi representasi kerangka (skeletal representation) image tersebut.

Secara umum, image-thinning berguna untuk mengurangi tresholdded citra output yang dihasilkan dari edge detector, menjadi garis dengan ukuran ketebalan satu pixel saja. Metode Thinning yang digunakan adalah metode hit and miss.

Transformasi hit-dan-miss adalah operasi morfologi yang umum yang dapat digunakan untuk memisahkan pola pixel-pixel foreground dan background pada suatu citra. Operasi hit -dan-miss dilakukan dengan mentranslasikan struktur elemen ke seluruh pixel pada citra, kemudian membandingkan struktur elemen dengan pixel dari citra di bawahnya. Jika pixel-pixel foreground dan background pada struktur elemen cocok (match) dengan pixel-pixel foreground dan background pada citra, maka pixel yang berada di bawah struktur elemen di-set menjadi warna foreground. Jika tidak cocok, maka pixel tersebut dijadikan warna background. Pixel foreground dinyatakan dengan angka 1 dan pixel background dinyatakan dengan angka 0.

Contoh: struktur elemen berikut:

	1	
0	1	1
0	0	

Gambar 1. contoh struktur elemen

dapat digunakan untuk menemukan posisi sudut kanan (*right angle convex corner*) dari suatu citra. Untuk dapat menemukan semua sudut dalam citra biner, kita harus melakukan transformasi *hit-dan-miss* sebanyak empat kali dengan empat elemen berbeda yang merepresentasikan empat jenis sudut yang ditemukan dalam citra biner tersebut. Empat bentuk elemen tersebut adalah:

	1	
0	1	1
0	0	

	1	
1	1	0
	0	0

	0	0
1	1	0
	1	

0	0	
0	1	1
	1	

Gambar 2 Macam-macam struktur elemen thinning

Object Detection

Fitur pejalan kaki yang digunakan adalah fitur bentuk dari kaki objek, sebelumnya kita harus mempelajari ciri dari bentuk kaki sebuah objek pejalan kaki.



Gambar 3. Fitur Kaki

Setelah kita mengetahui ciri dari bentuk bagian kaki, maka kita dapat melakukan pendeteksian pada bagian bawah objek, apakah memiliki bentuk yang sama dengan fitur kaki yang telah dipelajari.

Tracking

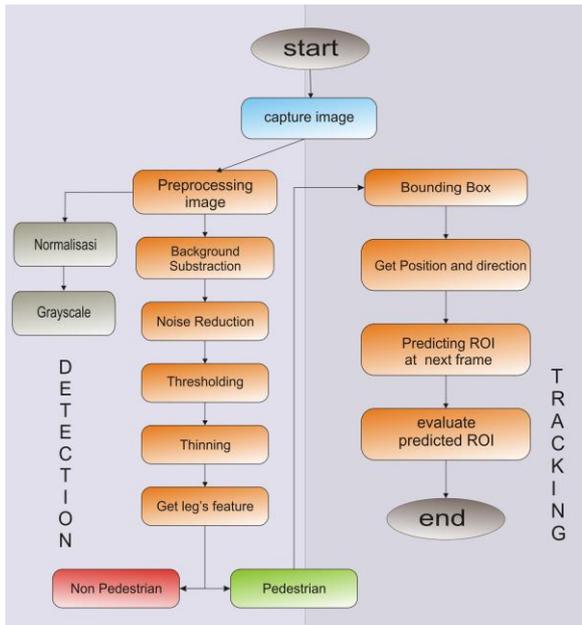
Pelacakan objek dalam hal ini adalah prediksi posisi objek pada frame berikutnya, posisi ditentukan dengan membuat persamaan linear dari dua koordinat pada kedua frame sebelumnya

$$AX + BY + C = 0 \dots\dots\dots(1)$$

Setelah diketahui posisi pada frame berikutnya (ROI) maka dilakukan pengecekan apakah ROI tersebut memiliki fitur pejalan kaki. Pelacakan ini berfungsi untuk mempercepat komputasi sehingga mampu bekerja secara realtime

3. Metodologi Penelitian

3.1 Skema perancangan sistem



Gambar 4. Skema perancangan sistem

3.2 Background Substraction

Setelah system menyimpan frame background maka dilakukan proses background subtraction untuk mendapatkan foreground. Dengan melakukan pengurangan nilai pixel baru dengan pixel background.



