

Pengukuran jarak objek pejalan kaki terhadap kamera menggunakan kamera stereo terkalibrasi dengan segmentasi objek *histogram of oriented gradient*

Tutur Urip, Kusworo Adi dan Catur Edi Widodo
Departemen Fisika, Fakultas Sains dan Matematika, Universitas Diponegoro, Semarang
E-mail: tutur.urip@st.fisika.undip.ac.id

ABSTRACT

The distance measurement of pedestrian object to camera using a calibrated stereo camera has been investigated. The study was conducted by comparing Data-A and Data-B with different conditions and environment data interpretation. The stereo vision methods include camera calibration, image rectification, disparity counting, three-dimensional reconstruction and object segmentation. The object segmentation is performed using the Histogram of Oriented Gradient feature to segment pedestrian object. Meanwhile, to determine the distance value is based on the information of the centroid of the bounding box segmented object. The calculations were performed using the Euclidian Distance calculation method to find the shortest distance between the centroid of the bounding box with both cameras. From the research results, the best accuracy was obtained with measurement error of 4%.

Keywords: *Measurement object, pedestrian, calibrated stereo camera, Histogram of Oriented Gradient, Euclidian Distance.*

ABSTRAK

Telah dilakukan penelitian pengukuran jarak objek pejalan kaki terhadap kamera menggunakan kamera stereo terkalibrasi. Penelitian dilakukan dengan membandingkan Data A dan Data B dengan kondisi dan keadaan lingkungan pengambilan data yang berbeda. Langkah-langkah menggunakan metode penglihatan stereo meliputi kalibrasi kamera, rektifikasi citra, penghitungan disparitas, rekonstruksi tiga dimensi dan segmentasi objek. Segmentasi objek yang dilakukan menggunakan fitur Histogram of Oriented Gradients untuk mensegmentasi objek pejalan kaki. Sedangkan untuk menentukan nilai jaraknya berdasarkan informasi nilai titik tengah dari bounding box objek yang tersegmentasi. Penghitungan jarak yang dilakukan menggunakan metode penghitungan jarak Euclidian Distance untuk mengetahui jarak terpendek antara titik tengah bounding box dengan kedua kamera. Dari hasil penelitian, akurasi terbaik didapatkan dengan kesalahan pengukuran sebesar 4%.

Kata kunci: *Pengukuran objek, pejalan kaki, kamera stereo terkalibrasi, Histogram of Oriented Gradient, Euclidian Distance.*

PENDAHULUAN

Pengukuran jarak objek menjadi informasi yang sangat dibutuhkan dalam beberapa bidang, seperti: sistem pengawasan dan keamanan, industri, navigasi, robotika maupun dalam aplikasi kendaraan pintar [1]. Secara umum ada dua cara pengukuran jarak objek, yaitu: secara aktif dan secara pasif. Metode pengukuran aktif mengukur jarak objek dengan mengirimkan beberapa sinyal ke objek. Beberapa metode yang sering digunakan, antara lain: radar, sensor infra

merah, sinyal radio, sensor ultrasonik, laser dan lain-lain [2]. Penggunaan sensor pada metode pengukuran jarak objek secara aktif sangat rentan terhadap kondisi lingkungan, seperti: temperatur, kabut, dan derau misalnya interferensi gelombang pada ruangan. Selain itu, metode tersebut hanya menghasilkan informasi jarak objek saja tanpa mampu mengukur geometri objek [3].

Metode yang kedua adalah pengukuran secara pasif yang hanya menerima informasi posisi objek tanpa memancarkan sinyal, biasanya dengan informasi intensitas cahaya

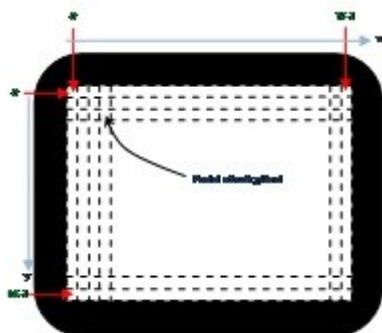
[2]. Hal ini merupakan salah satu dari perkembangan sistem visi komputer yang menggunakan informasi citra visual suatu objek.

Penghitungan jarak objek menggunakan informasi visual dengan sepasang citra dari dua kamera yang dinamakan *stereo image*. Korelasi antara posisi objek dalam dua citra kanan dan kiri memberikan informasi jarak yang dibutuhkan. Hal ini merupakan konsep dari *stereo vision* dan triangulasi [3].

DASAR TEORI

Citra

Citra merupakan data visual yang merepresentasikan distribusi spasial dari besaran fisis seperti intensitas cahaya dan frekuensi spasial dari suatu objek. Informasi tersebut direpresentasikan oleh komponen-komponen seperti tingkat kecerahan, warna dan tepian [4]. Citra digital terbentuk oleh kumpulan titik yang dinamakan piksel (*pixel* atau *picture element*). Tiap piksel pada sebuah citra direpresentasikan oleh data biner. Data ini dapat berupa satu nilai atau kombinasi dari beberapa nilai [5]. Setiap piksel mempunyai koordinat posisi. Sistem koordinat yang dipakai untuk menyatakan citra digital ditunjukkan pada Gambar 1.

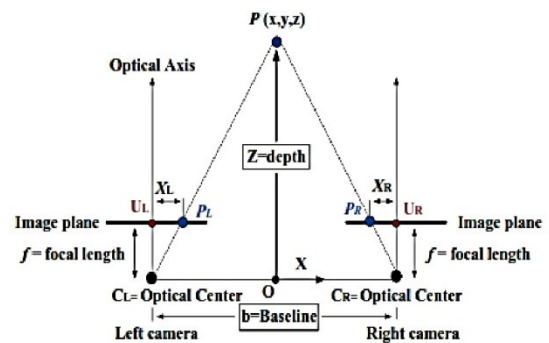


Gambar 1. Sistem koordinat citra berukuran MxN

***Stereo vision* dan geometri *stereo* kamera**

Sistem penglihatan *stereo* (*stereo vision*) merupakan bidang yang berkaitan dengan

penentuan struktur tiga dimensi pada suatu *scene* dari dua atau lebih citra digital yang didapat dari sudut yang berbeda [6]. Pengertian kamera stereo merupakan kamera dengan tipe dan spesifikasi yang sama dan dipasang dalam satu garis lurus baik pada bidang horisontal maupun vertikal. Pengukuran jarak dapat dilakukan ketika objek berada dalam sudut pandang yang saling overlap di antara dua kamera.



Gambar 2. Geometri epipolar dengan sumbu optis sejajar

Gambar 2 merupakan sistem koordinat *stereo* kamera yang dianggap berada dalam titik tengah antara sistem koordinat kamera kanan dan kamera kiri. Dari konsep kesebangunan segitiga $\Delta PC_L C_R$ dan $\Delta P_L P_R$ pada Gambar 2 tersebut, sistem koordinat *stereo* kedua kamera dapat ditulis pada Persamaan (2).

$$\frac{b}{Z} = \frac{(b + X_R) - X_L}{Z - f} \tag{1}$$

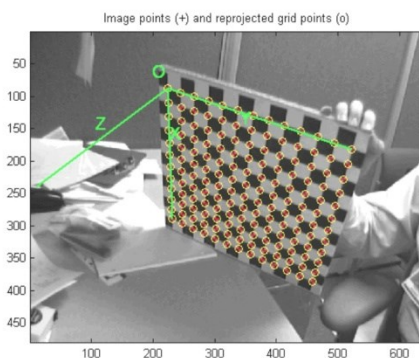
$$Z = \frac{b * f}{X_L - X_R} = \frac{b * f}{d} \tag{2}$$

dengan, $d = (X_L - X_R)$ merupakan disparitas, X_L adalah koordinat- x pada citra kiri, X_R adalah koordinat- x pada citra kanan, b adalah panjang garis dasar (jarak antara sumbu optis dua kamera) dan f merupakan panjang fokus kamera. Persamaan (2) menunjukkan bahwa kedalaman Z berbanding terbalik dengan disparitas. Semakin tinggi disparitas menunjukkan bahwa objek semakin dekat

dengan garis dasar kamera. Begitupun sebaliknya, semakin kecil nilai disparitas objek menunjukkan objek lebih jauh dari garis dasar [7].

Kalibrasi kamera

Fungsi kalibrasi kamera dalam *stereo vision* adalah menentukan parameter-parameter geometrik lensa dan mengevaluasi baik kinerja maupun stabilitas antara lensa dan kamera. Untuk mengkalibrasikan kamera, kita harus mengekstrak semua titik sudut dari tiap citra dan kemudian menghitung parameter intrinsik dan ekstrinsik kamera. Parameter intrinsik menunjukkan ciri-ciri internal kamera, seperti: panjang fokus, titik tengah citra, koefisien kemiringan, dan parameter penyimpangan lensa, sedangkan parameter ekstrinsik, seperti: orientasi rotasi dan translasi yang menghubungkan dua kamera [8]. Gambar 3 merupakan contoh proses kalibrasi kamera menggunakan papan catur.



Gambar 3. Proses kalibrasi kamera

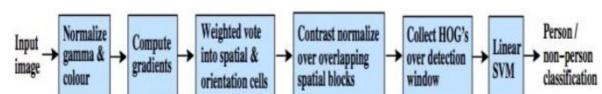
Rektifikasi Citra

Rektifikasi citra adalah proses untuk memproyeksikan dua atau lebih citra kedalam satu bidang citra. Proses ini memperbaiki penyimpangan lensa dengan mentransformasikan citra kedalam sistem koordinat standar [9]. Karena dua kamera menangkap citra dari sudut yang berbeda, citra yang dihasilkan pun berada pada bidang yang berbeda. Rektifikasi dapat dibedakan menjadi dua jenis, yaitu: rektifikasi pada kamera terkalibrasi dan rektifikasi pada hasil citra dari kamera yang tidak terkalibrasi.

Pada *stereo* kamera yang terkalibrasi, rektifikasi citra memproyeksikan dua citra kedalam satu bidang citra sehingga ciri-ciri yang sama pada masing-masing citra akan terlihat pada baris yang sama. Proyeksi citra ini membuat citra terlihat seperti dua kamera pada posisi sejajar. Pada rektifikasi dengan kamera tak terkalibrasi mengakibatkan parameter intrinsik dan ekstrinsik kamera tidak diketahui, sehingga proses ini membutuhkan beberapa titik yang berkorespondensi antara dua citra. Untuk menghasilkan titik berkorespondensi ini, kita harus mengumpulkan titik-titik piksel yang serupa dari kedua citra dan kemudian memilih penggabungan yang mungkin antara keduanya. Pencocokan citra dapat menggunakan fitur *SURF*, *SIFT*, *HARRIS*, *SAD* dan lain-lain.

Segmentasi citra menggunakan fitur HOG (*Histogram of Oriented Gradient*)

HOG (*Histogram of oriented gradient*) bertujuan untuk pengenalan obyek, informasi dalam bentuk lokal terdistribusi dari gradien intensitas atau arah tepi bahkan tanpa informasi yang tepat tentang lokasi tepi tersebut. Gambar dibagi menjadi bagian yang lebih kecil menjadi sub-gambar yang dinamakan sel. Sel tersebut dapat berbentuk persegi panjang (*R-HOG*) atau berbentuk melingkar (*C-HOG*). Histogram orientasi tepi dikumpulkan dalam sel tersebut. Masukan gabungan histogram digunakan sebagai fitur vektor yang menggambarkan objek. Untuk memberikan pencahayaan invarian yang lebih baik (pencahayaan, bayangan, dll) dilakukan normalisasi sel di daerah yang lebih besar dengan menggabungkan beberapa sel menjadi blok [10]. Alur proses deteksi menggunakan HOG dapat dijelaskan dengan diagram pada Gambar 4.



Gambar 4. Metode HOG

Penghitungan jarak dengan fungsi euclidian distance

Dalam menentukan nilai jarak objek terhadap kamera, diperlukan informasi *centroid bounding box* dari objek yang tersegmentasi. *Bounding box* adalah persegi panjang yang memuat semua region dengan elemen $[x \ y \ \text{lebar} \ \text{tinggi}]$. Properti *centroid* didefinisikan sebagai titik tengah atau pusat dari region. Elemen pertama dari centroid adalah koordinat horizontal koordinat $[x \ y]$ dari objek. Dalam menghitung nilai *centroid bounding box* menggunakan Persamaan (3).

$$\text{cent} = \left[a + \frac{c}{2} \quad b + \frac{d}{2} \right] \quad (3)$$

dengan, a dan b merupakan komponen dalam sumbu- x dan komponen dalam sumbu- y , c dan d merupakan komponen lebar *bounding box* dan komponen tinggi *bounding box* dalam satuan piksel.