Gambar 5. Layer Background



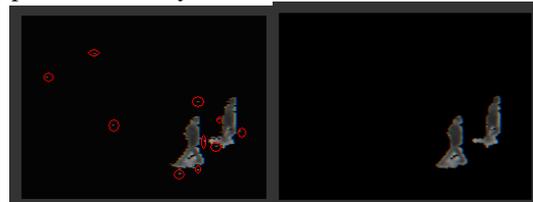
Gambar 6. hasil kamera



Gambar 7. Hasil Background Substraction

3.3 Noise Reduction

Agar pendeteksian dapat dilakukan dengan baik maka pixel-pixel noise harus dihilangkan, pixel yang dianggap noise adalah pixel yang tidak memiliki minimal nilai ketetangaan 8 pixel disekitarnya.



Gambar 8 sebelum dan sesudah proses noise reduction

3.4 Thresholding

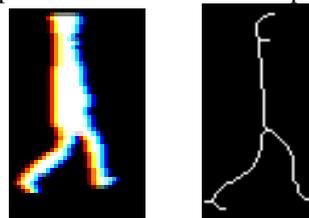
Thresholding bertujuan untuk mendapatkan figure utama atau siluet objek tanpa adanya bayangan objek, dengan meloloskan nilai pixel rendah dan menghilangkan nilai pixel tinggi



Gambar 9. sebelum dan sesudah proses thresholding

3.5 Thinning

Proses thinning digunakan untuk mendapatkan representasi 1 pixel rangka dari suatu objek tanpa memutuskan koneksi tiap rangkanya.



Gambar 9. hasil proses thinning

3.6 Detection

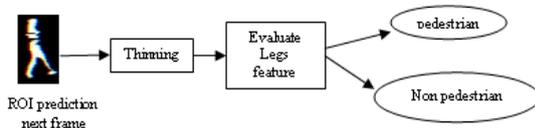
Setelah kita mengetahui ciri dari bentuk bagian kaki, maka kita dapat melakukan pendeteksian pada bagian bawah objek, apakah memiliki bentuk yang sama dengan fitur kaki yang telah dipelajari dengan membandingkan hasil thinning dengan model yang telah dimiliki.



Gambar 10 Siluet yang terdeteksi

3.7 Tracking

Pada tahap ini untuk daerah yang telah terdeteksi pejalan kaki (ROI) akan diambil posisi koordinatnya, dari koordinat yang didapat akan diketahui arah gerak sehingga dapat ditentukan kemungkinan ROI pada frame berikutnya setelah diketahui posisi ROI selanjutnya maka dilakukan evaluasi terhadap ROI tersebut.

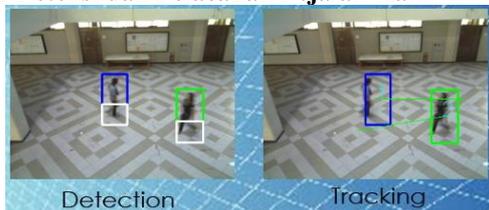


Gambar 11. Proses Tracking

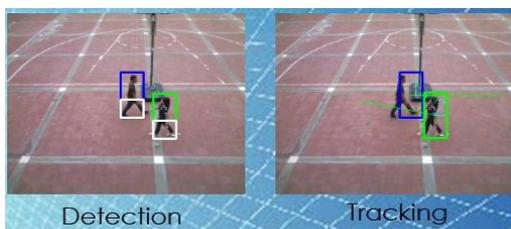
4. Hasil Penelitian

Berikut adalah hasil dari deteksi dan pelacakan pejalan kaki

4.1 Deteksi dan Pelacakan Pejalan Kaki



Gambar 12. hasil positif indoor



Gambar 13. hasil positif outdoor



Gambar 14. terdapat obyek yg tidak terdeteksi

4.2 Metode Perhitungan performance sistem

Tabel 1 Metode perhitungan sederhana

	Terdeteksi	Tidak Terdeteksi
Capture video terdapat Pejalan kaki	True Positive	False Negative
Capture video tidak terdapat pejalan kaki	False Positive	True Negatif

4.2.1 Presisi

Presentase tingkat ketepatan klasifikasi item didalam set jumlah gambar yang terdeteksi adanya object of interest dengan rumus $TP/(TP+FP)$. Ukuran ini mengindikasikan berapa banyak gambar yang benar-benar terdapat object of interest dari seluruh gambar seluruh gambar yang ditandai terdapat object of interest.