Setelah didapatkan posisi *centroid* dalam koordinat tiga dimensi, maka untuk menghitung jarak antara centroid dengan kamera digunakan persamaan *euclidean distance* dengan asumsi bahwa posisi kedua kamera berada pada sumbu origin atau pada koordinat $[x \ y \ z] = [0 \ 0 \ 0]$. *Euclidean distance* adalah metrika yang paling sering digunakan untuk menghitung kesamaan dua vektor. Jarak *euclidian* dapat dianggap sebagai jarak yang paling pendek antar dua poin dan pada dasarnya sama halnya dengan persamaan Pythagoras ketika digunakan di dalam 2 dimensi. Rumus *euclidean distance* adalah akar dari kuadrat perbedaan 2 vektor (*root of square differences between 2 vectors*) [11].

$$d_{ij} = \sqrt{\sum_{k=1}^n (x_{ik} - x_{jk})^2} \quad (4)$$

dengan, d_{ij} merupakan tingkat perbedaan (*dissimilarity degree*), n adalah jumlah vektor, x_{ik} adalah vektor koordinat *input* (koordinat

objek), x_{jk} adalah vektor koordinat citra pembanding (koordinat kamera).

METODE PENELITIAN

Alat dan Bahan

Bahan yang digunakan dalam penelitian ini adalah data primer beberapa citra untuk kalibrasi kamera dan video untuk penentuan nilai jarak pada objek pejalan kaki. Alat dalam penelitian ini diantaranya: komputer, laptop, dua buah kamera *webcam* dan perangkat lunak *Matlab*.

Prosedur Penelitian

Secara umum pengukuran jarak objek dapat dilakukan dengan beberapa tahapan. Tahapan penelitiannya meliputi tahap studi literatur, tahap akuisisi data, tahap perancangan dan implementasi algoritma sistem, tahap pengujian serta tahap analisis. Tahap akuisisi data bertujuan untuk mendapatkan materi dasar citra yang diinginkan. Akuisisi citra berfungsi saat perlakuan kalibrasi kamera maupun penentuan jarak objek. Pada saat kalibrasi, citra yang didapatkan adalah citra *frame* tunggal, yaitu: citra pola papan catur. Penentuan nilai jarak citra yang digunakan adalah citra *multiframe (video)*, yaitu: *video* orang yang bergerak. Penelitian ini dilakukan dengan membandingkan 2 data (Data A dan Data B) dengan parameter dan perlakuan yang berbeda. Pengambilan Data A bertempat di Laboratorium Elektronika dan Instrumentasi Universitas Diponegoro, sedangkan pengambilan Data B bertempat di Gedung *Corner Rusunawa* Universitas Diponegoro. Data A dan Data B masing-masing-masing terdiri atas 2 citra/video identik dari hasil akuisisi 2 kamera stereo dalam waktu yang bersamaan. Tahap perancangan dan implementasi algoritma sistem dilakukan setelah materi dasar citra dan video hasil akuisisi telah didapatkan. Tahapan ini menggunakan perangkat lunak *Matlab* dan terdiri atas beberapa tahapan diantaranya rektifikasi citra, penentuan disparitas,

rekonstruksi citra 3 dimensi, segmentasi citra menggunakan fitur HOG dan penghitungan jarak menggunakan fungsi *euclidian distance*.

HASIL DAN PEMBAHASAN

Akuisisi Citra Data A dan Data B

Akuisisi citra berfungsi saat perlakuan kalibrasi kamera maupun pengambilan video untuk diidentifikasi jarak objeknya. Parameter-parameter akuisisi untuk Data A dan Data B ditunjukkan pada Tabel 1.

Kalibrasi kamera

Kalibrasi kamera pada Data A dan Data B menghasilkan parameter ekstrinsik dan intrinsik kedua kamera yang ditunjukkan dalam Tabel 2 Parameter Kalibrasi Kamera Data A dan Tabel 3 Parameter kalibrasi Kamera Data B. Parameter ekstrinsik menjelaskan orientasi dan translasi kamera kanan (kamera 2) relatif terhadap kamera kiri (kamera 1), sedangkan parameter intrinsik untuk mengetahui transformasi pemetaan titik pada citra dari kamera ke koordinat piksel, pada setiap kamera dan beberapa parameter internal kamera. Kedua parameter ini digunakan dalam proses rektifikasi dan rekonstruksi citra tiga dimensi.

Tabel 1. Parameter pengambilan data Data A dan Data B

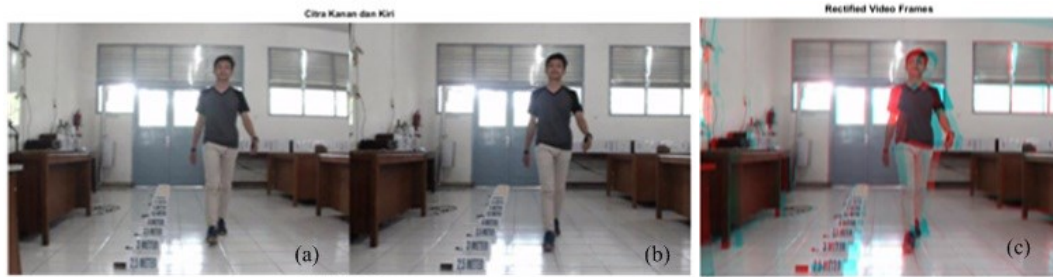
No	Parameter	Data A		Data B	
		Kamera Kanan	Kamera Kiri	Kamera Kanan	Kamera Kiri
1	Resolusi	640 x 480	640 x 480	320 x 240	320 x 240
2	Durasi	16 detik	16 detik	129 detik	129 detik
3	Total <i>frame</i>	160	160	1292 <i>frame</i>	1292 <i>frame</i>
4	<i>Frame rate</i>	10 fps	10 fps	10 fps	10 fps
5	Jarak kamera kanan dan kamera kiri	12 cm	12 cm	12 cm	12 cm
6	Jumlah citra kalibrasi	19 pasang	19 pasang	15 pasang	15 pasang
7	Intensitas cahaya	340 lux	340 lux	27 lux	27 lux
8	Jarak yang diamati	Jarak ketika objek sedang bergerak	Jarak ketika objek sedang bergerak	Jarak ketika objek sedang diam	Jarak ketika objek sedang diam
9	Kondisi ruangan	Ruangan tidak kosong	Ruangan tidak kosong	Ruangan kosong	Ruangan kosong

Tabel 2. Parameter intrinsik dan ekstrinsik dual kamera pada Data A

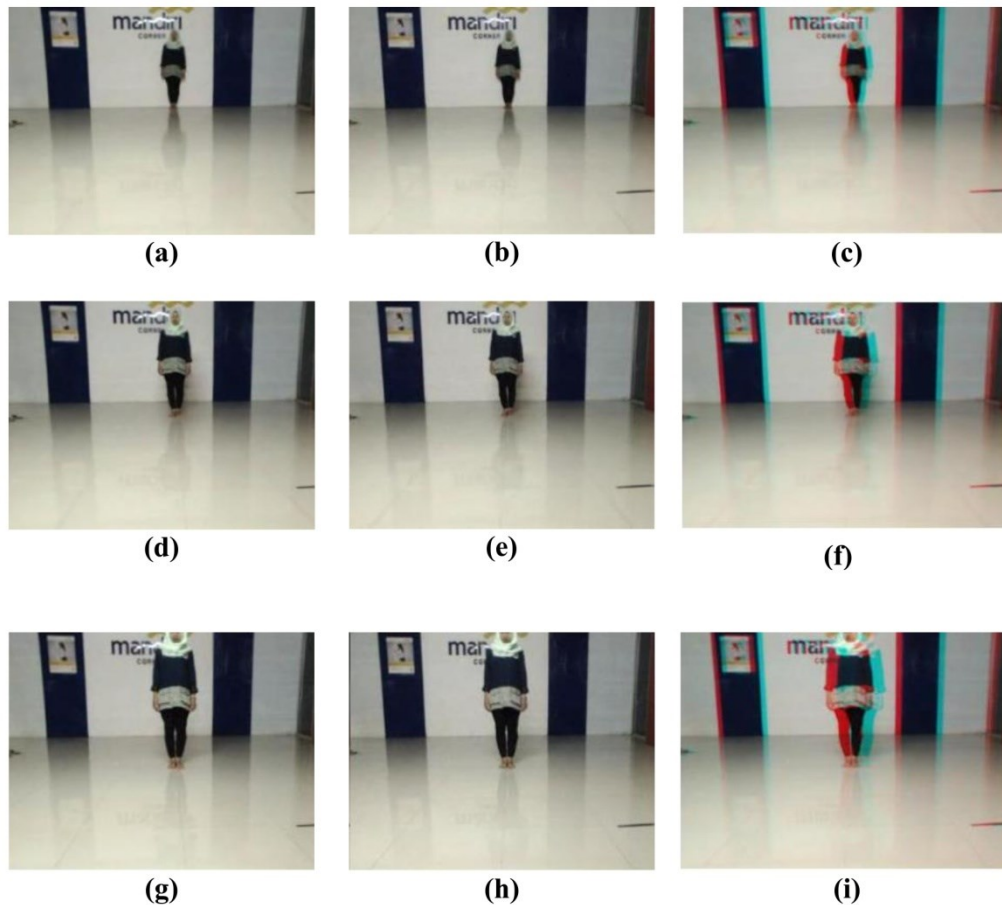
No	Parameter kamera	Hasil
1	Distorsi radial kamera kiri	[0,07 1,04]
2	Distorsi tangensial kamera kiri	[0 0]
3	Satuan	'mm'
4	Koefisien distorsi radial kamera kiri	2
5	Matriks intrinsic kamera kiri	[622,48 0 0; 0 622,66 0; 324,18 250,32 1]
6	Panjang fokus kamera kiri	[622,48 622,66]
7	Principal point kamera kiri	[324,18 250,32]
8	Kemiringan kamera kiri	0
9	<i>Mean ReprojectionError</i> kamera kiri	0,24
10	Distorsi radial kamera kanan	[0,04 0,11]
11	Distorsi tangensial kamera kanan	[0 0]
12	Satuan	'mm'
13	Koefisien distorsi radial kamera kanan	2
14	Matriks intrinsic kamera kanan	[623,49 0 0; 0 623,77 0; 336,16 233,35 1]
15	Panjang fokus kamera kanan	[623,49 623,77]
16	Principal point kamera kanan	[336,16 233,35]
17	Kemiringan kamera kanan	0
18	<i>Mean ReprojectionError</i> kamera kanan	0,25
19	Rotasi kamera kanan terhadap kamera kiri	[0,99 -0,007 0,05; 0,009 0,99 -0,021; -0,051 0,022 0,99]
20	Translasi kamera kanan terhadap kamera kiri	[-120,52 1,82 -2,27]
21	Fundamental matriks	[-1,94e-07 -1,008e-05 0,004; -5,74e-06 6,69e-06 0,19; -0,0015 -0,19 -1,16]
22	Matriks essensial	[-0,075 -3,91 0,87; -2,22 2,60 120,50; -1,86 -120,44 2,54]
23	Rata-rata kesalahan proyeksi kedua kamera	0,25

Tabel 3. Parameter intrinsik dan ekstrinsik dual kamera pada Data B.