4.2.2 Recall / Sensitivitas

Presentase dari jumlah item gambar yang terdeteksi adanya object of interest didalam seluruh set item gambar terdapat object of interest dengan rumus $TP/(TP+FN)$. Kuran ini mengindikasikan seberapa bagus filter dalam mengidentifikasi object of interest.

4.2.3 Spesifisitas / Kekhususan

Presentase dari jumlah item gambar tidak terdeteksi adanya object of interest didalam seluruh set item gambar yang tidak terdapat object of interest dengan rumus $TN/(TN+FP)$. Ukuran ini mengindikasikan seberapa bagus filter didalam mengidentifikasi object non of interest.

4.2.4 Akurasi

Presentase dari ketepatan hasil klasifikasi didalam seluruh jumlah item-item dengan ukuran $(TP+TN)/(TP+TN+FP+FN)$.

4.3 Performa Deteksi & Pelacakan Pejalan Kaki

Berikut ini adalah hasil performance dari frame video terdapat pejalan kaki (550 frame, ukuran 320x240 piksel), frame terdapat objek bukan pejalan kaki (260 frame, ukuran 320x240 piksel), dan frame tidak terdapat objek (250 frame, ukuran 320x240 piksel):

Tabel 2 Hasil performansi percobaan

	Positive identified	Negative Identified
Objek Pedestrian(+)	448	102
Objek Non Pedestrian(-)	32	228
Non Objek(-)	5	245

Tingkat keefektifan :

Presisi :

frame(+) & frame(-)-non pedestrian : $(448/(448+32)) * 100\% = 93\%$
 gambar(+) & gambar(-)-non objek : $(448/(448+5)) * 100\% = 98\%$
 gambar(+) & gambar(-)-non pedestrian & gambar(-)-non objek : $(448/(448+32+5)) * 100\% = 92\%$

Recall/sensitifitas: $(448/(448+102)) * 100\% = 81\%$

Spesifisitas:

gambar(+) & gambar(-)-non pedestrian : $(228/(228+32)) * 100\% = 87\%$
 gambar(+) & gambar(-)-non objek : $(245/(245+5)) * 100\% = 98\%$
 gambar(+) & gambar(-)-non pedestrian & gambar(-)non objek : $(228+245/(228+245+5+32)) * 100\% = 92\%$

Akurasi:

gambar(+) & gambar(-)-non pedestrian : $((448+228)/(448+228+102+32)) * 100\% = 83\%$
 gambar(+) & gambar(-)-non objek : $((448+245)/(448+245+102+3)) * 100\% = 86\%$
 gambar(+) & gambar(-)-no human & gambar(-)-human : $((448+228+245)/(448+228+245+3+32+102)) * 100\% = 87\%$

5. Kesimpulan

- 1 Penggunaan background subtraction sangatlah optimal dalam mengambil foreground pada kamera statis
- 2 Sistem akan dapat bekerja optimal jika intensitas cahaya stabil
- 3 Objek akan terdeteksi dengan baik jika memiliki tingkat keabuan yang berbeda jauh dengan warna background.
- 4 Sistem akan mampu mendeteksi jika objek tertangkap penuh di dalam kamera

6. Daftar Pustaka

Website

[://sourceforge.net/projects/opencvlibrary.](http://sourceforge.net/projects/opencvlibrary)

http

Riyanto Sigit, Achmad Basuki, Nana Ramadijanti, Dadet Pramadihanto 2003, Praktikum Pengolahan Citra, buku diktat PENS – ITS.

Eric J. Stollnitz Tony D. DeRose David H. Salesin 2001, Wavelets for Computer Graphics: A Primer, University of Washington

Kyousuke Uchida , Jun Miura, and Yoshiaki Shirai 2000, Tracking Multiple Pedestrians in Crowd, Dept. of Computer-Controlled Mechanical Systems Osaka University

Lih Zelnik-Manor 2008, Pedestrian Detection and Tracking, Video Analysis Course.

Viola, P.; Jones, M.; Snow, D 2003. Detecting Pedestrians Using Patterns of Motion and Appearance, Mitsubishi Electric Research Laboratories.

Fredrik Sikén, Andreas Andersson, Ali Hedayati, Maria Stegberg 2006, Tracking of Pedestrian, Department of Signals and Systems Chalmers.