No	Parameter kamera	Hasil
1	Distorsi radial kamera kiri	[0,1 -0,36]
2	Distorsi tangensial kamera kiri	[0 0]
3	Satuan	'mm'
4	Koefisien distorsi radial kamera kiri	2
5	Matriks intrinsik kamera kiri	[308,61 0 0; 0 308,34 0; 157,903 129,65 1]
6	Panjang fokus kamera kiri	[308,61 308,34]
7	Principal point kamera kiri	[157,91 129,65]
8	Kemiringan kamera kiri	0
9	<i>Mean ReprojectionError</i> kamera kiri	0,192
10	Distorsi radial kamera kanan	[0,05 -0,16]
11	Distorsi tangensial kamera kanan	[0 0]
12	Satuan	'mm'
13	Koefisien distorsi radial kamera kanan	2
14	Matriks intrinsik kamera kanan	[312,07 0; 0 311,31; 162,01 121,3]
15	Panjang fokus kamera kanan	[312,07 311,31]
16	Principal point kamera kanan	[162,01 1,21]
17	Kemiringan kamera kanan	0
18	<i>Mean ReprojectionError</i> kamera kanan	0,198
19	Rotasi kamera kanan terhadap kamera kiri	[0,99 -0,01 -0,02; 0,01 0,99 -0,03; 0,02 0,02 0,99]
20	Translasi kamera kanan terhadap kamera kiri	[-122,38 0,02 8,35]
21	Fundamental matriks	[1,19e-06 -1,19e-04 0,02; 8,68e-05 -3,16e-05 0,38; -0,01 -0,37 -1,58]
22	Matriks essensial	[0,11 -11,41 1,71; 8,34 -3,03 1,22e+02; -0,24 -122,09e+02 -3,19]
23	Rata-rata kesalahan proyeksi kedua kamera	0,195



Gambar 5. Citra asli dari stereo kamera dan citra hasil rektifikasi pada Data A (a) Citra asli *frame* ke 135 dari kamera kiri pada Data A (b) Citra asli *frame* ke 135 dari kamera kanan pada Data A (c) Hasil rektifikasi dua *frame* citra kamera kanan dan kiri pada *frame* ke-135 Data A



Gambar 6. Citra asli dan citra hasil rektifikasi *frame* ke-10, *frame* ke-366 dan *frame* ke-685 pada Data B. (a) Citra asli *frame* ke-10 dari kamera kiri, (b) Citra asli *frame* ke-10 dari kamera kanan, (c) Citra hasil rektifikasi dua *frame* citra ke-10, (d) Citra asli *frame* ke-366 dari kamera kiri, (e) Citra asli *frame* ke-366 dari kamera kanan, (f) Citra hasil rektifikasi dua *frame* citra ke-366, (g) Citra asli *frame* ke-685 dari kamera kiri, (h) Citra asli *frame* ke-685 dari kamera kanan, (i) Citra hasil rektifikasi dua *frame* citra ke-685

Rektifikasi citra

Citra asli dari kamera kanan dan kamera kiri pada Data A masing-masing ditunjukkan dalam matriks berukuran 480x640x3 piksel 8 bit. Setelah kedua citra kanan kiri digabungkan, ukuran citra ter-rectifikasi tersebut berubah dan direpresentasikan oleh matriks berukuran 408x553x3 piksel 8 bit. Gambar 5. menggambarkan citra asli dari stereo kamera dan citra hasil rektifikasi pada Data A(a) merupakan citra asli *frame* ke 135 dari kamera kanan, (b) citra asli *frame* ke 135 dari kamera kiri, (c) menunjukkan hasil rektifikasi citra dari kedua pasangan *frame* tersebut.

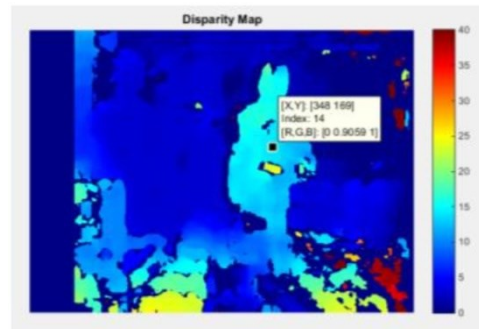
Khusus untuk Data B, penelitian ini membahas 3 jenis *frame* yaitu *frame* ke-10, *frame* ke-366 dan *frame* ke-685. Citra asli dari kamera kanan dan kamera kiri masing-masing ditunjukkan dalam matriks berukuran 240x320x3 piksel 8 bit. Setelah kedua citra kanan kiri digabungkan, ukuran citra ter-rectifikasi tersebut berubah dan direpresentasikan oleh matriks berukuran 219x294x3 piksel 8 bit. Gambar 6 menunjukkan citra asli dan citra hasil rektifikasi *frame* ke-10, *frame* ke-366 dan *frame* ke-685 dari kamera kanan dan kiri pada Data B.

Dari kedua Data A dan Data B, pada proses rektifikasi citra tidak ada suatu perbedaan yang mencolok dari hasil citra ter-rectifikasinya. Kedua video dapat ter-rectifikasi dengan baik. Perbedaan yang terjadi adalah pada ukuran matriks rektifikasi Data A dan Data B dikarenakan resolusi kedua video berbeda

Informasi kedalaman (*depth*) *rectified image*

Penggabungan dua citra *stereo* pada proses rektifikasi citra akan memunculkan selisih *corresponding point* pada kedua titik pasangan citra kanan dan kiri yang disebut dengan *disparity*. Peta disparitas dalam sistem *stereovision* merepresentasikan informasi kedalaman (*depth*) suatu objek. Nilai disparitas berbanding terbalik dengan nilai kedalaman. Semakin tinggi disparitas menunjukkan bahwa objek semakin dekat dengan garis dasar kamera. Begitupun sebaliknya, semakin kecil nilai

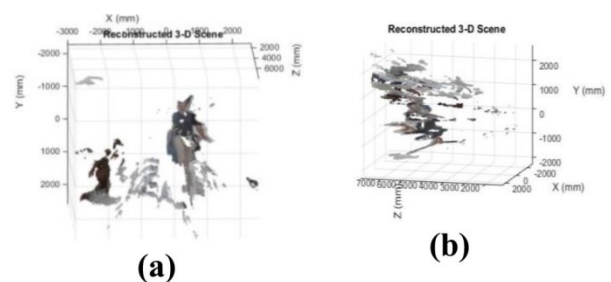
disparitas objek menunjukkan objek lebih jauh dari garis dasar kamera. Gambar 7 menunjukkan peta disparitas pada *frame* ke-135 Data A.



Gambar 7. Peta disparitas *frame* ke-135 Data A

Rekonstruksi tiga dimensi

Proses rekonstruksi citra tiga dimensi diberi masukan berdasarkan peta disparitas dan parameter kalibrasi kamera. Koordinat 3D sesuai dengan piksel masukan pada peta disparitas. Koordinat 3D relatif terhadap pusat optik kamera kiri dalam sistem *stereo* berdasarkan matriks fundamental dan matriks esensial.

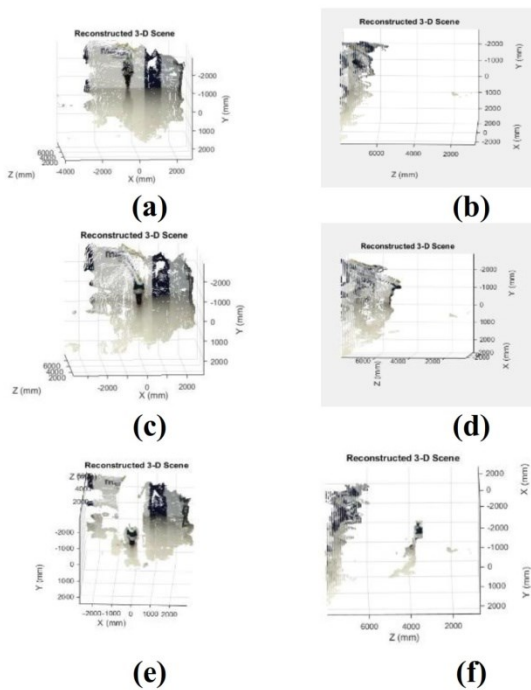


Gambar 8. Rekonstruksi 3D *frame* ke-135 Data A (a) Tampak depan, (b) Tampak samping

Gambar 8 menjelaskan rekonstruksi tiga dimensi dari *frame* ke-135 data A yang tampak dari sudut pandang dari sisi depan dan sisi samping. Terlihat dari tabel tersebut baik dari sisi depan dan sisi samping, sistem tidak

dapat merekonstruksi *frame* secara sempurna. Hal ini disebabkan masukan dari proses rekonstruksi tiga dimensi adalah peta disparitas. Sedangkan peta disparitas yang dihasilkan pada *frame* 135 Data A tidak dapat terlihat informasi kedalaman dengan baik dikarenakan faktor pencahayaan, faktor lingkungan dan parameter kalibrasi kamera.

hasil rekonstruksinya tidak dapat membentuk objek secara keseluruhan. Pada *frame* ke-366, objek manusia berjarak kira-kira 4000 dari sumbu (0,0,0) kamera kiri. Hasil rekonstruksi pada *frame* ke-685 tidak sempurna pada keseluruhan *scene*, namun dari hasil rekonstruksi tersebut dapat diperoleh informasi jarak objek (subregion manusia) diinginkan yaitu kurang lebih 4000 mm dari kamera kiri. Informasi koordinat tiga dimensi ini bukan merupakan jarak objek yang diamati. Karena dari koordinat tiga dimensi ini akan diambil informasi nilai pusat masa objek untuk diproses pada tahap selanjutnya.



Gambar 9. Rekonstruksi 3D *frame* ke-10, *frame* ke-366 dan *frame* ke-685 pada Data B
 (a) Rekonstruksi 3D *frame* ke-10 tampak depan,
 (b) Rekonstruksi 3D *frame* ke-10 tampak samping, (c) Rekonstruksi 3D *frame* ke-366 tampak depan, (d) Rekonstruksi 3D *frame* ke-366 tampak samping, (e) Rekonstruksi 3D *frame* ke-685 tampak depan, (f) Rekonstruksi 3D *frame* ke-685 tampak samping

Gambar 9 menunjukkan hasil rekonstruksi objek manusia dari dua citra yang digunakan pada *frame* ke-10, 366 dan 685 Data B terlihat dari sisi depan dan sisi samping. Pada *frame* ke-10 sistem tidak dapat memperoleh kedalaman objek pada peta disparitas, namun pada rekonstruksi tiga dimensi ini sistem dapat memperoleh informasi tiga dimensi objek terhadap kamera akan tetapi

Segmentasi objek menggunakan *Histogram of Oriented Gaussian Feature* dan klasifikasi *Support Vector Machines*

Segmentasi objek pada Data A

Tabel 3 merupakan fitur *HOG* yang digunakan untuk klasifikasi subregion manusia pada Data A. Model klasifikasi yang digunakan untuk pelatihan berukuran 128x64 piksel. Citra yang digunakan untuk melatih model mencakup piksel *background* di sekitar orang tersebut. Oleh karena itu, ukuran sebenarnya dari orang yang terdeteksi lebih kecil dari ukuran citra pelatihan.

Tabel 3.Fitur *HOG* dalam klasifikasi subregion manusia pada Data A

No	Jenis fitur <i>HOG</i>	Nilai
1	Model klasifikasi	[128 64]
2	<i>Threshold</i> klasifikasi	2
3	<i>Window Stride</i>	[8 8]
4	Faktor Penskala	1,05
5	Ukuran maksimum <i>bounding box</i>	[384 192]

Klasifikasi *threshold* yang digunakan bernilai 2. Ambang batas ini digunakan untuk mengontrol apakah sub-regional akan diklasifikasikan sebagai manusia. Semakin

tinggi nilai ambang batas, semakin ketat persyaratan untuk klasifikasi. Nilai *threshold* ini mempunyai pengaruh yang besar yaitu sekitar 80% sampai 90% terhadap keberhasilan proses deteksi yang ada. Parameter selanjutnya adalah *windowstrides* atau jendela geser, *windowstrides* yang digunakan adalah dengan ukuran [panjang lebar] = [8 8] yang "meluncur" melintasi gambar untuk pencarian subregion manusia. Semakin kecil *windowstride*, semakin banyak jendela yang perlu dievaluasi akan tetapi membutuhkan beban komputasi yang lebih tinggi.

Faktor penskalaan yang digunakan adalah 1,05. Penurunan faktor skala dapat meningkatkan akurasi deteksi. Namun, hal tersebut meningkatkan waktu komputasi. Kemudian ukuran maksimum subregion manusia yang ditampilkan dalam bentuk *bounding box* diatur dengan panjang 384 piksel dan lebar 192 piksel. Pada *frame* ke-135 subregion manusia yang terdeteksi berada pada *bounding box* dengan posisi dan ukuran [x y lebar tinggi] = [256 22 173 346]. Dengan menggunakan fitur-fitur *HOG*, pada Data A dari ke-160 *frame* tersebut, 152 *frame* diantaranya berhasil tersegmentasi menjadi subregion manusia dan 8 *frame* tidak berhasil mendeteksi manusia sehingga didapatkan akurasi segmentasi sebesar 95%.

Segmentasi objek pada Data B

Tabel 4 merupakan fitur *HOG* yang digunakan untuk klasifikasi subregion manusia pada Data B. Penelitian ini didapatkan 15 data objek statis yang terdiri atas 821 *frame*. Dengan menggunakan fitur-fitur *HOG* dari ke-821 *frame* tersebut, 584 *frame* diantaranya berhasil tersegmentasi menjadi subregion manusia, 234 *frame* tidak berhasil mendeteksi manusia, dan 3 *frame* menunjukkan adanya *double detection* yaitu ada objek yang terdeteksi namun bukan manusia. Sehingga didapatkan akurasi segmentasi sebesar 71,13%.

Tabel 4. Fitur *HOG* dalam klasifikasi subregion manusia pada Data B

No	Jenis fitur <i>HOG</i>	Nilai
1	Model klasifikasi	[128 64]
2	Threshold klasifikasi	0.5
3	<i>Window Stride</i>	[8 8]
4	Faktor Penskalaan	1,05
5	Ukuran maksimum <i>bounding box</i>	[400 200]

Penghitungan jarak menggunakan *Euclidian Distance*

Penghitungan jarak pada Data A

Nilai parameter tahapan penentuan jarak pada *frame* ke-135 ditunjukkan pada Tabel 5. Pada *frame* ke-135 segmentasi berhasil dilakukan pada nilai *bounding boxes* [256 22 173 346]. Nilai titik pusat *bounding box* berada pada koordinat [x y] = [343 195] mm dan nilai *centroid* indeks 139047. Nilai koordinat 3 dimensi *centroid* objek pada *frame* ke-135 adalah [x y z] = [2,19e+02 -35,06 3,19e+03] mm, Kemudian didapatkan nilai jaraknya sebesar 3,2 meter sedangkan nilai jarak sebenarnya sebesar 3 meter. Pada *frame* ke-135 ini tidak terjadi penyimpangan pengukuran terlalu jauh karena posisi objek tidak terlalu jauh dengan kamera.

Tabel 5. Parameter tahapan penghitungan jarak objek pada *frame* ke-135 Data A

No	Tahap penentuan jarak	Nilai
1	Ukuran <i>bboxes</i>	[256 22 173 346]
2	<i>Centroid</i> (titik pusat) <i>bboxes</i>	[343 195]
3	<i>Centroid</i> idx	139047
4	<i>Centroid</i> 3D	[2,19e+02 -35,06 3,19e+03]
5	Jarak	3.200 mm
6	Jarak sebenarnya	3.000 mm

Tabel 6 menunjukkan hasil pengukuran jarak objek dengan kamera menggunakan metode *stereovision* pada Data A. Dari tabel tersebut menunjukkan bahwa sistem pengukuran memiliki kesalahan rata-rata sebesar 125,74 %. Kesalahan pengukuran ini tidak dapat di tolerir untuk sebuah sistem pengukuran. Kesalahan terkecil saat jarak sebenarnya 3 meter yaitu sebesar 48,67 %. Sedangkan kesalahan terbesar sebesar 228,133 % pada saat jarak sebenarnya sebesar 7,5 meter. Dari data tersebut, ketika jarak objek relatif dekat dengan kamera maka kesalahan pengukuran sistem lebih kecil jika dibandingkan dengan pengukuran saat jarak objek jauh dengan kamera.

Tabel 6. Hasil pengukuran jarak objek pada Data A

No	Jarak sebenarnya (m)	Jarak terukur (m)	Selisih jarak (m)	Error (%)
1	7,5	24,61	17,11	228,13
2	7	21,53	14,53	207,57
3	6,5	13,68	7,18	110,46
4	6	15,9	9,9	165
5	5,5	10,44	4,94	89,81
6	5	11,79	6,79	135,8
7	4,5	7,53	3,03	67,33
8	4	6,63	2,63	65,75
9	3,5	8,36	4,86	138,85
10	3	4,46	1,46	48,66
Rata-rata kesalahan				125,74

Penghitungan jarak pada Data A

Tabel 7 menunjukkan hasil pengukuran jarak objek dengan kamera menggunakan metode *stereovision*. Dari tabel tersebut menunjukkan bahwa sistem pengukuran memiliki kesalahan rata-rata sebesar 4 %. Kesalahan terkecil pada data ke-11 yaitu sebesar 0,36 %. Sedangkan kesalahan terbesar sebesar 14,44 pada data ke-13.

Hasil keseluruhan akurasi pengukuran pada Data A dan Data B

Faktor utama penyebab kesalahan pengukuran pada Data A dan Data B adalah karena rendahnya prosesor yang mengakibatkan pada saat akuisisi video, kedua kamera tidak dimulai dan berakhir pada titik yang sama. Sehingga mengharuskan pemotongan *frame* secara manual yang menyebabkan beberapa *frame* tidak terpasangkan dengan *frame* seharusnya. Secara keseluruhan, kesalahan pengukuran Data B lebih baik jika dibandingkan dengan Data A. Kesalahan pengukuran pada Data A sebesar 125,74% sedangkan kesalahan pengukuran pada Data B sebesar 4 %. Data A dan Data B berada dalam kondisi lingkungan dan cara pengambilan data yang berbeda. Tingkat perbedaan tingkat akurasi dari Data A dan Data B dipengaruhi oleh beberapa faktor yaitu, metode pengambilan data, kondisi lingkungan, faktor pencahayaan, kesalahan reprojeksi pada saat kalibrasi kamera.

Cara pengambilan data merupakan salah satu faktor yang penting dalam pengukuran menggunakan sistem ini. Pada Data A pengukuran ketika objek bergerak menyebabkan kesalahan pengukuran karena dimensi objek berbeda sehingga menyebabkan kesalahan dalam penghitungan *centroid*, posisi koordinat pasangan yang berbeda, dan perbedaan intensitas cahaya pada citra kanan dan citra kiri yang mengakibatkan perbedaan nilai piksel pada kedua citra tersebut.

Keadaan lingkungan juga mempengaruhi tingkat akurasi dalam penelitian ini. Kondisi lokasi penelitian pada Data A tidak steril sehingga menyebabkan derau yang disebabkan adanya benda-benda asing. Derau oleh benda asing ini mempengaruhi proses segmentasi dan ketidaktepatan dalam penghitungan *centroid* karena piksel-piksel dari benda asing ini membur dengan piksel-piksel objek dalam

satu bounding box. Selain pengaruh benda-benda asing dalam lokasi penelitian, pencahayaan juga menjadi faktor penting dalam penelitian ini. Pada DataA intensitas yang terukur sebesar 340 lux sedangkan intensitas cahaya pada Data B sebesar 27 lux. Intensitas cahaya yang besar pada DataA ini disebabkan oleh kamera yang menghadap langsung dengan cahaya yang dihasilkan oleh jendela. Pengaruh pencahayaan yang terlalu terang akan mengakibatkan detail citra yang terbaur sehingga menyebabkan banyak derau yang terbentuk.

Tabel 7. Hasil pengukuran jarak objek pada Data B.

No	Jarak sebenarnya (m)	Jarak terukur rata-rata (m)	Selisih jarak (m)	Error (%)
1	6,67	6,44	-0,23	3,45
2	6,07	6,12	0,05	0,82
3	5,47	5,39	-0,08	1,46
4	4,87	4,68	-0,19	3,90
5	4,27	4,2	-0,07	1,64
6	3,67	3,44	-0,23	6,27
7	3,07	3,33	0,26	8,47
8	3,67	3,62	-0,05	1,36
9	4,27	4,28	0,01	0,23
10	4,87	4,93	0,06	1,23
11	5,47	5,45	-0,02	0,36
12	4,87	4,84	-0,03	0,61
13	3,67	4,2	0,53	14,44
14	3,07	3,33	0,26	8,47
15	3,67	3,94	0,27	7,35
Rata-rata kesalahan				4,00

Pada saat akuisisi citra Data A dan Data B mempunyai resolusi yang berbeda. Resolusi Data Video A sebesar 640x280 piksel sedangkan Data B sebesar 320x240 piksel. Perbedaan resolusi ini tidak berpengaruh terhadap nilai jarak tetapi hanya berpengaruh terhadap lamanya waktu komputasi, terutama saat segmentasi objek. *Windowstride* yang digunakan pada Data A dan B

keduanya sama-sama berukuran [8 8] tetapi *Windowstride* tersebut bekerja mencari objek manusia pada keseluruhan *frame*. Sehingga semakin luas ukuran resolusi citra maka semakin lama proses komputasi untuk melakukan segmentasi objek.

KESIMPULAN

Penelitian menunjukkan bahwa metode yang diusulkan dapat dilakukan untuk menghasilkan sistem pengenalan objek manusia yang akurat dan mampu memberikan informasi nilai jarak antara objek dan kamera saat posisi objek diam dengan kesalahan pengukuran sebesar 4%. Sebaiknya dalam penelitian selanjutnya menggunakan prosesor dengan kecepatan tinggi agar tidak perlu memotong video secara manual yang mengakibatkan kesalahan dalam memasang *frame* video. Diperlukan pengembangan metode *stereovision* agar dapat mengukur jarak dalam semua kondisi, baik objek bergerak maupun diam. Selain itu diperlukan pengembangan penelitian pengukuran jarak menggunakan *stereovision* secara real time agar metode ini dapat segera diaplikasikan dalam bidang robotik, *autocar* ataupun industri.

UCAPAN TERIMA KASIH

Penulis mengucapkan terimakasih kepada dosen pembimbing dan penguji yaitu Dr. Kusworo Adi, M.T., Dr. Catur Edi Widodo, M.T. dan Heri Sugito, M.Sc. atas segala dukungan dan bimbingannya. Penulis juga mengucapkan terimakasih kepada seluruh pihak di laboratorium elektronika dan Instrumentasi atas dukungan sarana prasana selama proses penelitian berlangsung.

DAFTAR PUSTAKA

- [1] Hou, A. L., Cui, X., Geng, Y., Yuan, J. dan Hou, J. (2011) *Measurement of Safe*

- Driving Distance based on Stereo Vision*, Sixth International Conference on Image and Graphics (ICIG), Hefei, Anhui, China, Hal. 902-907.
- [2] Sukhla, N. (2015) *A Review On Image Based Target Distance dan Height Estimation Technique Using Laser Pointer And Single Video Camera For Robot*, International Journal Of Engineering Research And Reviews, Vol. 3(1), Hal.128-135.
- [3] Nagar, S. dan Verma, J. (2015) *Distance Measurement Using Stereo Vision*, International Journal Of Electrical And Electronics Engineers, Vol. 07(01).
- [4] Jain, A. K. (1986) *Fundamental of Digital Image Processing*, Prentice Hall, inc, United State of America.
- [5] Kadir, A. dan Susanto, A. (2013) *Teori dan Aplikasi Pengolahan Citra*, Penerbit ANDI, Yogyakarta.
- [6] Daolei, W.dan Lim, K. B. (2011) *Obtaining Depth Map From Segment-Based Stereo Matching Using Graph Cuts*, Journal of Visual Communication and Image Representation, Vol. 22(4), Hal. 325–331.
- [7] Thaher, R. H.dan Hussein, Z. K. (2014) *Stereo Vision Distance Estimation Employing SAD with Canny Edge Detector*, International Journal of Computer Applications Vol. 07(3), Hal. 38–43.
- [8] Raajan, N. R., Ramkumar, M., Monisha, B., Jaiseeli, C. dan Venkatesan, S.P. (2012) *Disparity Estimation from Stereo Images*, Procedia Engineering. Vol. 38, Hal. 462-472.
- [9] Trucco, E. dan Verri, A. (1998) *Introduction Techniques for 3-D Computer Vision*. Prentice Hall, USA.
- [10] Dalal, N.dan Triggs, B. (2005) *Histograms of Oriented Gradients for Human Detection*. IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), Vol. 1, Hal. 886-893.
- [11] Wurdianarto, S. R. (2014) *Perbandingan Euclidean Distance dengan Canberra Distance pada Face Recognition*, Techno.Com, Vol. 13 No.1, Hal. 31-37.